Privacy-Preserving Image Retrieval and Sharing in Social Multimedia Applications

Kavadi Geetha

Department of Computer Science, Sri Venkateswara University, Tirupati

Abstract— *Every day social multimedia applications generate millions of images. To handle such huge amount of images, an optimal solution is using the public cloud, since it has powerful storage capability. Images usually contain a wealth of sensitive information, therefore social service providers need not only to provide services such as retrieval and sharing but also to protect the privacies of the images. In this paper, we propose a privacy-preserving scheme for content-based image retrieval and sharing in social multimedia applications. First, the users extract visual features from the images, and perform locality-sensitive hashing functions on visual features to generate image problem vectors. We then model the retrieval on the images as the equality search on the image problem vectors. To enable accurate and efficient retrieval, we design the secure index structure based on cuckoo hashing, which has constant lookup time. To meet the requirements of dynamic image updating, we enrich our service with image insertion and deletion. In order to reduce the key management overhead and the access control overhead in social applications, we process keys using secret sharing techniques to enable the users holding similar images to query and decrypt images independently. Finally we implement the prototype of the proposed scheme, and perform experiments over encrypted image databases.*

Keywords: Image retrieval, image sharing, multimedia, privacy-preserving.

I. INTRODUCTION

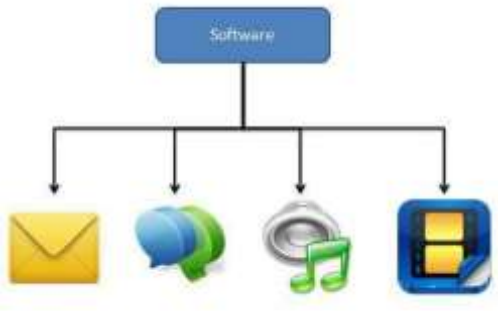
Due to the popularity of mobile devices with cameras, such as mobile phones, tablets, sensors, and etc., the number of images has grown tremendously. Specially, social multi-media applications, that provide platforms to post and share multimedia, generate massive amounts of images. According to Instagram, more than 100 million images are posted per day. Due to the high storage costs, social service providers prefer to outsource such massive number of images to public cloud platforms such as Amazon Cloud. Images usually contain sensitive information that could reveal personal privacies, and encryption is an effective method to protect privacy. Based on cryptography, a number of privacy-preserving schemes

1.1 Multimedia

Multimedia is an interactive media and provides multiple ways to represent information to the user in a powerful manner. It provides an interaction between users and digital information. It is a medium of communication. Some of the sectors where multimedia's is used extensively are education, training, reference material, business presentations, advertising and documentaries.

By definition Multimedia is a representation of information in an attractive and interactive manner with the use of a combination of text, audio, video, graphics and animation. In other words we can say that Multimedia is a computerized method of presenting information combining textual data, audio, visuals (video), graphics and animations. For examples: E-Mail, Yahoo Messenger, Video Conferencing, and Multimedia Message Service (MMS).

Multimedia as name suggests is the combination of Multi and Media that is many types of media (hardware/software) used for communication of information.



1.2 Components of Multimedia

Following are the common components of multimedia:

- **Text**- All multimedia productions contain some amount of text. The text can have various types of fonts and sizes to suit the profession presentation of the multimedia software.
- **Graphics**- Graphics make the multimedia application attractive. In many cases people do not like reading large amount of textual matter on the screen. Therefore, graphics are used more often than text to explain a concept, present background information etc. There are two types of Graphics:
    - **Bitmap images**- Bitmap images are real images that can be captured from devices such as digital cameras or scanners. Generally bitmap images are not editable. Bitmap images require a large amount of memory.
    - **Vector Graphics**- Vector graphics are drawn on the computer and only require a small amount of memory. These graphics are editable.
- **Audio**- A multimedia application may require the use of speech, music and sound effects. These are called audio or sound element of multimedia.Speech is also a perfect way for teaching. Audio are of analog and digital types. Analog audio or sound refers to the original sound signal. Computer stores the sound in digital form. Therefore, the sound used in multimedia application is digital audio.
- **Video**- The term video refers to the moving picture, accompanied by sound such as a picture in television. Video element of multimedia application gives a lot of information in small duration of time. Digital video is useful in multimedia application for showing real life objects. Video have highest performance demand on the computer memory and on the bandwidth if placed on the internet. Digital video files can be stored like any other files in the computer and the quality of the video can still be maintained.
- **Animation**- Animation is a process of making a static image look like it is moving. An animation is just a continuous series of still images that are displayed in a sequence. The animation can be used effectively for attracting attention. Animation also makes a presentation light and attractive. Animation is very popular in multimedia application

1.3 Applications of Multimedia

Following are the common areas of applications of multimedia.

