

APPLYING DIGITAL SIGNATURE TO PRINTOUT FOR TRUSTWORTHY VERIFICATION USING  
IMAGE HASH

Mr. Paradorn Athichitsakul



บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)  
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)  
are the thesis authors' files submitted through the University Graduate School.

A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Science Program in Computer Science and Information  
Technology

Department of Mathematics and Computer Science

Faculty of Science

Chulalongkorn University

Academic Year 2015

Copyright of Chulalongkorn University

การประยุกต์ลายมือชื่อดิจิทัลกับสิ่งพิมพ์ออกเพื่อการสอบทวนที่น่าเชื่อถือโดยใช้แฮชภาพ



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต  
สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ภาควิชาคณิตศาสตร์และวิทยาการ

คอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2558

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title	APPLYING DIGITAL SIGNATURE TO PRINTOUT FOR TRUSTWORTHY VERIFICATION USING IMAGE HASH
By	Mr. Paradorn Athichitsakul
Field of Study	Computer Science and Information Technology
Thesis Advisor	Assistant Professor Suphakant Phimoltares, Ph.D.
Thesis Co-Advisor	Atchara Mahaweerawat, Ph.D.

---

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

..... Dean of the Faculty of Science  
(Associate Professor Polkit Sangvanich, Ph.D.)

THESIS COMMITTEE

..... Chairman  
(Professor Chidchanok Lursinsap, Ph.D.)

..... Thesis Advisor  
(Assistant Professor Suphakant Phimoltares, Ph.D.)

..... Thesis Co-Advisor  
(Atchara Mahaweerawat, Ph.D.)

..... External Examiner  
(Saichon Jaiyen, Ph.D.)

ภราดร อธิจิตสกุล : การประยุกต์ลายมือชื่อดิจิทัลกับสิ่งพิมพ์ออกเพื่อการสอบทวนที่น่าเชื่อถือโดยใช้แฮชภาพ (APPLYING DIGITAL SIGNATURE TO PRINTOUT FOR TRUSTWORTHY VERIFICATION USING IMAGE HASH) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร. ศุภกานต์ พิมลธเรศ, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: อ. ดร. อัจฉรา มหาวีรวัฒน์, 48 หน้า.

กระบวนการออกใบรับรองหรือเอกสารสำคัญโดยส่วนมากจะอยู่ในรูปแบบการพิมพ์ลงบนกระดาษ บางกรณีปรับปรุงความน่าเชื่อถือของเอกสารกระดาษเหล่านั้นด้วยการใช้กระดาษหรือวิธีการพิมพ์ที่ป้องกันการปลอมแปลง เช่น การใส่ลายน้ำ การใส่ภาพสามมิติ ประเด็นปัญหาที่น่าสนใจคือการใช้กลไกความปลอดภัยนี้ไม่เพียงพอเพราะทุกวันนี้เครื่องพิมพ์ที่มีวิธีการพิมพ์ที่ป้องกันการปลอมแปลงดังกล่าวสามารถหาได้ง่ายมากขึ้นซึ่งหมายความว่าความปลอดภัยของเอกสารสามารถทำได้โดยไม่มี ความยุ่งยากเลย เพื่อแก้ปัญหาข้างต้นจึงได้เสนอกรอบกระบวนการความปลอดภัยที่ใช้ฟังก์ชันการแฮชภาพในการจับอัตลักษณ์ของเอกสารและสร้างลายมือชื่อดิจิทัลบนโครงสร้างพื้นฐานกุญแจสาธารณะโดยมีจุดประสงค์สำหรับการรับรองคุณภาพของอัตลักษณ์

กรอบกระบวนการนี้ประกอบด้วยสามองค์ประกอบหลักดังนี้ องค์ประกอบแรกคือฟังก์ชันการแฮชภาพสำหรับจับรูปลักษณ์ของเอกสารที่พิมพ์แล้วนำมาใช้เพื่อสอบทวนคุณภาพของเอกสารเหล่านี้ องค์ประกอบที่สองคือลายมือชื่อดิจิทัลบนโครงสร้างพื้นฐานกุญแจสาธารณะสำหรับป้องกันการปลอมแปลงแก้ไขค่าแฮชภาพ พิสูจน์ตัวตนผู้ลงลายมือชื่อบนเอกสาร และรับรองความรับผิดชอบของเนื้อหาจากผู้ลงลายมือชื่อบนเอกสาร องค์ประกอบที่สำคัญสุดท้ายของกรอบกระบวนการที่นำเสนอนี้คือบาร์โค้ดสองมิติสำหรับใช้เพื่อฝังผลที่ได้จากกระบวนการจัดทำลายมือชื่อดิจิทัลลงบนเอกสารที่พิมพ์เหล่านี้

ภาควิชา	คณิตศาสตร์และวิทยาการ	ลายมือชื่อนิติ	.....
	คอมพิวเตอร์	ลายมือชื่อ อ.ที่ปรึกษาหลัก	.....
สาขาวิชา	วิทยาการคอมพิวเตอร์และเทคโนโลยีลายมือชื่อ	อ.ที่ปรึกษาร่วม	.....
	สารสนเทศ		

ปีการศึกษา 2558

# # 5672605523 : MAJOR COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

KEYWORDS: DIGITAL SIGNATURE / IMAGE HASHING ALGORITHM / RADON TRANSFORM /  
WAVELET TRANSFORM / FAST FOURIER TRANSFORM

PARADORN ATHICHITSAKUL: APPLYING DIGITAL SIGNATURE TO PRINTOUT FOR  
TRUSTWORTHY VERIFICATION USING IMAGE HASH. ADVISOR: ASST. PROF.  
SUPHAKANT PHIMOLTARES, Ph.D., CO-ADVISOR: ATCHARA MAHAWEEERAWAT,  
Ph.D., 48 pp.

Most certificate or document issuing processes are in a form of printing document to a paper. Some might improve the trustworthiness of their paper via using secure paper or secure printing method such as watermark, hologram etc. The interesting problem is that this security mechanism is not enough because nowadays printer with secure printing module can be obtained easier than before, which means reproducing counterfeit document can be done uncomplicatedly. To solve this problem, a security framework using image hashing function to capture document fingerprint and perform a digital signature based on PKI infrastructures is proposed to guarantee this fingerprint integrity.

This framework consists of three important components. The first component is image hashing function for capturing perceptual appearance of printing document and it is then used to verify integrity of their document. The second component is digital signature based on PKI infrastructure for protecting image hash value from modification, authenticating document signer and guaranteeing content commitment from the document signer. The last important component of this proposed framework is 2D barcode for embedding result from the digital signature process that belongs to their printing document.

