IMAGE AND COMMENT PRIVACY BY USING WATERMARKING TECHNIQUES

K.Karthik[1], K. Ajith Kumar[2], B. Surya[3], K. Kesavan[4], G. Kavin Bharath[5]

[1]Asst. Professor, V. S. B Engineering College, Karur, TamilNadu- 639111.

[2,3,4,5] UG Scholars, V. S. B Engineering College, Karur, TamilNadu- 639111.

Abstract:

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. Social networking sites are varied and they incorporate a range of new information and communication tools such as availability on desktop and laptops, mobile devices such as tablet computers and smartphones, 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. Image over the social network is transferred or transmitted between servers and multiple users. Privacy of that data is very important as it belongs to personal sensitive information. In existing system, text based encryption can be implemented in social networks. There are many different approached of storing data securely over the social networks, using big data such as end-to-end encrypted data transmission, dynamic credential generation only for text data. In this project, can introduce a novel watermarking scheme with wavelet algorithm named as discrete wavelet transform in real time social network application. In this scheme can use images and stored in server in secure format. And also extend the project, categorize the picture as sensitive or normal. If it is sensitive means, perform copyrights algorithms. Then provide the permission to the receiver end for download the images in secure manner. And also implement protection controls to block mouse operations and print screen options. Then 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.

1 INTRODUCTION

A social networking sites could be a new world to create social relations among folks that share data like text, image, videos, events, interests, backgrounds or daily-life connections. Communications over the Social Networks aren't secure. The Social networking sites are Facebook, Google plus, Linked-In, etc [1]. Many attacks and violation of privacy are recently faced in our most popular networking sites. We use the social networking sites for chatting with our friends and sharing digital data like text, images, video and etc. When we share a digital data to our friends; the information may face several attacks from the attackers and/or unauthorized users. For example, Alice needs to share an image with Alice's friends Bob, Don, etc. During this communication alternative accredited users or third parties shouldn't be concerned. Any unauthorized users (like, those are not friends with Alice or attacker) makes an attempt to attack a communication, that is trying to access the image for editing or misusing. The attacker's ultimate aim is to make crime using the private digital data from social networking sites. The attacker tries to attack the communication in many ways ie, violate the privacy, data attacking from the servers, etc. So, our aim is to protect our extremely private, confidential or secret data from unauthorized users [2]. Here, privacy protection is an important issue of many social networking sites. And our work using Reversible Data Hiding (RDH) Techniques, goes to attain its importance attributable to the exponential growth and secret communication of potential user over the web.

Digital Watermarking is a technology of embedding watermark with intellectual property rights into images, videos, audios and other multimedia data by a certain algorithm. This kind of watermark contains the author and the user's information, which could be the owner's logo, serial number or control information. In fact, It's making use of the ubiquitous redundancy and randomness in data, and adding to the data information which is difficult to be detected but can be distinguished to protect product copyright and data integrity. Finally, the watermarks will have exactly the same transformation experience as the works, that means you can get the information of transformation by looking at the watermarks.

1.1 Level of Social networks

In general, social networks are self-organizing, mergent, and complex, such that a globally coherent pattern appears from the local interaction of the elements that make up the system. These patterns become more apparent as network size increases. However, a global network analysis of, for example, all interpersonal relationships in the world is not feasible and is likely to contain so much information as to be uninformative. Practical limitations of computing power, ethics and participant recruitment and payment also limit the scope of a social network analysis. The nuances of a local system may be lost in a large network analysis, hence the quality of information may be more important than its scale for understanding network properties. Thus, social networks are analyzed at the scale relevant to the researcher's theoretical question. Although levels of analysis are not necessarily mutually exclusive, there are three general levels into which networks may fall: micro-level, meso-level, and macro-level.

1.1.1 Micro level:

At the micro-level, social network research typically begins with an individual, snowballing as social relationships are traced, or may begin with a small group of individuals in a particular social context.

1.1.2 Dyadic level:

A dyad is a social relationship between two individuals. Network research on dyads may concentrate on structure of the relationship (e.g. multiplexity, strength), social equality, and tendencies toward reciprocity/mutuality.

