Fuzzy Rule based Novel Approach to Spam Filtering

G.Santhi
CRD
Prist University
Thanjavur, India

S.Maria Wenisch
Department of IST,
Anna University
Chennai, India.

Dr. P. Sengutuvan
VMKV Engineering College
Salem,
India.

ABSTRACT
The spam mails are used by spammer which amounts to be a headache to the internet users and organizations using online. Rapid growth rate of the use of the internet has dramatically increased the spam mails. More methods are adopted for filtering spam. This approach is to identify the spam mails using spam word ranking and fuzzy rules. This work classifies the emails with the help of word ranking database and sender's mail address database. And the ranks are used based on the degree of the threat that each word possess. For this purpose the work has considered the subject and content of the email. In addition this effort includes the sender's mail address feature for the classification of the spam mail.

Keywords
Email, Spam Word Ranking, Spam Classification, Fuzzy Rule, Fuzzy Inference, Linguistic Variables.

1 INTRODUCTION
Today the internet world is facing a rapid growing threat of spam. Email is an invaluable tool for communication for the society. The word spam denotes nothing but unsolicited commercial email. Spammer has a practice of sending unwanted identical email messages in bulk to the internet users. Spam mail has become a very serious issue among the internet users when the internet was opened to the public in the mid-1990s. Spam mails are used with profit motive to promote a product or service. Currently many spam prevention techniques are used by the internet service provider at the server level. Even though these filters are efficiently function spam mails are flooded in the user's inbox. This paper describes the research that aims to classify spam mails in the inbox level. U.S.A is leading spam relaying country with 18.3%. Figure1 reveals the spam rate of top twelve countries. International telecommunication union report reveals that the internet will be used by 39% of the world population by 2013.

Email communication is a powerful communication channel used for information interchange. This communication is seriously affected by the spam mails. It is inevitable to develop new methods to efficiently filter spam. Spammers are used three techniques to harvest the email address. They are collecting the users email address by using spam bots, (a program) by performing Dictionary attack, and by purchasing address lists from organizations or individuals. Spammers are using a technique automated bulk mailers program to send the spam mails. This program consumes small amount of time to send large volume of mails. Spammers can also use zombie networks to send spam mails. Spammers could hide the sender mail address. Spammers used web spiders (a program) to find email addresses from web pages. Spammers have also get email addresses directly from Google search results.

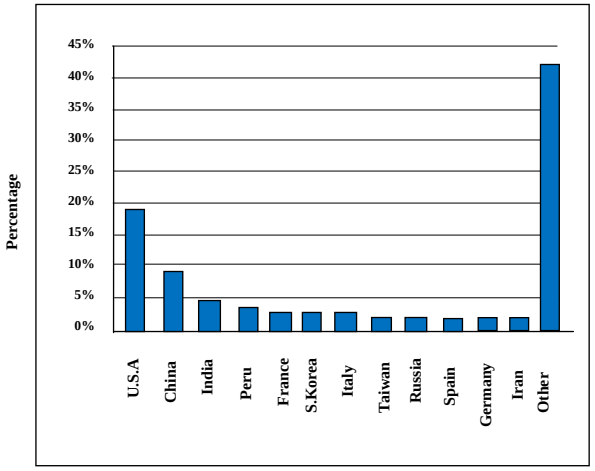

Fig 1: Spam Rate (December 2012 to Feb. 2013)
(Courtesy Sophos Lab)

There are many types of spam techniques used to send spam mails. They are email appending, image spam, blank spam and backscatter spam. Email appending or e-appending is a marketing practice in which they use known customer address matching it against a vendor's database to obtain email addresses. The blank spam may be originated either intentionally or unintentionally. Obfuscation method is used in an image spam in which the text of the message is stored as Gif or jpeg format and displayed in the mail. To avoid OCR tool detection the spammers used new technique to contort the shape of the letters in the image. This method prevents the spam detection from text based filters. Backscatter spam is an incorrect automated bounce messages sent by mail servers, typically as a side effect of incoming spam.

The images and spam written in html in the mails which contains an embedded object called web bug. This web bug helps the spammer to identify the valid mail address of the user and IP address of the computer. It is difficult for the user to detect the object embedded in the image. Every internet user's inbox is flooded with spam and therefore they should spend time to read and delete the spam mails. This spam mails causes waste of time of user and the organization, steal the bandwidth of network and causes financial loss to the user. There are many anti spam techniques are adopted by the end users and administrators of email systems to filter spam. Many spam mails are being sent to undisclosed recipient address. Spammers are used to send attachments with virus.

Naive Bayesian classification, Support Vector Machines for text categorization, K-NNC for classifying nearest neighbor test pattern, and other clustering methods are employed for filtering spam mails. Spammers are increasingly developing innovative techniques to send their spam mails. Some organizations and many researchers have tried to filter spam mails by applying various methods at different levels. Nowadays spam filtering is necessary to protect the internet users which is quite challenging. Spam filter is a program or software used to filter spam mails.

