IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

Tracing Origin of Social Media Posts

Mr. Sabin TT[1], Anupama V[2], Bharathi N[3], Deekshitha R[4], Nithya Shree V N[5]
Assistant Professor, Department of Information Science and Engineering[1]
Students Department of Information Science and Engineering[2,3,4,5]
SJC Institute of Technology Chikkballapur, Karnataka, India

Abstract: *This problem statement should result in a solution that can identify perpetrators who first posted sexually explicit abuse content on social media platforms. Images and videos of individuals of a sexually explicit nature are quite often posted online intentionally by perpetrators on various social media platforms with the sole intent of causing harassment, humiliation and distress to the victim. Social media platforms are constantly trying to create and improve automatic mechanisms to identify and reject such content at the time of upload itself. However, there are times when such content containing nudity and of a sexually explicit nature, fails to be detected at the time of upload and gets published. Such content usually has a tendency to instantly become viral, and others thereafter share it, or download and post and it again from their social media accounts. Social media platforms take down such posts when reported but by the time this is done, several copies have already been made and go in circulation. It quickly becomes ambiguous as to who posted it first, and taking advantage of this ambiguity, the original perpetrator evades detection. It is therefore critical to identify the person who first posted such distressing content online. Given a piece of text, image or video snippet as input, build a solution that can identify the person who was the first one to post it online on a particular social media platform. Please bear in mind that people could have copied it and made minor modifications before re-posting it from their accounts. Participants are expected to obtain suitable data required to work on this problem statement on their own.*

Keywords: Social Media

I. INTRODUCTION

The global connectivity and the increasing capability of the hardware and technology has helped the world connect over a network. Sharing images, documents, videos and other data is now just a click away on social media and online platforms. Data and information is now available at ease, but with pros there are always cons. With the facility to share images, documents, videos within lesser time, there is always a risk of fake graphics and media being shared globally and that too at rapid paces. The sharing a viral or fake image creates a huge risk and problems in the societies, resulting into hatred, violence and other severe consequences. The law enforcements and investigators require a huge support from the social media platforms to help them identify the source of fake images being sent over the platform.

1.1 Social Media and Social Networking

Television channels, Papers, Articles, Magazine etc. were the media before the existent of internet. With the availability and development of World Wide Web, the definition of media had a drastic change. Capability of interaction with mass and in-person was available to everyone within the network and connection. Social media basically is a very wide domain containing various types of media and documents; such as videos, vlogs, blogs, pages etc. It is a transportation medium and a great platform for communication if used ethically and in a standard manner. Social networking, is another such domain. It is a platform to connect to other people, community and a wide group of people. Social networking is a sub-domain of social media, where a user profile is created to interact and share their day-to-day activities, thoughts, views about religion, politics, nation, discoveries and other such domains.

1.2 Metadata

Metadata is the “data describing the data”. Meta which is meant as "description". Metadata can be very much helpful in finding various details and data related to the images, videos, webpages, documents, etc. For example, the creator, date/time of creation, type of document, extensions, versions etc. Filtering and extracting these data can help many

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

investigators and agencies to trace down and get conclusive evidences. Metadata and tags can also be used for search engine optimization and listings.

Metadata can be generated manually as desired by codes or manipulation, or by automated information processing. Manual addition is more accurate, allowing the user to add any data they feel is needed to describe the file. Automated metadata creation is primary, usually only containing data such as size, extension, creation time and creator.

