National Conference on Research Challenges and Opportunities in Digital and Cyber Forensics
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2020 IJSRCSEIT | Volume 4 | Issue 11 | ISSN : 2456-3307 (www.ijsrcseit.com)

Effective Video Copy Detection Technique of Multimedia Content in Cloud Environment

Shubham Dilip Vyawahare[1], Dr. Avinash Kapse[2]
[1]Department of Information Technology, Anuradha Engineering College, Chikhli, Maharashtra, India
[2]Associate Professor & HOD, Department of Information Technology, Anuradha Engineering College, Chikhli, Maharashtra, India

ABSTRACT

In a view of large-scale multimedia content protection systems and the charges to provide cloud infrastructures to provide cost efficiency, rapid deployment, scalability, and elasticity to accommodate varying workloads. We proposing a system that can be used to protect different multimedia content types, including 2-D videos, 3-D videos, images, audio clips, songs, and music clips. It can be implemented on private and/or public clouds. We design a system with two method processing: (i) Creating digital signatures, and (ii) Comparison database to recognized modifications. The signature method creates robust and representative signatures of contents. Comparison with storage that is real in cloud with the available content. The high accuracy and scalability of the proposed system include high database and storage content. In addition, we compared our system to the protection system used by some videos channels.

Keywords : Cloud storage, Digital signatures, scalability, protection channels and database.

I. INTRODUCTION

The existing mechanism a new significant privacy issue introduced in the case of shared data with the use of the leakage of identity privacy to public verifiers. The traditional approach for checking data correctness is to retrieve the entire data from the cloud, and then verify data integrity by checking the correctness of signatures.

To securely introduce an effective third party auditor (TPA), the following two fundamental requirements have to be met:

1) TPA should be able to efficiently audit the cloud data storage without demanding the local copy of data, and introduce no additional on-line burden to the cloud user;

2) The third party auditing process should bring in no new vulnerabilities towards user data privacy.

To support a privacy-preserving mechanism .

To make public auditing on shared data stored in the cloud using various encrytion algorithm.

II. Literature Survey

Public Auditing Mechanism
A new public auditing mechanism for shared data with efficient user revocation in the cloud. When a user in the group is revoked, we allow the semi-trusted cloud to re-sign blocks that were signed by the revoked user with proxy re-signatures. The group can save a significant amount of computation

and communication resources during user revocation. [3]

A secure and efficient RDC scheme for network coding-based distributed storage systems that rely on untrusted server. RDC-NC scheme can be used to ensure data remains intact when faced with data corruption, replay, and pollution attacks. The RDC-NC is inexpensive for both clients and servers. [4]

Short Group Signatures: Signatures in our scheme are approximately the size of a standard RSA signature with the same security. The group signature is based on the Strong Diffie-Hellman assumption and a new assumption in bilinear groups called the Decision Linear. [5]

Storing Shared Data on the Cloud via Security-Mediator: We believe is the right approach to achieve anonymity in storing data to the cloud with publicly-verifiable data-integrity in mind. The de couples the anonymous protection mechanism from the provable data possession mechanism via the use of security mediator. They minimize the computation and bandwidth requirement of this mediator, but also minimize the trust placed on it in terms of data privacy and identity privacy. [6]

III. Module

User Registration:

For the registration of user with identity ID the group manager randomly selects a number. Then the group manager adds into the group user list which will be used in the traceability phase. After the registration, user obtains a private key which will be used for group signature generation and file decryption.

Registration

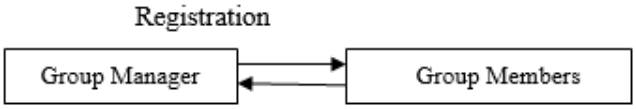


Fig 1: Key Distribution

Public Auditing:

Homomorphic authenticators are unforgeable verification metadata generated from individual data blocks, which can be securely aggregated in such a way to assure an auditor that a linear combination of data blocks is correctly computed by verifying only the aggregated authenticator. Overview to achieve privacy-preserving public auditing, we propose to uniquely integrate the Homomorphic authenticator with random mask technique. In our protocol, the linear combination of sampled blocks in the server's response is masked with randomness generated by a pseudo random function (PRF). The proposed scheme is as follows:

- Setup Phase
- Audit Phase

Sharing Data:

The canonical application is data sharing. The public auditing property is especially useful when we expect the delegation to be efficient and flexible. The schemes enable a content provider to share her data in a confidential and selective way, with a fixed and small cipher text expansion, by distributing to each authorized user a single and small aggregate key.