- **Multimedia in Business**- Multimedia can be used in many applications in a business. The multimedia technology along with communication technology has opened the door for information of global wok groups. Today the team members may be working anywhere and can work for various companies. Thus the work place will become global. The multimedia network should support the following facilities:
    - Voice Mail
    - Electronic Mail
    - Multimedia based FAX
    - Office Needs
    - Employee Training

- Sales and Other types of Group Presentation
- Records Management

- **Multimedia in Marketing and Advertising**- By using multimedia marketing of new products can be greatly enhanced. Multimedia boost communication on an affordable cost opened the way for the marketing and advertising personnel. Presentation that have flying banners, video transitions, animations, and sound effects are some of the elements used in composing a multimedia based advertisement to appeal to the consumer in a way never used before and promote the sale of the products.
- **Multimedia in Entertainment**- By using multimedia marketing of new products can be greatly enhanced. Multimedia boost communication on an affordable cost opened the way for the marketing and advertising personnel. Presentation that have flying banners, video transitions, animations, and sound effects are some of the elements used in composing a multimedia based advertisement to appeal to the consumer in a way never used before and promote the sale of the products.
- **Multimedia in Education**- Many computer games with focus on education are now available. Consider an example of an educational game which plays various rhymes for kids. The child can paint the pictures, increase reduce size of various objects etc apart from just playing the rhymes.Several other multimedia packages are available in the market which provide a lot of detailed information and playing capabilities to kids.
- **Multimedia in Bank**- Bank is another public place where multimedia is finding more and more application in recent times. People go to bank to open saving/current accounts, deposit funds, withdraw money, know various financial schemes of the bank, obtain loans etc. Every bank has a lot of information which it wants to impart to in customers. For this purpose, it can use multimedia in many ways. Bank also displays information about its various schemes on a PC monitor placed in the rest area for customers. Today on-line and internet banking have become very popular. These use multimedia extensively. Multimedia is thus helping banks give service to their customers and also in educating them about banks attractive finance schemes.
- **Multimedia in Hospital**- Multimedia best use in hospitals is for real time monitoring of conditions of patients in critical illness or accident. The conditions are displayed continuously on a computer screen and can alert the doctor/nurse on duty if any changes are observed on the screen. Multimedia makes it possible to consult a surgeon or an expert who can watch an ongoing surgery line on his PC monitor and give online advice at any crucial juncture.

  In hospitals multimedia can also be used to diagnose an illness with CD-ROMs/ Cassettes/ DVDs full of multimedia based information about various diseases and their treatment.Some hospitals extensively use multimedia presentations in training their junior staff of doctors and nurses. Multimedia displays are now extensively used during critical surgeries.
- **Multimedia Pedagogues**- Pedagogues are useful teaching aids only if they stimulate and motivate the students. The audio-visual support to a pedagogue can actually help in doing so. A multimedia tutor can provide multiple numbers of challenges to the student to stimulate his interest in a topic. The instruction provided by pedagogue have moved beyond providing only button level control to intelligent simulations, dynamic creation of links, composition and collaboration and system testing of the user interactions.
- **Communication Technology and Multimedia Services**- The advancement of high computing abilities, communication ways and relevant standards has started the beginning of an era where you will be provided with multimedia facilities at home. These services may include:
  - Basic Television Services
  - Interactive entertainment
  - Digital Audio
  - Video on demand
  - Home shopping
  - Financial Transactions

- Interactive multiplayer or single player games
- Digital multimedia libraries
- E-Newspapers, e-magazines

