ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 9, September 2016

# CONTENT FILTERING USING ARITIFICIAL INTELLIGENCE

**Tanvi Samel, Darshan Vira, Anushka Patil, Durgesh Pandey, Smita Patil**

**Abstract—** ***Today content is consumed at the speed of light. But to reach the required content, we need to scan through other useless content. In order to effectively filter out spam messages, a methodology is put forth in this paper. The aim is to make lives a bit simpler by letting only the important content consume the time of the user. The proposed software using Artificial Intelligence algorithm, Naive Bayes which can think and process over content and thus effectively filter the spam which is not worth the user's time, ultimately giving access to the important content faster. The aim here is to train the software on 25% data to increase the efficiency level. After training, testing phase can be done on rest of the 75% data. Training the software on 25% data will make the software aware about spam words which then will be filtered out in testing phase successfully.***

***Index Terms*—Content, Data, Filter, Spam- Filter, Spam-Detection, Ham, Bayesian.**

## I. INTRODUCTION

The internet has become an inseparable part of our technologically advanced life and e-mail has become a powerful tool for communicating with each other and even sharing information with each other. Along with the growth of the Internet and e-mail, there has been a prodigious growth in spam in recent years. Daily we get hundreds of mails out of which 70% to 80% are spams. The sources of spam can be from any location across the globe where Internet access is available. We get spam emails and messages which takes our time and bandwidth and even affects the system of the user if it contains any virus. Despite the development of anti-spam services and technologies, the number of spam messages continues to increase rapidly. In order to address the growing issue, each organization must analyze the tools available to determine how to effectively counter spam. Tools, such as the corporate e-mail system, e-mail filtering gateways, contracted anti-spam services, and end-user training, provide an important arsenal for any organization. However, users cannot avoid the very serious problem of attempting to deal with tremendous amount of spam, as they become more banal day by day. If there are no anti-spam activities, spam will be pernicious to the network systems, killing employee productivity, stealing bandwidth, and be extant even tomorrow. We aim to implement an all in one system to fight spam. [1] In this paper we are focusing on content filtering using Naïve Bayes.

We have surveyed various papers including multiple spam techniques like multi-neural networks [4], k-nearest neighbor [1], distributed adaptive blacklist [1], genetic algorithm [5] and stacking classifiers [6]. We have even compared Naïve Bayes with artificial neural networks [8] and found Naïve Bayes better as it provides 92% spam detection rate.

The remainder of the paper is organized as follows. In section 2, we have discussed about history of spam and spam filtrering by Naïve Bayes. Section 3 talks about the working of the system we propose. Section 4 discusses various features of our system followed by advantages and applications in section 5 and section 6 respectively. Section 7 is the conclusion that we have reached and section 8 gives the references.

## II. HISTORY

Though the first spam was sent in 1978 it began to be written about spam as a problem in scientific literature only from 1982. One of the first papers where this problem is considered is the Peter J. Denning's article. [3] The first mathematical apparatus applied to spam filtering systems is the Bayes' algorithm, which was used first by Sahami et.al in 1996 and then by other researchers. Bayes' classifier relies on famous Bayes theorem and the first papers about it was made as early as 1960.[3] During more than 40 year history Naive Bayes Classifier (NBC) was used for the solution of various types of tasks: from classification of texts in news agencies till primary diagnosis of diseases in medicine. For the problems where NBC is used there is usually selected presence or absence of words in the text as a characteristic, i.e. the set of characteristics T is a set off all words in documents. [3] Hereby, if the word $t_i$ is present, the weight of characteristics $w_i=1$ otherwise, $w_i=0$. In case of e-mail filters where spam classification is used, there taken into account the area where the word had been met: heading, subject and body of the e-mail. Considering amount of spam messages coming to e-mail boxes it is possible to assume that spammers operate not alone, there are global, organized, virtual social networks of spammers. They attack e-mails of not only users, even whole corporations and countries. [3]

## III. WORKING

### *A*) ***BASIC CONCEPT***

E-mail spam, known as unsolicited bulk Email (UBE), junk mail, or unsolicited commercial email (UCE) [1], is the practice of sending unwanted e-mail messages, frequently with commercial content, in large quantities to an indiscriminate set of recipients. The technical definition of spam is "An electronic message is "spam" if (A) the recipient's personal identity and context are irrelevant because the message is equally applicable

to many other potential recipients; and (B) the recipient has not verifiably granted deliberate, explicit, and still-revocable permission for it to be sent". [1] The major risk that stands in the way of spam filtering is legitimate mails getting rejected sometimes or marked as spam. The risks of not filtering spam are the constant flow of spam clogs networks and adversely impacts user inboxes, but also drain valuable resources such as bandwidth and storage capacity, productivity loss and interfere with the expedient delivery of legitimate emails.[1] The basic advice to tackle spam is, avoid giving your "real" email address to all but close associates, setup different accounts (Google, hotmail etc.) for registering with certain unknown web sites or for communicating with people you do not know, make your contacts aware about exercising caution with email address, do not open junk email.[1]