Integrity Checking:

Hence, supporting data dynamics for privacy-preserving public risk auditing is also of paramount importance. Now we show how our main scheme can be adapted to build upon the existing work to support data dynamics, including block level operations of modification, deletion and insertion. We can adopt this technique in our design to achieve privacy-preserving public risk auditing with support of

IV. Algorithm Used

Definition

The Advanced Encryption Standard (AES) is an encryption algorithm for securing sensitive but unclassified material by U.S. Government agencies

and, as a likely consequence, may eventually become the de facto encryption standard for commercial transactions in the private sector. (Encryption for the US military and other classified communications is handled by separate, secret algorithms.)In January of 1997, a process was initiated by the National Institute of Standards and Technology (NIST), a unit of the U.S. Commerce Department, to find a more robust replacement for the Data Encryption Standard (DES) and to a lesser degree Triple DES.

The specification called for a symmetric algorithm (same key for encryption and decryption) using block encryption (see block cipher) of 128 bits in size, supporting key sizes of 128, 192 and 256 bits, as a minimum.

The algorithm was required to be royalty-free for use worldwide and offer security of a sufficient level to protect data for the next 20 to 30 years. It was to be easy to implement in hardware and software, as well as in restricted environments (for example, in a smart card) and offer good defenses against various attack techniques. The entire selection process was fully open to public scrutiny and comment, it being decided that full visibility would ensure the best possible analysis of the designs.

In 1998, the NIST selected 15 candidates for the AES, which were then subject to preliminary analysis by the world cryptographic community, including the National Security Agency. On the basis of this, in August 1999, NIST selected five algorithms for more extensive analysis. These were:

- MARS, submitted by a large team from IBM Research.
- RC6, submitted by RSA Security.
- Rijndael, submitted by two Belgian cryptographers, Joan Daemen and Vincent.
- Rijmen Serpent, submitted by Ross Andersen, Eli Biham and Lars Knudsen
- Two fish, submitted by a large team of researchers including Counterpane's.
- Respected cryptographer, Bruce Schneier.

Implementations of all of the above were tested extensively in ANSI C and Java languages for speed and reliability in such measures as encryption and decryption speeds, key and algorithm set-up time and resistance to various attacks, both in hardware- and software-centric systems.

Once again, detailed analysis was provided by the global cryptographic community (including some teams trying to break their own submissions). The end result was that on October 2, 2000, NIST announced that Rijndael had been selected as the proposed standard. On December 6, 2001, the Secretary of Commerce officially approved Federal Information Processing Standard (FIPS) 197, which specifies that all sensitive, unclassified documents will use Rijndael as the Advanced Encryption Standard.

Also see cryptography, data recovery agent (DRA)RELATED GLOSSARY TERMS: RSA algorithm (Rivest-Shamir-Adleman), data key, greynet (or graynet), spam cocktail (or anti-spam cocktail), fingers canning (fingerprint scanning),munging, insider threat, authentication server, defense in depth, nonrepudiation.

High-level description of the algorithm:

1. Key Expansion—round keys are derived from the cipher key using Rijndael's key schedule.
2. Initial Round

AddRoundKey— each byte of the state is combined with the round key using bitwise xor.

3. Rounds

- Sub Bytes—a non-linear substitution step where each byte is replaced with another according to a lookup table.
- Shift Rows—a transposition step where each row of the state is shifted cyclically a certain number of steps.
- Mix Columns—a mixing operation which operates on the columns of the state, combining the four bytes in each column.
- Add Round Key

