An Optimized Technique for Image Archive and Image Search Using Heterogeneous APKET Technique

Karthiga K.R, Student,Dept.Computer Science And Engineering ,Vandayar Engineering College,

Anna University,Chennai.E-mail-shakthitulips@gmail.com Contact-+91-9585562069

ABSTRACT- Increasing popularity of storing and managing personal multimedia data using online services. Preserving confidentiality of online personal data for offering efficient functionalities becomes important pressing research issue. In system study the problem of content based search of image data archived online while preserving content confidentiality. The problem has different settings from those typically considered in the secure computation literature, as it deals with data in rank-ordered search, and has a different security-efficiency requirement. Secure computation techniques, such as homomorphic encryption, can potentially be used in this application, at a cost of high computational and communication complexity. Efficient techniques based on randomizing visual feature and search indexes have been proposed recently to enable similarity comparison between encrypted images. This focuses on comparing these two major paradigms of techniques, namely, homomorphism encryption-based techniques and feature/index randomization-based techniques, for confidentiality preserving image search. It develops novel and systematic metrics to quantitatively evaluate security strength in this unique type of data and applications with the comparison of these two paradigms of techniques in terms of their search performance, security strength, and computational efficiency. The insights obtained through this project and comparison will help design practical algorithms appropriate for privacy-aware cloud multimedia systems.

Keywords: Content based image retrieval, secure search, secure cloud computing, homomorphic encryption

1. INTRODUCTION

Nowadays the need grows to securely outsource data to an untrusted system. Think, for instance, of a remote database server administered by somebody else. If you want your data to be secret, you have to encrypt it. The problem then arises how to query the database. The most obvious solution is to download the whole database locally and then perform the query. This of course is terribly inefficient. A single-server solution for remote querying of encrypted relational databases on untrusted servers is presented. The approach is based on the use of B+ tree indexing information attached to the relations. The designed indexing mechanism can balance the trade-off between efficiency requirements in query execution and protection requirements due to possible inference attacks exploiting indexing information.

It is desirable to store data on data storage servers such as mail servers and file servers in encrypted form to reduce security and privacy risks. But this usually implies that one has to sacrifice functionality for security. For example, if a client wishes to retrieve only documents containing certain words, it was not previously known how to let the data storage server perform the search and answer the query without loss of data confidentiality

Our techniques have a number of crucial advantages. They are provably secure: they provide provable secrecy for encryption, in the sense that the untrusted server cannot learn anything about the plaintext when only given the ciphertext; they provide query isolation for searches, meaning that the untrusted server cannot learn anything more about the plaintext than the search result; they provide controlled searching, so that the untrusted server cannot search for an arbitrary word without the user's authorization; they also support hidden queries, so that the user may ask the untrusted server to search for a secret word without revealing the word to the server. The algorithms system present are simple, fast (for a document of length the encryption and search algorithms only need stream cipher and block cipher operations), and introduce almost no space and communication overhead, and hence are practical to use today.

The processing and encryption of multimedia content are generally considered sequential and independent operations. In certain multimedia content processing scenarios, it is, however, desirable to carry out processing directly on encrypted signals. The field of secure signal processing poses significant challenges for both signal processing and cryptography research; only few ready-to-go fully integrated solutions are available. This study first concisely summarizes cryptographic primitives used in existing solutions to processing of encrypted signals, and discusses implications of the security requirements on these solutions. The study then continues to describe two domains in which secure signal processing has been taken up as a challenge, namely, analysis and retrieval of multimedia content, as well as multimedia content protection. In each domain, state-of-the-art algorithms are described. Finally, the study discusses the challenges and open issues in the field of secure signal processing.

