

CAPTCHA AUTHENTICATED UNDESIREED MESSAGE FILTERING TECHNIQUE FOR SOCIAL NETWORKING SITE

Fathimath Shahistha M.¹, Prabhakara B. K.²

¹ Department of Computer science and Engineering, SCEM, Mangalore

² Associate Professor, Department of Computer science and Engineering, SCEM, Mangalore

Abstract-The authentication and security mechanism for any application is important. The major function of any security system is to protect the data from unauthorized access. The existing password mechanism based on textual passwords provides low level of security against the unwanted access. One of the alternatives to improve the security and authentication is a graphical based password. One fundamental issue in today's Social Networks Site(SNSs) is to give users the ability to control the messages posted on their own private wall . OSNs provide little support to this requirement. In this paper a system with hybrid graphical password system based on Captcha technology. Captcha is a standard Internet security technique to protect online emails and other services. Captcha as graphical passwords known as CaRP. CaRP is a click-based graphical passwords. In which a sequence of clicks on an image is used to derive the password. It also include rule-based system, which allows users to customize the filtering criteria and a Machine Learning-based soft classifier automatically labeling messages with content-based filtering.

Keywords-Online Social Network, Short Text Classification, Policy-based Personalization, Information Filtering, CaRP, Graphical password, Filtering Rule.

I. INTRODUCTION

The passwords is used for the authentication purpose. The major function of any security system is to protect the data from unauthorized access. Passwords are the secrets that are provided for the user. Conventional password scheme uses textual passwords or alphanumeric characters and this password scheme is easy but the problem is that if user give small password then the user can easily guess the password and these small passwords are hacked by different attackers. When the text password length is bigger, then it is hard to remember and if the passwords are not frequently used then the passwords are easily forgotten. Graphical passwords are alternative to the textual passwords where graphical passwords use pictures as passwords instead of textual passwords. Graphical based password system is the combination of recognition and recall based techniques where in the recognition method user need to identify the images or objects which are displayed to him/her and in the recall based method the user need to recall the images or redraw the images which he has already selected. The main advantage of using graphical password is that it can be easily remembered. Existing system consist only images for password or draw a secret pattern for password, but there was a problem of shoulder surfing. One person can easily steal the password by seeing and access his secret information. The hybrid graphical password system based on Captcha is the security system based on hard Artificial Intelligence (AI) problems where the Captcha presents the user with a challenge or a puzzle. This proposed system is based on the click method (portions of the image where the users are likely to select the click-points), Captcha is a standard Internet security technique to protect online emails and other services. CaRP is a click-based graphical passwords with a sequence of clicks on an image is used to derive the password. CaRP scheme

is classified into two categories first is the recognition and the second is the recall based methods. The images used in the CaRP are Captcha challenges. For every login attempt a new CaRP image is generated. CaRP resolves the number of security issues such as dictionary attacks on passwords, online guessing attacks etc. CaRP requires a Captcha challenge for every login. The problem with this password system is that the login process is slow. CaRP is robust to shoulder surfing- attack.

Online Social Networks (SNS) is mainly used to share contents like text, image, audio, and video data ,a considerable amount of human life information. SNS is a platform to build social networks (or) social relations among people for sharing interest, picture, text and real time connections. A social network service consists of each user having his own profile, social links, blogs and additional services. Web based service allows individuals to create a public profile to create a list of users with whom to share connection and to view the connection within the system. Some of the social networks which are mainly used to connect with friends for example applications like: Face book, Google+, YouTube, Twitter etc. Web content Mining is used to discover useful and relevant information from a large amount of Data. In SNS, information filtering can be used with a different purpose. In SNS there is the possibility of posting (or) commenting other posts on particular public (or) private areas called Walls. Information filtering is mainly used to give user the ability to control the message written on their own walls by filtering out unwanted messages. The aim of the present work is to propose and experimentally evaluate an automated system, called Filtered Wall (FW) which is able to filter unwanted messages from SNS user walls. We introduce a Machine Learning text categorization techniques to automatically assign with each short text message a set of categories based on its contentment. Efforts in creating a short text classifier are concentrated in the extraction and selection of a set of characterizing and discriminates features. Data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information. Data recovery software is one of a number of analytical tools for examine data. It allows users to study data from various different views and methods. Data extraction is the method of finding correlations or patterns among dozens of fields in large relational databases.

