Filtering of Unwanted Messages and Phish Links on OSN Wall

M.G. Devikar[1] Vaishali Bhagat[2] Tejashree Dhole[3] Neha Hivarkar[4] Varsha Bhosale[5]
[1]Professor
[1,2,3,4,5]Department of Computer Engineering
[1]Savitribai Phule University of Pune [2,3,4,5]Modern Education Society's, College of Engineering, Pune

Abstract— OSN system is a powerful tool to prevent users to avoid their walls getting overwhelmed by unwanted data. Thus, it provides a classification mechanism to sort unwanted text and images. On contrary it also supports phish link detection mechanism. The links that possess catastrophic results are detected and alerted. OSN system is mainly used to provide security environment on internet. To detect phish links an anti-phishing algorithm is used called ObURL which has six different steps to filter the links.

Key words: OSN, phishing, filtered wall, ObURL

I. Introduction

OSN being an important media of communications and building relationship on Internet. It is generally being influenced by many factors some of those are, OSN user has to deal with unwanted post or messages on their walls which are vulgar or of no meaning. It may also contain some political, casteism related rumors which may cause riots in the social media. Links which are been posted on wall may be phishing and may mislead the user. To overcome the above problems a new proposed system “Filtering of Unwanted Messages and Phish Links on OSN wall” an online application is proposed. There are two major task proposed in this system. The first one is the message filtering and the second is the phishing link detection. The message filtering technology helps the OSN wall user to avoid from getting unwanted message from all the users on social networks. The phishing link detection technology is use to secure the OSN user from getting phishing links and avoid from further attacks.

II. Existing System

Today OSNs provide very little support to prevent unwanted messages on user walls. For example, Facebook allows users to state who is allowed to insert messages in their walls (i.e., friends, friends of friends, or defined groups of friends). However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them. In the current paper “[1] A System to Filter Unwanted Messages from OSN User Walls” the use of neural learning which is today recognized as one of the most efficient solutions in text classification. “[2] Intelligent Phishing Website Detection And Prevention System” this paper is implemented for filtering of the phishing links posted on the user walls to prevent user credentials information leaking. “[3] Anti-Phishing Technique to Detect URL Obfuscation” this paper provide us with the ObURL Algorithm, its depth knowledge and different test perform on the links to get more accuracy in detecting the phish links.

A. Drawbacks:

- User cannot avoid unwanted messages delivered on OSN walls.
- Phishing link are not detected leading user redirect to phishing site that may cause stealing of user’s credential data.
- Short-text classification does not provide sufficient word occurrences.

III. Literature Survey

A. Background of the project:

On-line Social Networks (OSNs) have become a popular interactive medium to communicate, share and disseminate a considerable amount of human life information. Daily and continuous communication implies the exchange of several types of content, including free text, image, audio and video data. The huge and dynamic character of these data creates the premise for the employment of web content mining strategies aimed to automatically discover useful information dormant within the data and then provide an active support in complex and sophisticated tasks involved in social networking analysis and management

Filtered Wall (FW), able to filter out unwanted messages and links from social network user walls. We believe that the proposed strategy is a key service for social networks in that in today social networks users have little control on the messages displayed on their walls. For example, Facebook allows users to state who is allowed to insert messages in their walls (i.e., friends, friends of friends, or defined groups of friends). However, no content-based preferences are supported. For instance, it is not possible to prevent political or vulgar messages. In contrast, by means of the proposed mechanism, a user can specify what contents should not be displayed on his/her wall, by specifying a set of filtering rules.

B. Domain of the study:

The project basically comes under information security domain. The key concept of information security in project is that the system prevents the OSN users from fraudulent attackers. Even the messages which are not vital for a user are filtered. In addition, the system provides the support for user defined blacklist management, that is, list of users that are temporarily prevented to post messages on a user wall.

C. Motivation of the project:

Due to the growing use of social networking it becomes mandatory to prevent the user walls from unwanted data. There are possibilities of riots arising due to unwanted posts and images. The innocent users get trapped by attacker’s phished link and there can be loss of confidential data as well as property. The aim of the present work is to propose and experimentally evaluate an automated system, called

Filtered Wall (FW), able to filter out unwanted messages from social network user walls.