In several application scenarios, however, it is desirable to carry out signal processing operations directly on encrypted signals. Such an approach is called *secure signal processing*, *encrypted signal processing*, or *signal processing in the encrypted domain*. For instance, given an encrypted image, can system calculate the mean value of the encrypted image pixels? On the one hand, the relevance of carrying out such signal manipulations, that is, the algorithm, directly on encrypted signals is entirely dependent on the security requirements of the application scenario under consideration. On the other hand, the particular implementation of the signal processing algorithm will be determined strongly by the possibilities and impossibilities of the cryptosystem employed. Finally, it is very likely that new requirements for cryptosystems will emerge from secure signal processing operations and applications. Hence, secure signal processing poses a joint challenge for both the signal processing and the cryptographic community.

2. RELATED WORK

Most of the existing systems for rank-ordered search over encrypted text documents, so that documents can be returned in the order of their relevance to the query term. In that work, several protocols are studied to address different operational constraints such as different communication cost allowed performing the secure search. Secure text retrieval techniques can also be applied to keyword based search of image data.

However, keyword search relies on having accurate text description of the content already available, and its search scope is confined to the existing keyword set. In contrast, content-based search over an encrypted image database provides more flexibility, where by sample images are presented as queries and documents with similar visual content in the database are identified. An emerging area of work related to confidentiality preserving image retrieval is secure signal processing, which aims at performing signal processing tasks while keeping the signals being processed secret.

Erkin et al.[1][2] provided a review of related cryptographic primitives and some applications of secure signal processing in data analysis and content protection. However, applying cryptographic primitives to the task of content-based image retrieval is not straightforward.

Effective image retrieval typically relies on evaluating the similarity of two documents using the distance between their visual features, such as color histograms, shape descriptors, or salient points. By design, traditional cryptographic primitives do not preserve the distance between feature vectors after encryption. Given the much larger data volume for image data than that of text and other generic data, efficiency and scalability are critical for image retrieval but can be difficult to achieve using cryptographic primitives alone. Another work by Shashank et al.[3] addresses the problem of protecting the privacy of the query image when searching over a public database, where the images in the database are not encrypted.

2.1 Classification-based methods in optimal image interpolation Encryption

In this work introduce two new approaches to optimal image interpolation which are based on the idea that image data falls into different categories or classes, such as edges of different orientation and smoother gradients. Both these methods work by classifying the image data in a window around the pixel being interpolated, and then using an interpolation alter designed for the selected class. The first method, which system calls Resolution Synthesis (RS),[4][5] performs the classification by computing probabilities of class membership in a Gaussian mixture model. The second method, which system calls Tree-based Resolution Synthesis (TRS),[6][7] uses a regression tree. Both of these methods are based on stochastic models for image data whose parameters must have been estimated beforehand, by training on sample images. System demonstrates that under some assumptions, both of these methods are actually optimal in the sense that they yield minimum mean-squared error (MMSE) estimates of the target-resolution image, given the source image. System also introduces Enhanced Tree-based RS, which consists of TRS interpolation followed by an enhancement stage.

During the enhancement stage, system recursively add adjustments to the pixels in the interpolated image. This has the dual effect of reducing interpolation artifacts while imparting additional sharpening. System present results of the above methods for interpolating images which are free of artifacts. In addition, system present results which demonstrate that RS can be trained for high-quality interpolation of images which exhibit certain characteristic artifacts, such as JPEG images and digital camera images. System also present results of a new interpolative image coding method which uses RS along with the well-known JPEG compression scheme. These results demonstrate that for relatively low bit rates, the RS-based compression scheme can improve upon JPEG compression used alone, in terms of subjective image quality (for an approximately fixed bit-rate), and in terms of better rate-distortion tradeof.