II. LITERATURE SURVEY

L. von Ahn, M. Blum, N. J. Hopper, and J. Landford proposed a Captcha that is used to solve the AI problems[1]. There are two types of Captcha: Text captcha and Image-Recognition Captcha [2]. Captcha-based Password Authentication (CbPA) Protocol is used for Captcha authentication [3]. Here the user need to input a valid pair of user ID and then the user is provided with the Captcha challenge to resolve. Captcha can also be used with the recognition-based graphical password to resolve the spyware problem. In this method a text Captcha is displayed below each image and user has to enter the characters at specific locations of the captcha below each image selected as his password during authentication.

R. Lin, S.-Y. Huang, G. B. Bell, and Y.-K. Lee proposed Recognition-based CaRP based on ClickAnimal of Captcha Zoo .The Captcha Zoo is a Captcha scheme which uses 3D models of horse and dog with different colors, poses,lightings etc. It uses passwords in a sequence of animal names where for each animal one or more3D models are built[4]. This method uses 2D animal with different textures and then the resulting 2D animals are arranged on different backgrounds. Some animals may closer to another animal in the image.

M. Chau and H. Chen proposed the Web surfing, it has become increasingly difficult to search for relevant information using search engines[5]. Topic-specific is an alternative way search engines that

support efficient information retrieval on the Web by providing more precise and customized searching in various domains. Developers of topic-specific search engines need to address two issues: how to locate relevant documents (URLs) on the Web and how to filter out irrelevant documents from a set of documents collected from the Web. Machine-learning-based approach is proposed which combines Web content analysis and Web structure analysis. Each Web page is represented by a set of content-based and link-based features which can be used for machine learning algorithms as the input. Implementation can be used as both a feed forward/back propagation neural network and a support vector machine. The experiments are designed and conducted to compare the proposed Web-feature approach with two existing Web page filtering methods - a keyword-based approach and lexicon-based approach. It has performed better than the benchmark approaches, when the number of training documents was small. This can be approached in Web applications such as Web content management and topic specific search engine .

Macro Vanetti, Elena Ferrari, and Moreno Carullo proposed a system that provides the user to have a straight rule over their own private wall to avoid the unwanted messages[6]. Here we provide users to have a straight control over messages posted on their own private space. Automated system called Filtered wall (FW) is used ,which can filter unwanted messages .The system can blocks only the unwanted messages send by the user. Drawback of the system is that the user will not be blocked; only the content posted by the user will block .The system supports content based message filtering and short text classification.

Mohamed Sylla proposed Combination Drag Pattern Graphical Password[7]. Here in the System one graphical keyboard is provided to user for selection of a password. During password selection the user has to choose set of characters from the graphical keyboard. The characters will be shown in textbox. User have to follow the sequence for creation of password. Then system checks password if it is not strong then system suggests different character between passwords. And to create a password user draws a pattern.

III. PROPOSED SYSTEM

Our system is based on Recognition Technique. In this three group of images are used, each group contains 25 images. User has to select at least one image from each group during registration phase. During login the user has to click on that images which is selected during registration phase. The system provide protection against attacks like shoulder surfing, dictionary and brute force attack using text password as well as graphical password. And also to propose and evaluate an automated system, called Filtered Wall (FW) , able to filter unwanted messages from SNS user walls. In Machine Learning text categorization techniques to automatically assign with each short text message a set of categories. The efforts in developing a strapping short text classifier (STC) are concentrated in the extraction and selection of a set of characterizing and unique features. So we introduce an extension of those adopted from which we inherit the learning model and the elicitation procedure for generating pre-classified data. FRs can support a variety of different filtering criteria that can be combined and customized according to the user requirement. Filtering Rules (FRs) utilizes user indicates, user associations as well as the creation of the Machine Learning categorization process to state the filtering criteria to be implemented.