1.4 How Multimedia In Social Media?

Social media platforms are filled with multimedia. This hasn't always been the case and many people have experienced it as a gradual transition, the kind which can easily escape your notice as the service you use every day subtly changes in character. The reasons for this are relatively clear. Mobile internet speeds and access have rapidly increased and it's simply much more viable to send and receive video than was previously the case. Platforms like Instagram (2010), Snapchat (2011), Vine (2012) and TikTok (2016) have multimedia at the heart of what they do. Partly in response to these challengers, established platforms like Facebook and Twitter have made multimedia much more prominent than was previously the case. The result is a social media in which video, audio and images play a much more prominent role than they did only a few years ago. An internet which had one been primarily based around text is increasingly awash with multimedia.

This creates a challenge for social researchers. It's exciting to see so many podcasts and videocasts being produced in order to communicate research. It can easily seem as if there have never been more academic voices in the public sphere, speaking in their own words in powerful and innovative ways. However there are questions which remain about this and they are ones which funders will undoubtedly ask in time, if they are not doing so already. How many viewers should we expect a video to achieve? How many listeners should a podcast aim to attract? What constitutes a satisfactory return on investment? It can be off putting to raise these issues when you are someone who is enthusiastic about the creative engagement we are seeing taking shape all around us. But asking these strategic questions can help us secure this activity and make it a sustainable part of research communication, as opposed to a brief flash of creativity which faded away when people started querying its financial viability.

The danger when asking how many views we should expect a video to receive is that we choose the wrong benchmark. No one expects that academics video would receive the billions of views that have been enjoyed by the most popular YouTube videos of all time. What is sometimes called the Billion View club began with Psy's Gagnam Style in 2012 and has grown rapidly since then. But even if we recognise that academic videos will tend to be fringe features on YouTube, with the possible exception of TED and comparable series which can often feature in the tens of millions, it's nonetheless easy for our sense of scale to be defined by such high profile offerings. Social media is often defined by large numbers, not least of all by the firms themselves who rely on these metrics to demonstrate continued growth to their investors. For these reasons we need to be conscious of what we take numbers to mean, where we get our benchmarks from and what this means for how we see our digital engagement.

The vast majority of academics videos I've watched have a few hundreds views or less. Is this a problem? I suspect funders might come to see it as one, at least in terms of the investment required to produce the typical video. If we don't think it's a problem then we need to develop a language in which we can talk convincingly about digital engagement in qualitative terms. If we don't see success for social researchers in terms of numbers, what would success look like? I'm convinced the answer lies in the *dissemination mechanisms*, the cultural initiatives which enable individual items of content to be shared online, as well the relation to community which they entail. Blogs, YouTube channels, Facebook pages and Twitter feeds that are successful tend to have a lived engagement with a community which has coalesced around a shared interest in the project in question. Patreon, Substack, Flattr and Buy Me A Coffee condense the relationship down into a formalised gift relationship. These projects enable us to increase the quantitative success of our multimedia, making it easier to ensure an audience by pre-aggregating it in one place. But they also open up the possibility of a much richer, qualitative engagement which moves beyond the quantitative measurement of audience.

II. EXISTING SYSTEM

To handle such huge number of images, an optimal solution is using the public cloud, since it has powerful storage capability. Images usually contain a wealth of sensitive information, therefore social service providers need not only to provide services such as retrieval and sharing but also to protect the privacies of the images. Images usually contain sensitive information that could reveal personal privacies, and encryption is an effective method to protect privacy. Based on cryptography, a number of privacy-preserving schemes that support remote image retrieval and sharing have been proposed. Considering that social

multimedia application is becoming one of the main platforms for image retrieval and sharing, this paper proposes a privacy-preserving image retrieval and sharing scheme for social multimedia applications.

2.1 Disadvantages

- The privacy of the image not in this system.
- In the large dataset, the security isn't given accurately.
- We don't retrieval image by the way of security.

