SPAM DETECTION IDENTIFICATION ON MALICIOUS USER ATTACK OVER NETWORK

Ms.THEERTHANAP
DepartmentofMCA
Dhanalakshimi SrinivasanEngineeringcollege,
college,Perambalur,TamilNadu,India.
Email-Theerthana1997@gmail.com

prof. M. Bavithara
Department ofMCA
DhanalakshimiSrinivasanEngineering
Perambalur, TamilNadu,India.

Abstract— the increasing volume of unsolicited bulk e-mail (also known as spam) has generated a need for reliable anti-spam filters. Machine learning techniques very useful for detecting a spam email .An effective design and implementation of a tool is important to monitor and detect spam attacks in a real time network. Mail may be delivered through text, verbal, communication. Terrorist activities communicate over application and mail programs over internet. It also uses these mail applications over the internet for getting their message to younger generation and making all of type's terrorists.

Keywords— Bag of words, Term Frequency, Inverse Document Frequency, Bayes Law, Natural Language Tool Kit

technology had been increasing more. The law looking for solutions to detect these discussion forums for all possible criminal activities and download suspected Postings as evidence for investigation. Active Mail Monitoring System which will tackle with this problem. It has used a data mining algorithm to detect criminal activities, legal and illegal postings. In this system will use text data mining technique. Active Mail Monitoring System will let us help to analyze online plain text sources from selected discussion forums and will classify the text into groups and system will decide which post is legal and illegal accordingly to their points. It will help us to reduce and minimize many criminal activities which are held on socialite such as face book, Twitter, Tinder, etc.

I.   INTRODUCTION

Email is a service which allows us to send the message in electronic mode over the internet. It offers an efficient, inexpensive and real time mean of distributing information among people. Thus the mail process will be referred as mail over internet. Mail refers to the process of communicating, interacting and/or exchanging messages over the Internet. It involves two or more individuals that communicate through a mail-enabled service or software.Mail may be delivered through text, verbal, communication. Terrorist activities communicate over application and mail programs over internet. It also uses these mail applications over the internet for getting their message to younger generation and making all of type's terrorists. The mail monitor system is an important application that could allow for secure mails along with terrorism related mail detection that helps track down spread of terrorist networks and locate the activities using IP addresses. The internet Mail application is a dedicated Mail application for free internet Mail as well as tracking down on spread of terrorism online. Communication provides effective areas for illegal activities such as threatening messages. This system we have produced called as Active Mail Monitoring & Suspicious Mail Detection over Internet which will tackle with these issues.Internet

II.   EXISTINGSYSTEM

Online chat might allude to any sort of correspondence over the Internet that offers a continuous transmission of instant messages from sender to beneficiary. Chat messages are by and large short keeping in mind the end goal to empower different members to react rapidly. Along these lines, an inclination like a talked discussion is made, which recognizes chatting from other content based online correspondence structures, for example, Internet gatherings and email.

Online chat might deliver point-to point correspondences and in addition multicast interchanges from one sender to numerous collectors and voice and videochat, or might be a component of a web conferencing administration. As information goes through server it constantly filters it for any suspicious watchwords. The customary examination of Internet chat room dialogs puts an asset trouble on the knowledge group due to the time required to screen a huge number of persistent chatsessions.

DISADVANTAGES

- No such smart or basic system to identify the activities going on onlineforum

- Considering the criminal attack, if the discussion is held on a forum then there is no evidence of discussion held between two parties via online forum as there is no track ofrecords.

III.   PROPOSED SYSTEM

The proposed System will analyze online plain text sources from selected discussion forums and will classify the text into different groups and system will decide which post is legal and illegal. This system will ensures that the admin may not watch all the chat at a time, so in order to stop chatting illegally, the keywords are set by the admin as suspicious words which will be blocked or it cannot be able to view by the other person. Admin will detect the suspicious mail and block the mail id of the suspicious person and report to the authority. after the suspicious mail is detected using the identification of their location. Four major attributes are specified they are: MAC address, IP address, latitude and longitude.

ADVANTAGES

- The mail can be monitored and suspicious chats are identified
- Illegal mail user will be blocked by the admin and they will be reported to theauthority.
- The user will be protected from the unknownuser
- The terrorist location can be identified using latitude and longitudevalues.
- The MAC address and the IP address isalso predicted

III EXPERIMENT AND RESULT

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. System architecture can comprise system components, the externally visible properties of those components, the relationships (e.g. the behavior) between them.

