Fuzzy Logic for E-Mail Spam Deduction

P.SUDHAKAR¹, G.POONKUZHALI², K.THIAGARAJAN³, R.KRIPA KESHAV⁴, K.SARUKESI⁵

¹Vernalis systems Pvt Ltd, Chennai- 600116
²,⁴ Department of Computer Science and Engineering, Rajalakshmi Engineering College, Affiliated to Anna University- Chennai, Tamil Nadu
³ Department of Science and Humanities, KCG College of Technology Affiliated to Anna University-Chennai, Tamil Nadu
⁵ Hindustan Institute of Technology and Science-Chennai,Tamil Nadu INDIA

¹ sudhakar.asp@gmail.com, ²poonkuzhalis@rajalakshmi.edu.in,
³ vidhyamannan@yahoo.com, ⁴kripa_keshav@yahoo.co.in, ⁵profsaru@gmail.com

Abstract: - In this Information era, most of the communication is happening through e-mails. Many e-mails contain web spam, as the transaction through this internet is affected by Passive attacks and Active attacks. Recently, several algorithms are developed partially for detecting spam emails. Therefore there is a need for improving the performance of the spam-detecting algorithm. In this proposed work fuzzy rules are defined and applied to all incoming emails. Based on the result of the various rules against user attitude, input email is classified as spam or ham. This method outperforms the existing spam-detecting algorithms in terms of accuracy and user friendly.

Key-Words: - Attitude, Email, Email spam, Fuzzy, Fuzzy logic, Spam, Spam deduction

1 Introduction

E-mail spam, also known as junk e-mail or unsolicited bulk e-mail (UBE), is a subset of spam that involves nearly identical messages sent to numerous recipients by e-mail. Definitions of spam usually include the aspects that e-mail is unsolicited and sent in bulk. One subset of UBE is UCE (unsolicited commercial e-mail). E-mail spam has steadily grown since the early 1990s. Botnets, networks of virus-infected computers, are used to send about 80% of spam. Since the cost of the spam is borne mostly by the recipient, it is effectively postage due advertising.

The legal status of spam varies from one jurisdiction to another. Spammers collect e-mail addresses from chatrooms, websites, customer lists, newsgroups, and viruses which harvest users' address books, and are sold to other spammers. They also use a practice known as "e-mail appending" or "e-pending" in which they use known information about their target (such as a postal address) to search for the target's e-mail address. Much of spam is sent to invalid e-mail addresses. Spam averages 78% of all e-mail sent. According to the Message Anti-Abuse Working Group, the amount of spam email was between 88-92% of email messages sent in the first half of 2010. Most of the inbox is flooded with these Spams which occupies lot of memory space. There are several algorithms available for detecting and filtering spam emails. Among the existing algorithms, Bayesian filtering produces best result, still it does not detect all the spam emails. Most of the existing algorithms considers content alone for filtering the spam emails. To detect all the spam emails, existing spam filtering methods has to be enhanced. In this proposed work, a new algorithm is devised with various fuzzy rules and fuzzy variables. Each fuzzy rule will produce Attack Factor value which are consider for arriving result. Each rule Attack Factor value was arrived by comparing input parameter against Black list and White list. Black list contains predetermined spam content. White list contains acceptable contents. This final result from above calculated Attack Factor will decide the input email content to be spam or ham or to be sent hold state. The final result of the algorithm was obtained by summing up each rule result value and decision was taken based on the result of the individual rules.
1.1 Related works
Xavier Carreras et al.[1] proposed a Boosting algorithm for Anti Spam filtering. Even though Boosting algorithm delivers good result, possibility of misclassification costs persist inside the AdaBoost learning algorithm.

William W. Cohen et al.[2] suggested Speech act theory for email filtering. The outcome of Speech act theory highly depend on the learning and this approach shows new projection for classifying email spam content.

Harris Drucker et al.[6] developed support vector Machines for Spam Categorization. Even though support vector approach outperforms well, switching from training model need user intervention. Addition to that, reply emails are considered as no spam.


1.2 Outline of the document
Section 2 contains Generation of Fuzzy rules. Section 3 projects Fuzzy rule implementation. Section 4 delivers results and discussion on results. Section 5 propose future work and concludes.