1.1.3 Triadic level:

Add one individual to a dyad, and you have a triad. Research at this level may concentrate on factors such as balance and transitivity, as well as social equality and tendencies toward reciprocity/mutuality. In the balance theory of Fritz Heider the triad is the key to social dynamics. The discord in a rivalrous love triangle is an example of an unbalanced triad, likely to change to a balanced triad by a change in one of the relations. The dynamics of social friendships in society has been modeled by balancing triads. The study is carried forward with the theory of signed graphs.

1.1.4 Actor level:

The smallest unit of analysis in a social network is an individual in their social setting, i.e., an "actor" or "ego". Ego network analysis focuses on network characteristics such as size, relationship strength, density, centrality, prestige and roles such as isolates, liaisons, and bridges. Such analyses are most commonly used in the fields of psychology or social psychology, ethnographic kinship analysis or other genealogical studies of relationships between individuals.

1.1.5 Subset level:

Subset levels of network research problems begin at the micro-level, but may cross over into the meso-level of analysis. Subset level research may focus on distance and reachability, cliques, cohesive subgroups, or other group actions or behavior.

1.1.6 Meso level:

In general, meso-level theories begin with a population size that falls between the micro- and macro-levels. However, meso-level may also refer to analyses that are specifically designed to reveal connections between micro- and macro-levels. Meso-level networks are low density and may exhibit causal processes distinct from interpersonal micro-level networks.

1.1.7 Macro level

Rather than tracing interpersonal interactions, macro-level analyses generally trace the outcomes of interactions, such as economic or other resource transfer interactions over a large population.

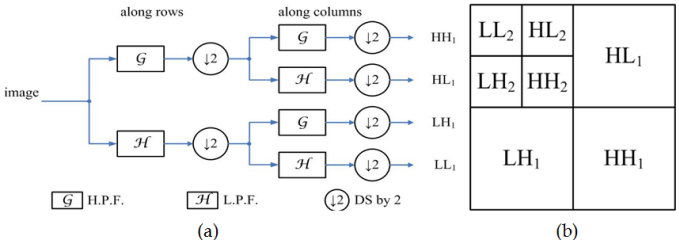


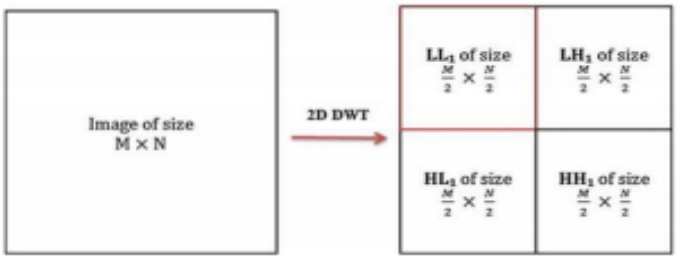
Fig. 2D DWT decomposition of an input image using filtering approach.

2 WATER MARKING TECHNOLOGY

Due to the recent development in internet technology, redistribution of digital content have become very easy. It leads to the powerful need of safe and authentic environment for the processing of digital content. This drawback can be overcome by using watermarking technology. Digital Image Watermarking embeds distinctive information in a picture in such a way that it cannot simply be removed. The original (host) image which is used for the embedding confidential data is called the cover image. The watermarking methodology is fruitful if it is indiscernible and sturdy to usual image distortions like cropping scaling, shearing, filtering etc. Digital watermarking is viewed as an effective way to deter content users from illegal distributing. In essence, watermarking intentionally embeds digital information into the software for purposes such as identification and copyright. Such information could be the author's name, company name or other messages highly related to the owner and/or the legal users of the software. If necessary, this information can be used in court to prove authorship of the software or proof of legal users entitled to distribute copies. Therefore, proposed properties are shown that for watermarked media several requirements must be satisfied. According to the toughness of the watermark embedding, the invisible watermark is classified in three types i.e. the robust, the semi fragile and the fragile watermark. In robust watermarking technique, the watermark is embedded with very high embedding strength so that it can withstand a huge number of intentional or unintentional attacks. In fragile watermarking technique, the watermark embedding is done with a very low embedding strength, therefore it is damaged even with a very light amount of distortion . The semi-fragile watermark lies in the middle of the robust and fragile watermark. Watermarking technique can be divided into two main groups: spatial domain watermarking and frequency domain watermarking. Techniques that work in spatial domain can suffer from signal compression and hostile attacks. Frequency domain techniques are much more robust against compression and geometrical transformations than spatial domain techniques. The frequency domain watermarking approaches are the most popular for robust image watermarking. In these schemes, the image is transformed via some common frequency transform.

