Volume 4, Issue 11, November 2014 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

An Evaluation of Spam Detection by Pattern Based and Facial Authentication System by Using Security

R S Devika, (Pursing M.tech), **K Jagannath** (Associate Professor)
Dept of CSE, Kuppam Engineering College,
Kuppam, India

Abstract— Network Security Consist of the provisions and policies adopted by a network administrator to prevent and monitor unauthorized access. E-mail is the main communication link now a day everyone uses/have mail access all officials company sent on by a mail communication. In this mail communication we will have a spam mails. Spam E-mails/many E-mails consists URL's to a websites or WebPages leads to virus or hacking. So we already have a method of detecting the spam mails but it won't detect the entire spam mails. Spamming is the use of Electronic messages to send/receive unsolicited bulk messages especially advertising indiscriminately. Where as in this method we are going to detect the entire spam by email scanning before it read by the users, blocking the domain irrespective of the users E-mail ID, keyword based blocking by monitoring the subjects, manipulating the difference between public and private domain before blocking, password security by bio-metric, Facial Recognition, Fractal detection (face scanning) and recognition is an unique method to identify every human being. We use brute force string match algorithm. It shows the candidate images of face scanning recognition system could be recognized efficiently using inter dependence of pixels arising from facial codes of images.

Keywords- Security evaluation, face recognition, spam filter, facial-authentication.

I. INTRODUCTION

Electronic mail, most commonly referred to as email or e-mail. It is a method of exchanging digital messages from a sender to one or more receivers. Gmail is a free email service produced by Google. Users may access Gmail as reliable webmail via POP3 or IMAP4 protocols. Spam can be defined as unsolicited email for a recipient or any email that the user do not wanted to have in his inbox. It is also defined as "Internet Spam is one or more unwanted messages, sent as a part of larger set of messages, all having considerably similar content." There are significant problems from the spam mails, wastage of network resources, delay, destruction to the PC's & laptops due to viruses & the ethical issues such as the spam emails advertising pornographic sites which are harmful to the young generations [1].

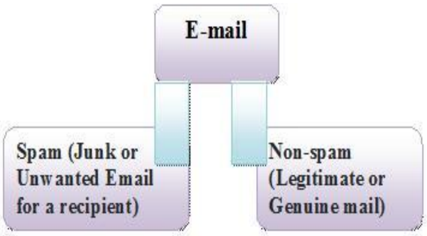

E-mail
Spam (Junk or Unwanted Email for a recipient)
Non-spam (Legitimate or Genuine mail)

Fig. 1: Email types

Spam contents in the web are not only utilizes valuable resources inside the web but can also mislead the users to unsolicited websites and award undeserved search engine rankings to spammer's campaign websites. Most of the research in anti-spam filtering focuses on the origin of spam content, only a few have investigated the identification of spam content on the web, and filtering mechanism. There is open research area in identifying the individual person's emails by manipulating through an automated supervised machine learning solution which utilizes web navigation behavior to detect the possible spam. The existing approaches need an effective representation of e-mail. Large sets of reported spam has to be stored in the known spam database, the storage size of e-mail abstraction should be small. Moreover, the email abstraction should capture the near-duplicate phenomenon of spam, and should avoid accidental deletion of non spam e-mails.

Password hijacking or Password exposure is one of the major problem in the area of Password the scenarios. We don't have proper system for this major issue. Multimodal biometric systems for personal identity recognition are more popular in the past few years. It has been shown that combining information coming from different biometric traits can overcome the limits and the weaknesses inherent in every individual biometric, resulting in a higher accuracy. Moreover, it is commonly believed that multimodal systems also improve security against spoofing attacks, which consist of claiming a false identity and submitting at least one fake biometric trait to the system.

II. BACKGROUND AND PREVIOUS WORK

Here we analyzed previous work, highlighting the concepts that will be utilized in our system.