### *B) WAY OF WORKING*

NAÏVE BAYES TO DEAL WITH SPAM FILTERING: There are various spam filtering techniques which can filter spam with various levels of efficiency namely, rule based filtering, adaptive blacklist, k-nearest neighbor, Bayesian classifier and neural networks. [1] Our paper focuses on the technique of Bayesian Classifier to filter out the spam content effectively. Naive Bayes classifier uses the Bayes' theorem of conditioned probability to recognize an email to be spam or not. [2] Conditioned Probability is given as

$P(c_j|d) = P(d \mid c_j)\, P(c_j) / P(d)$…….(I)

Considering each attribute and class label as a random variable and given a record with attributes (A1, A2,… An), the goal is to predict class C. Specifically, we want to find the value of C that maximizes P (C| A1, A2,…An). The approach taken is to compute the posterior probability P (C| A1, A2,…An) for all values of C using the Bayes theorem. [2]

P (C| A1A2…An) = [P(A1A2….An |C) P(C)] / P(A1A2….An)……..(II)

So you choose the value of C that maximizes P(C| A1,A2,…An). This is equivalent to choosing the value of C that maximizes P(A1, A2, …An | C) P(C). Naïve Bayesian prediction requires each conditional probability be nonzero. Otherwise, the predicted probability will be zero. [2] From (I) AND (II)

$$P(X|C_i) = \prod_{k=1}^{n} P(x_k|C_i)$$

In order to overcome this, we use probability estimation from one of the following:

Original: $P(A_i|C) = N_{ic} / N_c$
Laplace: $P(A_i|C) = (N_{ic}+1) / (N_c+1)$
M-Estimate: $P(A_i|C) = (N_{ic}+mp) / (N_c+m)$

Where, c: number of classes
p: prior probability
m: parameter
The complete process of content classification can be divided into three phases, as shown in FIG. 1:

1. **FIRST PHASE**: Creation of rules which is stored in database as a set of tokens for classifier. Based on these axioms the classifier will classify the content as spam or ham (not spam). The authenticity of these rules will determine how accurate the overall system is.

2. **SECOND PHASE**: Training phase. In this phase the user manually trains the system by providing content which is spam and even the content which is ham, keywords from which get stored in the database as fresh tokens for classifier.

3. **THIRD PHASE**: Using knowledge of tokens, the filter classifies new input. Probability of new word is calculated and updated in the database and it is classified as ham or spam. Boasts of self- learning algorithms and words misclassified, if any can be rectified by the user.

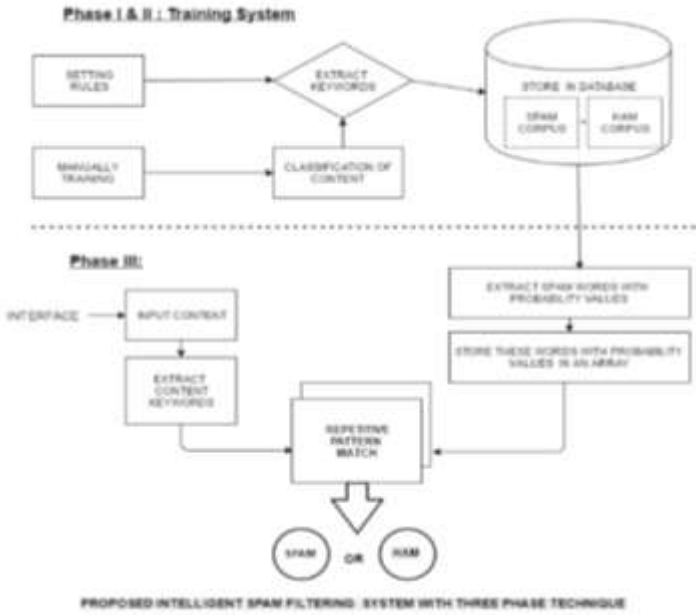

FIG 1.

### *C) REQUIREMENTS*

i. **HARDWARE**:

1. Dual core or higher version processor
2. 4GB RAM or more
3. 1 TB HDD Space

ii. **SOFTWARE**:

1. Windows 2007 or higher version of OS
2. Eclipse IDE
3. MS SQL Server

## IV. FEATURES

1. Virus Detection: Because a good deal of spam emails contain viruses and other malware, virus detection software can scan an email before it reaches user's email box and delete it accordingly. Also, if an email contains suspicious attachments, it will detect and delete them.
2. Auto-spam Detection: All incoming mail on the server can be scanned to automatically detect common spam threats and delete them before they ever reach the inbox.

3. Email Recovery and Quarantine: Spam blocking software can also move a suspect file into a temporary folder for user to judge its legitimacy. Generally, these folders will hold the suspect emails for a few days and then delete them – with the option of user recovering whatever user feels user have missed or didn't check.

