Image and Comment Privacy Using Watermarking and Text Classification in OSN Framework

K. Kousalya
Department of CSE
TRP Engineering College
Trichy

M. Ruba, S. Jesiema A. Arulmozhi, K.Geetha
Department of CSE,
TRP Engineering College,
Trichy.

***Abstract* - A social networking service is an online platform that is used by people to build social networks or social relations with another persons who are share their own details or career interests, activities, backgrounds or real-life connections. Social networking sites are varied and they incorporate a range of new information and various tools such as availability personal computers, mobile devices such as tablet computers and smart phones, digital photo/video/ sharing and web logging diary entries online (blogging). While Online Social Networks (OSNs) enable users to share photos easily, they also expose users to several privacy threats from both the OSNs and external entities. The current privacy controls on social networks are far from adequate, resulting in inappropriate flows of information when users fail to understand their privacy settings or OSNs fail to implement policies correctly. Social networks may be complicated because of privacy expectations when they reserve the right to analyze uploaded photos using automated watermarking technique. A user who uploads digital data such as image to their home page may wish to share it with only mutual friends, which OSNs partially satisfy with privacy settings. It solve the privacy violation problem occurred when images are published on the online social networks without the permission. Accordingto such images are alwaysshared after uploading process. Therefore, the digital image watermarking based on DWT co-efficient. Watermark bits are embedded in uploaded images. Watermarked images are shared in user home page. So images can be difficult to misuse by other persons. It provide protection approach to design the flexible policies for uploaded data. And also extend the work to implement information filtering approach to be used to give users the ability to automatically monitor the messages written on their own walls, by filtering out unwanted messages and comments about images. This concept can be implemented in real time for sending mobile intimation at the time of user in offline mode about negative comments. So user can easily guard the system from privacy violations.**

Keyword: Watermark-DWT, blocking, STC, filtering, mobile intimation.

I. INTRODUCTION

A social networking service (also social networking site, SNS or social media) is an online platform that is used by people to build social networks or social relations with other people who share similar personal or career interests, activities, backgrounds or real-life connections. The variety of stand-alone and built-in social networking services currently available in the online space introduces challenges of definition; however, there are some common features: (1) social networking services are Web 2.0 internet-based applications (2) user-generated content (UGC) is the lifeblood of SNS organisms, (3) users create service-specific profiles for the site or app that are designed and maintained by the SNS organization, and (4) social networking services facilitate the development of online social networks by connecting a user's profile with those of other individuals and/or groups. Most social network services are web-based and provide means for users to interact over the Internet, such as by e-mail and instant messaging and online forums.

Social networking sites are varied and they incorporate a range of new information and communication tools such as availability on desktop and laptops, mobile device such as tablet computers and smartphone ,digital photo/ video/sharing and "weblogging" diary entries online (blogging). Online community services are sometimes considered a social network service, though in a broader sense, social network service usually means an individual-centred service whereas online community services are group-centred. Social networking sites allow users to share ideas, digital photos, and videos, posts, and inform others about online or real world activities and events with people in their network. While in-person social networking, such as gathering in a village market to talk about events has existed since the earliest developments of towns, the Web enables people to connect with others who live in different locationsranging from across a city to across the world. Depending on the social media platform, members may be able to contact any other member. In other cases, members can contact anyone they have a connection to, and subsequently anyone that contact has a connection to, and so on. LinkedIn, a career social networking service, generally requires that a member personally know another member in real life before they contact them online. Some services require members to have a pre-existing connection to contact other members.

The main types of social networking services are those that contain category places (such as former school year or classmates), means to connect with friends (usually with self-description pages), and a recommendation system linked to trust. Social network services can be split into three types: socializing social network services are primarily for socializing with existing friends (e.g., Facebook); networking social network services are primarily for non-social interpersonal communication (e.g., LinkedIn, a career

and employment-oriented site); and social navigation social network services are primarily for helping users to find specific information or resources. There have been attempts to standardize these services to avoid the need to duplicate entries of friends and interests.