Message passing through emails is one the well-known way of today's world since it is more effective and fast than any other sources. Authentication is the major part often involves verifying the validity of at least one form of identifications of the users. Normally authentication for logging in to the Gmail service by means of username and password characters is applicable in the existing system. Security type of authentication such as logging in to the Gmail service using the secret code received to the mobile device of the user is also applicable. This in turn less effectual since anybody who accesses the user's mobile can log on to the service or there is no option in case of mobile theft. Spam is a typical message passing that floods the Internet with many copies of the same message, which tried to force the message on people who would not otherwise choose to receive it. Spam keyword filtering is the way used in existing system to get rid of spam emails. Frequent mails from a mail id can be spammed if it is tested against spam filter but the domain cannot be filtered under the spam filter. Hence any number of E-mail ID can be created by the spammers to send spam mail under the same domain. Automatic email content examining and spamming is not possible. Many a time the concept of spamming is false positive in this system.

In Pattern classification method used in biometric authentication system, network intrusion detection system, and spam filtering system. They evaluated *security* of pattern classifiers, such as the performance degradation under potential attacks they may incur during operation. They developed a frame work and that used in three application examples, such as spam filtering system, biometric authentication system, and network intrusion detection system. P. Fogla, M. Sharif, R. Perdisci, O. Kolesnikov, and W. Lee[3] discussed their initial research efforts focused on the detection of malicious insiders who exploit internal organizational web servers. The objective of the research is to apply lessons learned in network monitoring domains and enterprise log management to investigate various approaches for detecting insider threat activities using standardized tools and a common event expression framework. The author is emphasizing insider threat detection in network monitoring domain. Scanning through the web server log and identify the threat is the main factor. Here, multiple events and single events from multiple hosts and single hosts is identified. In [4], they extract spam/non-spam email and detect the spam email efficiently. Here representation of data is done using a vector space model and clustering is the technique used for data reduction. In this paper an email clustering method is proposed and implemented to efficient detects the spam mails. The BIRCH clustering, decisions made without scanning the whole data & BIRCH utilizes local information (each clustering decision is made without scanning all data points). BIRCH is a better clustering algorithm requiring a single scan of the entire data set thus saving time. So, it cannot work in any other algorithm.

In this work [5], they present the design and implementation of Pyramid-like Face Detection (P-FAD),which is a real-time face detection system constructed on general embedded devices. It is motivated by the observation that the computation overhead increases proportionally to its pixel manipulation-FAD propose a hierarchical approach to shift the complex computation to the promising regions. They introduce the hierarchical framework for face detection on embedded smart camera briefly. And focus on tackling the challenging issues in constructing the hierarchical scheme.

III. MOTIVATIONS

Major issue of the existing system is identification of patterns to avoid spams. Nowadays, spams are considered as one of the major technical problem for most of the users and we don't have proper solution in manipulating the following key issues like,

- Email scanning before it's read by the users
- Blocking the domain irrespective of the users E-mail ID
- Keyword based blocking by monitoring the particular subjects
- Blocking the users URL
- Manipulating the difference Between the public and the Private domain before blocking

Password visualizing or Password trap is one of the major flaws that are available in systems that too for the public domains.

VI. PROPOSED SYSTEM

In our proposed system, high effective authentication with the purpose of log on to the Gmail service securely and efficient spamming are taken into consideration .Here authentication in the form of fractal detection and recognition after contour detection of the face using the image of the user is introduced. Since fractal detection and recognition is an unique method to identify every human

being, this concept is more effective in terms of authenticating into the service. Pattern classifiers such as Keywords and URL's for data check, tag construction and keyword identity, automatic reading of mails is the concepts used in this system. Administrator of the email service uses the pattern classifiers and maintains a repository to filter out spam domains and keywords. Hence this perception spams the frequent surplus mails from same domain with different mail id.

Automatic reading of mails to examine the spammed keyword is an intriguing conception introduced in this system to overcome many flaws in case of spam filtering. Hence the authentication by means of fractal recognition and pattern classifier based spam filtering in the email service turn this proposed system more thriving.