There are several anti spam algorithms available for spam filtering. But spammers try to break the anti spam filters by applying obfuscation and various Techniques. Shalini puri et al (2012) proposed a study on different fuzzy similarity related algorithms and methodologies. They concluded that by using fuzzy logic and fuzzy sets which results good effect on text mining and text classification and paved a better way for text categorization.

P.Divya et al (2012) proposed a method to extract email abstractions from HTML content by using SAG procedure (Structure Abstraction Generation). Sp trees and sp table are used to store reported spam abstractions from emails. User registration, mail composer, spam detection module, block list, and report module are the five modules used to find and block spam. Every user must register in the domain and send mail to register user only. If they received any spam mail that is reported to the administrator then the administrator will block the spam and also reported to the database. This system works only for registered users but it would not work for whole. Christina et al (2010) proposed a study on email spam filtering techniques. They discussed about various problems aroused by spam, different filtering methods and techniques are used to filter spam.

Here this proposed work applied fuzzy logic to classify spam. This work used a fuzzy inference system to classify spam words in an email. This work has a list of spam words and spam mail addresses in the database. This method extracts features from the subject, content, and sender address fields of the emails which, this work compares them against a list of spam words and spam mail addresses stored in the database ranked with its values and categorize the words and addresses in accordance to the ranking. Fuzzy inference system finally classified the values of input as least dangerous or moderate or most dangerous spam mail.

The rest of this paper is organized as follows: Section 2 explains the related work. Section 3 describes the proposed work. Section 4 explains the implementation of this work Section 5 discusses the expected result. Finally Section 6 concludes the Paper.

2 RELATED WORKS

MD. Rafiqul Islam et al [1] discussed about different machine learning algorithms for spam filtering and presented a comparative study of spam filters. Their research includes a study of automated filtering and machine learning techniques like rule based, content based, personalized, collaborative, support vector machine and kernel based algorithms for filtering spam. They presented a comparative analysis on different filtering techniques and its advantages

Ni Zhang et al [2] developed a method for filtering spam mails from the Internet service providers in its heavy traffic. Finger print method is used to detect the similar earlier mails and sets a parameter for the email category. Mail database and finger print database are used to store information. By simply adding the entry in the MD and delete the unimportant mails. They explained about the three advantages of BMTC. They are automatic hand-free deployment and online update mechanism, high accuracy in identifying emails, and handling a large amount of data with small memory and reasonable CPU time.

Seongwook Youn et al [3] proposed a comparative study for Email classification. Neural Network, SVM, Naive Bayesian and J48 classifiers are used to filter spam from the datasets of Emails. J48 is a decision tree creates a binary tree used for classification of legitimate and spam. They suggested J48 and NB classifiers obtained a better result and accuracy than SVM and NN classifiers.

Ali Cıltık et al [4] proposed a method of spam e-mail filtering methods with high accuracies and low time complexities. They took Turkish mails for their research. They used PC-KIMMO system, a morphological analyzer to extract root forms of words as input and produce parse of words as output. This method is based on the n-gram approach and a heuristics. They developed two models, a class general model and an e-mail specific model. The general model classifies the mail as spam or legitimate by using bayes rule. The second model determines the correct class of a message by comparing it with the similar previous message for matching. The third model is a combined perception refined model. It is a combination of above two models. Free word order is used for ordering the word in fixed order for n gram model. This spam filtering method is based on classifying text contents and raw contents of emails obtaining results from the categorization of data sets. They faced the increase of time complexity problem when handling the larger number of words. Adaboost ensemble algorithm is used to compare with its previous work. They performed extensive tests on various number datasets sizes and initial words. They have obtained a result of high success rates in both Turkish language and English.

A.G. López-Herrera et al [5] developed a multiobjective evolutionary algorithm for filtering spam. They evaluated the concepts of dominance and pareto–set. SPAM-NSGA-II-GP is used for filtering spam mails. MOEA is used to learn a set of queries with good precision and recall. PUI datasets are used for spam filtering. SPAM-NSGA-II-GP with very strong filtering rules are (high recall and low precision) used to block all the legitimate emails and labeled as spam. They used the weak filtering rules (high precision and low recall) for labeling a minimum portion of spam emails.

Liu Pei-yu et al [6] suggested the method of improved bayesian algorithm for filtering spam. KNN algorithm, SVM, decision tree, and improved Bayesian algorithm are used for classifying texts. KNN algorithm is a simple and accurate method for spam filtering by using the k nearest neighbor. SVM is also used for filtering spam and finds hyper plane to classify the legitimate and spam mails. It works with smaller training set. Decision tree is used for faster and simple classification it gives higher accuracy of judgment. Bayesian algorithm is a base and simple classification method classifies the mail as C_{legal} and spam $C_{rubbish}$. In the bayesian method one feature is treated as independent of other. Improved naive bayesian algorithm is a combination of bayesian algorithm with boosting method, developed to reduce the rate of misjudgment and improve the accuracy of classification. Boosting is a universal learning algorithm. They treated the naive bayesian algorithm as weaker learning algorithm and make stronger by boosting it with boosting algorithm. They obtained better result by applying this boosted naive Bayesian algorithm for filtering spam.