2.2 Limits on super-resolution and how to break them

Nearly all super-resolution algorithms are based on the fundamental constraints that the super-resolution image should generate the low resolution input images when appropriately warped and down-sampled to model the image formation process. (These reconstruction constraints are normally combined with some form of smoothness prior to regularize their solution.) In the first part of this paper, systems derive a sequence of analytical results which show that the reconstruction constraints provide less and less useful information as the magnification factor increases. System also validate these results empirically and show that for large enough magnification factors any smoothness prior leads to overly smooth results with very little high-frequency content (however many low resolution input images are used.) In the second part of this paper, system proposes a super-resolution algorithm that uses a different kind of constraint, in addition to the reconstruction constraints. The algorithm attempts to recognize local features in the low resolution images and then enhances their resolution in an appropriate manner. System calls such a super-resolution algorithm a *hallucination* or *recogstruction* algorithm. System tried our hallucination algorithm on two different datasets, frontal images of faces and printed Roman text. System obtained significantly better results than existing reconstruction-based algorithms, both qualitatively and in terms of RMS pixel error.

2.3 Sparse Bayesian learning and the relevance vector data security

This proposed introduces a general Bayesian framework for obtaining sparse solutions to regression and classification tasks utilizing models linear in the parameters. Although this framework is fully general, system illustrate our approach with a particular specialisation that system denote the 'relevance vector machine' (RVM)[8], a model of identical functional form to the popular and state-of-the-art 'support vector machine' (SVM)[9]. System demonstrate that by exploiting a probabilistic Bayesian learning framework, system can derive accurate prediction models which typically utilise dramatically fewer basis functions than a comparable SVM while offering a number of additional advantages. These include the benefits of probabilistic predictions, automatic estimation of 'nuisance' parameters, and the facility to utilise arbitrary basis functions (e.g. non-'Mercer' kernels).

System details the Bayesian framework and associated learning algorithm for the RVM, and give some illustrative examples of its application along with some comparative benchmarks. System offers some explanation for the exceptional degree of sparsity obtained, and discusses and demonstrates some of the advantageous features, and potential extensions, of Bayesian relevance learning.

2.4 Empirical filter estimation for sub pixel interpolation and matching

System studies the low-level problem of predicting pixel intensities after sub pixel image translations. This is a basic subroutine for image warping and super-resolution, and it has a critical influence on the accuracy of subpixel matching by image correlation. Rather than using traditional frequency-space filtering theory or ad hoc interpolators such as splines, system take an empirical approach, finding optimal subpixel interpolation filters by direct numerical optimization over a large set of training examples. The training set is generated by subsampling larger images at different translations, using subsamplers that mimic the spatial response functions of real pixels. System argues that this gives realistic results, and design filters of various different parametric forms under traditional and robust prediction error metrics. System systematically study the performance of the resulting filters, paying particular attention to the influence of the underlying image sampling regime and the effects of aliasing (“jaggies”). System summarizes the results and gives practical advice for obtaining subpixel accuracy.

3. PROPOSED DESIGN

Secure image retrieval through feature encryption and index encryption are closely related. Image features themselves can be considered as a special form of search index, where each image is represented by its feature vectors and during retrieval, the query image's feature is compared to all features in the database. On the other hand, modern indexing schemes are built upon image features and allow efficient retrieval by reducing the number of images that need to be compared. Since the encrypted features preserve the capability of similarity comparison, they can be used to build efficient indexing schemes.

The content owner has the flexibility either to provide the server with encrypted features and let the server perform the time consuming index generation, or to generate the secure index on his/her side to reduce the amount of information that needs to be sent to the server. Therefore, the two kinds of approaches represent different trade-offs between user-side computational complexity and communication overhead.

So this proposed technique for image search and secure storage options will be well handled by the **APKET (Asynchronous Private Key based Encryption Technique)** that shows the complete solution to the problem of image retrieval from the cluster of random images. This system feature protection schemes that enable similarity comparison between features in the encrypted domain. The encrypted features along with encrypted images can protect image content privacy against untrustworthy service providers and malicious intruders. The ability to generate encrypted indexes on the user side provides an alternative for secure retrieval with reduced communication overhead. In the second part of this section, system discusses the protection of search indexes by exploiting the visual words representation of images. Visual words method hierarchically clusters features into a vocabulary tree, following which each image is indexed based on this vocabulary tree and represented as a bag of visual words.