II. METHODOLOGY

Here, the unique proposed method is described graphically and theoretically for social media sites and networks, which can be adapted to trace out the origin of the user who shared a particular image on the network based on the credentials used to login on the platform. The main idea here is to use credential or signup information (Mobile Number or E-mail address) of the user. The mobile number or E-mail ID used as credential or sign-up information is always unique and can be helped to identify a user. Now the service provider has a trace of the credentials, which is not publically displayed. The hash value of the credential of the user should be stored in the data base of the service provider or the social media company. The hashing type MD5 should be used as it is fast and reliable creating a 16-digit irreversible unique hash value of the unique credential information. Now, whenever a user login into the app the credentials should be stored and hashed on local system. After login, while using the application if the user wishes to upload any image from the device and share to others; the hash generated by the credentials should be added as creator meta tag in the image as shown in the experimental sections. If the user wishes to forward or share the image received on the platform itself, then the image should be forwarded with the existing meta tags. Thus, meta tag named creator should be generated and edited in an image if the user uploads from the device, else only forwarding / sharing within the app keeps the meta tags as previous; which serves the purpose of keeping the creator tag as that of the user who uploaded the image on the platform. The pictorial representation is shown in following diagram:

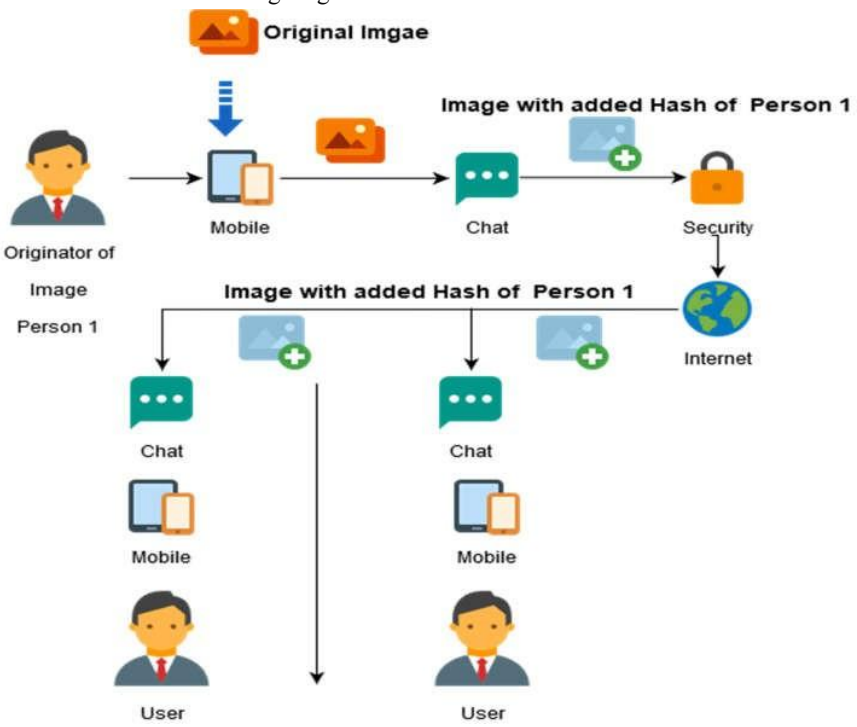

Figure 1: Sample image sharing scenario with proposed mechanism

The image shows a scenario where user 1 login into the application or social media platform using his mobile phone, the credentials are mobile number and password. The user wishes to upload and send an image over social media network; thus when he uploads the image the image's metadata is updated by adding a Meta tag of creator with hash value of users mobile number, registered during sign-up. The image is then transferred over the internet adding its other security features and processes. The receiver receives the image which has the metadata of the user who initially

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

uploaded the image. Here, no changes would be made in the Meta tagcreator, if the recipient forwards or share the image from the platform itself. Thus the image always contains the Meta tag of the user who uploads an image. Later sections shows the experimental tests done on images to add the Meta tags.