Department: Mathematics and Student's Signature .....

Computer Science Advisor's Signature .....

Field of Study: Computer Science and Co-Advisor's Signature .....

Information Technology

Academic Year: 2015

## ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisors, Assistant Professor Dr. Suphakant Phimoltares and Dr. Atchara Mahaweerawat. Their guidance helped me in all the time of research and writing of this thesis. Without my advisors' assistance, this thesis definitely cannot be complete.

I would also like to thank to my thesis committee Professor Dr. Chidchanok Lursinsap and Dr. Saichon Jaiyen for their valuable questions and comments to improve this thesis.

Besides my advisors and my thesis committee, I would like to thank Electronic Transactions Development Agency (Public Organization) colleagues, especially Research and Development team and Mr. Thitikorn Trakoonsirisak, who reviewed and commented my research.

Finally, my sincere thanks also goes to my family and Miss Patthanan Thummachadee for their consistently supports all over the time since I started my thesis.



## CONTENTS

	Page
THAI ABSTRACT .....	iv
ENGLISH ABSTRACT .....	v
ACKNOWLEDGEMENTS .....	vi
CONTENTS .....	vii
CONTENTS OF TABLES .....	x
CONTENTS OF FIGURES .....	xi
Chapter 1. Introduction .....	1
1.1 Objective.....	3
1.2 Problem formulation.....	3
1.3 Scope of thesis and constraints.....	3
1.4 Expected outcome .....	3
Chapter 2. Theoretical Background.....	4
2.1 Image Hashing Function.....	4
2.2 Digital Signature.....	4
Chapter 3. Related works.....	8
3.1 Image Hashing Function.....	8
3.1.1 Radon Transform .....	8
3.1.2 Wavelet Transform.....	11
3.2 Enhanced Digital Signature .....	12
3.2.1 CMS Advanced Electronic Signature (CADES).....	13
3.2.1.1 CADES - BES .....	13
3.2.1.2 CADES - EPES.....	14

	Page
3.2.1.3 CAdES - T .....	15
3.2.1.4 CAdES - LT .....	16
3.2.2 JSON Web Signature (JWS).....	17
3.3 2D Barcode .....	18
3.3.1 PDF 417.....	18
3.3.2 QR Code .....	19
Chapter 4. Proposed Methods.....	21
4.1 Framework for Creating and Verifying Trustworthy Printout.....	21
4.2 Image Hashing Function Design.....	23
4.3 Printable Digital Signature Design .....	25
Chapter 5. Experimental Results .....	29
5.1 Image Hash Evaluation.....	29
5.1.1 Image Hash Algorithm Performance .....	30
5.1.2 Image hash Algorithm Accuracy .....	33
5.1.2.1 Accuracy of the Proposed Algorithm .....	33
5.1.2.2 Accuracy of Wu, Zhou, and Niu's Algorithm.....	34
5.1.2.3 Accuracy of DCT-Based Algorithm .....	35
5.1.2.4 Accuracy of pHash's DCT-Based Algorithm .....	36
5.2 Printable Digital Signature Performance.....	38
5.2.1 Digital Signature Size .....	39
5.3 Printout and Printout Verification.....	40
Chapter 6. Analysis and Discussion.....	42
Chapter 7. Conclusion and Future Research .....	44



	Page
REFERENCES .....	45
VITA.....	48



## CONTENTS OF TABLES

Table 1 Radon coefficients matrix of different images.....	10
Table 2 Relationship of QR code version, error correction level and data capacity from version 3-10.....	20
Table 3. SignedData parameters description.....	26
Table 4. EncapsulatedContentInfo parameters description.....	26
Table 5. SignerInfo parameters description.....	26
Table 6. SignedAttributes parameters description.....	26
Table 7. Performance of Hashing under Attacks.....	32
Table 8. Performance and Accuracy of the Proposed Algorithm.....	33
Table 9. Performance and Accuracy of the Approach Proposed by [17].....	34
Table 10. Performance and Accuracy of DCT-Based Algorithm.....	35
Table 11. Performance and Accuracy of pHash's DCT-Based Algorithm [27].....	36
Table 12. Size comparison of digital signature with content data.....	39
Table 13. Size comparison of digital signature without content data.....	39
Table 14. Distance between specified regions obtained from original image and scanned image.....	41
Table 15. FRR of each algorithm at 15.25% FAR.....	42

## CONTENTS OF FIGURES

Figure 1. Fingerprint capturing process of printing document via OCR and cryptographic hashing function.....	2
Figure 2. Fingerprint capturing process of printing document via image hashing function.....	2
Figure 3. Process of digitally sign any digital content .....	5
Figure 4. Process of digital signature validation .....	5
Figure 5. Certificate validation Model.....	6
Figure 6. Radon transform result of straight line (a) Straight line image (b) Result from applying radon transform to straight line image (a) .....	9
Figure 7. Radon transform result of straight line with noise (a) Straight line image with noise (b) Result from applying radon transform to straight line image with noise (a) .....	9
Figure 8. Process of 1-D wavelet transform. ....	11
Figure 9. Process of 2-D wavelet transform .....	11
Figure 10. CADES-BES signature structure .....	14
Figure 11. CADES-EPES signature structure .....	14
Figure 12. CADES-T signature structure .....	15
Figure 13. CADES-LT signature structure .....	16
Figure 14. JWS digital signed content data structure .....	17
Figure 15. PDF 417 Symbol Structure .....	18
Figure 16. QR Code Symbol Structures (a) QR Code version 6 symbol structure (b) QR Code version 7 symbol structure.....	19
Figure 17. Printout creation process. ....	21
Figure 18. Printout verification process .....	22

Figure 19. Image hash value computation process .....	23
Figure 20. Output of each image hash value computation process. (a) Cropped Image, (b) Image from Radon transform, (c) Resized Image, (d) Image from Wavelet transform, (e) Image from Fourier transform, (f) Binary hash string. ....	23
Figure 21. CMS digital signature data structure .....	25
Figure 22. Proposed digital signature data structure .....	27
Figure 23. Process of barcode creation.....	28
Figure 24. Example test data in this experiment .....	29
Figure 25. Example images of each attacks .....	31
Figure 26. Accuracy of the proposed algorithm.....	33
Figure 27. Accuracy of Wu, Zhou, and Niu's algorithm .....	34
Figure 28. Accuracy of DCT-based algorithm.....	35
Figure 29. Accuracy of pHash's DCT-based algorithm. ....	36
Figure 30. Comparison of our algorithm and the other existing image hash algorithms in terms of FAR and FRR. ....	37
Figure 31. 256Bit ECC based Certificate .....	38
Figure 32. Image hash data set over 4 regions .....	38
Figure 33. Digital signature embedded printout.....	40
Figure 34. Scanned barcode .....	41
Figure 35. Large detail image (a) Original image (b) Edit image.....	43

## Chapter 1. Introduction

Printing documents such as photocopy of ID card or passport, academic certificate etc. are generally used as evidence or attachment in many kinds of processes in both private and public sectors. Benefit of using printing documents is not just convenient but also easy to verify alteration of document. Unfortunately, nowadays printing technologies are significantly improved. Printer with secure printing feature such as micro printing, watermark printing etc. can be obtained easier than before. Resulting in document counterfeiting, altering, and reproducing can be done uncomplicatedly. Establishing a trustworthy framework that can solve or mitigate this problem is a motivation of this thesis.