Unlike dedicated servers, cloud servers can be run on a hypervisor. The role of a hypervisor is to control the capacity of operating systems so it is allocated where needed. With cloud hosting there are multiple cloud servers which are available to each particular client. This allows computing resource to be dedicated to a particular client if and when it is necessary. Where there is a spike in traffic, additional capacity will be temporarily accessed by a website, for example, until it is no longer required. Cloud servers also offer more redundancy. If one server fails, others will take its place.

3.1 CLIENT REQUEST METRIC

Clients are used to create requests, create transactions, send requests through an HTTP handler, and return a response. This exchange of messages is an example of inter-process communication. To communicate, the computers must have a common language, and they must follow rules so that both the client and the server know what to expect. The language and rules of communication are defined in a communications protocol. All client-server protocols operate in the application layer. The application-layer protocol defines the basic patterns of the dialogue. To formalize the data exchange even further, the server may implement an API such as a web service. The API is an abstraction layer for such resources as databases and custom software. By restricting communication to a specific content format, it facilitates parsing. By abstracting access, it facilitates cross-platform data exchange.

3.2 64-BIT OCTET ENCRYTION PROCESS

In many applications of public-key cryptography, user security is ultimately dependent on one or more secret values. Since a image is not directly applicable as a key to any conventional cryptosystem, however, some processing of the image is required to perform cryptographic operations with it. Moreover, as images are often chosen from a relatively small space, special care is required in that processing to defend against search attacks. It is difficult for an opponent to recompute all the keys corresponding to a dictionary of images, or even the most likely keys. If the salt is 64 bits long, for instance, there will be as many as 2^{64} keys for each image. An opponent is thus limited to searching for images after a image-based operation has been performed and the salt is known.

3.3 OTC ALLOCATION

Passwords are used by almost all business applications for authentication. However static passwords have lots of limitations and can get hacked; careless employee may write down passwords somewhere; system with saved passwords may be used by various users or a malicious user may reset all passwords just to create havoc. Hence it is advisable to move to a more dynamic password scheme like one time passwords or OTC. OTC is way more secure than static passwords as there are no chances to forget or reuse passwords. Each time a new password is generated for each login session. Authentication by one time passwords are more reliable and user friendly as well. OTC generation can be done by OTC generation algorithms for generating strings of passwords.

3.4 IMAGE RETRIVAL USING APKET

In order to overcome these difficulties APKET (Asynchronous Private Key based Encryption Technique) was introduced. APKET (Asynchronous Private Key based Encryption Technique) is the application of computer vision to the image retrieval problem. In this approach instead of being manually annotated by textual keywords, images would be indexed using their own infomation contents .The information contents may be color, texture and shape. This approach is said to be a general framework of image retrieval .There are three fundamental bases for Content Based Image Retrieval which are information feature extraction, multidimensional indexing and retrieval system design. The color aspect can be achieved by the techniques like averaging and histograms. The texture aspect can be achieved by using transforms or vector quantization shown in the figure 1.