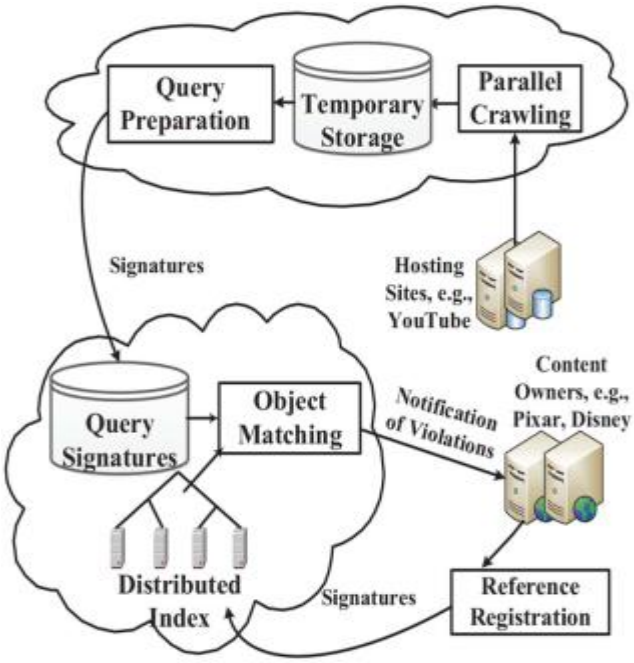
4. Final Round (no Mix Columns)

• Sub Bytes

• Shift Rows

• Add Round Key

Diagrams



Examples

In this appendix, twenty examples are provided for the MAC generation process. The underlying block cipher is either the AES algorithm or TDEA. A block cipher key is fixed for each of the currently allowed key sizes, i.e., AES-128, AES-192, AES-256, two key TDEA, and three key TDEA. For each key, the generation of the associated sub keys is given, followed by four examples of MAC generation with the key.

The messages in each set of examples are derived by truncating a common fixed string of 64 bytes. All strings are represented in hexadecimal notation, with a space (or a new line) inserted every 8 symbols, for readability. As in the body of the Recommendation, K1 and K2denote the sub keys, M denotes the message, and T denotes the MAC. For the AES algorithm examples, Tlen is 128, i.e., 32 hexadecimal symbols, and K denotes the key. For the TDEA examples, Tlen is 64, i.e., 16 hexadecimal symbols, and the key, K, is the ordered triple of strings, (Key1, Key2, and Key3). For two key TDEA, Key1 = Key3.

D.1 AES-128

For Examples 1–4 below, the block cipher is the AES algorithm with the following 128 bit key: K 2b7e1516 28aed2a6 abf71588 09cf4f3c.

Sub key Generation CIPHK (0128) 7df76b0c 1ab899b3 3e42f047 b91b546f K1 fbeed618 35713366 7c85e08f 7236a8de

K2 f7ddac30 6ae266cc f90bc11e e46d513b

Example Explanations

The Advanced Encryption Standard (AES) specifies a FIPS-approved cryptographic algorithm that can be used to protect electronic data. The AES algorithm is a symmetric block cipher that can encrypt (encipher) and decrypt (decipher) information. Encryption converts data to an unintelligible form called cipher text; decrypting the cipher text converts the data back into its original form, called plaintext.

The AES algorithm is capable of using cryptographic keys of 128, 192, and 256 bits to encrypt and decrypt data in blocks of 128 bits.

Ring signature (digital Signature):

In cryptography, a ring signature is a type of digital signature that can be performed by any member of a group of users that each have keys. Therefore, a message signed with a ring signature is endorsed by someone in a particular group of people. One of the security properties of a ring signature is that it should be difficult to determine which of the group members' keys was used to produce the signature.

Ring signatures are similar to group signatures but differ in two key ways.

Suppose that a group of entities each have public/private key pairs, (PK1, SK1), (PK2, SK2), (PKn, SKn). Party i can compute a ring signature σ on a message m, on input (m, SKi, PK1, PKn). Anyone can check the validity of a ring signature given σ, m, and the public keys involved, PK1, PKn.

If a ring signature is properly computed, it should pass the check. On the other hand, it should be hard for anyone to create a valid ring signature on any message for any group without knowing any of the secret keys for that group.

V. Conceptualization

Cloud Computing:

Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network (typically the Internet). The name comes from the use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote services with a user's data, software and computation.

Cloud computing, or something being in the cloud, is an expression used to describe a variety of different types of computing concepts that involve a large number of computers connected through a real-time communication network such as the Internet. Cloud computing is a term without a commonly accepted unequivocal scientific or technical definition.