With existing technology, digital signature is a widely used method to make digital content discriminable between original one and alteration content. Reliability of the digital signature is based on “Public Key Infrastructure (PKI)” with appropriate PKI implementation [1, 2]. Also, this digitally signed content comes up with following feature, “Integrity guarantee of signed content”, “Authentication of content signer”, and “Non-repudiation of content signer”. With these features PKI digital signature is compiled and accepted as a reliable electronic signature in Thailand Electronic Transaction ACT [3]. Implementing digital signature to a printing document, one publication uses document content to be a source of fingerprint, and uses OCR to extract document content for fingerprint validation process [4]. Unfortunately, using OCR to extract document content from scanning process might face a problem from OCR recognition accuracy, which can be influenced by many factors such as unsupported language, unsupported font, etc.

In this thesis, another approach is presented for capturing printing document fingerprint. Image hashing algorithm is another alternative content fingerprinting method. Different from ordinary hashing algorithm, image hashing is used to extract certain features from multimedia content and calculate a hash value relied on these

features. Compared to using OCR to extract document and calculate the content's fingerprint, image hashing is a lot more flexible due to its robustness against some image variation. Process of extracting document fingerprint using OCR and image hashing can be illustrated in comparison as shown in Figure. 1 and Figure. 2.

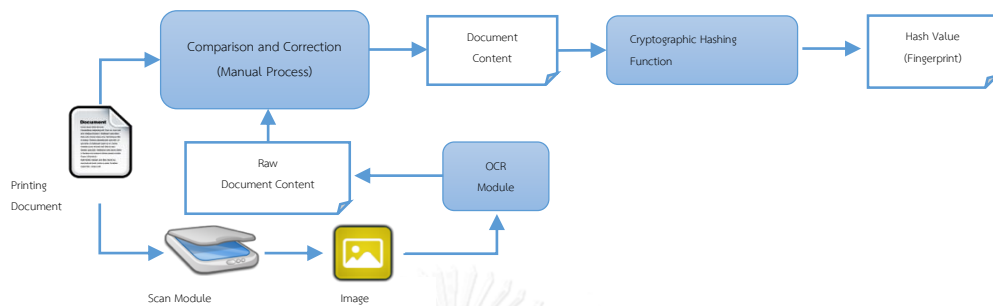


Figure 1. Fingerprint capturing process of printing document via OCR and cryptographic hashing function



Figure 2. Fingerprint capturing process of printing document via image hashing function

Reliable image hashing not only has to be robust to an acceptable change, but also has to be composed of these following properties: [5] One-way function, Collision-resistance and Key-dependence. Concept of using image hash instead of cryptographic hash in the digital signature process was published in many papers [5-9].

Using image hashing to extract a fingerprint value from printing document and digitally sign this fingerprint leads us to get trustworthy element that can verify integrity of printing document. To embed this trustworthy element with their belonging document, 2D barcode is a suitable solution that can transform some digital data into a printable format including a trustworthy element that can guarantee their own integrity and authenticity of the document. Resulting in counterfeit document or

altering document can be easily discriminated and printing document reliability can be significantly increased.

### 1.1 Objective

1. To create trustworthy verification framework for printing document.
2. To provide practical image hashing algorithm for printing document

### 1.2 Problem formulation

1. How to discriminate the authentic document from the counterfeits using image hash?
2. How to identify document issuer from any issuing document?

### 1.3 Scope of thesis and constraints

1. The language used in a document can be Thai or English
2. The document is in image format of the following extensions, .jpg, .png, .tiff, .bmp
3. A whole or partial document considered to be certified.

### 1.4 Expected outcome

1. A trustworthy procedure and framework of document issuing and verification that can verify document integrity, originality, and issuer identity

## Chapter 2. Theoretical Background

### 2.1 Image Hashing Function

Image hashing function is used to calculate hash value from content feature. As long as perceptual for the content is still unchanged, calculation through the same function will give the same hash value. Image hashing function can be described by Eq. 1.

$$h = H(I(x, y)) \quad (1)$$

where  $I(x, y)$  is an image,  $H(\cdot)$  is an image hashing function. According to its property, the hash value can be used to create indexing table, so the functions are widely used for searching and indexing image in the database and image retrieval application [10]. Moreover, the function is used for a security purpose such as authentication, digital watermarking including digital signature [5-9, 11].

### 2.2 Digital Signature

In order to ensure originality of any digital document, Digital Signature is the one of most suitable procedure. Digital Signature based-on Asymmetric key cryptography theory which used two key that created with mathematical dependency. Using one key to encrypt any content, only using public key

After integrity was ensured via perceptual hashing, enveloping this value from modification or reproduction needs to be applied in order to make this document secure and trustable. Digital Signature is the most appropriate solution for this issue, not only encrypt and secure hash value from editing, but also authenticate identity of the signer and signer cannot repudiate their signed content. Process of digitally sign the content can be shown as Figure 3.



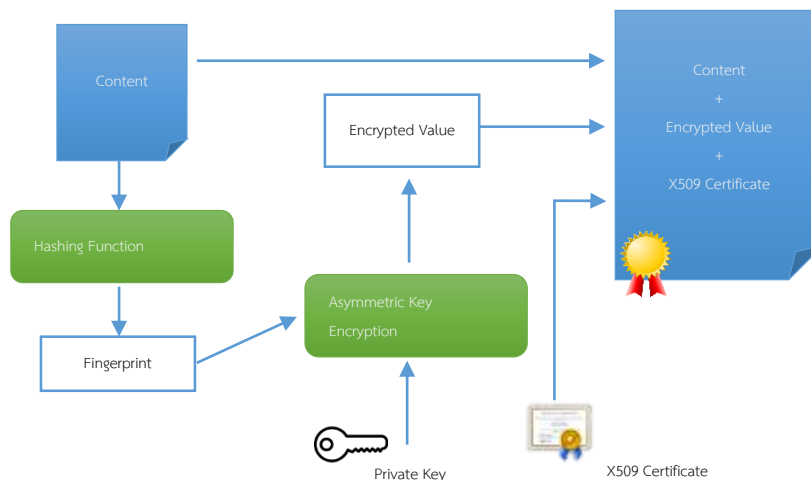


Figure 3. Process of digitally sign any digital content

As long as encrypted value can be decrypted with their public key, the value from decryption process can be fully trusted and signer is the owner of a public key used in the decryption process [2]. Figure. 4 illustrate process of digital signature validation.