Alaa El-Halees et al [7] developed to filter spam messages in mixed Arabic and English. Six classifiers are used for filtering spam messages and compared the results obtained from these classifiers. Maximum entropy, decision trees, artificial neural sets, naive bayes, support system machines and k- nearest neighbor are used for spam filtering. Recall and precision are the two ways for presenting the system performance. SVM is used as the best classifier for English and ME performed better than NB in Arabic messages. They suggested increasing the parameter will improve the performance.

Jan Gobel et al [8] proposed a method to filter spam mail in a proactive way which intercepts the communication held between spambot and the intended server and redirects its communication with local mail server at the gateway. They collect spam messages at the gateway and obtain the current spam messages sent by spambotnet. They clean the machine system using a software based restoring system to execute next spambot. They have collected the spam messages by resetting the honeypot. The next process is filtering the message in a proactive way. Longest common string algorithm is used for the extraction of emails and single raw template. They generate templates taking subject, xmailer and complete body of the message for consideration. Longer emails are processed first as they took the first email from the list sorted based on text length and considered it as first raw template named α. They took second email from the list and merge it with α to form a second raw template named β which was more specific than previous one. Then they compared both α and β and determine the amount of text that was replaced by the placeholders. If the removed text percentage is below a predefined threshold Ø then they are treated β as their new α. The email used to form β was removed from the list and they continued with third step. If the changed text percentage is above Ø the current β is too generic and is therefore discarded. They used template generation process for detecting the spam rate.

M. Basavaraju et al [9] proposed the text based clustering method for spam detection. Preprocessing of data, methodology of classification, vector space model, and data reduction are the methodologies used for spam filtering. The Porters stemming and stopping algorithm are used for preprocessing of data. Hierarchical and partition clustering algorithms are used for partitioning and clustering. They used BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) method to clustering the documents. NNC and K-NNC are the two classifiers used to classify the neighbors. K-NNC classifier is used to classify the patterns. The vector space model is used to calculate the inverse document frequency of each word i.e. tf-idf test patterns. After clustering of training patterns the non-spam data are stored in the centroids. Test patterns and centroids are passed in to the classification module for spam and non-spam detection.

Alireza Nemaney Pour et al [10] proposed for Minimizing the time of spam detection by relocating the filter to the sender messages. They used DSPAM and TREC anti-spam software for filtering spam. They used four steps to detect spam. The first two steps are used to check the IP validity of both the sender and receiver. The Sender IP validity is checked by the mail server and receiver IP validity is checked by the DNS server respectively. The third step checks the category of mail belongs to white list or black list or grey list. In the Final step they applied rule and content based filters for detection of spam. This research helps them to preserve the network resources such as bandwidth, time and memory and also minimize the time.

Dhananjay et al [11] developed an adaptive neural fuzzy inference system classifier which includes both the neural networking and fuzzy logic concept to detect the spam message on social networking websites. ANFIS classifier is used to identify the spam from input vector. They identified five input vector parameters. Number of associated user pages, number of times marked as spam, text priority, presence of URL or Hyperlink and the number of common timestamps are the parameters used to classify the spam. They developed the fuzzy inference system with three parts. They are input member function, output member function and the rule set linking the two member functions. Input member function has five parameters like the number of associated user pages, number of times marked as spam by user, presence of Hyperlink or URL, the number of instances of common timestamps and the priority of text in the message. Seven fuzzy rules are used and the output member functions produce a result which equals the rules in the fuzzy rule set. The researcher suggested increasing the parameters would decrease the false positives and improve the detection rate of spam. This system is not designed for a particular specialized social networking website.

Sudhakar.P et al [12] developed fuzzy logic concept for spam detection. They applied five fuzzy rules on five fuzzy parameters. The 5 fuzzy parameters are sender address, sender IP, subject words, content words, and attachments. All the five parameters are compared against the black list and white list. If match was found they considered the parameter as spam or ham. This approach consumes large amount of time to identify spam words.

Subhodini gupta et al [13] suggested a fuzzy filtration module for spam detection. They developed two modules. The first module applied stemming, stop-word elimination and tokenization process on the extracted email words. Fuzzy rules are applied on the document set to verify it for spam or ham. In this method five fuzzy parameters are used. They are sender address, sender IP, subject words, content words and attachments. These extracted parameters are passed through the fuzzy rules for detecting spam. This method applied only for plain text used in subject and body content.

Dr. Sonia et al [14] developed a vector space model to classify the mail. It converts the mail into matrix and inverse frequency is calculated. They calculated the similarity coefficient by using term frequency and inverse message frequency. The fuzzy decision maker used to take the sc as input for fuzzification. Fuzzification classifies the input as legitimate or spam mail.