In science, cloud computing is a synonym for distributed computing over a network and means the ability to run a program on many connected computers at the same time. The phrase is also more commonly used to refer to network-based services which appear to be provided by real server hardware, which in fact are served up by virtual hardware, simulated by software running on one or more real machines.

Such virtual servers do not physically exist and can therefore be moved around and scaled up (or down) on the fly without affecting the end user arguably, rather like a cloud.

Cloud computing relies on sharing of resources to achieve coherence and economies of scale similar to a utility (like the electricity grid) over a network. At the foundation of cloud computing is the broader concept of converged infrastructure and shared services.

There are many types of public cloud computing

- ✓ Infrastructure as a service (IaaS)
- ✓ Platform as a service (PaaS)
- ✓ Software as a service (SaaS)
- ✓ Storage as a service (STaaS)
- ✓ Security as a service (SECaaS)
- ✓ Data as a service (DaaS)
- ✓ Test environment as a service (TEaaS)
- ✓ Desktop as a service (DaaS)
- ✓ API as a service (APIaaS)

The business model, IT as a service (ITaaS), is used by in-house, enterprise IT organizations that offer any or all of the above services. Using software as a service, users also rent application software and databases. The cloud providers manage the infrastructure and platforms on which the applications run.

End users access cloud-based applications through a web browser or a light-weight desktop or mobile app while the business software and user's data are stored on servers at a remote location. Proponents claim that cloud computing allows enterprises to get their applications up and running faster, with improved manageability and less maintenance, and enables IT to more rapidly adjust resources to meet fluctuating and unpredictable business demand.

Application of Cloud computing:

- ✓ Autonomic computing — Computer systems capable of self-management.
- ✓ Client–server model — Client–server computing refers broadly to any distributed application that distinguishes between service providers (servers) and service requesters (clients).
- ✓ Grid computing — "A form of distributed and parallel computing, whereby a 'super and virtual computer' is composed of a cluster of networked, loosely coupled computers acting in concert to perform very large tasks."
- ✓ Mainframe computer — Powerful computers used mainly by large organizations for critical applications, typically bulk data processing such as census, industry and consumer statistics, police and secret intelligence

services, enterprise resource planning, and financial transaction processing.

- ✓ Utility computing — The "packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility, such as electricity."
- ✓ Peer-to-peer — Distributed architecture without the need for central coordination, with participants being at the same time both suppliers and consumers of resources (in contrast to the traditional client–server model).
- ✓ Cloud gaming - Also called on-demand gaming is a way of delivering to games to computers. The gaming data will be stored in the provider's server, so that gaming will be independent of client computers used to play the game.

VI. Characteristics

Agility improves with users' ability to re-provision technological infrastructure resources.

Application Programming Interface (API) accessibility to software that enables machines to interact with cloud software in the same way the user interface facilitates interaction between humans and computers. Cloud computing systems typically use REST-based APIs.

Cost is claimed to be reduced and in a public cloud delivery model capital expenditure is converted to operational expenditure.

This is purported to lower barriers to entry, as infrastructure is typically provided by a third-party and does not need to be purchased for one-time or infrequent intensive computing tasks.

Pricing on a utility computing basis is fine-grained with usage-based options and fewer IT skills are required for implementation (in-house).The e-FISCAL project's state of the art repository contains several articles looking into cost aspects in more detail, most of them concluding that costs savings depend on the type of activities supported and the type of infrastructure available in-house.

Device and location independence enable users to access systems using a web browser regardless of their location or what device they are using (e.g., PC, mobile phone). As infrastructure is off-site (typically provided by a third-party) and accessed via the Internet, users can connect from anywhere.

Virtualization technology allows servers and storage devices to be shared and utilization be increased. Applications can be easily migrated from one physical server to another.

Centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.)

Peak-load capacity increases (users need not engineer for highest possible load-levels)

Utilization and efficiency improvements for systems that are often only 10–20% utilized.

Reliability is improved if multiple redundant sites are used, which makes well-designed cloud computing suitable for business continuity and disaster recovery.[30]

Scalability and elasticity via dynamic ("on-demand") provisioning of resources on a fine-grained, self-service basis near real-time, without users having to engineer for peak loads.