III. PROPOSED SYSTEM

We propose a privacy-preserving scheme for content-based image retrieval and sharing in social applications. First, the users extract visual features from the images, and perform locality-sensitive hashing functions on visual features to generate image problem vectors. We then model the retrieval on the images as the equality search on the image problem vectors. To enable accurate and efficient retrieval, we design the secure index structure based on cuckoo hashing, which has constant lookup time. To meet the requirements of dynamic image updating, we enrich our service with image insertion and deletion. In order to reduce the key management overhead and the access control overhead in social applications, we process keys using secret sharing techniques to enable the users holding similar images to query and decrypt images independently. Finally, we implement the prototype of the proposed scheme, and perform experiments over encrypted image databases.

The contributions of this paper can be summarized as follows

- We propose a privacy-preserving content-based image retrieval and sharing scheme for social multimedia applications. The scheme allows the image owners to out- source images to the public cloud server, and outsource secure index to the social service provider.
- Our scheme is constructed under a stronger and more realistic threat model that the social service provider is not a completely trusted entity. The social service provider is allowed to acquire a part of the users' information, but it cannot be trusted to store the keys.
- Our scheme greatly simplices access control and key management in multi-user social applications. We allow the users holding similar images to submit the retrieval queries, and decrypt the query results on their own.
- We design a secure index that provides efficient retrieval service for large image set. The index requires constant lookup time and supports dynamic image updating.

3.1 Advantages

- High security and more effective.
- Communication and search efficiency.
- The image should be store and retrieval is more secure compared to other methods.
- It's given more performance

3.2 Algorithm

3.2.1 Collaborative filtering:

Collaborative filtering (CF) is a technique used by recommender system. Collaborative filtering has two senses, a narrow one and a more general one.

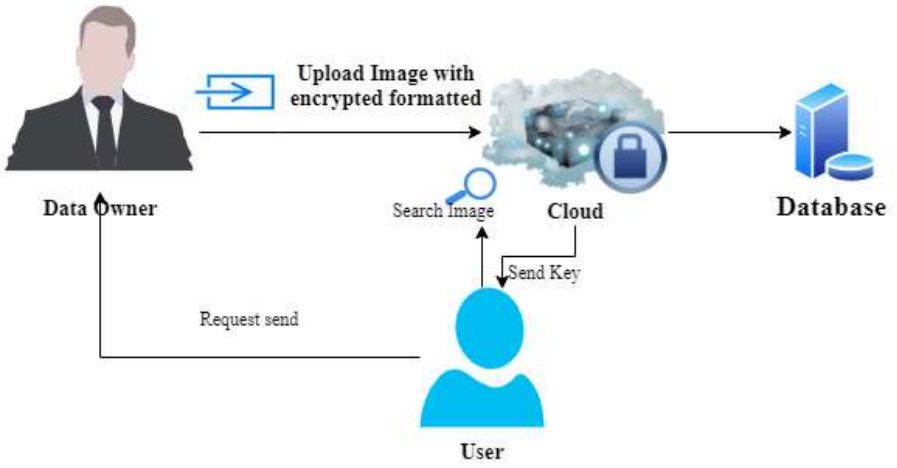
3.2.2 Ciphertext:

Ciphertext is also known as encrypted or encoded information because it contains a form of the original plaintext that is unreadable by a human or computer without the proper cipher to decrypt it. Decryption, the inverse of encryption, is the process of turning ciphertext into readable plaintext.

3.2.3 AES:

The **algorithm** described by **AES** is a symmetric-key **algorithm**, meaning the same key is used for both encrypting and decrypting the data.