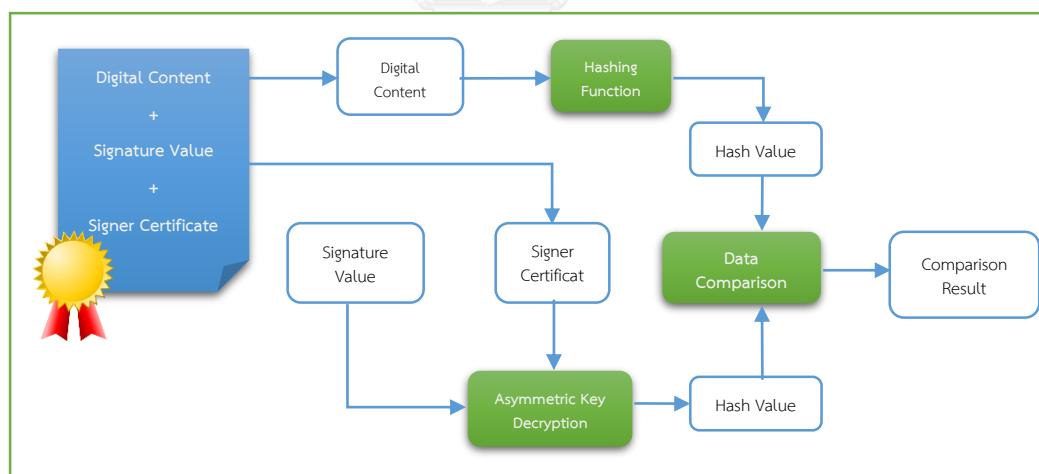


Figure 4. Process of digital signature validation

This research uses X509 Certificate v.3 [12] instead of using public key directly because X509 certificate not only contains public key but also contains information of certificate owner. In PKI environment after key pair was created, public key will send to CA (Certification Authority) to binding certificate owner information such as, common name, organization, key usages, key validity, and location to checking revocation

information. Trustworthiness and validity of certificate can be validated against their CA except Root CA which issues the certificate by itself, resulting in the validation of root CA that has to be done manually. Certificate Validation process can be shown as Figure. 5.

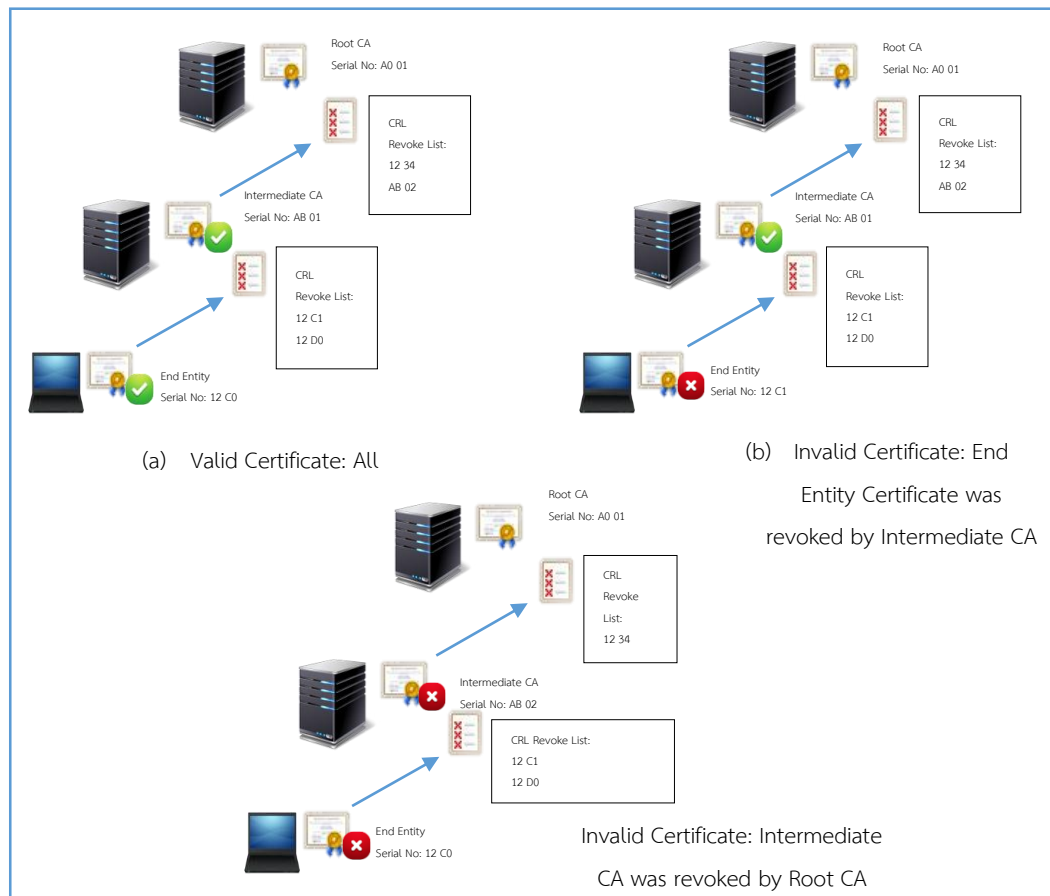


Figure 5. Certificate validation Model

This proposed framework is based on PKI infrastructure which means digital signature from this framework cannot re-produce in any case since no one can obtain same identity through PKI infrastructure. In case private key got hacked by any intruder, certificate's owner can report to issuer CA then CA will revoke that certificate immediately, this affects hacked private key and certificate become unusable. In the

other word, with PKI Infrastructure, this research can claim that printing document from this framework is highly secure and cannot made any modification or reproduction.



## Chapter 3. Related works

Related works are separated into three parts: image hashing function, digital signature, and 2D barcode.