2.1 Flowchart and Algorithm

Here, the basic flowchart of the process and an algorithm based on the method is showwith variables as follows:

- Login credential unique mobile number or E-mail ID: ***a***
- New Image to be uploaded: ***x***

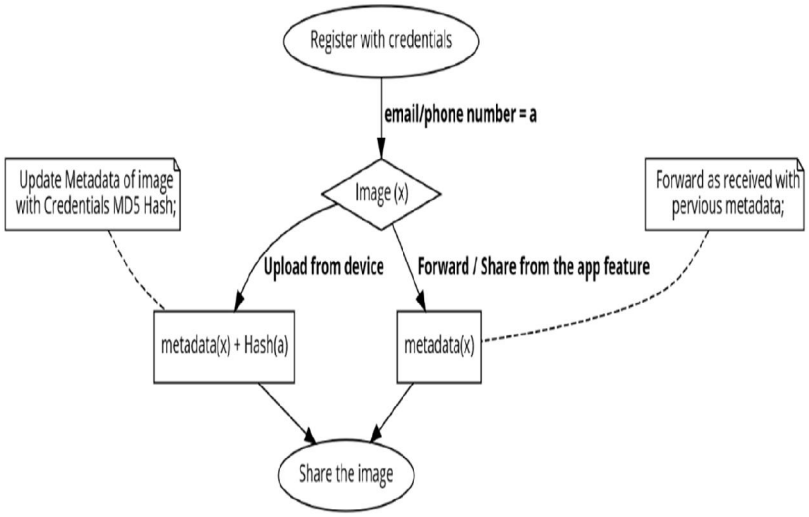

Figure 2. Flowchart for proposed mechanism

Algorithm:

Start

Register with credentials [email/phone number = a];If (Image (x)) {

// Update Metadata of image with Credentials MD5 Hash;

Metadata(x) + Hash (a);

}Else{

//Forward as received with pervious metadata;

Metadata(x);

}

Share the image;

End

2.2 Use of MD5 Hash And Credentials

Here, MD5 is basically used as MD5 hashing as a security feature to keep the credentialshidden, even if the hash is fetched, the irreversible feature of hashing will not allow the unauthorised persons to know the credentials. It becomes easy and less space consuming to store small hashes.

Now, credentials are always unique, Mobile number or email id are never allotted same; username, address, id, password, device type, browser, time cannot be used as identifier as they might be same in cases or they are constantly changing in nature, thus using credentials as the identifier for metadata seems to be the optimal solution.

III. CONCLUSION

The main conclusion of the technique is to find out the originator of the image based on unique identifier. With the increase in number of users, if MD5 hashes seems to collide, then other SHA based algorithms can be used depending

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

on the storage and processing space. More security features can also be added with XOR or logical operations with third party keys for users to mark the identification of images or documents transferred over the social network.

REFERENCES

[1]. Marra, Francesco & Gragnaniello, Diego & Cozzolino, Davide & Verdoliva, Luisa. (2018). Detection of GAN-Generated Fake Images Over Social Networks. 384-389. 10.1109/MIPR.2018.00084.

[2]. Wu, Bing & Zhang, Chenyan. (2016). Enterprise Social Media Research. 10.2991/aiie- 16.2016.123.

[3]. Lubna, Jahanara & Chowdhury, S.. (2020). Detecting Fake Image: A Review for Stopping Image Manipulation. 10.1007/978-981-15-3666-3_13.

[4]. Gonzalez-Marquez, Monica & Bergmann, Christina. (2020). Meta-Data Inventory Project description (in progress). 10.13140/RG.2.2.31714.48323.

[5]. Kester, Quist-Aphetsi & Senkyire, Isaac. (2019). Validating of Digital Forensic Images Using SHA-256. 118-121. 10.1109/ICSIoT47925.2019.00028.

[6]. Quick, Darren & Choo, Kim-Kwang Raymond. (2016). Big forensic data reduction: Digital forensic images and electronic evidence. Cluster Computing. 19. 10.1007/s10586- 016-0553-1.

[7]. Kessler, Gary. (2016). The Impact of SHA-1 File Hash Collisions on Digital Forensic Imaging: A Follow-up Experiment. Journal of Digital Forensics, Security and Law. 10.15394/jdfsl.2016.1433.