Literature Review on a secure hashing scheme for Image Authentication

Samruddhi S.kuralkar, Dr. P. R. Deshmukh

Abstract— This paper presents a robust and secure image hash algorithm. The algorithm extracts robust image features in the Radon transform domain. A randomization mechanism is designed to achieve good discrimination and security. The hash value is dependent on a secret key .

We evaluate the performance of the proposed algorithm and compare the results with those of one existing Radon transform-based algorithm. We show that the proposed algorithm has good robustness against content preserving distortion. It with stands JPEG compression, filtering, noise addition as well as moderate geometrical distortions. Additionally, we achieve improved performance in terms of discrimination, sensitivity to malicious tampering and receiver operating characteristics. We also analyze the security of the proposed algorithm using differential entropy and confusion /diffusion capabilities. Simulation shows that the proposed algorithm well satisfies these metrics.

***Index Terms*— Hash Value, Hash Strength, Image Fusion, Visual Cryptography**

I.INTRODUCTION

In order to efficiently identify digital images, perceptual hash techniques have been used. With the rapid growth of multimedia applications, protection of intellectual property is becoming more prominent. Image hashing is useful in protection of intellectual property[1]. In this technique an image hashing function takes an image as input and computes a short hash key which is a random value in some large set.

The performance of a perceptual image hashing technique primarily consists of robustness, discrimination, and security. Robustness means the technique always generates the same (or similar) hash values for similar image contents[2]. Discrimination means different image inputs must result in independent (different) hash values.

Hash key is a measure of the effectiveness of a password in resisting guessing and brute-force attacks. In its usual form, it estimates how many trials an attacker who does not have direct access to the hash key would need, on average, to guess it correctly. The strength of a hash key is a function of length, complexity, and unpredictability [3]. Using strong hash key lowers overall risk of a security breach, but strong hash key do not replace the need for other effective security controls.

Since an image can be stored under different digital representations, a perceptual hash value is expected to be resilient to content-preserving manipulations. perceptual hash technique are also useful for secure applications such as image content authentication. an image hash technique is also required to make the hash output dependent on a secret key[4]. Only the entity knowing the key can generate the correct hash value.

The effectiveness of a hash key of a given strength is strongly determined by the design and implementation of the authentication system software, particularly how frequently hash key guesses can be tested by an attackher and how securely information on user hash key is stored and transmitted[5,6]. Risks are also posed by several means of breaching computer security which are unrelated to hash key strength. There are two factors to consider in determining hash key strength: the average number of guesses the attacker must test to find the correct hash key and the ease with which an attacker can check the validity of each guessed hash key.

A number of media-specific hash functions have been proposed for multimedia authentication [4,7]. A multimedia hash is a content-based string. To generate a multimedia hash, a secret key is used to extract certain features from the data. These features are further processed to form the hash. The hash is transmitted along with the media either by appending or embedding it to the primary media data. At the receiver side, the authenticator uses the same key to generate the hash values, which are compared to the ones transmitted along with the data for verifying its authenticity.

In addition to content authentication, multimedia hashes are used in content-based retrieval. To search for multimedia content, naïve methods such as sample-by-sample comparisons are computationally inefficient. Moreover, these methods compare the lowest level of content representation and do not offer robustness in such situations as geometric distortions. Robust image hash functions can be used to address this problem [8]. Image hash functions have also been used in applications involving The hash functions have also been used as image Dependent keys.

There are two important design criteria for image hash functions, namely, robustness and security. By robustness, we mean that when the same key is used, perceptually similar images should produce similar hashes[8,9]. The security of image hash functions is

Samruddhi S. Kuralkar *,computer science and engg.,sant gadage baba university.,.Amravati,India*
9561620747

Dr. Prashant R. Deshmukh *,computer science and engg.,sant gadage baba university.,.Amravati,India* 09423610594

ISSN: 2278 – 1323

introduced by incorporating a secret key in generating the hash. Without the knowledge of the key, the hash values should not be easily forged or estimated.