Online Social Networks (OSNs) have become part of daily life for millions of users. Users building explicit networks that represent their social relationships and often share a wealth of personal information to their own benefit. The potential privacy risks of such behavior are often underestimated or ignored. The problem is exacerbated by lacking experience and awareness in users, as well as poorly designed tools for privacy-management on the part of the OSN. Furthermore, the centralized nature of OSNs makes users dependent and puts the Service Provider in a position of power. Because Service Providers are not by definition trusted or trustworthy, their practices need to be taken into account when considering privacy risks. Aside from allowing users to create a network to represent their social ties, many OSNs facilitate uploading of multimedia content, various ways of communication and sharing many aspects of daily life with friends. People can stay in touch with (physically remote) friends, easily share content and experiences and stay up to date in the comfort of their own home or when on the move. However, benefits aside, potential threats to user privacy are often underestimated. For example, due to the public nature of many OSNs and the Internet itself content can easily be disclosed to a wider audience than the user intended. Users often have trouble revoking or deleting information, and information about a user might even be posted by others without their consent. Privacy in OSNs is a complicated matter and is not always intuitive to users, especially because it is not always similar to how privacy works in real-life interactions.

II. EXISTING SYSTEM

Social networking has been around for many years. People of all walks of life depend on Internet for obtaining various kinds of information. When sensitive information is disclosed that might be misused by unknown people. Moreover the security settings provided by social networks are inadequate. An inference attack is the attack used to obtain private and sensitive information from the known data. This can be prevented by proposing new sanitization techniques. And then implement graph based and risk model can be implemented for preserving privacy. In general, OSNs have three main types of entities: users, their connections, and the information that users are generating and diffusing. Each entity has its own characteristics. As the first kind of entity, online users can build connections with each other and can generate their own content, which leads to the emergence of the other two kinds of entities. For the second kind of entity, similar to one's daily social life, the connections among online users are usually topic-dependent and time-sensitive. People are posting images of their social events, gatherings, vacations, graduation ceremonies etc. These images not just include them and their families, but other people on the network too, and tagging them on these social networking websites is an unsolicited disclosure and privacy violations. Most of the content sharing websites have a set of privacy settings for the user to manage, but, unfortunately, these confidentiality system settings are not just adequate, especially with images. The reason is mostly the amount of information that is being carried by an image, essentially because of the unknown fact that if the image is even reliable or processed using some of the image processing software's.

DISADVANTAGES

- Only analyzed image privacy which are posted by users
- Fixed policies are used and limited privacy settings such as Public Post or Private post
- Private friends may be misuse the uploaded images

.

III. PROPOSED SYSTEM

Images on the social networks, execute three major security characteristics: Confidentiality, Integrity and Authenticity. In this project, we will implement watermarking approach to hide default pattern into image. Water mark bits are embedded into image. So unauthorized users only get watermark data only. Based in inverse DWT, we will get the seen water mark that can be restored into customary image. In the interface aspect, we will exchange the color of textual content pixels into color of photograph pixels. Person can set privacy settings to dam the pictures to down load by way of third parties. So unauthorized users most effective get watermark information handiest. Then utilizing disable options in mouse right click on and print reveal options. Snapshot privacy is maintained in social networks. Then using disable options in mouse right click and print screen options. Image privacy is maintained in social networks. Furthermore, the concept of blacklists and their administration are not believed by any of these access control models. The application of content-based filtering on messages posted on OSN user walls poses additional challenges given the short length of these messages other than the wide range of topics that can be discussed. Short text categorization has acknowledged up to now few attentions in the scientific community. This classifier will be used in hierarchical strategy. The first level task will be classified with positive and negative labels. The second level act as a negative, it will develop gradual membership. This grade will be used as succeeding phases for filtering process. Short text classifier includes text representation, machine learning based classification.

ADVANTAGES

- Provide privacy to uploaded images
- Complexity is less
- There is no predefined policies to images

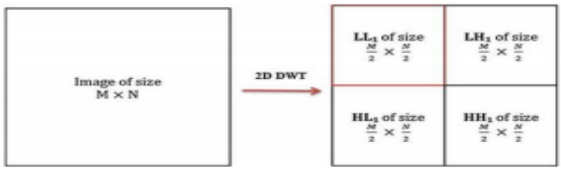
Can be implement in real time environments.