M. Muztaba Fuad et al [15] proposed a method of trainable fuzzy filters for filtering spam mails automatically. A trainable fuzzy classification module consists of a set of fuzzy rules and fuzzy inference system used for the classification of spam. The messages from the corpus parsed and features are extracted. It extracts the features from, to, cc, subject, and header fields. In the fuzzification five fuzzy sets are used in which two sets are used for feature extraction from the header part and others for the body features. In the fuzzification it determines the degree of input that belongs to which fuzzy set. Then the rule antecedents are evaluated by fuzzy AND operation and the consequences are combined by OR operation and passed the output for defuzzification process. It will produce a crisp output and it is compared with threshold value and predicts the output value as a spam or ham. They suggested that this method can eliminate a large amount of spam from inbox of the user.

Mehdi Samiei yeganeh et al [16] developed a model for fuzzy logic based machine learning approach for filtering spam. They discussed the methods of automatic spam filters like naive bayes classifier, artificial immune classifier and fuzzy logic. They have enhanced the functionality of the model and also enhanced the feature identification of emails and deletion of spam mails on its own. They suggested that the fuzzy logic is adaptable for spammer tactics.

Begol et al [17] proposed a fuzzy system method to detect the edges of image. Two types of pixel vicinities used for edge detection. They are four and eight pixel vicinities. In an image grey pixels are treated as noise. The noises are omitted from the image. Canny and sobel method did not detect the edges properly. But fuzzy technique is used to detect the original image edges correctly and eliminates the noises in the image.

3 PROPOSED FRAME WORK

The paper's related work consist of methods like naive bayesian, K-NNC, SVM, J48, and fuzzy logic filtering methods to detect the emails as spam or ham. The proposed work relies on fuzzy rule based filtering approach which includes fuzzy inference system with fuzzy rules for classifying spam mails. This work focused on classifying the spam words and spam mail addresses by applying fuzzy rule. The spam words and spam mail addresses are assigned different values by using fuzzy rules. Collection of spam words list from the paper's related works, spam words list spam mail addresses are available in the website and spam mails in the inbox. This assigned value helps us to rank the spam features as shown in figure 2 and finally fuzzy inference system classifies the input rank values and produces the output.

gets attracted by these trapped words and contacts the sender immediately and gets deceived. So the words are treated as very strong spam words. The database contains ID, spam words, spammer mail address and rank values for both spam words and spam mail address fields. The actual features are extracted from the user inbox and compared against the spam list in the database.

The actual words and Actual sender mail address are matched against the list of database of spam words and spam mail addresses. If it is matched then the rank value assigned to the detected spam words and mail address in the list. Fuzzy inference system takes the ranked value as input for classifying the spam mails. In order to classify the spam fuzzy inference system has been designed which take the ranked input value and produce the output. Figure 4 explains the classification of spam procedure. The output data are classified in to three linguistic variables i.e. Least dangerous, Moderate dangerous and Most dangerous.

Step1: It is used to read the email from the inbox of user

Step2: The features are extracted from email

Step3: It is used to count the number of spam words

Step4: The features are compared against a list of ranked spam words and spam mail address in the database.

Step5: The spam words and spam mail address are classified accordance to the rank.

Step6: Input the rank values to the input and produce the result.

The mamdani fuzzy inference model is followed in the implementation. This model's simplicity helps anyone to understand the concept easily. Figure 5 shows the process of ranking and classification of spam words.

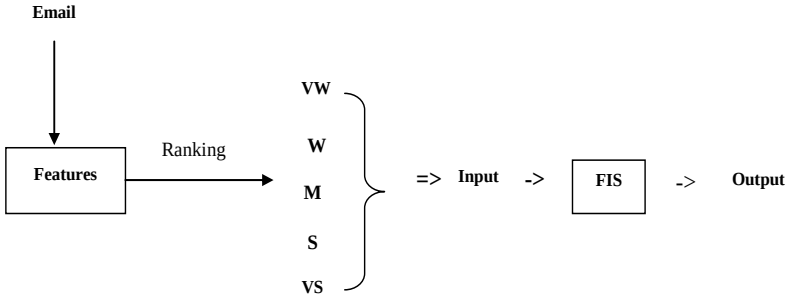

Fig.2 Concept of classification of spam words

3.1 Fuzzy Classification Module

The fuzzy classification contains fuzzy inference system and a set of rules. The proposed fuzzy rule based classification shown in figure 3.This work has a list spam words and spam mail addresses are in the database with its ranked value. The spam words and sender mail address are extracted from the subject, content and from address fields of an email. The spam words and spam mail address are assigned a value and categorized in to five i.e. weak (W), very weak (VW), moderate (M), strong (S), and very strong (VS). Email contains many spam words. This work extracted spam words and sender address of spammers from 100 mails. Attention, Dear Lucky Winner, Information, Job Opportunities, Notification, Congratulations, and Business Partnership are the few most attracted words used by the spammers in the subject and content fields to cheat the users. The internet user