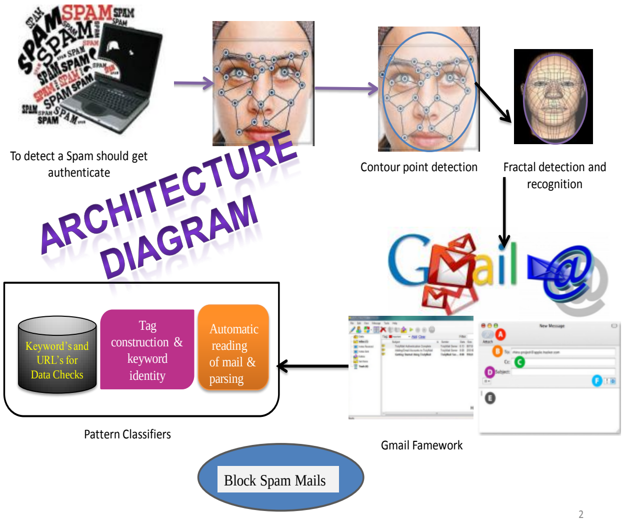
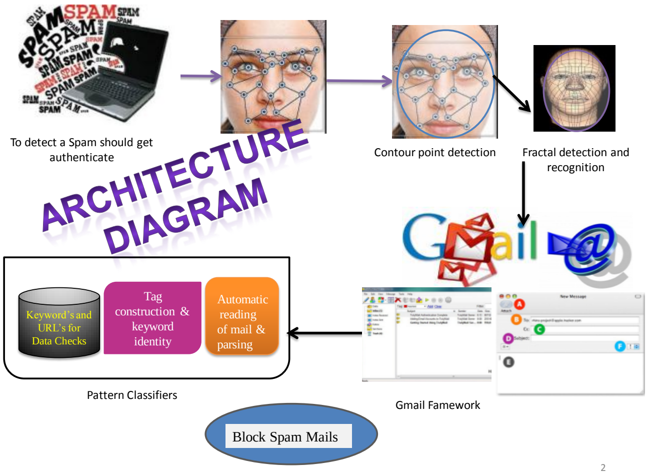

To detect a Spam should get authenticate
ARCHITECTURE DIAGRAM
Contour point detection
Fractal detection and recognition
Keyword's and URL's for Data Checks
Tag construction & keyword identity
Automatic reading of mail & parsing
Pattern Classifiers
Gmail Famework
Block Spam Mails
2

Fig.2: Architecture diagram


To detect a Spam should get authenticate
ARCHITECTURE DIAGRAM
Contour point detection
Fractal detection and recognition
Keyword's and URL's for Data Checks
Tag construction & keyword identity
Automatic reading of mail & parsing
Pattern Classifiers
Gmail Famework
Block Spam Mails
2

V. OBJECTIVE

High effective authentication with the purpose of log on to the email service securely and efficient spamming are taken into consideration. Authentication in the form of fractal detection and recognition after contour detection of the face using the image of the user is introduced. Pattern classifiers such as Keywords and URL's for data check, tag construction and keyword identity, automatic reading of mails are the concepts used in this system.

VI. HUMAN FACE RECOGNITION BASED ON FRACTAL IMAGE CODING

Human face recognition is an important area in the biometrics field. It has been an active area of research for many years. Human face is a biometric form. Biometrics are a set of measurable physiological and/or behavioral characteristic properties of the human body that can be used to infer a person's identity[5].One of the first step of the face recognition is to detect and extract face from an image. There are many face detection algorithms. These algorithms are mainly classified into four. Those are

- knowledge based methods
- feature invariant approaches
- template matching methods
- Appearance-based methods.

Knowledge-based methods use a set of rules that developed from what humans know about the appearance of a face. Feature invariant approaches use structural features that are invariant to changes in pose, expression. Template matching methods use a set patterns representative of the face, which are then correlated with the input image. Appearance-based methods perform face detection using
models or templates learnt from a set of representative training images.[6].