IV.ALGORITHM AND IMPLEMENTATION

Discrete Wavelet Transform:

Discrete Wavelet transform (DWT) is a mathematical tool for hierarchical decomposition of an image. The transformation is based on decomposing a signal into wavelets or small waves, having varying frequency and limited duration. The properties of wavelet decompose an original signal into wavelet transform coefficients which

contains the position information. The original signal can be reconstructed completely by performing Inverse Wavelet Transformation on these coefficients. DWT decomposes an image into sub images or sub bands, three details and one approximation. The bands are LL, LH, HL and HH.



The figure shows the sub bands in DWT. LL contains low frequencies both in horizontal and vertical direction. HH contains high frequencies both in horizontal and vertical direction. HL contains high frequencies in horizontal direction and low frequencies in vertical direction. LH contains low frequencies in horizontal direction and high frequencies in vertical direction. The low frequency part comprises of the coarse information of the signal while high frequency part comprises of the information related to the edge components. The LL band is the most significant band as it contains most of the image energy and represents the approximations of the image. Watermarks can be embedded in the high frequency detail bands (LH, HL and HH) as these regions are less sensitive to human vision. Embedding into these bands increases the robustness of the watermark without having additional impact on the quality of the image. At each level of decomposition, first DWT is performed in the vertical direction, followed by the DWT in the horizontal direction. The first level of decomposition yields four subbands: LL1, LH1, HL1, and HH1. The LL sub band of the previous level is used as the input for every successive level of decomposition. This LL sub-band is further decomposed into four multi resolution sub-bands to acquire next coarser wavelet coefficients. This process is repeated several times based on the application for which it is used. DWT has excellent spatio-frequency localization property that has been extensively utilized to identify the image areas where a disturbance can be more easily hidden. Also this technique does not require the original image for watermark detection. Digital image watermarking consists of two processes first embedding the watermark with the information and second extraction.

Watermark Embedding:

In this process 2D DWT is performed on the cover image that decomposes the image into four sub-bands: low frequency approximation, high frequency diagonal, low frequency horizontal and low frequency vertical sub- bands. Similarly 2D DWT is performed on the watermark image that has to be embedded into the cover image. Here we have used Haar wavelet. The technique used for inserting watermark is alpha blending. The decomposed components of cover image and watermark are further multiplied by a particular scaling factor and are added. During the embedding process the size of the watermark should be smaller than the cover image but the frame size of both the images should be made equal. The watermark embedded in this paper is perceptible or visible in nature, so we embedded it in the low frequency approximation component of the cover image.

Watermark Extraction

In this process the steps applied in the embedding process are applied in the reverse manner. First discrete wavelet transform is applied to both cover image and the watermarked image. After this the watermark is recovered from the watermarked image by using inverse discrete wavelet transform.

$R' \leftarrow$ WatermarkEmb(R,w, $kemb1$, $kemb2$, $kemb3$)

1. For each image $c \in R$

1) Divide c into $s \times s$ sized nonoverlapping blocks. Choose the low frequency blocks using DWT. The watermark is a sequence of binary bits denoted as $w = w1,w2, ...,wNw$. A set of blocks $\{BKi\}Nwi=1$ are chosen by a pseudorandom function as $kemb1$. Each block will carry one bit of the watermark.

2) For each watermark bit wi, $i \in [1, ...,Nw]$,

a) The pixels in block $BKi$are divided into two sets $S0$ and $S1$ according to a pseudorandom function with the water mark text $kemb2$;

b) If $wi = 0$, flip the bits of pixels in $S0$. Otherwise, flip the pixel bits in $S1$. In order to preserve the image quality, we make less flipping on higher bit-planes. We denote the ratios of flipped bits on 8 bit-planes as $\epsilon = [\epsilon 1, \epsilon 2, ..., \epsilon 8]$. That is to say, for the i-th bit-plane, there are $Nw \times s2 \times \epsilon i/2$ bits will be flipped randomly. The flipped positions are determined by $kemb3$ using Inverse DWT. Flip the water mark text color into image color.

2. Output the watermarked image set R'.

$wt \leftarrow$ WatermarkExtra(mt,mo, $kemb1$, $kemb2$, $kemb3$)

1. Divide mtinto nonoverlapping blocks with the size $s \times s$ using DWT.

2. Locate the set of blocks $\{BKi\}Nwi=1$ that carries the watermark bits$w = w1,w2, ...,wNw$according to the secret key $kemb1$.