- **Ranking of Words**

If spam word <=0.9 And > =0.7

It is a Very Strong spam word

Elseif spam word < 0.7 And >=0.5

It is a Strong spam word

Elseif spam word < 0.5 And >=0.4

It is a Moderate spam word

Elseif spam word <0.4 And >=0.2

It is a Weak spam word

Else Very Weak spam word

- **Classification of Spam emails**

If spam words <=0.9And >=0.5

It is most dangerous spam mail

If spam words <0.5And >=0.4

It is moderate dangerous spam mail

If spam words <0.4And >=0

It is least dangerous spam mail

X→ spam word

0<= X<0.2 → Very weak spam word

0.2<=X<0.4 →Weak spam word

0.4<=X<0.5 → Moderate spam word

0.5<=X<0.7 → Strong spam word

0.7<=X<=0.9 → Very strong spam word

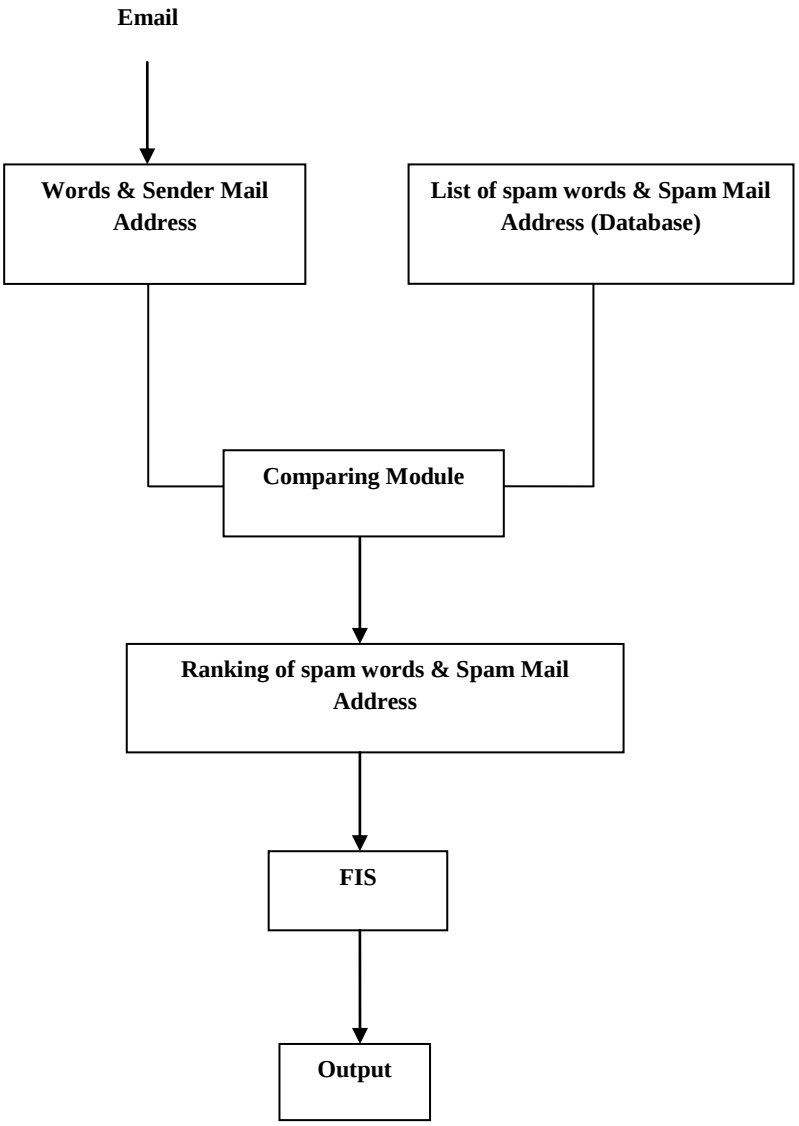

Fig. 3 Spam classification model

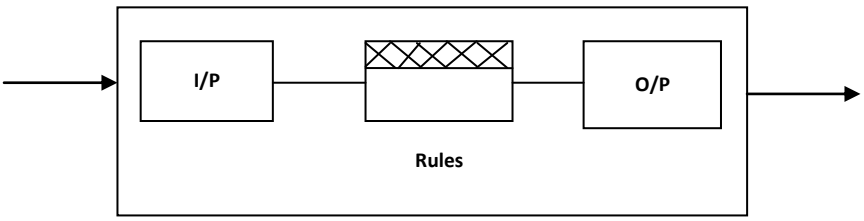

Fig. 4 Fuzzy rule based classification

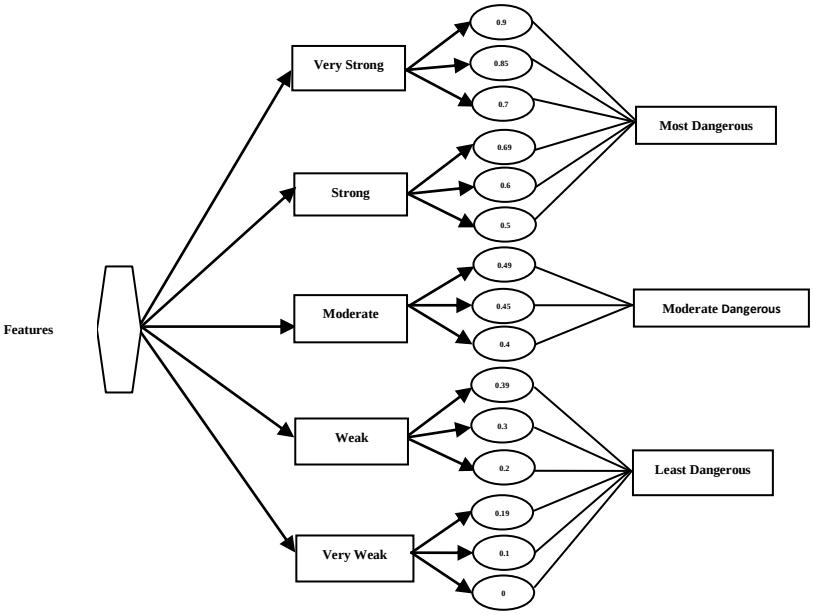

Fig.5 Ranking and classification of spam

4 IMPLEMENTATION

For implementing this work Vb.Net is used for ranking and matlab for classification of spam mails. This work is compared the spam words and spam addresses with the available database and produced the ranked value of spam words and spam mail address for the classification of spam mail. Fuzzy inference system is used to map input space to an output space by using fuzzy logic. Matlab is used for the fuzzy classification of spam mails. In matlab Fuzzy inference process comprises of five parts. The process begins with

1. Fuzzification of input variables
2. Application of the Fuzzy operator (AND, OR,) in the antecedents.
3. Implication from the antecedent to consequent
4. Aggregation of the consequents across the rules, and