It can provide a plan from which products can be procured, and systems developed, that will work together to implement the overall system. There have been efforts to formalize languages to describe system architecture; collectively these are called architecture description languages (ADLs).

Various organizations define systems architecture in different ways, including:

- An allocated arrangement of physical elements which provides the design solution for a consumer product or life-cycle process intended to satisfy the requirements of the functional architecture and the requirementsbaseline.
- Architecture comprises the most important, pervasive, top-level, strategic inventions, decisions, and their associated rationales about the overall structure (i.e., essential elements and their relationships) and associated characteristics and behavior.
- If documented, it may include information such as a detailed inventory of current hardware, software and networking capabilities; a description of long-range plans and priorities for future purchases, and a plan for upgrading and/or replacing dated equipment and software
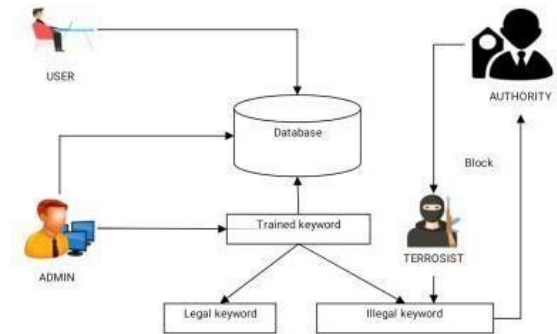- The composite of the design architectures for products and their life-cycleprocesses.

FIG. 1 SYSTEM ARCHITECTURE

IV.CONCLUSION

In this work a methodology has been proposed to detect the most common and frequent list of words which are leading a mail to be spam filtered. With the help of these methodologies we are successful in determining the spam behavioral word and Emails.

To stay away from spam/unessential sends we'd like successful spam separating systems. Bayesian classifier is a standout amongst the most essential and generally utilized classifier and furthermore it's the easiest order strategy because of its controlling capacities of tokens and related probabilities as indicated by the users" grouping choice and exact execution. In this undertaking, we executed the framework to break down every single mail. And further more give security-based location framework to encode the

messages by utilizing the Digest based framework. Later on, work we have an arrangement to execute other calculation to our grouping strategy to accomplish a better execution.

V. REFERENCE

[1] Mengjun Xie; Haining Wang; , "A Collaboration-based Autonomous Reputation System for Email Services," INFOCOM, 2010 Proceedings IEEE , vol., no., pp.1-9, 14- 19 March2010

[2] Nikita Spirin and Jiawei Han. 2012. Survey on web spam detection: principles and algorithms.SIGKDD Explor. Newsl. 13, 2 (May 2012),50-64.

[3] Amin, Rohan, Julie Ryan, and Johan van Dorp. "Detecting targeted malicious email." IEEE Security & Privacy 10.3(2012).

[4] Crawford, Michael, et al. "Survey of review spam detection using machine learning techniques." Journal of Big Data 2.1(2015).

[5] Sahın, Esra, Murat Aydos, and Fatih Orhan. "Spam/ham e-mail classification using machine learning methods based on bag of words technique." 2018 26th Signal Processing and Communications ApplicationsConference (SIU). IEEE,2018.

[6] Rathod, Sunil B., and Tareek M. Pattewar. "Contentbased spam detection in email using Bayesianclassifier." International Conference on. IEEE, 2015

[7] Trivedi, Shrawan Kumar. "A study of machine learning classifiers for spam detection." Computational and Business Intelligence (ISCBI), 2016 4th International Symposium on. IEEE,2016.

[8] Alurkar, Aakash Atul, et al. "A proposed data science approach for email spam classification using machine learning techniques." Internet of Things Business Models, Users, and Networks, 2017. IEEE,2017.

[9] Wang, Alex Hai. "Detecting spam bots in online social networking sites: a machine learning approach." IFIP Annual Conference on Data and Applications Security and Privacy. Springer, Berlin, Heidelberg,2010.

[10] Harisinghaney, Anirudh, et al. "Text and image-based spam email classification using KNN, Naïve Bayes and Reverse DBSCAN algorithm." 2014 International.