3. For each $i \in [1,Nw]$,

1) Divide the pixels in BKiinto two sets $S0$ and $S1$ according to locations $kemb2$;

2) Flip the pixels in $S0$ and $S1$ respectively according to $[\epsilon i]8i=1$ and $kemb3$ to get two blocks $BK0i$ and $BK1i$. Construct the corresponding block $BKi$from the original image with the secret key $kemb1$. Calculate $\delta 0 = \Sigma pj \in BKi; p0j \in BK0i$ $(p0j\ pj)2$ and $\delta 1 = \Sigma pj \in BKi; p1j \in BK1i$ $(p1j - pj)2$. If $\delta 0 < \delta 1$, the watermark bit is extracted as '0'. Else, the watermark bit is extracted as '1'.

4. Output the extracted watermark wt.

Short text classification:

A hierarchical two level classification is advantageous to short text classification as per the suggestion. The first level of a classifier labels the message into neutral and non-neutral. In second level non neutral messages are estimated into one or more of the conceived categories.

Filtering rule - A filtering rule is a tuple (auth,CreaSpec,ConSpec,action)

1. auth is the user who state the rule.
2. CreaSpec is the Creator specification.
3. ConSpec is a boolean expression.
4. action is the action performed by the system.

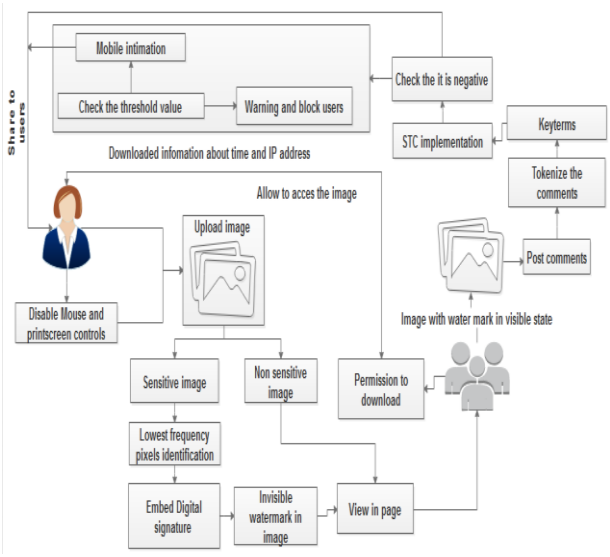
Filtering rules will be applied, when a user profile does not hold value for attributes submitted by a FR. This type of

situation will dealt with asking the owner to choose whether to block or notify the messages initiating from the profile which does not match with the wall owners FRs, due to missing of attributes.

Blacklist:

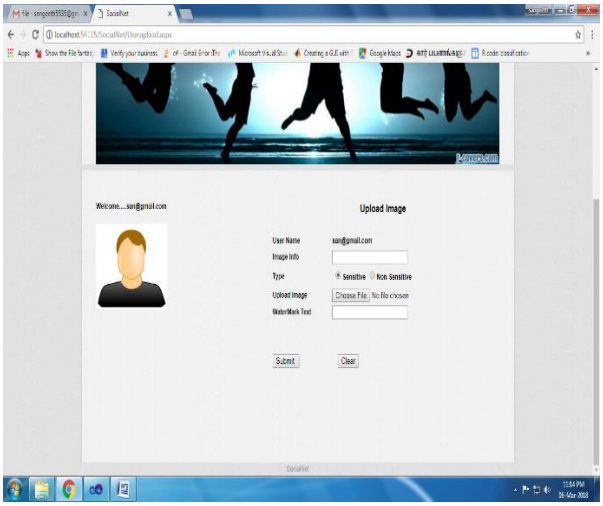
The main implementation of our paper is to execute the Blacklist Mechanism, which will keep away messages from undesired creators. BL are handled undeviating by the system. This will able to decide the users to be inserted in the blacklist. And it also decides the user preservation in the BL will get over. Set of rules are applied to improve the stiffness, such rules are called BL rules. By applying the BL rule, owner can identify which user should be blocked based on the relationship in OSN and the user's profile. The user may have bad opinion about the users can be banned for an uncertain time period. We have two information based on bad attitude of user. Two principles are stated. First one is within a given time period user will be inserted in BL for numerous times, he /she must be worthy for staying in BL for another sometime. This principle will be applied to user who inserted in BL atleast once. Relative Frequency is used to find out the system, who messages continue to fail the FR. Two measures can be calculated globally and locally, which will consider only the message in local and in global it will consider all the OSN users walls.