2.1 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.

3 SYSTEM ANALYSIS

3.1 EXISITNG 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.

3.2 LIMITATIONS:

- Only analysed 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
- Difficult to predict misbehave users

3.3 PROPOSED SYSTEM:

Images on the social networks, execute three major security characteristics: Confidentiality, Integrity and Authenticity. Confidentiality means that only the entitled persons have the access to the particular images, hence tagging.

- Integrity means the picture has not been modified by non-authorized person.
- Authenticity is the proof that image has indeed the exact people as shown, or is a modified version using the various image processing softwares.

The increment in the development and use of software image editors has accompanied the increase in the tampering of these

basic characteristics. Above all, the flourishing use of social networks has made the sharing and distribution of images pretty convenient. The integrity and authenticity is the compelling question as, among other fields, these images are also being used as evidence in the courts of law. It is very critical to verify the integrity of these images and is often desirable to identify if an image has been manipulated from the time of recording. To understand, how things go in the background of a jpeg image, 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. So photo may also be considered as undeniable content. Person can set privateness 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 of screenshots in interface system. We use a related idea to classify the users to which a filtering rule applies. For itself, one of the key elements of our scheme is the availability of an explanation for the message contents to be exploited by the filtering mechanism as well as by the language to express filtering rules. In distinguish no one of the access control models previously cited exploit the content of the resources to enforce access control. We consider that this is an essential difference. 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. Aim of the short text classifier is to recognize and eradicate the positive sentences and categorize the negative sentences in step by step, not in single step. 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.

3.3.1 ADVANTAGES:

- Provide privacy to uploaded images
- Complexity is less
- There is no predefined policies to images
- Can be implement in real time environments

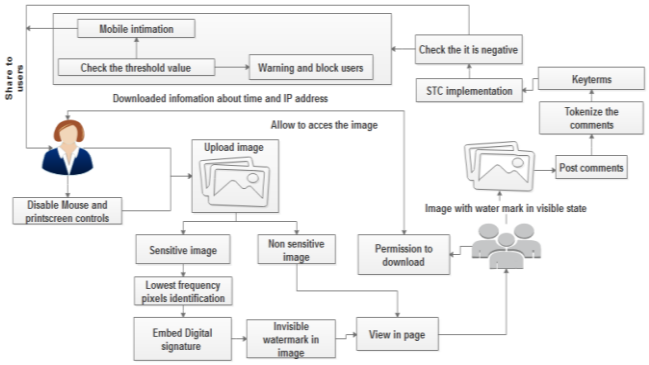


Fig. Block diagram

4 SYSTEM IMPLEMENTATION

4.1 MODULES:

- Social network creation
- Upload image
- Embed the watermark
- Privacy settings
- Protection system
- STC Implementation
- Filtered rules implementation

4.2 Social network creation:

Social network refers to interaction among people in which they create, share, and/or exchange information and ideas in virtual communities and networks. In this module, we can have three types of users such as image owner, image users and image server. Image owner can be upload the image into system and image server stores the images in database. Image users use images which are shared by image owner. We can social network application as android application for image owner. Server page can be designed as PHP page.

4.3 Upload image:

The first stage of any sharing system is the image acquisition stage. In this module, we can upload various images such as natural images, face images and other images. Uploaded images can by any type and any size. In this module, specify the image as sensitive or non-sensitive image. Sensitive image is referred as personal image. Non-sensitive image can be referred as forwarded image.

4.4 Embed the watermark:

In this module, we can embed the watermark text into images. Watermarking ensures authenticating ownership, protecting hidden information, prevents unauthorized copying and distribution of images over the internet and ensures that a digital picture has not been altered. We can implement Discrete Wavelet Transform (DWT) domain image watermarking system for real time image. In the embedding process, the watermark may be encoded into the cover image

using a specific location. This location values is used to protect the images. The output of the embedding process, the watermarked image, is then transmitted to the OSN home page.