Performance is monitored, and consistent and loosely coupled architectures are constructed using web services as the system interface.

Security could improve due to centralization of data, increased security-focused resources, etc., but concerns can persist about loss of control over certain sensitive data, and the lack of security for stored kernels. Security is often as good as or better than other traditional systems, in part because providers are able to devote resources to solving security issues that many customers cannot afford. However, the complexity of security is greatly increased when data is distributed over a wider area or greater number of devices and in multi-tenant systems that are being shared by unrelated users. In addition, user access to security audit logs may be difficult or impossible. Private cloud installations are in part motivated by users' desire to retain control over the infrastructure and avoid losing control of information security.

Maintenance of cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places.

Virtualization

Virtualization (or virtualization) is the creation of a virtual (rather than actual) version of something, such as a hardware platform, operating system (OS), storage device, or network resources.

While a physical computer in the classical sense is clearly a complete and actual machine, both subjectively (from the user's point of view) and objectively (from the hardware system administrator's point of view), a virtual machine is subjectively a complete machine (or very close), but objectively merely a set of files and running programs on an actual, physical machine (which the user need not necessarily be aware of). Virtualization can be viewed as part of an overall trend in enterprise IT that includes autonomic computing, a scenario in which the IT environment will be able to manage itself based on perceived activity, and utility computing, in which computer processing power is seen as a utility that clients can pay for only as needed.

The usual goal of virtualization is to centralize administrative tasks while improving scalability and overall hardware-resource utilization. With virtualization, several operating systems can be run in parallel on a single central processing unit (CPU). This parallelism tends to reduce overhead costs and differs from multitasking, which involves running several programs on the same OS.

Feasibility Study

Feasibility study is the test of a system proposal according to its workability, impact on the organization, ability to meet user needs, and effective use of recourses. It focuses on the evaluation of existing system and procedures analysis of alternative candidate system cost estimates. Feasibility analysis was done to determine whether the system would be feasible.

The development of a computer based system or a product is more likely plagued by resources and delivery dates. Feasibility study helps the analyst to decide whether or not to proceed, amend, postpone or cancel the project, particularly important when the project is large, complex and costly. Once the analysis of the user requirement is complement, the system has to check for the compatibility and feasibility of the software package that is aimed at. An important outcome of the preliminary investigation is the determination that the system requested is feasible.

Technical Feasibility:

The technology used can be developed with the current equipment's and has the technical capacity to hold the data required by the new system.

• This technology supports the modern trends of technology.

• Easily accessible, more secure technologies.

Technical feasibility on the existing system and to what extend it can support the proposed addition. We can add new modules easily without affecting the Core Program. Most of parts are running in the server using the concept of stored procedures.

Operational Feasibility:

This proposed system can easily implemented, as this is based on JSP coding (JAVA) & HTML .The database created is with My Sql server which is more secure and easy to handle. The resources that are required to implement/install these are available. The personal of the organization already has enough exposure to computers. So the project is operationally feasible.

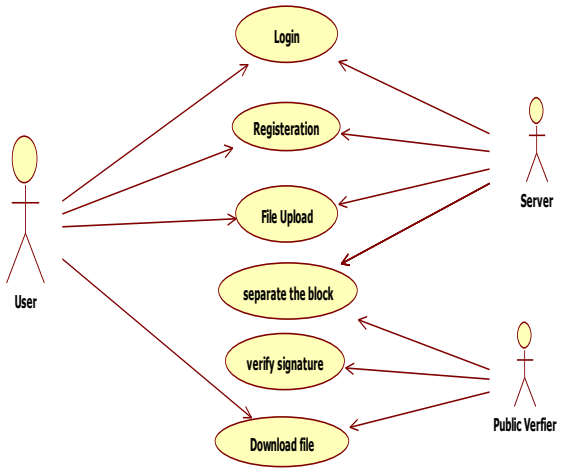
Economic Feasibility:

Economic analysis is the most frequently used method for evaluating the effectiveness of a new system. More commonly known cost/benefit analysis, the procedure is to determine the benefits and savings that are expected from a candidate system and compare them with costs. If benefits outweigh costs, then the decision is made to design and implement the system.