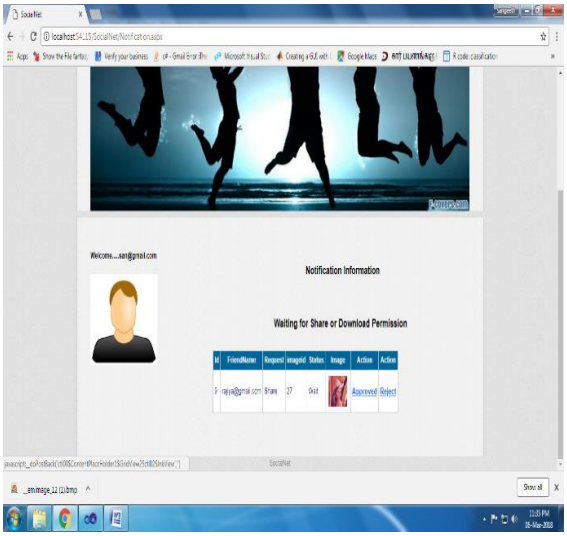
V. ARCHITECTURE DIAGRAM

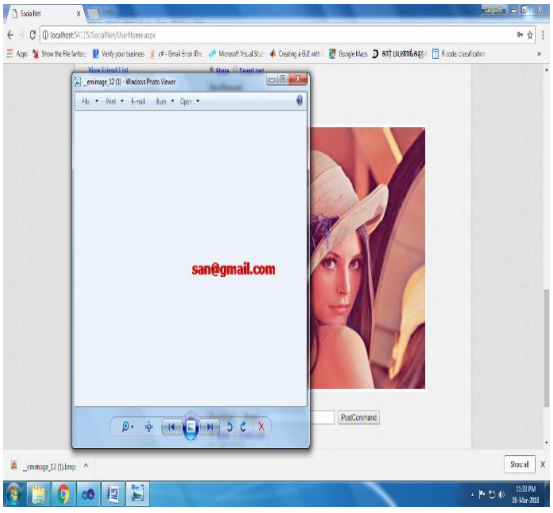


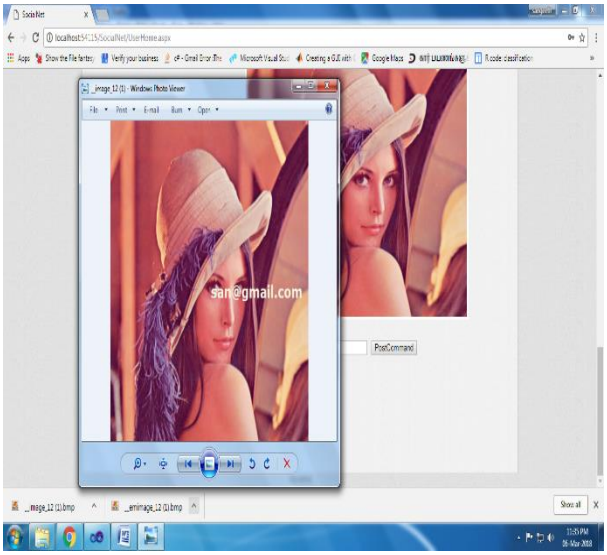
The user uploads the image in own home page, it prompts whether it should be sensitive or non-sensitive. If non sensitive it will be viewed as it is, if it's sensitive then low pixel frequency is identified and the digital signature is embedded into the image .An image with invisible watermark will be viewed by their authorized users(friends).when it is in non-sensitive state the authorized user can download the image as it is. If the image is in sensitive state then the authorized user need to take the permission from the user, if accepted they can download else not. When the unauthorized user tries to download ,they will get only the watermark and the print screen and mouse save option will be disabled. Authorized user can comment for the post, the comment is tokenized into keywords and checked if it is negative with STC algorithm. If the negative comment by a single user exceeds the threshold value the user will be blocked automatically. Mobile intimation of the IP address who have downloaded and the user who has been blocked is also intimated to the user.