II.LITERATURE REVIEW AND RELATED WORK

In this paper we outline the traditional approach to the hashing problem. Many researchers suggested that Hash functions, that are useful in various field for secure mapping binary strings .

Sheng Tang, Jin-Tao Li, and Yong-Dong Zhang[1] suggested research Article on Image hashing is an alternative approach to many applications accomplished with watermarking. In this paper, they propose a novel image hashing method in the DCT Domain which can be directly extended to MPEG video without DCT transforms. A key goal of the method is to produce randomized hash signatures which are un-predictable for unauthorized users.

Sujoy Roy Qibin Sun[2] suggested that An image hash is a short signature of the image that preserves its semantic information under allowable changes made to it while at the same time differentiates it from a different image. That is, it should be robust to allowable modifications (like small rotations, compression, scaling, addition of noise etc) and sensitive to distinct images or illegal manipulations to the original like tampering. Hashes find application in verifying the authenticity of protected content.

R. Venkatesan1, S.-M. Koon, M. H. Jakubowski, and P. Moulin[3] suggested that they introduce a novel algorithm that utilizes a wavelet representation of images and new randomized processing strategies for hashing. Constructions based on error correcting codes serve to reduce the length of the hash value while keeping collision probability low. The statistical properties of the hash values on distinct images are similar to the properties used in cryptography to minimize collision probabilities . An extension of this algorithm was extensively tested and found to be robust against common image processing and malicious attacks.

Ashwin Swaminathan, , Yinian Mao, and Min Wu[4] suggested in JUNE 2006 that the proposed hash function is resilient to content-preserving modifications, such as moderate geometric and filtering distortions. They introduce a general framework to study and evaluate the security of image hashing systems. Under this new framework, we model the hash values as random variables and quantify its uncertainty in terms of differential entropy. Using this security framework, they analyze the security of the proposed schemes and several existing representative methods for image hashing.

Ammar M. Hassan , Yassin M. Y. Hasan and Mohamed A. A. Wahab[5] suggested that In this paper, secure and robust image (visual) hashing based on the discrete wavelet transform (DWT) and non-negative matrix factorization (NMF) is proposed. In the proposed method, the considered image is undergone pre processing manipulations. Then, the DWT is applied on the pre processed image to generate image features that are largely invariant under perceptually insignificant manipulations.

Zhenjun Tang, Shuozhong Wang, Xinpeng Zhang, Weimin Wei, and Shengjun Su[6]suggested in MAY 2008 that proposed scheme is robust against perceptually acceptable modifications to the image such as Gaussian filtering, moderate noise contamination, JPEG compression, re-scaling, and watermark embedding. Hashes of different images have very low collision probability. Tampering to local image areas can be detected by comparing the Hamming distance with a predetermined threshold, indicating the usefulness of the technique in digital forensics.

Chun-Shien Lu · Chao-Yong Hsu suggested in 17 October 2005 that they work on Geometric distortion-resilient image hashing scheme and its applications on copy detection and authentication scheme[9]. A robust mesh-based image-hashing scheme for content management of digital images has been proposed in this paper. The scheme is mainly composed of two components: (i) mesh-based robust hash generation and (ii) hash database construction for error-resilient and fast searching. In comparison with the existing methods, the main contribution of there approach is that it significantly improves the resistance of image hashing to geometric distortions[10]. Furthermore, they had investigated several media hashing issues, including robustness and discrimination, error analysis, complexity, granularity, and scalability. they had also demonstrated application of the robust mesh-based image hashing system to both copy detection and content authentication.