Once face image has been detected, recognition can be carried out. Here we present fractal methods for face recognition. It is shown that candidate images of face recognition system could be identified, systematically, using interconnection of pixels emerge from fractal codes of images. The interdependence of the pixels is inherent within the fractal code in the form of chain of pixels. The mathematical principal behind the application of fractal image codes for recognition is, An Image Xf can be represented as

$$Xf = A \times Xf + B$$

Where A and B are fractal parameters of image Xf. Dissimilar fractal codes can be proposed for any arbitrary image, with the definition of a fractal transformation.

$$T(X) = A (X- Xf) + Xf$$ [4].

In this system contains the following steps:

- Normalization of the face image.
- Feature extraction using fractal coding of the normalized face image.

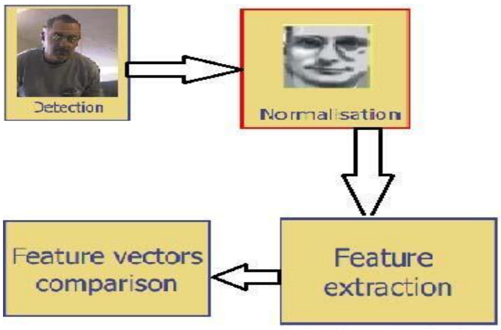


Fig. 3: Face recognition process

Face localization is an uncomplicated form of face detection. If any two face feature points are detected then that face can be located and normalized. Therefore, face localization can also be tackled via facial feature detection. The eyes, nose and mouth were identified using straight image processing techniques. Assume that the nose's horizontal position was also determined and an accurate locus for the nose tip is available. The identification of the loci of these feature points (eyes, nose and mouth) gives an estimate of the pose of an individual's face. Here we used geometrical normalization, where find some points in the face image (contour points) then find the orientation of points.

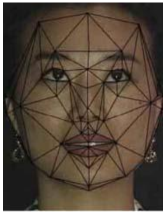


Fig.4: Face Normalization

Features often used are color and texture [8,9]. Furthermore, several authors have suggested using shape properties [9], or relative position of objects within an image called spatial similarity.

Algorithm1: Fractal Coding

Input: Normalized image

Output:

- Partition the image into non-overlapping range blocks Ri using quad-tree partitioning method.
- Cover the image with sequence of overlapping domain blocks Dj.
- For each range blocks, find the domain block and corresponding transformation that will match the range block.
- Set the geometrical positions of the range blocks and matching domain blocks as well as matching transformation as fractal code of face image.

Quad tree partitioning method utilizes the image processing technique based on recursive splitting of selected image quadrants. The resulting partition is represented by a tree structure in which each non-terminal node has four successors. The task of a fractal encoder is to find a domain block D of the same image for every range block R such that a transformation of this block W (D) is a good approximation of the range block. The main step in fractal image coding is the mapping of domain block to range blocks. For each range block, the algorithm compares transformed versions of the domain blocks to the range block. The transformations here used are typically affine transformation. The transformations are a combination of a geometrical transformation and luminance of the transformation.

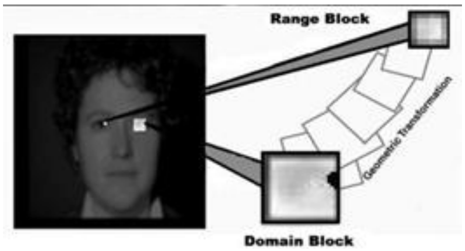


Fig.5: An illustration of domain and range blocks

VII. PATTERN BASED SPAM DETECTION

Major issue of the existing system is identification of patterns to avoid spams. Nowadays, spams are considered as one of the major technical problem for most of the users and we don’t have proper solution in manipulating the following key issues like,

- Email scanning before it’s read by the users.
- Blocking the domain irrespective of the users email id.
- Keyword based blocking by monitoring the subjects.
- Manipulating the difference between the public and private domain before blocking

it. Spam is most often considered to be electronic junk Spam is flooding the Internet with numerous copies of the similar messages. This tried to push the message on people who would not otherwise choose to receive mail.