### 3.1 Image Hashing Function

Many practical image hashing algorithms are based on image processing such as Discrete Cosine Transform-based hashing algorithms [6, 11, 13, 14], Radon-based hashing algorithms [15-17], Wavelet-based hashing algorithms [8, 9], however, algorithm that is robust to attacks from rotation, shearing, luminance, noise, and other factors from print-scan scenario has not been proposed. The following subsection explores feature and benefit of previous proposed image hashing algorithm especially radon transform and wavelet transform that are widely used and adapted in image hashing function, in order to remove drawback and combine strong aspect of each algorithm together.

#### 3.1.1 Radon Transform

When compared to the other algorithms, radon-based algorithm seems to be the most suitable algorithm for print, scan situation because the transform is based on the parameterization of straight lines and the evaluation of integrals of an image along these lines [16], so using radon transform as a perceptual feature extraction function will eradicate problem from noise, luminance, and color. The general form of Radon transform is formulated as

$$\tilde{I}(\rho, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I(x, y) \delta(\rho - x \cos\theta - y \sin\theta) dx dy \quad (2)$$

where the  $\delta(\cdot)$  is the Dirac function. Different pattern of lines and spots can be recognized easily by radon transform. The radon coefficients of line represent as a peak value and spot or noise shown in a sinusoidal curve. Figure 6 and Figure 7 show

representation of radon coefficients against straight line image and straight line with noise image.

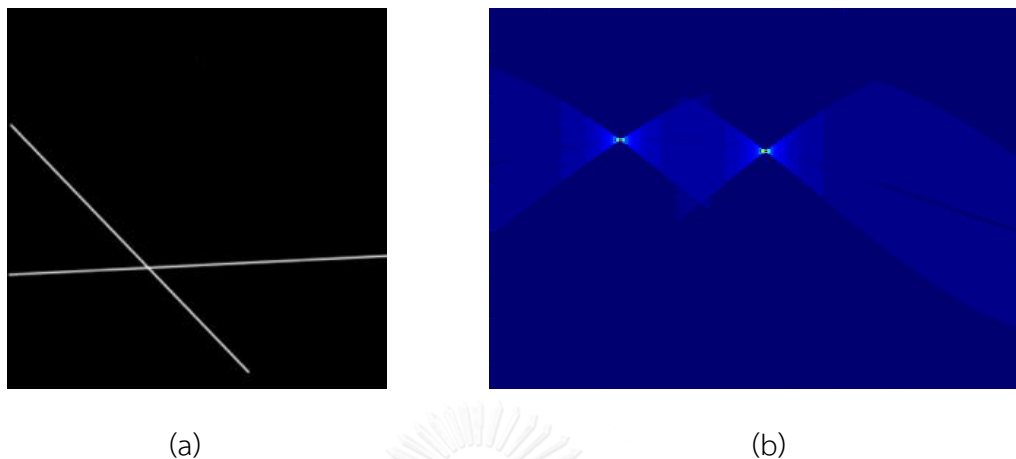


Figure 6. Radon transform result of straight line (a) Straight line image (b) Result from applying radon transform to straight line image (a)

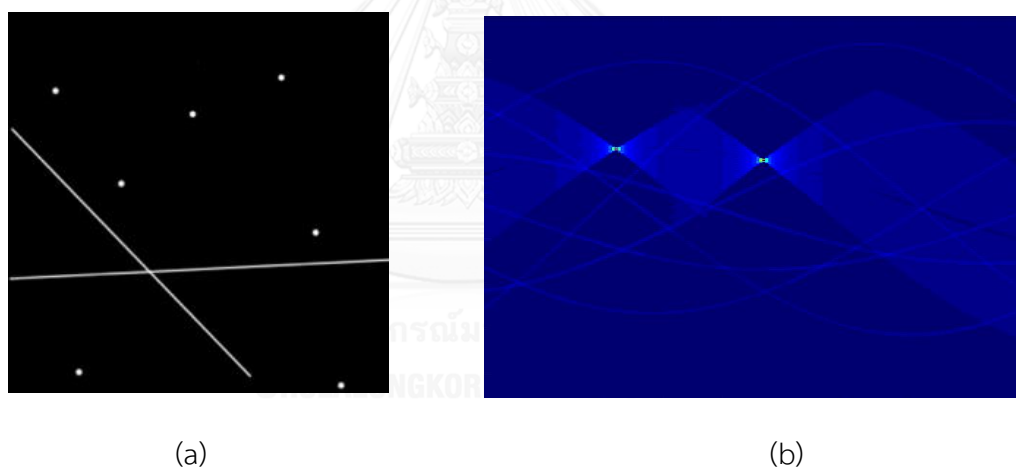
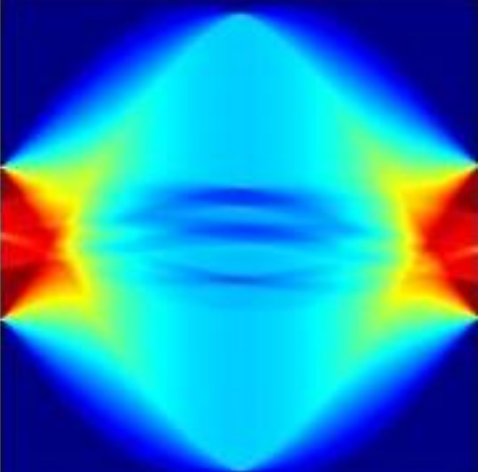
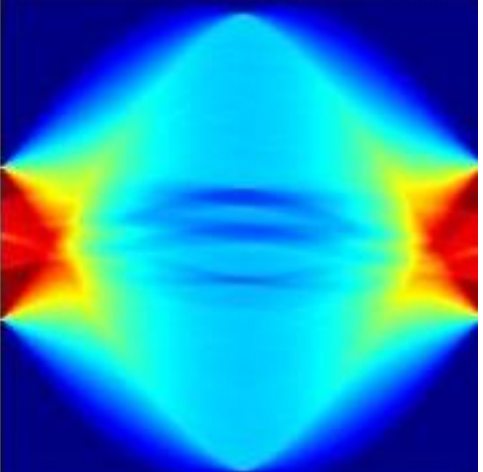
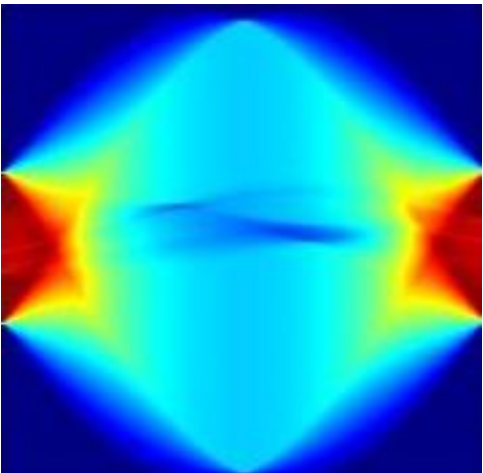


Figure 7. Radon transform result of straight line with noise (a) Straight line image with noise (b) Result from applying radon transform to straight line image with noise (a)

This results in which shape extracted from image obtained from the transform is robust to noise, luminance, and color. Using radon as a shape extraction method is proposed in many previous researches [15-19]. Performance and discriminability of radon transform can be shown in table 1, resulted from original image, noise-added image, and text-changed image. It clearly shows that radon coefficients of original image and noise-added image is almost the same compared to radon coefficients of

text-changed image. Difference between radon coefficients of text-changed image and other images is very noticeable.