III.TRADITIONAL APPROACH FOR HASHING

1. Human-generated hash key

People are notoriously remiss at achieving sufficient entropy to produce satisfactory hash key. Some stage magicians exploit this inability for amusement, in a minor way, by divining supposed random choices (of numbers, say) made by audience members. Thus, in one analysis of over 3 million eight-character hash key, the letter "e" was used over 1.5 million times, while the letter "f" was only used 250,000 times. A uniform distribution would have had each character being used about 900,000 times. The most common number used is "1", whereas the most common letters are a, e, o, and r. Users rarely make full use of larger characters sets in forming hash key. For example, hacking results obtained from a MySpace phishing scheme in 2006 revealed 34,000 passwords, of which only 8.3% used mixed case, numbers, and symbols.[8]Note that the full strength associated with using the entire ASCII character set (numerals, mixed case letters and special characters) is only achieved if each character in the password is chosen randomly from that set. Capitalizing a letter and adding a couple of numbers and a special character to a hash key will not achieve the same strength. If the numbers and special character are added in predictable ways, say at the beginning and end of the hash key, they could even lower

hash key strength compared to an all letter random hash key of the same length.

- *NIST Special Publication 800-63*

NIST Special Publication 800-63 suggests the following scheme to roughly estimate the entropy of human-generated hash key: The entropy of the first character is four bits;

- The entropy of the next seven characters are two bits per character;
- The ninth through the twentieth character has 1.5 bits of entropy per character;
- Characters 21 and above have one bit of entropy per Character.
- A "bonus" of six bits is added for passwords of length 1 Through 19 characters following an extensive dictionary Check to ensure the password is not contained within a large dictionary. hash key s of 20 characters or more do not receive this bonus because it is assumed they are pass-phrases consisting of multiple dictionary words.

Using this scheme, an eight-character human-selected hash key without upper case letters and non-alphabetic characters is estimated to have 18 bits of entropy. The NIST publication concedes that at the time of development, little information was available on the real world selection of hash key. Later research into human-selected hash key entropy using newly available real world data has demonstrated that the NIST scheme does not provide a valid metric for entropy estimation of human-selected hash key.[10]

- Usability and implementation considerations

Because national keyboard implementations vary, not all 94 ASCII printable characters can be used everywhere. This can present a problem to an international traveler who wished to log into remote system using a keyboard on a local computer. *See* keyboard layout. Many hand held devices, such as tablet computers and smart phones, require complex shift sequences to enter special characters. Authentication programs vary in which characters they allow in hash key. Some do not recognize case differences (e.g., the upper-case "E" is considered equivalent to the lower-case "e"), others prohibit some of the other symbols. In the past few decades, system have permitted more characters in hash key, but limitations still exist. Systems also vary in the maximum length of hash key allowed.

- ***3*** *Bit strength threshold*

As a practical matter, hash key must be both reasonable and functional for the end user as well as strong enough for the Intended purpose. hash key that are too difficult to remember may be forgotten and so are more likely to be written on paper, which some consider a security risk.[1] In contrast, others argue that forcing users to remember hash key without assistance can only accommodate weak passwords, and thus poses a greater security risk. According to Bruce Schneider, most people are good at securing their wallets purses, which is a "great place" to store a written password. Some basic benchmarks have been established for brute force searches in the context of attempting to find keys used in encryption. The problem is not the same since these approaches involve astronomical numbers of trials, but the results are suggestive for password choice.

In 1999, an Electronic Frontier Foundation project broke 56-bit DES encryption in less than a day using specially designed hardware.[7] In 2002, distributed.net cracked a 64-bit key in 4 years, 9 months, and 23 days. As of October 12, 2011, distributed.net estimates that cracking a 72-bit key using current hardware will take about 45,579 days or 124.8 years. Due to currently understood limitations from fundamental physics, there is no expectation that any digital computer (or combination) will be capable of breaking 256-bit encryption via a brute-force attack. Whether or not quantum computers will be able to do so in practice is still unknown, though theoretical analysis suggests such possibilities.

As a result, there can be no exact answer to the somewhat different problem of the password strength required to resist brute force attack in practice. NIST recommends 80-bits for the most secure passwords, which can nearly be achieved with a 95-character choice (e.g., the original ASCII character set) with a 12-character random password (12 x 6.5 bits = 78).[2] A 2010 Georgia Tech Research Institute study also recommended a 12-character random password, but as a minimum length requirement.[3]

There are some more drawbacks in existing system

- In existing system hash key is generated from text so it is easily hack by unauthorized users.
- Sequence of the hash key is easily remember by unauthorized users.