A spam filter is a program that is used to detect unsolicited and unwanted email and prevent those messages from getting to a user's inbox. Here spam filter checks all incoming emails to your email accounts against mail filter rules. Outline the tags and automatic parsing of the HTML tags and read the data from each tuple. Compare the tags with the keywords and URL’s located in our repository using top down parsing keyword search algorithm. If the data in the tags matched, then the email will be identified as the spam email. We are notifying all the Private Domain related to the spam emails and URL’s received in the email will be reposited for future email rejections. A very good repository of URL’s received in the emails and domain related to the spam user is created and can be maintained by repository of URL’s received in the emails and domain related to the spam user is created and can be maintained by an administrator. There is no specific algorithm for statistically determining whether or not a given e-mail spam message.

The obvious method for pattern matching is just to check, for each possible position in the text at which the pattern could match, whether it does in fact match. There are number of pattern matching algorithms. Here we used Brute Force algorithm for top down key word parsing. The brute-force pattern matching algorithm compares the pattern *Pt* with the text *Txt* for each possible shift of *Pt* relative to *Txt*, until either n a match is found. Brute-force pattern matching runs in time *O* (*nm*).

Algorithm 2: Brute-Force Pattern Matching

```
Input:   text Txt of size n and pattern Pt of size m
Output: starting index of a

        Substring of Txt equal to Pt or -1 if no such substring exists

        for (i = 0; i< n ; i ++) j = 0;

        while (j <m && Txt[i + j] = = Pt[j]) j = j + 1;

        if ( j== m) return I;
        return -1;
```

VIII. CONCLUSION

In this paper, we have introduced a new method for high effective authentication with the purpose of log on to the email service securely and efficient spamming. Pattern classifiers such as Keywords and URL’s for data check, tag construction and keyword identity, automatic reading of mails is the concepts used in this system.

In the future enhancement the user perspective like forgetting the password will be implemented. Voice recognition concept can also be implemented to make the system more users interactive.

REFERENCE

[1] Wanli Ma, Dat Tran, Dharmendra Sharma, Member, IAENG,” *An Extendable Software Architecture for Spam Email Filtering”,* IAENG International Journal of Computer Science, 34:1, IJCS_34_1_18 An International Multi Conference of Engineers and Computer Scientists, Hong Kong, 21-23 March, 2007

[2] Enrico Blanzieri and Anton Bryl, *“Evaluation of the Highest Probability SVM Nearest Neighbor Classifier with Variable Relative Error Cost,”* Proc. Fourth Conf. Email and Anti-Spam (CEAS), 2007.

[3] Ho-Yu Lam, Dit-Yan Yeung” *A Learning Approach to Spam Detection based on Social Networks”* CEAS 2007-Fourth Conference on Email and Anti-Spam, August 2-3, 2007, Mountain View, California USA.

[4] M. Basavaraju , Dr. R. Prabhakar *“A Novel Method of Spam Mail Detection using Text Based Clustering Approach”* , International Journal of Computer Applications (0975 – 8887) Volume 5– No.4, August 2010.

[5] Hasan shojaa alkahtani, Paul gardner-stephen, and robert goodwin,” *A taxonomy of email spam filters*” Computer Science Department, College of Computer Science and Information Technology, King Faisal University, Al-Hassa 31982, Kingdom of Saudi Arabia.

[6] Shalendra Chhabra, Willam .S. Yerazunis, and Christian Siefkes, “*Spam Filtering Using a Markov Random Field Model with Variable Weighting Schemas,”* supported by the German Research Society (DFG grant no. GRK 316)

AUTHOR BIOGRAPHIES

R S Devika received B.Tech degree in Information Technology from Kuppam Engineering College, Kuppam affiliated to JNTUniversity (Ananthapur)A.P in 2012 currently Pursuing M.tech in Computer Science and Engineering in Kuppam Engineering College at Kuppam affiliated to JNTUniversity (Ananthapur)A.P

K Jagannath received M.Tech degree in Computer Science and Engineering, from Dr.M.G.R. University, Chennai in 2008, Received B.Tech degree in Computer Science and Engineering from Kuppam Engineering College, affiliated to JNTUniversity (Ananthapur) A.P in 2005. Curently rendering his service as Associate Professor at Kuppam Engineering College, in Kuppam affiliated to JNTUniversity (Ananthapur) A.P