5. Defuzzification.

Fuzzy inference system takes the ranked values as input. The rules are generated by this work and are evaluated in parallel by fuzzy reasoning. The results of the rules are combined and distilled in defuzzification process produces crispy output i.e. linguistic variables.

4.1 Structure of Fuzzy Inference System

Name : 'three7'

Type : 'mamdani'

AndMethod : 'min'

OrMethod : 'max'

DefuzzMethod : 'centroid'

ImpMethod : 'min'

AggMethod : 'max'

Input : [1x3 struct]

Output : [1x1 struct]

Rule : [1x32 struct]

4.2 Input Membership Functions

Input Membership functions of three FIS variable and output membership function are shown in the following figures 6, 7, 8 and 9. This work has used triangular membership function for both the inputs and outputs.

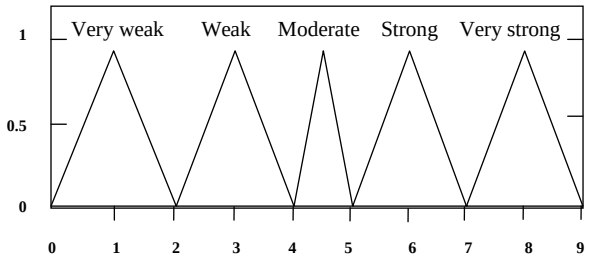

Fig.6 Input membership function Sender address

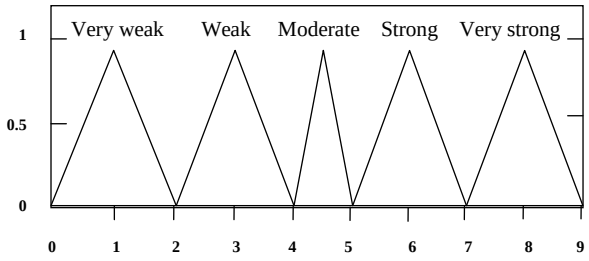

Fig.7 Input membership function – Subject

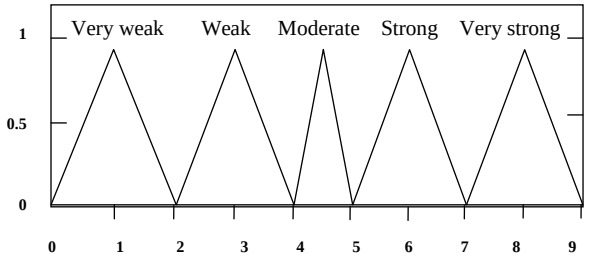

Fig.8 Input membership function – Content

4.3 Output Membership Functions

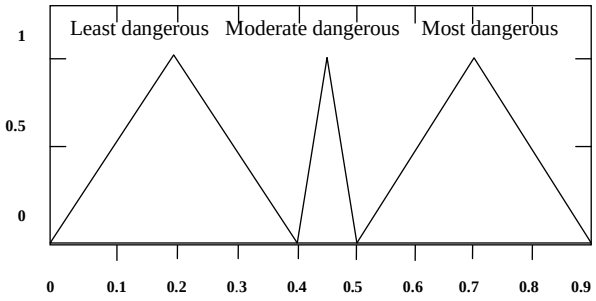

Fig.9 Output membership function

4.4 Rule Viewer

In the rule viewer the first three columns of plots shows the membership functions referenced by the antecedent or if part of each rule. The last column of plot shows the membership function referenced by the consequent or then part of each rule. This rule viewer interprets the entire fuzzy inference process of classification of spam. The thick line passing through aggregate fuzzy set shows the defuzzified output value. Figure 10 shows the output of this method.