Rapid growth of multimedia application ,protection of intellectual property is becoming more prominent . In this method we are trying to come back from this drawback and design a strong hash key that breach security. Following are the objectives of our hash key technique

- Hash key is not generated by text but it is generated from images.
- We perform image fusion that combines all selected images into a single image that's why no one can guess the way of generation of hash key.
- Once the images are fused , we will apply tow shares cryptography that encrypt the image so it become difficult to hack to unauthorized users.
- The Key is Mix up all the generated sections with permutations so that every time & in every round an Unique hash key will generate.

ISSN: 2278 – 1323

REFERENCES

[1]C-Y Lin and S-F Chang, "A Robust image authentication method", IEEE Trans circuits syst. Video technol. 11(2):153-168,200 distinguishing JPEG Compression from malicious manipulation

[2]Sujoy Roy Qibin Sun:Robust hash for detecting and localizing image tampering.

[3]R. Venkatesan1, S.-M. Koon, M. H. Jakubowski, and P. Moulin " Robust and secure image hashing", proc.IEEE ICIP ,Vancouver, Canada September 2000.

[4]S.S Kozat ,K .Mlhcak and R Venkatesan, "Robust Perceptual image hashing via matrix invariance ",In proc IEEE Conf. On image Processing ,Oct,2004

[5]Swaminathan, A., Mao, Y., Wu, M.: Robust and secure image hashing. In: IEEETransactions on Information Forensics and Security, Vol.1, No. 2. (June 2006).

[6] Ammar M. Hassan , Yassin M. Y. Hasan and Mohamed A. A. Wahab :Robust visual hashing for image authentication 2000.

[7].Zhenjun Tang, Shuozhong Wang, Xinpeng Zhang, Weimin Wei, and Shengjun Su: Robust image hashing for tamper detection using Non-Negative Matrix Factorization in journal of ubiquitous convergence and technology, vol.2,no.1,may 2008

[8].Gerold Laimer and Andreas Uhl: Key-Dependent JPEG2000-Based Robust Hashing for Secure Image Authentication Received 31 May 2007; Accepted 12 December 2007

[9]Vishal Monga, Arindam Banerjee, and Brian L. Evans, "Clustering Algorithms forPerceptual Image Hashing", Proc. IEEE Work. on Digital Signal Processing, Aug.1-4, 2004, pp. 283-287, Taos, NM.

[10]J. Fridrich and M. Goljan, "Robust hash functions for digital watermarking", Proc. IEEE Int. Conf. Info. Tech.: Coding and Comp., Mar. 2000.

[11]F. Lefebvre, J. Czyz, and B. Macq, "A robust soft hash algorithm for digital image signature", In: Proceedings of IEEE ICIP 2003, Barcelona, II:495-498, Sept. 2003.

[12]F. Lefebvre, B. Macq, and JD Legat, "RASH: RAdon Soft Hash Algorithm", In11th European Signal Processing Conference, Toulouse, France, Sept. 2002.

[13] Mao, Y., Wu, M.: Unicity distance of robust image hashing. In: IEEE Transactions on Information Forensics and Security, Vol. 2, No. 3. (Sep. 2007)

Authors Information

Samruddhi S. Kuralkar : Assistant Professor at Sipna College of Engineering & Technology, Amravati. She did her B.E. in Computer Science and Engineering from Amravati University. And She is also pursuing her M.E. in Department of Computer Science and Engineering from Sipna College of Engineering & Technology. And her area of interest is Digital image processing.

Dr. P.R. Deshmukh : Professor at Sipna College of Engineering & Technology, Amravati. He did his B. E. in 1988, M. E in 1997 and Doctor of Philosophy in Engineering (Ph.D.) in 2005. He has more than 20 years of Teaching Experience and various Microprocessor/Microcontroller based/VHDL based projects undertaken. He has published many research papers in National as well as International journal and conferences. And also he is a member of IEEE. His area of interest is Digital Imag Processing.