*Table 1 Radon coefficients matrix of different images*

Input Image	Result Radon Coefficients
Original image <div style="border: 1px solid black; padding: 10px; text-align: center; margin: 10px auto; width: 100px;"> <b>B+</b> </div>	
Noise-added image <div style="border: 1px solid black; padding: 10px; text-align: center; margin: 10px auto; width: 100px;"> <b>B+</b> </div>	
Text-change image <div style="border: 1px solid black; padding: 10px; text-align: center; margin: 10px auto; width: 100px;"> <b>A</b> </div>	

### 3.1.2 Wavelet Transform

Originally, wavelet transform was designed and used as a decomposition function. With single-level 1-D discrete wavelet transform, a discrete signal  $x[n]$  can be decomposed into two signals obtained by filtering signal by low-pass filter and high-pass filter, followed by applying a downsampling function, respectively. Figure 8. Illustrates process of 1-D wavelet transform.

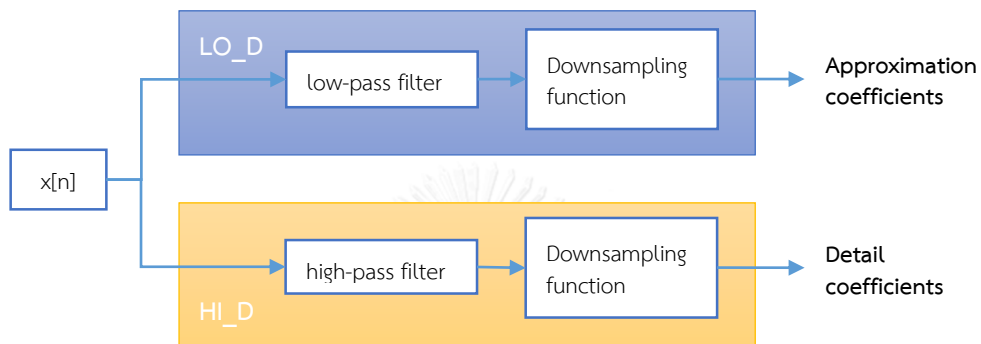


Figure 8. Process of 1-D wavelet transform.

In applying wavelet transform to an image, since the image is in 2-D form resulting an existing 1-D wavelet transform has to be applied twice in horizontal dimension (row) and vertical dimension (column). Such process is known as 2-D wavelet transform. Figure 9. Illustrates process of applying 2-D wavelet transform to an image.

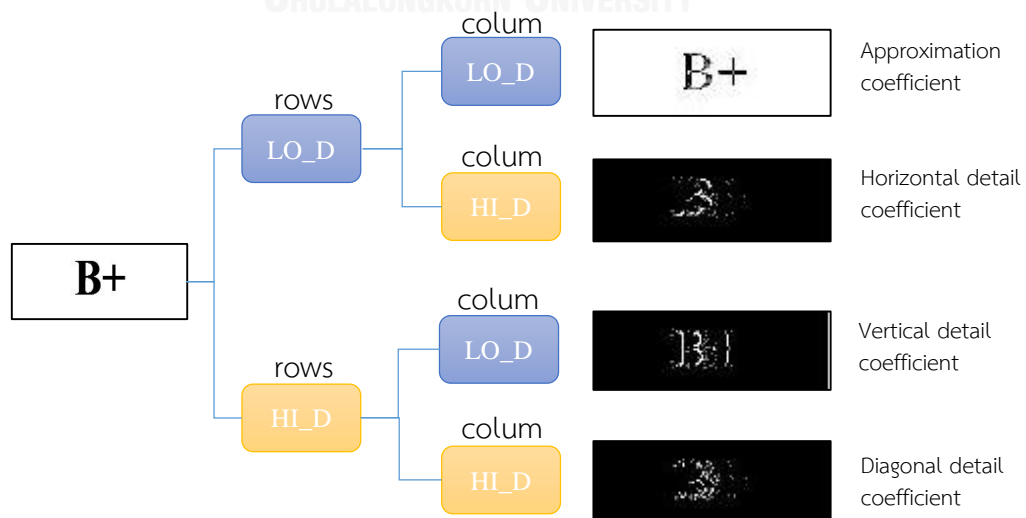


Figure 9. Process of 2-D wavelet transform

Decomposition from simple 2-D wavelet transform contains four coefficients: approximate coefficient, vertical detail coefficient, horizontal detail coefficient and diagonal detail coefficient, which can be used as extracted features [8, 9, 16, 18, 19]. In this study, image hash value can be calculated using horizontal detail coefficient.

### 3.2 Enhanced Digital Signature

With a digital signature based on standard implementation, digitally signed document can *verify integrity of content* and *authenticate content signer*. Verification of integrity of digitally sign document can be done without question via a mathematical prove of asymmetric key cryptography and cryptographic hashing function. The subsequent problem arises from a process of *signer authentication* and *trustworthy verification of content signer*. Ordinary digital signature such as CMS or PKCS#7 neither declares format and requires element of signer information nor declares practice of signer verification process, resulting in verification mechanism depending on implementation of application. To overcome trustworthy verification problem, standardization of formatting of digital signature has to be declared and implemented. CMS Advance Electronic Signature (CADES) [20] is one of standardization announced to improve trustworthy and usability of digital signature. The following statement explores detail of this standard.

Another enhancement on digital signature is implementation of digital signature on JavaScript Object Notation (JSON) message. JSON is a light-weight data-interchange format that well various used by a lot of applications such as JavaScript Programming, Stateless web service API, Representational state transfer (REST, RESTful). Implementation of digital signature on JSON message established a small message that still secure and trustworthy.

In this session detail of two standards including CADES and JSON Web Signature (JWS) will describe.