IV. Proposed System

The aim of the proposed system is merging the above two existing system. In this the use of neural learning which is today recognized as one of the most efficient solutions in text classification. It evaluates an automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls. In this it exploit Machine Learning (ML) text categorization techniques to automatically assign with each short text message a set of categories based on its content. The major efforts in building a robust short text classifier are concentrated in the extraction and selection of a set of characterizing and discriminate features. The solutions investigated are an extension of those adopted in a previous work from which we inherit the learning model and the elicitation procedure for generating pre-classified data. The original set of features, derived from endogenous properties of short texts, is enlarged here including exogenous knowledge related to the context from which the messages originate.

It also includes filtering of both messages and links, links are tested whether phishing or not, applying the ObURL algorithm for detecting the phishing link and alerting the user, thus securing the user from attacks or mislead by the hackers.

A. *Objectives of the Proposed System:*

- Filtering unwanted information from OSN wall as per user requirement.
- Alert user from phishing links and sites.
- Making OSN more reliable, secure, trustworthy and comfortable for the user.

V. Phishing Prevention

```
v_link: visual link;
a_link: actual_link;
v_dns: visual DNS name;
a_dns: actual DNS name;
sender_dns: sender'sDNS name.
int link (v_link, a_link)
 {
  v_dns = GetDNSName (v_link);
  a_dns = GetDNSName (a_link);
  if ((v_dns and a_dns are not
  empty) and (v_dns! = a_dns))
 return PHISHING;
 if (a_dns is dotted decimal)
 return POSSIBLE_PHISHING;
 if (a_link or v_link is encoded)
 {
    v_link2 = decode (v_link);
    a_link2 = decode (a_link);
   return link(v_link2, a_link2);
 }
 /* analyze the domain name for possible phishing */
 if (v_dns is NULL)
 return AnalyzeDNS (a_link);
int AnalyzeDNS (actual_link) {
/* Analyze the actual DNS name according to the blacklist and whitelist*/
 if (actual_dns in blacklist)
 return PHISHING;
 if (actual_dns in whitelist)
 return NOTPHISHING;
 return PatternMatching(actual_link);
}
int PatternMatching (actual_link){
 if (sender_dns and actual_dns are different)
 return POSSIBLE_PHISHING;
 for (each item prev_dns in seed_set)
 {
      bv = Similarity(prev_dns, actual_link);
      if (bv == true)
      return POSSIBLE_PHISHING;
 }
 return NO_PHISHING;
}
float Similarity (str, actual_link) {
 if (str is part of actual_link)
 return true;
 int maxlen = the maximum string
 lengths of str and actual_dns;
 int minchange = the minimum number of changes needed to transform str
 to actual_dns (or vice verse);
 if (thresh<(maxlen-minchange)/maxlen<1)
 return true
 return false;
}
```

A. *ObUrl algorithm works as follows:*

First it will extract the DNS names from actual link and visual link, compare it and if both are not same then detect it as a phishing link. If actual link present in the dotted decimal IP address then there is a possibility of phishing attack. If actual link and virtual link is present into encoded form then this encoded link is firstly decoded and again performs above tests to get the result.

When DNS of visual link is NULL or destination information is not present then ObUrl calls analyzeDNS function for analyzing actual DNS which works as follows,
We are using two lists as blacklist which contains phishing links and whitelist which contain non phishing links. So this function

VI. System Implementation

The system implementation contains the following modules that are essential to build the system:-

A. *Modules:*

(1) User Registration (Sign In / Sign Up)
(2) Adding/Inviting Friends
(3) Chatting/Messaging
(4) Post on User Wall
(5) Filtering rules
(6) Online setup assistant for FRs thresholds
(7) Phishing prevention on links posted on user walls
(8) Blacklists

1) *User Registration:*

In this module first user register with our application by adding his personal information like his name, password, address and his hobbies etc. After registering with our application he can login with us using user id and password. System authenticates this information and allows user to login into the system.