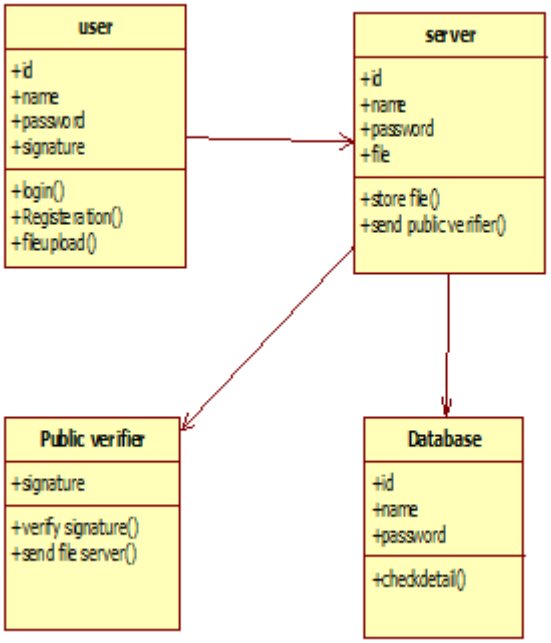
An entrepreneur must accurately weigh the cost versus benefits before taking an action. This system is more economically feasible which assess the brain capacity with quick & online test.

VII. System Design and deployment

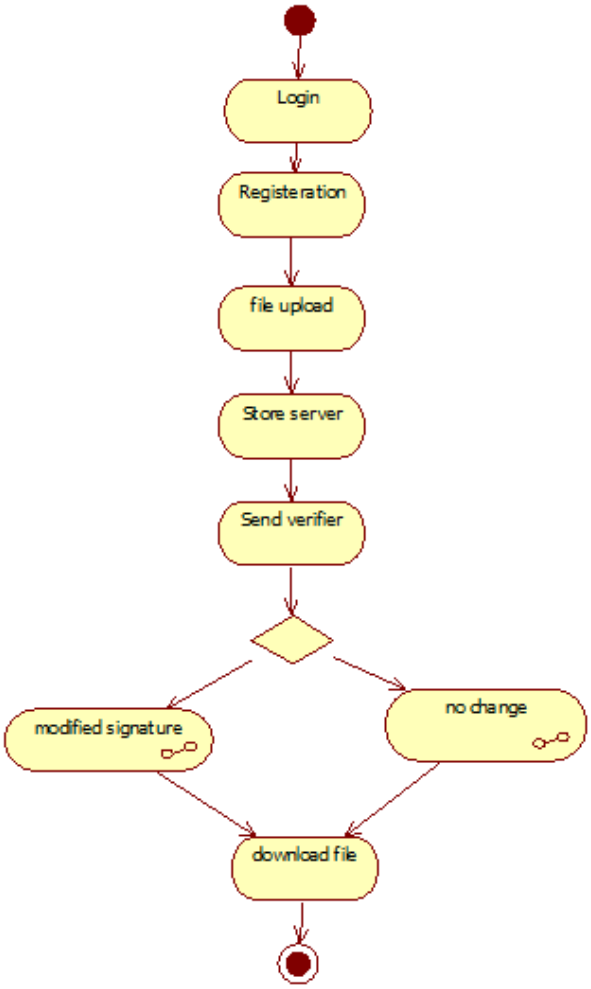
Use Case Structure



Class Diagram



Activity Diagram



VIII. EXISTING SYSTEM

The existing mechanism a new significant privacy issue introduced in the case of shared data with the use of the leakage of identity privacy to public verifiers. The traditional approach for checking data correctness is to retrieve the entire data from the cloud, and then verify data integrity by checking the correctness of signatures.

To securely introduce an effective third party auditor (TPA), the following two fundamental requirements have to be met: 1) TPA should be able to efficiently audit the cloud data storage without demanding the local copy of data, and introduce no additional on-line burden to the cloud user; 2) The third party auditing process should bring in no new vulnerabilities towards user data privacy

IX. LIMITATIONS

As users no longer physically possess the storage of their data, traditional cryptographic primitives for the purpose of data security protection cannot be directly adopted.