### 3.2.1 CMS Advanced Electronic Signature (CAAdES)

CAAdES [20] is a precise profiles of CMS developed and announced by European Telecommunications Standards Institute (ETSI) and described in “ETSI TS 101 733 Electronic Signature and Infrastructure (ESI) – CMS Advanced Electronic Signature (CAAdES) document”. This standard specifies a data structure and requires information of CMS signature in order to make it compile with EU Regulation on electronic identification and trust services for electronic transactions in the European internal market (eIDAS). From this reason, signature itself can describe linkage to signatory, signatory validation information or signatory validation result. CAAdES not only fulfills gap in signature and signatory linkage but CAAdES also establishes mechanism to make signature remains valid even if signer or verifying party later cannot be verified, terminated or invalidly attempted

CAAdES defines a number of digital signature format regard to trust level. Each profile was designed to encapsulate the previous profile with additional information that brings specified profile to higher-level trust. The following details show some remarked profile used in the proposed framework.

#### 3.2.1.1 CAAdES - BES

CAAdES-BES stands for CAAdES Basic Electronic Signature. This profile is a based-line of all electronic signature profiles defined in ETSI-TS-101-733 document. CAAdES-BES contains three following components as follows:

1. Signed user data (Signer's document)
2. Signed attributes
3. Digital signature value on Signed user data and Signed attributes

The structure of CAAdES-BES signature is illustrated as shown in Figure. 10.

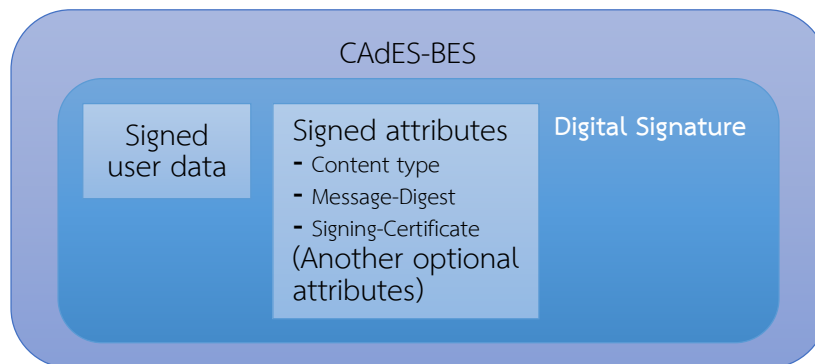


Figure 10. CAAdES-BES signature structure

### 3.2.1.2 CAAdES - EPES

A CAAdES-**Explicit Policy-based Electronic Signature** (CAAdES-EPES) is extended cades structure that required signature-policy-identifier to be specify in signed attributes. This value is an object identifier of signature verification rules that unambiguous signature validation process for example signature rules can specify if signature does not contain signingTime that signature is cannot be trust and marked as invalid. Structure of CAAdES-EPES extends from CAAdES-BES and include signature-policy-identifier in signed attributes as a required element. This structure is illustrated as shown in Figure. 11.

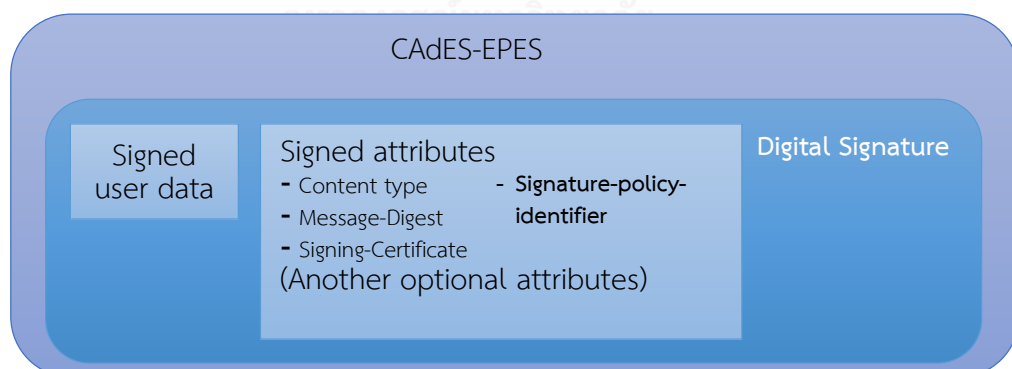


Figure 11. CAAdES-EPES signature structure

### 3.2.1.3 CADES - T

Electronic Signature with Time (CADES-T) is an extension signature format from both CADES-BES and CADES-EPES. This signature profile required to include time-stamp token that define in “RFC 3161-Time-Stamp Protocol (TSP)” from trusted time-stamp authority in the signature. With inclusion of time-stamp token, CADES-T can guarantee integrity of content, signatory identification and signing time from trusted time-stamp authority. A time-stamp token is added to the CADES-BES or CADES-EPES as an unsigned attribute. Structure of CADES-T is illustrated as shown in Figure. 12.

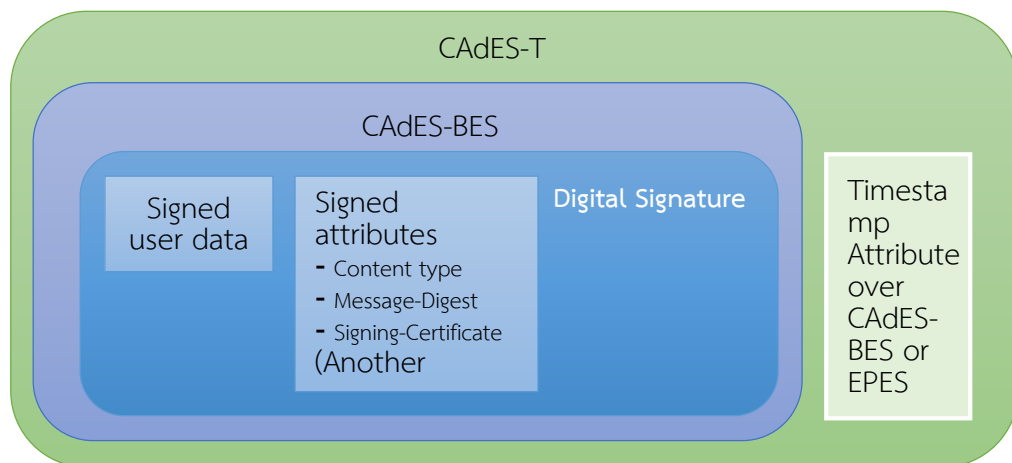


Figure 12. CADES-T signature structure

### 3.2.1.4 CADES - LT

CADES-LT stands for CADES Long Term form. CADES-LT is a profile aiming on making signature to be archived and still valid in long term. Extending properties from CADES-T, CADES-LT requires complete validation data on whole signer's certificate chain and timestamp over whole CADES-LT signature. Validity of signature is not only extended to become valid over life span of signer or issuing party but CADES-LT is also capable to protect signature against vulnerable hashing algorithm or signature algorithm. Rather than CADES-T, CADES-LT signature can also build on CADES-C, CADES-X Long, CADES-X Long Type 1 or 2 by adding one or more long-term-validation attributes. The structure of CADES-LT signature is illustrated as shown in Figure.13.