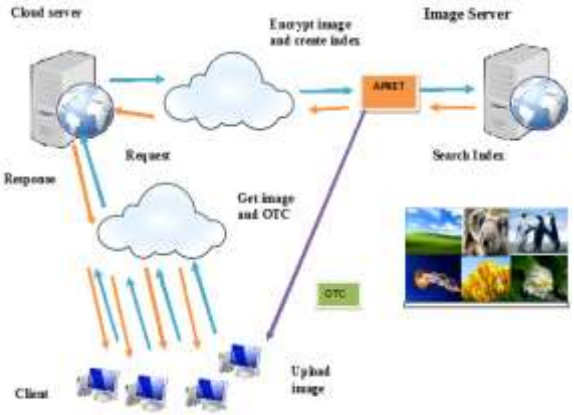

Figure 1 shows the architecture of proposed model

4. RESULTS AND DISCUSSIONS

Security-related aspects are one of the main factors hampering further cloud adoption. The inherent multi-tenancy of the environment, and the existence of three different actors in a normal scenario the end user, the application provider, and the cloud provider, raise a significant number of security issues. A cloud environment presents additional security challenges. Instead of operating in a single data center, controlled by one company, cloud applications can operate over a multi-tenant infrastructure from different providers. There is a need for extending security best practices to a federated environment. Federated identity and trust have been some of the key elements of research and industry since the early days of service-oriented architecture applications. These advances can be applied to both the applications accessing the infrastructure, and also to inter-cloud communications. Cross cloud solutions also need to potentially consider the networking aspect of the federated infrastructure. Data must not only be stored securely, but communications need to be protected with the adequate means.

For geo-distributed applications, a fully centralized management plane presents some difficulties. Decisions must consider all the relevant runtime information collected at the distributed instances. Therefore, decision algorithms can either be run on the distributed environment, potentially incurring on significant penalty from the network component, or run on a centralized location, after having moved all the information to a single points. In any of these cases, the decision will require the use of parallel computing techniques for large-scale services. The cited papers report total computation time in the order of hours and days, making it not well suited for a highly dynamic environment. A decentralized management scheme can process most of the monitoring information locally, and only exchange a subset of that information with its peers. Some research studies show how decentralized decision making can be taken to geo-distributed application, combining the ability to enforce user-defined policies with an added level of resiliency over a central solution.

CONCLUSION

The main differentiating aspect of cross-cloud applications is the nature of their physical distribution as an overlay across networks and data centers. This characteristic makes the network element critical for an effective deployment and management. In the project system have highlighted how. **APKET (Asynchronous Private Key based Encryption Technique)** enabled networks can become a key element for fully realizing the vision behind these applications, with the ability to provide virtual links with certain guarantees, support to seamlessly integrate network-wide services and support for the low-level activities related to WAN-scale virtual machine migration. System believe a networked cloud marketplace might provide incentives for the different infrastructure stakeholders to collaborate, including the potential for better infrastructure management based on using information about the real needs of the applications that use the infrastructure. On the application side, the location dimension greatly impacts application architecture and management. Decisions become substantially more complex; a cross-cloud infrastructure provides the required means for applications to achieve satisfactory performance for a changing workload of users around the world. The architecture of these applications can resemble a dynamic, distributed overlay that raises several challenges regarding how to manage the internal application state, provide service reliability, and ensure security requirements.

FUTURE WORK

As there are other algorithms that may well be candidates for software implementation in the kernel. It should be quite interesting to implement all of these algorithms and to test them on equal footing, using the same hardware, rule-bases, and traffic load. Furthermore, it would be interesting to do this comparison with real rule-bases, in addition to synthetic Perimeter-model rules. The system leaves such a “bake-off” for future work. As for GEM itself, the system would like to explore the algorithm’s behavior when using more than 4 fields, e.g., matching on the TCP flags, Meta data, interfaces, etc. The main questions are: How best to encode the non-range fields? Will the space complexity still stay close to linear? What will be the best order of fields to achieve the best space complexity? Another direction to pursue is how GEM would perform with of IPv6, in which IP addresses have 128 bits.

REFERENCES:

[1] D. Song, D. Wagner, and A. Perrig, ``Practical techniques for searches in encrypted data,'' in *Proc. IEEE Symp. Res. Sec. Privacy*, Feb. 2000, pp. 44_55.

[2] R. Brinkman, J. M. Doumen, and W. Jonker, ``Using secret sharing for searching in encrypted data,'' in *Proc. Workshop Secure Data Manag.Connected World*, 2004, pp. 18_27.

[3] D. Boneh, G. Crescenzo, R. Ostrovsky, and G. Persiano, ``Public-keyencryption with keyword search,'' in *Proc. Eur.*, 2004, pp. 506_522.

[4] A. Swaminathan, Y. Mao, G.-M. Su, H. Gou, A. L. Varna, S. He, *et al.*,``Con_dentiality preserving rank-ordered search,'' in *Proc. ACMWorkshopStorage, Sec., Survivability*, 2007, pp. 7_12.