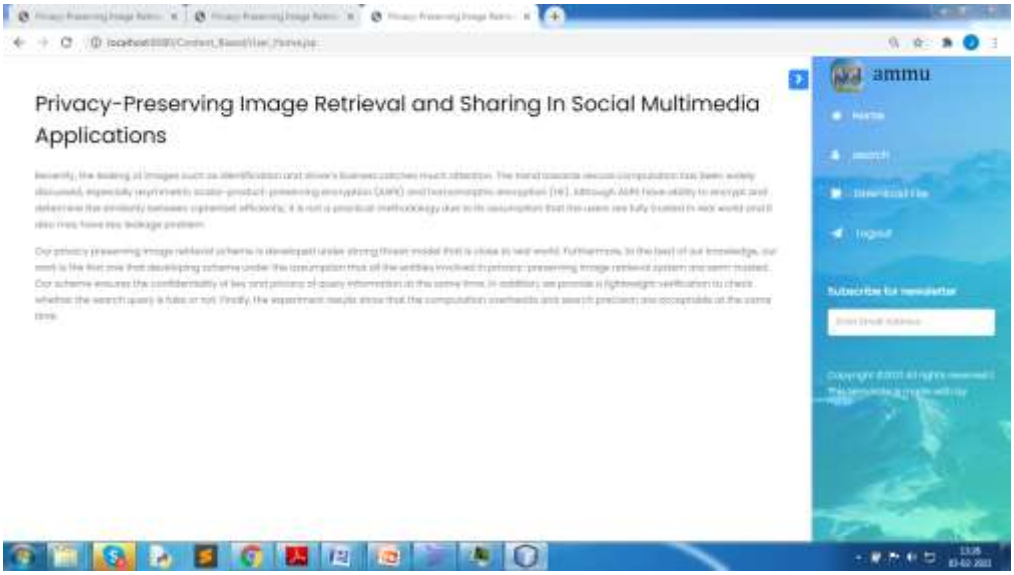
IV. SYSTEM ARCHITECTURE



Cloud server

Data owner

Data user



V. CONCLUSION

In this paper, we proposed and implemented a privacy- preserving content-based image retrieval and sharing scheme, which can be used for friend recommendation in social multimedia applications. We measured image similarity through image visual features. We used *leash* to reduce the dimensionality of visual features and realize similarity search on visual features. We designed the index based on cuckoo hashing to speed up the similarity search. Based on secret sharing, we allowed the user to query and recover images on his own, which eliminates key management overhead and access control overhead compared with other schemes. Finally, we implemented a prototype to evaluate the efficiency of our proposed scheme. The results showed that our scheme achieves practical performance under the UK Bench database.

REFERENCES

[1] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, ``ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564_2571.

[2] D. Xiaoding Song, D. Wagner, and A. Perrig, ``Practical techniques for searches on encrypted data," in *Proc. IEEE Symp. Secur. Privacy.*, May 2000, pp. 44_55.

[3] X. Yuan, X. Wang, C. Wang, A. C. Squicciarini, and K. Ren, ``Towards privacy-preserving and practical image-centric social discovery," *IEEE Trans. Dependable Secure Comput.*, vol. 15, no. 5, pp. 868_882, Sep. 2018.

[4] K. W. Ahmed, M. Z. Hasan, and N. Mohammed, ``Image-centric social discovery using neural network under anonymity constraint," in *Proc. IEEE Int. Conf. Cloud Eng. (ICE)*, Apr. 2017, pp. 238_244.

[5] Facebook. (2019). *Keeping Passwords Secure*. [Online]. Available: https://newsroom.fb.com/news/2019/03/keeping-passwords-secure/

[6] W. Pieters and Q. Tang, ``Data is key: Introducing the data-based access control paradigm," in *Proc. IFIP Annu. Conf. Data Appl. Secur. Privacy*. Berlin, Germany: Springer, 2009, pp. 240_251.

[7] Boldyreva, T. Tang, and B. Warinschi, ``Masking fuzzy-searchable public databases," in *Proc. Int. Conf. Appl. Cryptography Netw. Secur.* Cham, Switzerland: Springer, 2019, pp. 571_591.

[8] P. Indyk and R. Motwani, ``Approximate nearest neighbors: Towards removing the curse of dimensionality," in *Proc. 13th Annu. ACM Symp. Theory Comput. (STOC)*, 1998, pp. 604_613.

[9] W. Dong, Z. Wang, W. Josephson, M. Charikar, and K. Li, ``Modeling LSH for performance tuning," in *Proc. 17th ACMConf. Inf. Knowl. Mining (CIKM)*, 2008, pp. 669_678.

[10] Q. Lv,W. Josephson, Z.Wang, M. Charikar, and K. Li, ``Multi-probe ISN: Ef_cient indexing for high-dimensional similarity search," in *Proc. 33rd Int. Conf. Very Large Data Bases*, 2007, pp. 950_961.

[11] Gionis, P. Indyk, and R. Motwani, ``Similarity search in high dimensions via hashing," *Proc. VLDB*, vol. 99, no. 6, pp. 518_529, 1999.