2) *Posting messages on user walls:*

User can update his status on his wall, all his friends can see this status and post there view about your status. So these messages should get filtered. So we are implementing here filtering rules.

3) *Filtering rules:*

In defining the language for FRs specification, we consider three main issues that, should affect a message filtering decision. First of all, in OSNs like in everyday life, the same message may have different meanings and relevance based on who writes it. As a consequence, FRs should allow users to state constraints on message creators. Creators on which a FR applies can be selected on the basis of several different criteria; one of the most relevant is by imposing conditions on their profile's attributes. Given the social network scenario, creators may also be identified by exploiting information on their social graph. This implies to state conditions on type, depth and trust values of the relationship creators should be involved in order to apply them the specified rules.

The same message on OSNs may have different meanings and relevance based on who writes it. It is necessary to apply constraints on messages. Constraints can be selected on several different criteria's. User can state what contents should be blocked or displayed on filtered wall by means of Filtering rules. Filtering rules are specified on the basis of user profile as well as user social relationship.

4) *Blacklist:*

A further component of the system is a BL mechanism to avoid messages from undesired creators, independent from their contents. BLs is directly managed by the system, which should be able to determine who are the users to be inserted in the BL and decide when users retention in the BL is finished. To enhance flexibility, such information given to the system through a set of rules, hereafter called BL rules. Similar to FRs, our BL rules make the wall owner able to identify users to be blocked according to their profiles as well as their relationships in the OSN. Therefore, by means of a BL rule, wall owners are for example able to ban from their walls users they do not directly know (i.e., with which they have only indirect relationships), or users that are friend of a given person as they may have a bad opinion of this person. This banning can be adopted for an undetermined time period or for a specific time period. More precisely, among possible information denoting users bad behavior we have focused on two main measures. The first is related to the principle that if within a given time interval a user has been inserted into a BL for several times, say greater than a given threshold, he/she might deserve to stay in the BL for another while, as his/her behavior is not improved. This principle works for those users that have been already inserted in the considered BL at least one time.

5) *Phishing prevention for links posted on user walls:*

Phishing is a new word produced from 'fishing', it refers to the act that the attacker allure users to visit a faked Web site by sending them faked e-mails (or instant messages), and stealthily get victim's personal information such as user name, password, and national security ID, etc. This information then can be used for future target advertisements or even identity theft attacks (e.g., transfer money from victims' bank account).

So we are providing here an anti-phishing environment for the links posted on user wall.

VII. SYSTEM ARCHITECTURE

In this block we are demonstrating the overall flow of the project or proposed idea. Sender is the one who post messages/links or both on the user wall, for that sender should be friend of user. Before posting the post on the user wall, System will check if the user is blocked user or not. If it is the blocked user then the messages or post will be discarded and would not reach to the user wall, if it is not a blocked user the filtering criteria will be applied on the message or post. Messages or post has to pass through Short Text Classifier and Content Based Message Filtering. In Short Text Classifier the message will be classified as Neutral and Non-neutral according to the stored dataset. Non-neutral messages will be further filtered for Content Based Message Filtering to show the behaviour and relationship between the users and further the sender from where the message arrives is blocked for a particular period of time.