2. Generation of Fuzzy Rules
Input variable : {Sender’s Address, SenderIP, SubjectWords, ContentWords, Attachment}
Fuzzy set : {positive, Zero, Negative}
Linguistic set : (highpositive, lowpositive,positive, highNegative,LowNegative,Negative, Zero}

Rule 1:
   a: IF ∃ SenderAddress ∈ spammer list → 
       AttackFactor=-0.25;
   b: IF ∃ SenderAddress ∈ to Ham list → 
       AttackFactor=0.25;
   c : IF ∃ Sender Address ∈ Spammer list & 
       ∃ Sender address ∈ Ham addresslist 
       →AttackFactor=0;

Explanation:
   Rule 1.a : If there exist a sender address belongs to spammer list, then Attack Factor of this rule should be set to -0.25;
   Rule 1.b : If there exist a sender address belongs to Ham list then, Attack Factor of this rule should be set to 0.25;
   Rule 1.c : If there exist a sender address that doesn’t belong to spammer list and Ham list then, Attack Factor of this rule should be set to 0;

Rule 2:
   a: IF ∃ SenderIP ∈ SpammerIPlist 
       →AttackFactor= -0.25;
   b: IF ∃ SenderIP ∈ HamIPlist 
       →AttackFactor=0.25;
   c: IF ∃ SenderIP ∉ SpammerIPlist & 
       HamIPlist →AttackFactor=0;

Explanation:
   Rule 2.a : If there exists a sender IP address belongs to Spammer list, then Attack Factor of this rule should be set to -0.25;
   Rule 2.b : if there exists a sender IP address belongs to Ham list, then Attack Factor of this rule was set to 0.25;
   Rule 2.c : If there exists a sender IP address doesn’t belongs to Spammer list and Ham List then Attack Factor of this rule was set to 0;

Rule 3:
   a: IF ∀ Subject words ∈ Spam words 
       →AttackFactor= -0.50;
   b: IF ∃ Subjectword ∈Spamwords → -0.50<AttackFactor< 0.50

Explanation:
   Rule 3.a: If all Subject words belongs to Spam words then, Attack Factor of this rule should be set to -0.50;
   Rule 3.b : If there exists a subject word that belongs to spam word then Attack Factor of this rule is varies from -0.50 to +0.50;

Rule 4:
   a: IF ∀ Content words ∈Spamwordlist 
       →AttackFactor= -0.50;
b: IF ∃ Content words € SpammersList → 0.50 < AttackFactor < 0.50;

**Explanation:**
Rule 4.a : If all email content words belongs to Spam words then, Attack Factor of this rule should be set to -0.50;
Rule 4.b : If there exists an email content word that belongs to spam word then Attack Factor of this rule is varies from -0.50 to +0.50;

**Rule 5 :**
a: IF ∀ Attachment € VirusList → AttackFactor = 1.0;
b: IF ∃ Attachment € VirusList → AttackFactor = -1.0;

**Explanation:**
Rule 5.a : If all attachment doesn’t belong to virus list then, Attack Factor of this rule is set to 1.0;
Rule 5.b : If there exist an attachment belongs to virus list, then Attack Factor of this rule is set to -1.0;

### 3. Fuzzy Rule implementation

![Diagram](image.png)

Figure 1. Architecture of proposed system

When an email is arrived, fuzzy input parameters are extracted. Rule 1 was applied on Fuzzy input parameter- Sender address. Based on Rule 1, Sender address was extracted from email and compared against the Black list which has spammer email address list. If any match was found then, Attack Factor for this rule was set to -0.25. If sender address was not found in the black list, then it was compared against the White list which contains all good and acceptable email addresses. If match was found, then attack factor for this rule was set to 0.25. If sender address was not found in both Black and White list, then attack factor for this rule was set to 0. Set this rule result in R1.

Rule 2 was applied on Fuzzy Input parameter – Sender IP. IP Address of the sender was compared against the IP Address Black List. If match was found, then Rule 2 Attack Factor was set to -0.25. If not found, then Sender IP Address was compared against White List IP Address. If match found then attack factor of Rule 2 was set to 0.25. If not found then Attack Factor of the Rule 2 was set to 0. Assign resultant value in R2.

Rule 3 was applied on Fuzzy input parameter – Subject words. An Email may contain one or more words in subject line. All words are taken and compared against the Black list words. Following are Algorithm to compute Attack Factor of Rule 3

**Algorithm for Email Subject Attack Factor Calculation:**
Step 1 : Split the Subject content into words say W_i where n ≥ i ≥ 1
Step 2 : assign to T_w = n
Step 3 : Calculate word Impact Factor W_i where W_i = 0.5 / T_w
Step 4 : Perform comparison for each word W_i in Black list
Step 5 : If match found then update the update W_i = - W_i else W_i = W_i; where i ≤ T_w;
Step 6 : Calculate Attach Factor = ∑ W_i
Step 7 : Calculate R3 = ∑ W_i

Rule 4 was applied on Fuzzy Input variable- ContentWords. Every email body may contain one or more words. Every word are taken and compared against the Block list words. Following are the Algorithm to compute Attack Factor of Rule 4.
Algorithm for Email Content Attack Factor Calculation:

Step 1: Split the email body content to words say \( W_i \) where \( i \geq 1 \)

Step 2: Count the total number of words in email Body and assign to \( T_w \)

Step 3: If \( T_w > 0 \) then continue Step 4.