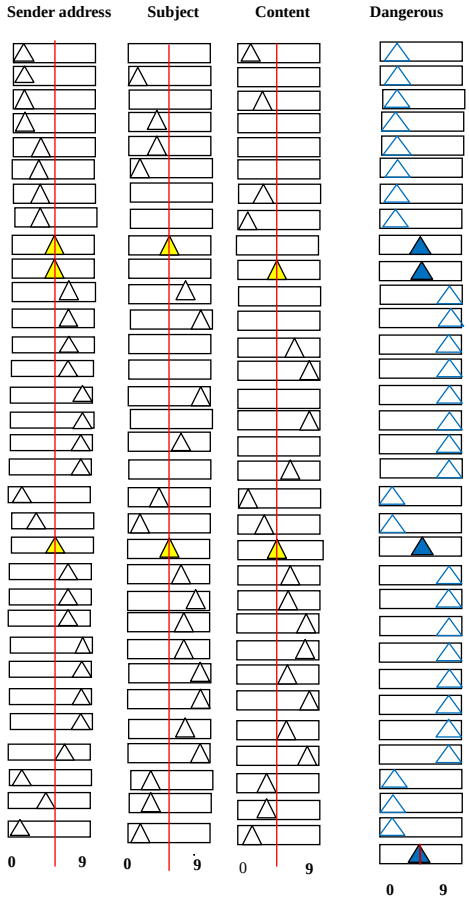

Fig.10 Rule viewer

4.5 Surface Viewer

Three dimensional Surface viewer helps in plotting of two inputs and one output. Figure 11 shows the sender address, subject inputs and an output. Figure 12 shows the plot of sender address, content inputs and an output. Surface viewer displays the entire output set based on the input set.

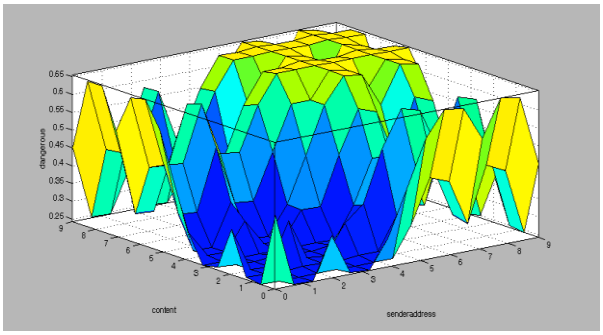

Fig.11 Surface Viewer

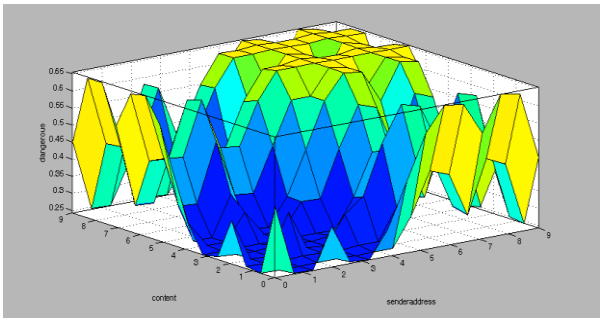

Fig.12 Surface Viewer

5 RESULTS

This work used 100 emails for experiment. This method extracted the words and sender mail addresses and used fuzzy inference system with fuzzy rules for classification of spam mails. It is a novel approach used for classification of spam. This approach helps the end-users to identify the spam mails by using the linguistic variables i.e., least dangerous, moderate dangerous, and most dangerous. The user can easily distinguish the spam mails and delete the spam mails in the inbox level. Figure.13 &14 displays the output of this work.