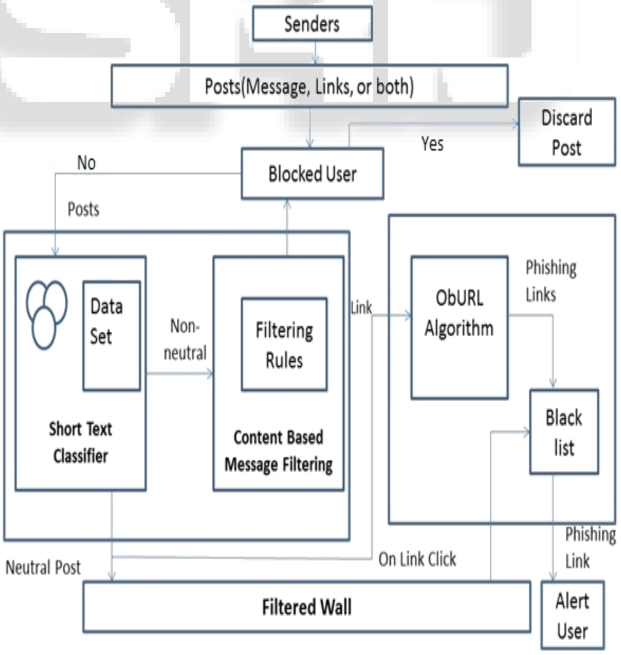

Fig. 1: Block diagram of Filtering of Unwanted Messages and Phishing Links on OSN wall

If the messages is neutral then it is posted on the user wall, if the post contains a link it is processed by ObURL Algorithm for checking if the link is phishing one or not. If the link found to be phishing is stored in the blacklist and if user tries to access this phishing link the alert will be generated by the system to prevent user from further loss.

VIII. FUTURE WORK

In future work, we plan to address this problem by investigating the use of online learning paradigms able to include label feedbacks from users. Additionally, we plan to enhance our system with a more sophisticated approach to decide when a user should be inserted into a Black Lists.

IX. CONCLUSIONS

In this paper, we have presented a system to filter undesired messages and links from OSN walls. The system develops a ML soft classifier to implement customizable content-dependent FRs. In particular, we aim at investigating a tool able to automatically recommend trust values for those contacts user does not individually identified. We do consider that such a tool should propose expectation assessment based on users procedures, performances, and reputation in OSN, which might involve enhancing OSN with assessment methods. Though, the propose of these assessment based tools is difficult by several concern s, like the suggestions an assessment system might have on users' confidentiality and/or the restrictions on what it is possible to audit in present OSNs. An introduction work in this direction has been prepared in the context of expectation values used for OSN access control purposes. However, we would like to remark that the system proposed in this paper represents just the core set of functionalities needed to provide a sophisticated tool for OSN message filtering. Still if we have balanced our system with an online associate to set FR thresholds, the improvement of a absolute system effortlessly exploitable by average OSN users is a wide topic which is out of the scope of the present paper.

X. ACKNOWLEDGEMENT

We take this opportunity to express our profound gratitude and deep regards to our respected Principal Dr. A. A. Keste , H.O.D. Prof N. F. Shaikh and our project guide Prof. M. G. Devikar for their guidance, monitoring and constant encouragement throughout the course of this thesis .The blessing, help and guidance given by them time to time shall carry us a long way in the journey of life on which we are about to embark.

REFERENCES

[1] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, Moreno Carullo "A System to Filter Unwanted Messages from OSN User Walls" Department of Computer Science and Communication University of Insubria 21100 Varese, Italy. IEEE Transactions on knowledge and Data Engineering Vol:25 YEAR 2013**,** E-mail: moreno.carullog@uninsubria.it.

[2] M. Madhuri, K. Yeseswini, U. VidyaSagar, "Intelligent Phishing Website Detection And Prevention System"**,** International Journal of Communication Network Security, ISSN:2231–1882, Volume-2, Issue-2,2013. Email:madhurimitai01@gmail.com,yeshukandagaddala@gmail.com, engg.sagar@gmail.com

[3] Jigar Rathod, Debalina Nandy, "Anti-Phishing Technique to Detect URL Obfuscation " Internatinal Journal of Engineering Research and Applications www.ijera.com , ISSN : 2248-9622, Vol. 4, Issue 5(Version 2) , May 2014.

[4] International Journal of Advanced Research in Computer Science and Software Engineering ", Research Paper Available online at: www.ijarcsse.com, 33 Volume 4, Issue 2ISSN: 2277 128X , February 2014 Filtering of Unwanted Messages and Phish Links on OSN wall Department of Computer Engineering 33

[5] Bimal Viswanath, M. Ahmad Bashir, Mark Crovella, Saikat Guha, "Towards Detecting Anomalous User Behavior in Online Social Networks", Research Paper Available online.