Step 4: Calculate word impact factor \( W_f \) where

\[
W_f = \frac{0.5}{T_w}
\]

Step 5: Perform comparison for each word \( W_i \) in Black list

Step 6: If match found then update the update \( W_{fi} = -W_i \) else \( W_{fi} = W_i \); where \( i \leq T_w \);

Step 7: Calculate Attach Factor = \( \sum W_{fi} \)

Step 8: Calculate \( R_4 = \sum W_{fi} \);

Rule 5 was applied to calculate Attack Factor for email containing attachment. If email does not contain Attachment, then Attack Factor was set to zero. If any one of the attachment content was identified in virus list then Attack Factor was set to -1. If none of the content was identified in virus list, then Attack Factor was set to 1. Rule 5 result was assigned to \( R_5 \).

Result value of each was arrived by sum up previous rule results.

\[
R_1 = R_1; \\
R_2 = R_1 + R_2; \\
R_3 = R_3 + R_2 + R_1; \\
R_4 = R_4 + R_3 + R_2 + R_1; \\
R_5 = R_5 + R_4 + R_3 + R_2 + R_1;
\]

4. Results and Discussion

All fuzzy rules are applied over fuzzy input variables: Sender’s Address, SenderIP, SubjectWords, ContentWords, Attachment for sample emails namely A, B, C, D and E. Results of the each fuzzy rules are distributed in the table. The results of the email can be obtained by plotting each rule result.

Table 1: Results based on Fuzzy Rules

<table>
<thead>
<tr>
<th>Email</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.25</td>
<td>0.5</td>
<td>1</td>
<td>1.5</td>
<td>2.5</td>
</tr>
<tr>
<td>B</td>
<td>0.25</td>
<td>0</td>
<td>-0.5</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>C</td>
<td>0.25</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>-0.25</td>
<td>-0.5</td>
<td>-1</td>
<td>-1.5</td>
<td>-2.5</td>
</tr>
</tbody>
</table>

The following graph was plotted for tabulated result where the X axis represent the Rules and Y axis predicts the Results. When line move towards the Positive value without any axis interception, then the identified email is Ham. When line starts with negative and lies in the negative region without axis interception, then the email is definitely spam. Whenever the result enters in the negative region, then the email is set to spam. If it lies in the axis, then the email is set to hold state.

![Graphical Representation](image-url)
5. Conclusion and Future Work
In this proposed work, Fuzzy rules are constructed for 5 input parameters namely Sender’s Address, SenderIP, SubjectWords, ContentWords and Attachment for common user to deduct the spam emails. The proposed simplistic approach out performs in terms of accuracy and faster in deduction spam emails than the existing approaches. In future, this approach can be extended based on user attitude. User can tighten or relax their spam levels by moving x axis towards positive / negative region.

Acknowledgment
The authors would like to thank Dr. Ponnammal Natarajan worked as Former Director – Research, Anna University- Chennai, India and currently an Advisor, (Research and Development), Rajalakshmi Engineering College and Dr. K. Ravi, Associate Professor, Department of Mathematics, Sacred Heart College- Tirupattur, India for their intuitive ideas and fruitful discussions with respect to the paper’s contribution.

References:
Engineering at Anna University – Chennai, India. She has presented and published 10 research papers in international conferences & journals and authored 5 books. She is a life member of ISTE (Indian Society for Technical Education) , IAENG (International Association of Engineers), and CSI (Computer Society of India).

K. Thiagarajan working as Senior Lecturer in the Department of Mathematics in KCG College of Technology - Chennai-India. He has totally 14 years of experience in teaching. He has attended and presented research articles in 33 National and International Conferences and published one national journal and 26 international journals. Currently he is working on web mining through automata and set theory. His area of specialization is coloring of graphs and DNA Computing.

R. Kripa Keshav currently undergraduate student of Rajalakshmi Engineering College. He is the Member of computer society of India. He has presented one paper in national conference and won the best paper award and one paper published in international journal.

Dr. K. Sarukesi has a very distinguished career spanning of nearly 40 years. He has a vast teaching experience in various universities in India and abroad. He was awarded a commonwealth scholarship by the association of common wealth universities, London for doing Ph.D in UK. He completed his Ph.D from the University of Warwick – U.K in the year 1982. His area of specializations is Technological Information System. He worked as expert in various foreign universities. He has executed number of consultancy projects. He has been honored and awarded commendations for his work in the field of information technology by the government of TamilNadu. He has published over 40 research papers in international conferences/journals and 40 National Conferences/journals.