They do not perform the multiple auditing task in simultaneously.

X. PROPOSED SYSTEM

The propose system a privacy-preserving public auditing mechanism for shared data in the cloud. We utilize ring signatures to construct homomorphism authenticators, so that a public verifier is able to audit shared data integrity without retrieving the entire data, yet it cannot distinguish who is the signer on each block.

To improve the efficiency of verifying multiple auditing tasks, we further extend our mechanism to support batch auditing. There are two interesting problems we will continue to study for our future work. One of them is traceability, which means the ability for the group manager to reveal the identity of the signer based on verification metadata in some special situations

XI. ADVANTAGES

The proposed system can perform multiple auditing tasks simultaneously. They improve the efficiency of verification for multiple auditing tasks. High security provide for file sharing.

XII. CONCLUSION

In this paper, we propose the system, the first privacy-preserving public auditing mechanism for shared data in the cloud for protecting the multimedia content. With this system, the public verifier is able to efficiently audit the integrity of shared data, yet cannot distinguish who is the signer on each block, which can preserve identity privacy for users. An interesting problem in our future work is how to efficiently audit the integrity of shared data with dynamic groups while still preserving the identity of the signer on each block from the third party auditor.

XIII. REFERENCES

[1]. Abdelsadek, "Distributed index for matching multimedia objects," M.S. thesis, School of Comput. Sci., Simon Fraser Univ., Burnaby, BC, Canada, 2014.

[2]. Abdelsadek and M. Hefeeda, "Dimo: Distributed index for matching multimedia objects using MapReduce," in Proc. ACMMultimedia Syst. Conf. (MMSys'14), Singapore, Mar. 2014, pp. 115–125.

[3]. M. Aly, M. Munich, and P. Perona, "Distributed Kd-Trees for retrieval from very large image collections," in Proc. Brit. Mach. Vis. Conf. (BMVC), Dundee, U.K., Aug. 2011.

[4]. J. Bentley, "Multidimensional binary search trees used for associative searching," in Commun. ACM, Sep. 1975, vol. 18, no. 9, pp. 509–517.

[5]. P. Cano, E. Batle, T. Kalker, and J. Haitsma, "A review of algorithms for audio fingerprinting," in Proc.

[6]. Privacy-Preserving Public Auditing for Secure Cloud Storage (I. Agudo, D. Nu˜ nez, G. Giammatteo, P. Rizomiliotis, and C. Lambrinoudakis, "Cryptography goes to the cloud," in Secure and Trust Computing, Data Management, and Applicat., 2011, pp. 190– 197.)

[7]. Efficient and Secure Multi-Keyword Search on Encrypted Cloud Data (G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in Proc. 14th ACM conf. Compu. Commun. Security (CCS), 2007, pp. 598–609.)

[8]. Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud (G. Ateniese, A. Faonio, and S. Kamara, "Leakage-resilient identification schemes from zero-knowledge proofs of storage," in IMA Inte. Conf. Cryptography and Coding, 2015, pp. 311–328)

[9]. Panda: Public Auditing for Shared Data with Efficient User Revocation in the Cloud (G. Ateniese,

R.D. Pietro, L.V. Mancini, and G. Tsudik, “Scalable and efficient provable data possession,” in Proc. 4th Int. Conf. Secure and Privacy in Commun. Netw. (SecureComm), 2008, pp. 1–10.)

[10]. Remote Data Checking for Network Coding-based Distributed Storage Systems(K. D. Bowers, A. Juels, and A. Oprea, “Proofs of retrievability: theory and implementation,” in Proc. 2009 ACM Workshop Cloud Computing Security (CCSW), 2009, pp. 43–54.)

[11]. Short Group Signatures(L. Chen, “Using algebraic signatures to check data possession in cloud storage,” Future Generation Computer Systems, vol. 29, no.7, pp. 1709–1715, 2013.)

[12]. Storing Shared Data on the Cloud via Security-Mediator (Y. Dodis, S. Vadhan, and D. Wichs, “Proofs of retrievability viahardness amplification,” in Proc. Theory Cryptography Conf. (TCC), 2009, pp. 109–127.,)