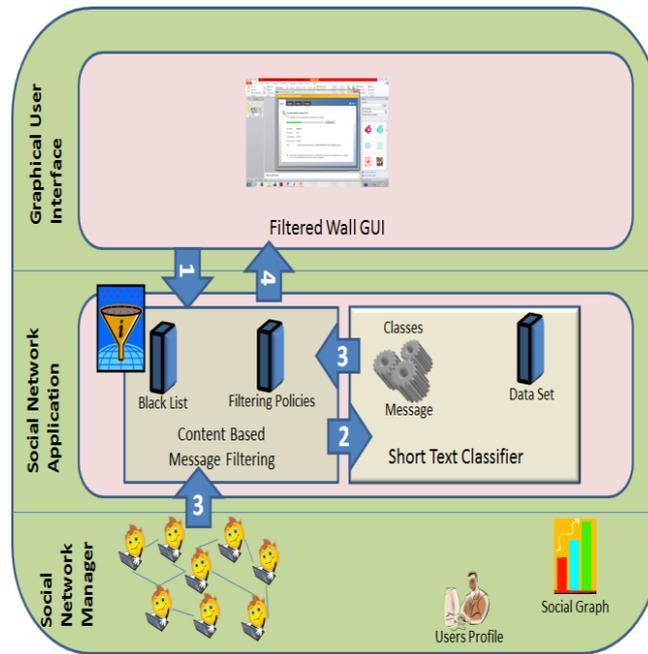


Figure 1: System architecture

The OSN service is a three-tier structure. The first layer called Social Network Manager (SNM) which aims to provide the basic SNS functionalities (like profile and relationship management). The second layer provides the support for external Social Network Applications (SNAs). The supported SNAs require an additional layer in turns for their Graphical User Interfaces (GUI). According to this architecture, the proposed system is placed in the second and third layers. The users interact with the system by means of a GUI to set up and manage their FRs and the GUI provides users with a FW. The core components of the proposed system are the Content-Based Messages Filtering (CBMF) and the Short Text Classifier (STC) modules. The first component exploits the message categorization provided by the STC module to enforce the FRs specified by the user.

The path followed by a user can be summarized as follows:

- 1) After entering the private wall of his/her contacts, the user tries to post a message, which is intercepted by FW.
- 2) A ML-based text classifier extracts metadata from the content of the message.
- 3) FW uses metadata provided by the classifier, together with data extracted from the social graph and user's profiles to enforce the filtering rules.
- 4) Depending on the result of the previous step, the message will be published or filtered by FW.

Advantages of proposed system:

- A system to automatically filter unwanted messages from SNS user walls on the basis of both message content and the message creator relationships and characteristics.
- The current paper extends for what concerns both the rule layer and the classification module.
- Most of the distinguishes include various features for filtering rules to better fit the considered domain.

IV. REFERENCES

- [1] L. Von ahn, M. Blum, N. J. Hopper, and J. Landford, “Captcha: Using hard AI problem for security”, in Proc.Eurocrypt, 2003, pp.294-311.
- [2] Ragavi V, Dr. G. Geetha “Captcha celebrating its Quattuordecennial – A Complete Reference”, Department of Computer Applications, Sathyabama University, Chennai, India.
- [3] B. Pinkas and T. Sander, “Securing passwords against dictionary attacks”, in Proc. ACM CCS, 2002, pp.161-170.
- [4] R. Lin, S. Y. Huang, G. B. Bell, and Y. K. Lee, “A new CAPTCHA interface design for mobile devices”, in Proc. 12th Austral. User Inter. Conf., pp. 3-8.
- [5] M. Chau and H. Chen, “A Machine Learning Approach to Web Page Filtering Using Content and Structure Analysis”, Decision Support System, vol.44, no.2, pp.482-492, 2008.
- [6] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, an Moreno Carullo, “A system to Filter Unwanted Messages from OSN User Walls”, 2013.
- [7] Mohammed Sylla, Muhammad, Kaleem Habib and Jamaludin, Ibrahim, “Combinatoric Drag-Pattern Graphical Password”, Journal of Emerging Trends in Computing Information Sciences, Vol.4, NO.12, Dec 2013.