4.5 Privacy settings:

Each user images are first categorized into privacy policy. Then privacy policies of each images can be categorized and analyzed for predict the policy. So we adopting two stages approach for policy recommendation than applying the common one-stage data mining approaches to mine both image features and policies together. The two-stage approach allows the system to employ the first stage to classify the policy as with privacy or without privacy. In the second stage, we can set without privacy means, prefer the user list details.

4.6 Protection system:

In this module, we can set the protection or blocking system to avoid third party aces without knowledge of image owners. This module is used to set the image with privacy. If user set with privacy settings means, all users are considered as third parties. Based on this setting, unauthorized user only views the image and can't be used. If he downloads means, only get water mark values. Finally provide hardware control system such as screenshot controls. Then disable the screenshot options. Device controls values are extracted and to provide coding implementation to disable the coding at the time protection. We can implement this concept in all browsers.

4.7 STC implementation:

In this module, we designan automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls. The architecture in support of OSN services is a three-tier structure. The first layer commonly aims to provide the basic OSN functionalities (i.e., profile and relationship management). Additionally, some OSNs provide an additional layer allowing the support of external Social Network Applications (SNA). Finally, the supported SNA may require an additional layer for their needed graphical user interfaces (GUIs). The major efforts in building a robust short text classifier (STC) are concentrated in the extraction and selection of a set of characterizing and discriminant features. In order to specify and enforce these constraints, we make use of the text classification. From STC point of view, we approach the task by defining a hierarchical two-level strategy assuming that it is better to identify and eliminate "neutral" sentences, then classify "non-neutral" sentences by the class of interest instead of doing everything in one step.

6.1.7 Filtered rules implementation:

The filtering rules should allow users to state constraints on message creators. Thus, creators on which a filtering rule applies should be selected on the basis of several different criteria; one of the most relevant is by imposing conditions on user profile's attributes. In such a way it is, for instance, possible to define rules applying only to young creators, to creators with a given religious/ political view, or to creators that we believe are not expert in a given field (e.g. by posing constraints on the work attribute of user profile). This means filtering rules identifying messages according to constraints on their contents. And block the users who are post the negative comments more than five times and also send mobile intimation to users at the time offline.

5 CONCLUSION

The appearance 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] M. Cheung, J. She, and Z. Jie, "Connection discovery using bigdata of user-shared images in social media," Multimedia, IEEE Transactions on, vol. 17, no. 9, pp. 1417–1428, 2015.

[2] M. Cheung and J. She, "Evaluating the privacy risk of user-shared images," ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 12, no. 4s, p. 58, 2016.

[3] M. Cheung, J. She, and X. Li, "Non-user generated annotation on user shared images for connection discovery," in 2015 IEEE International Conference on Data Science and Data Intensive Systems.IEEE, 2015, pp. 204–209.

[4] M. Douze, H. J´egou, H. Sandhawalia, L. Amsaleg, and C. Schmid, "Evaluation of gist descriptors for web-scale image search," in Proceedings of the ACM International Conference on Image and Video Retrieval. ACM, 2009, p. 19.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.

[6] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, “Return of the devil in the details: Delving deep into convolutional nets,” arXiv preprint arXiv:1405.3531, 2014.
[7] Y. Jia, E. Shelhamer, J. Donahue, S.Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional architecture for fast feature embedding,” in Proceedings of the ACM International Conference on Multimedia. ACM, 2014, pp. 675–678.
[8] E. M. Jin, M. Girvan, and M. E. Newman, “Structure of growing social networks,” Physical review E, vol. 64, no. 4, p. 046132, 2001.
[9] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, “Measurement and analysis of online social networks,” in Proceedings of the 7th ACM SIGCOMM conference on Internet measurement. ACM, 2007, pp. 29–42.
[10] J.-D. Zhang and C.-Y. Chow, “igslr: personalized geo-social location recommendation: a kernel density estimation approach,” in Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM, 2013, pp. 334–343