Output

0.1867

0.1871

0.1880

0.1865

0.1868

0.1870

0.1869

0.1865

0.1865

0.1880

Plot T

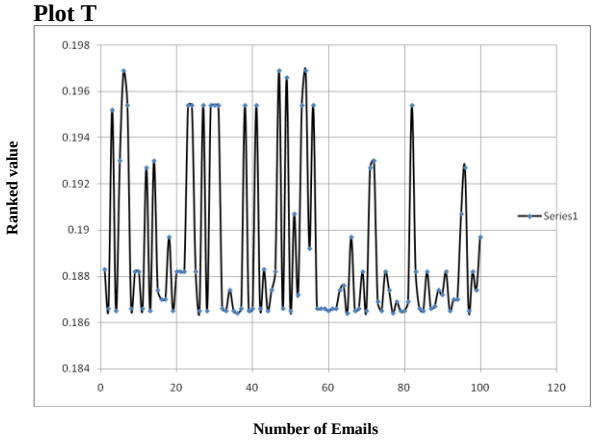

Fig.13 Classified spam emails

Plot Y

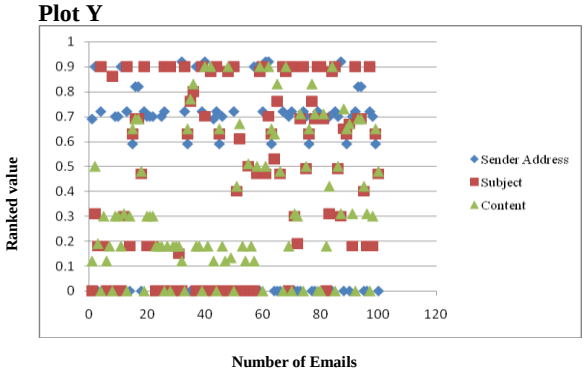

Fig.14 Ranked emails

6 CONCLUSION AND FUTURE WORK

This paper presented a rule based classification of spam mails with fuzzy word ranking. There are many classifiers and filters available for classifying and filtering spam mails. This study analyzed the previous related works. The proposed work used five input variables and three output variables for ranking and classifying spam mails. This method has extracted only the features from the sender mail address and subject and content of an email instead of extracting all the features from the mail.

Comparing the actual words and sender mail address with spam words and spammer mail address helps to detect spam mails. This work categorized the words and sender's mail addresses in accordance to its rank. The input value passed to the fuzzy inference system. FIS classifies the spam and produce the output. This work obtained a better result from ranking and classifying of spam words. The future work aims at classification of spam words in the attachments.

7 REFERENCES

[1] MD. Rafiqul Islam and Morshed U. Chowdhury, 2005, Spam Filtering Using ML Algorithms, IADIS International Conference on WWW/Internet, 419-426.

[2] Ni Zhang, Yu Jiang, Binxing Fang, Xueqi Cheng and Li Guo, 2006, Traffic Classification-Based Spam Filter, IEEE International Conference on Communications, Vol.5, *pp.* 2130 – 2135.

[3] Youn, Seongwook, and Dennis McLeod, 2007, A Comparative Study for Email Classification, Editor. Khaled Elleithy, Advances and Innovations in Systems, Computing Sciences and Software Engineering, 387-391.

[4] Çıltık, Ali, and Tunga Güngör, 2008, Time-Efficient Spam E-mail Filtering Using n-gram Models, Pattern Recognition Letters, 19–33.

[5] López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F., 2008, A Multiobjective Evolutionary Algorithm for Spam E- Mail Filtering, Proceedings of 2008 3rd International Conference on Intelligent System and Knowledge Engineering, 366-371.

[6] Liu Pei-yu, Zhang Li-wei and Zhu Zhen-fang, "Research on Email Filtering Based on Improved Bayesian", Journal of Computers, Vol. 4, No. 3, March 2009, pp. 271-275.

[7] El-Halees, Alaa, 2009, Filtering Spam E-Mail from Mixed Arabic and English Messages: A Comparison of Machine Learning Techniques, The International Arab Journal of Information Technology, 52-59.

[8] Göbel, Jan, Thorsten Holz, and Philipp Trinius, 2009, Towards Proactive Spam Filtering, editors. Ulrich legel and Danilo Bruschi, in Detection of Intrusions and Malware and Vulnerability Assessment, 6th international conference, DIMVA, proceedings, 38-47.

[9] M. Basavaraju and Dr. R. Prabhakar, 2010, A Novel Method of Spam Mail Detection using Text Based Clustering Approach, IJCA,15-25.

[10] Alireza Nemaney Pour, Raheleh Kholghi and Soheil Behnam Roudsar, 2012, Minimizing the Time of Spam Mail Detection by Relocating Filtering System to the Sender Mail Server, International Journal of Network Security & Its Applications, 53-62.

[11] Dhananjay Kalbande, Harsh Panchal, Nisha Swaminathan and Preeti Ramaraj, 2012, ANFIS Based Spam Filtering Model for Social Networking Websites, IJCA, 32-36.

[12] Sudhakar, P., G. Poonkuzhali, K. Thiagarajan, R. Kripa Keshav, and K. Sarukesi, 2011, Fuzzy Logic for E-mail Spam Deduction, In Proceedings of the 10th WSEAS International Conference on Applied Computer and Applied Computational Science, 83-88.

[13] Subhodini gupta, Parekh .B.S and Jaimine N.Undavia, 2012, A Fuzzy Approach for Spam Mail Detection Integrated with Wordnet Hypernyms Key term Extraction, IJERT,1-5.

[14] Dr. Sonia, 2010, Spam Filter: VSM based Intelligent Fuzzy Decision Maker, International Journal of Computer Science and Technology, 48-52.

[15] M.Muztaba Fuad, Debzani Deb and M. Shahriar Hossain, 2004, A Trainable Fuzzy Spam Detection System, In Proc. of the 7th International Conference on Computer and Information Technology.

[16] Mehdi Samiei yeganeh, Li Bin and G. Praveen Babu, 2012, A Model for Fuzzy Logic Based Machine Learning Approach for Spam Filtering, IOSR Journal of Computer Engineering , 07-10.

[17] Begol, Moslem, Maghooli and Keivan, 2011, Improving Digital Image Edge Detection by Fuzzy Systems, World Academy of Science, Engineering and Technology,76-79.