## V. ADVANTAGES

1. The Bayesian method considers the whole content – It recognizes keywords that identify spam, but it also recognizes words that are present in valid mail. For example: not all emails that contain the word "free" and "sale" are spam. The benefit of the Bayesian method is that it considers the most interesting words and comes up with a probability that a message is spam. In other words, Bayesian filtering is an intelligent approach because it examines all aspects of a message, as opposed to keyword checking that classifies a mail as spam on the basis of a single word.
2. A Bayesian filter is constantly learning - By learning over a period of time from first occurrence of certain type of spam mails and new valid mails, the Bayesian filter evolves and learns to new spam techniques. For example, when spam emails started using "s-a-l-e" instead of "sale" they succeeded in skipping the keywords checking until "s-a-l-e" was also included in the keyword database. On the other hand, the Bayesian filter automatically notices such strategies; in fact if the word "s-a-l-e" is found, it is an even better spam indicator, since there is very less probability of such form of words to occur in a ham mail.
3. Bayes' classifier adapts with individual user- A Bayesian Classifier takes into account the recent spam and ham emails of the user and takes decision accordingly. For example, for some user communicating over email with a friend and receives a mail asking," are you free this evening?" then it is not considered as spam and Bayesian Classifier successfully classifies spam and ham mails even in such cases.

## VI. APPLICATIONS

The applications of the system will increase tenfold with the advent of Internet of Things (IoT) where for example even your washing machine, microwave, refrigerator and other appliances and electronic items will give you notification about certain important things to be taken care of. With the current scenario and with the current technologies the following are some of the application is that it's an easy to use software to detect spam content even when they are not in regular format because of which they bypass the network level firewall and the system need not be configured each time, once configured it will work forever unless customization or modification is required. User will never miss anything that needs urgent attention. It can be used for Email filtering to normal filtering of content that can occupy notable time of the user.

## VII. CONCLUSION

We through our paper aimed at improving the efficiency of filtering techniques based on Naïve Bayesian Algorithm which has proved to be effective to tackle the increasing problem of spam emails. In this study, we use an approach based on Naïve Bayesian, which is a good machine learning algorithm where we consider all types of input content. The results show that our approach to classify content is reasonable and an effective one, with accuracy going up to 92%. We hope to add in further enhancements like multilingual formats, and a better understanding of content to aid the user in all aspects.

.

## VIII. REFERENCES

[1] Christina V, Karpagavalli S and Suganya G, "A Study on Email Spam Filtering Techniques" International Journal of Computer Applications (0975 – 8887), Volume 12– No.1, December 2010.
[2]Savita Pundalik Teli, Santoshkumar Biradar,"Effective Email Classification for Spam and Non-Spam" International Journal of Advance Research in Computer Science and Software Engineering, Volume 4, Issue 6, June 2014.
[3] Saadat Nazirova,"Survey on Spam Filtering Techniques" Communications and Network, 2011, 3, 153-160.
[4] Ann Nosseir, Khaled Nagati, Islam Taj-Eddin,"Intelligent Word-Based Filter Detection Using Multi-Neural Networks" International Journal of Computer Science Issues, Volume 10,Issue 2,No 1,March 2013.
[5] Jitendra Nath Shrivastava, Maringanti Hima Bindu"E-mail Spam Filtering Using Adaptive Genetic Algorithm" Intelligent Systems and Applications,2014,02,54-60.
[6] Georgious Paliouras, Vangelis Karkaletsis, Constantine D Spyropoulos, Panagiotis Stamatopoulos"Stacking Classifiers for Anti-Spam Filtering of E-mail" ResearchGate,June 2001.
[7] Qin Luo, Bing Liu, Junhua Yan" Research of a Spam Filtering Algorithm Based on Naïve Bayes and AIS" International conference on Computational and Information Science(ICCIS) 2010.
[8] Md. Saiful Islam,Shah Mostafa Khaled,Khalid Farhan,Md. Abdur Rahman, Joy Rahman," Modeling Spammer Behavior: Naïve Bayes vs. Artificial Neural Networks." International Conference on Information and Multimedia Technology (ICIMT) 2009.
[9]Chen Liang, Yuen-Fu Chen, Guo Tan Liao, Bo-Chan Cheng"Anti-spam Email System in Facebook" IEEE 2009.

**Tanvi Samel** currently pursuing final year of Bachelor of Engineering at Atharva College of Engineering from University of Mumbai. She has written another paper titled NoSQL which is under review of IJARCET for issue 6, October 2016. She has completed Oracle Certified Java Professional examination.

**Darshan Vira** currently pursuing final year of Bachelor of Engineering at Atharva College of Engineering from University of Mumbai. He has written another paper titled NoSQL which is under review of IJARCET for issue 6, October 2016. He has completed Oracle Certified Java Professional examination.

**Anushka Patil** currently pursuing final year of Bachelor of Engineering at Atharva College of Engineering from University of Mumbai. She has written another paper titled NoSQL which is under review of IJARCET for issue 6, October 2016.

**Durgesh Pandey** currently pursuing final year of Bachelor of Engineering at Atharva College of Engineering from University of Mumbai.

**Prof. Smita Krishna Patil** Department of Information Technology at Atharva College of Engineering, University of Mumbai, smitasukrut@gmail.com.