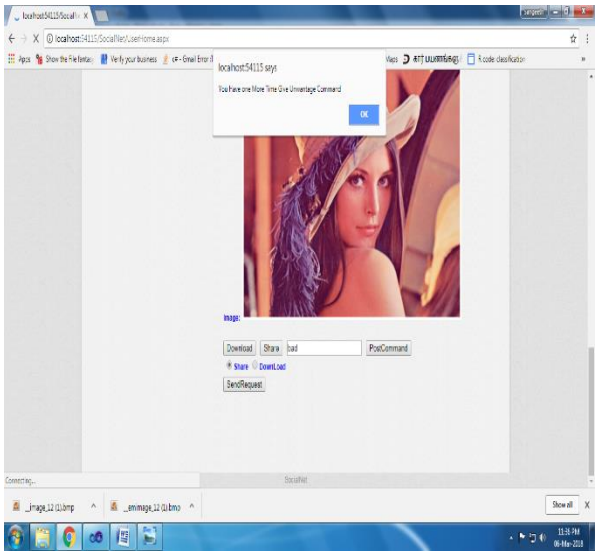
VI. IMPLEMENTATION











VII. CONCLUSION

The apperence of well-known online social networking has triggered within the compromise of conventional notions of privateness, certainly in visual media. With a view to facilitate useful and principled protection of picture privateness online, we have got supplied the design, implementation, and evaluation of photo shield gadget that successfully and successfully protects client's photo privateness across famous OSNs. The digital watermarking approach based fully on DWT coefficients modification for social networking offerings has been presented on this paper. In the embedding manner, the coefficients in LL sub-band had been used to embed watermark. Within the extraction process, normal coefficient prediction based on imply clear out is used to boom the accuracy of the extracted watermark. On extending the Machine Learning (ML) text categorization techniques to automatically assign with each short text message a set of categories based on its content. Then exploiting a flexible language to specify Filtering Rules (FRs), by which users can state what contents, should not be displayed on their walls. FRs can support a variety of different filtering criteria that can be combined and customized according to the user needs.As part of future work, to implement cryptographic techniques and various filtering techniques to secure OSN home page. And also extend the work in privacy based uploaded video content sharing sites. The experimental outcome confirmed a larger overall efficiency in specific time application.

REFERENCES

[1] H. Cheng, X. Zhang, J. Yu, and F. Li, "Markov process based retrieval for encrypted jpeg images," in Proc. of 10th International Conference on Availability, Reliability and Security. IEEE, 2015, pp. 417–421.

[2] A. Rial, M. Deng, T. Bianchi, A. Piva, and B. Preneel, "A provably secure anonymous buyer–seller watermarking protocol," Information Forensics and Security, IEEE Transactions on, vol. 5, no. 4, pp. 920–931, 2010.

[3] J. Zhang, Y. Xiang, W. Zhou, L. Ye, and Y. Mu, "Secure image retrieval based on visual content and watermarking protocol," The Computer Journal, vol. 54, no. 10, pp. 1661–1674, 2011.

[4] T. Bianchi and A. Piva, "Secure watermarking for multimedia content protection: A review of its benefits and open issues," IEEE Signal Processing Magazine, vol. 30, no. 2, pp. 87–96, 2013.

[5] A. Piva, T. Bianchi, and A. De Rosa, "Secure client-side st-dm watermark embedding," Information Forensics and Security, IEEE Transactions on, vol. 5, no. 1, pp. 13–26, 2010.

[6] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Advances in Cryptology-Eurocrypt. Springer, 2004, pp. 506–522.

[7] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. of IEEE Symposium on Security and Privacy. IEEE, 2000, pp. 44–55.

[8] E.-J. Goh et al., "Secure indexes." IACR Cryptology ePrint Archive, vol. 2003, p. 216, 2003.

[9] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in Proc. of 13th ACM conference on Computer and communications security. ACM, 2006, pp. 79–88.

[10] J. Shen, H. Tan, J. Wang, J. Wang, and S. Lee, "A novel routing protocol providing good transmission reliability in underwater sensor networks," Journal of Internet Technology.