[5] Z. Erkin, A. Piva, S. Katzenbeisser, R. L. Lagendijk, J. Shokrollahi,G. Neven, *et al.*, ``Protection and retrieval of encrypted multimedia content:When cryptography meets signal processing,'' *EURASIP J. Inf. Sec.*, vol. 7,no. 2, pp. 1_20, 2007.

[6] R. Datta, D. Joshi, J. Li, and J. Z.Wang, ``Image retrieval: Ideas, in_uences,and trends of the new age,'' *ACM Comput. Surveys*, vol. 40, no. 2, pp. 1_5,2008.

[7] J. Shashank, P. Kowshik, K. Srinathan, and C. Jawahar, ``Private contentbased image retrieval,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*,Jun. 2008, pp. 1_8.

[8] M.-L. Yiu, G. Ghinita, C. S. Jensen, and P. Kalnis, ``Outsourcing searchservices on private spatial data,'' in *Proc. IEEE 25th Int. Conf. Data Eng.*,Apr. 2009, pp. 1140_1143.

[9] W. K. Wong, D. W.-L. Cheung, B. Kao, and N. Mamoulis, ``Secure kNNcomputation on encrypted databases,'' in *Proc. 35th SIGMOD Int. Conf.Manag. Data*, 2009, pp. 139_152.

[10] Z. Erkin, M. Franz, J. Guajardo, S. Katzenbeisser, I. Lagendijk, andT. Toft, ``Privacy-preserving face recognition,'' *Privacy Preserving Tech-nol., LNCS*, vol. 5672, pp. 235_253, Aug. 2009.

[11] A.-R. Sadeghi, T. Schneider, and I. Wehrenberg, ``Ef_cient privacypreservingface recognition,'' in *Proc. 12th Int. Conf. Inf. Sec. Cryptol.*,2009, pp. 229_244.

[12] M. Osadchy, B. Pinkas, A. Jarrous, and B. Moskovich, ``Sci__A systemfor secure face identi_cation,'' in *Proc. IEEE Symp. Sec. Privacy*,May 2010, pp. 239_254.

[13] W. Jiang, M. Murugesan, C. Clifton, and L. Si, ``Similar document detectionwith limited information disclosure,'' in *Proc. IEEE 24th Int. Conf.Data Eng.*, Apr. 2008, pp. 735_743.

[14] D. E. Yan Huang, L. Malka, and J. Katz, ``Ef_cient privacy-preservingbiometric identi_cation,'' in *Proc. 18th Network Distrib. Syst. Sec. Symp.*,2011, pp. 1_9.

[15] W. Lu, A. L. Varna, A. Swaminathan, and M.Wu, ``Secure image retrievalthrough feature protection,'' in *Proc. IEEE Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 1533_1536.

[16] W. Lu, A. Swaminathan, A. L. Varna, and M. Wu, ``Enabling search overencrypted multimedia databases," *Proc. SPIE*, vol. 7254, pp. 7254_7318,Jan. 2009.
[17] Y. Mao and M.Wu, ``A joint signal processing and cryptographic approachto multimedia encryption," *IEEE Trans. Image Process.*, vol. 15, no. 7,pp. 2061_2075, Jul. 2006.
[18] M. Grangetto, E. Magli, and G. Olmo, ``Multimedia selective encryption bymeans of randomized arithmetic coding," *IEEE Trans. Multimedia*, vol. 8,no. 5, pp. 905_917, Oct. 2006.
[19] H. Kim, J. Wen, and J. D. Villasenor, ``Secure arithmetic coding," *IEEETrans. Signal Process.*, vol. 55, no. 5, pp. 2263_2272, May 2007.
[20] D. Lowe, ``Distinctive image features from scale-invariant keypoints," *Int.J. Comput. Vis.*, vol. 60, no. 2, pp. 91_110, 2004.