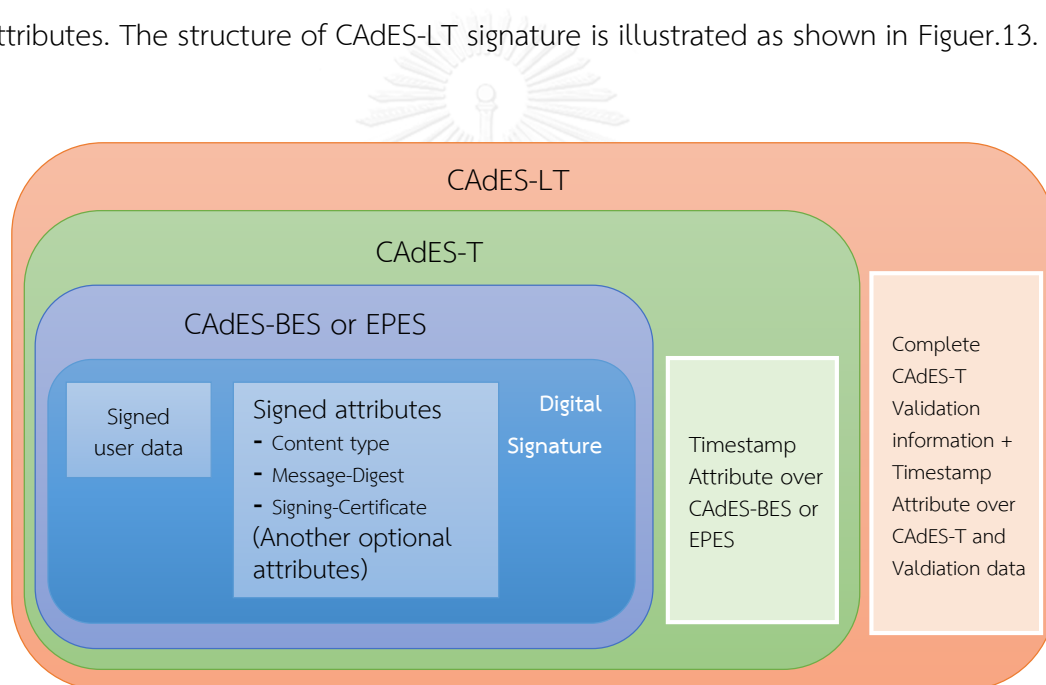


Figure 13. CADES-LT signature structure

### 3.2.2 JSON Web Signature (JWS)

JWS is a structure defined to secure content with digital signature or message authentication codes based on JavaScript Object Notation (JSON) format developed by IETF-JOSE working group. It is announced their proposed specification in “RFC 7515-JSON Web Signature (JWS)” [21]. Unlike CAdES aiming to develop a standard of CMS signature to meet requirement of eIDAS regulation, JWS is developed to improve security of JSON data. JSON is originally developed to be a lightweight data exchange format that is widely implemented in RESTful web service. With this reason, JWS was designed to secure data while the data is still lightweight.

JWS separates content of signed document into 3 parts “JOSE Header”, “Payload” and “Digital signature value”. The structure of digital signed content in JWS format is illustrated as shown in Figure. 14.

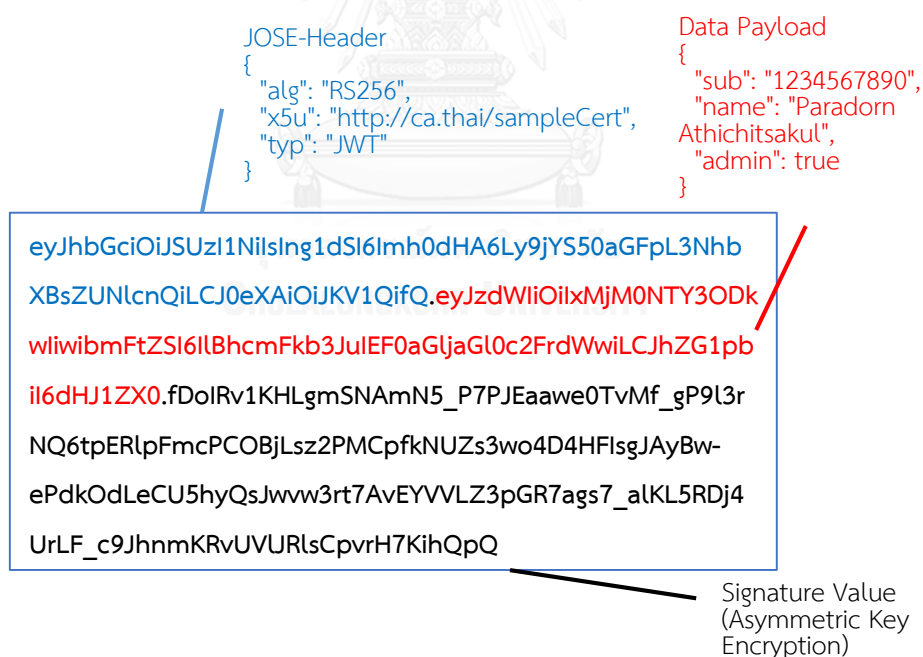


Figure 14. JWS digital signed content data structure

### 3.3 2D Barcode

Only Image hashing and digital signature seem to be enough to establish a framework that can verify integrity of document meanwhile identify document issuer. However, the other problems are how to transfer digital signature from the document issuer to the recipients and how recipients match the signature element to the printing document. Many publications [4, 22, 23] solve these problems via embedding digital signature element to print document via 2D barcode such as PDF 417 or QR Code.

#### 3.3.1 PDF 417

PDF 417 is a kind of 2D barcode containing a lots of features inside such as, binary data container, error correction, macro mode, etc. PDF 417 specification is defined in document “ISO 15483 – Information technology— Automatic identification and data capture techniques — PDF417 bar code symbology specification” [24]. The PDF 417 widely adopts in various type of application. The structure of PDF 417 can be illustrated as shown in Figure 15.

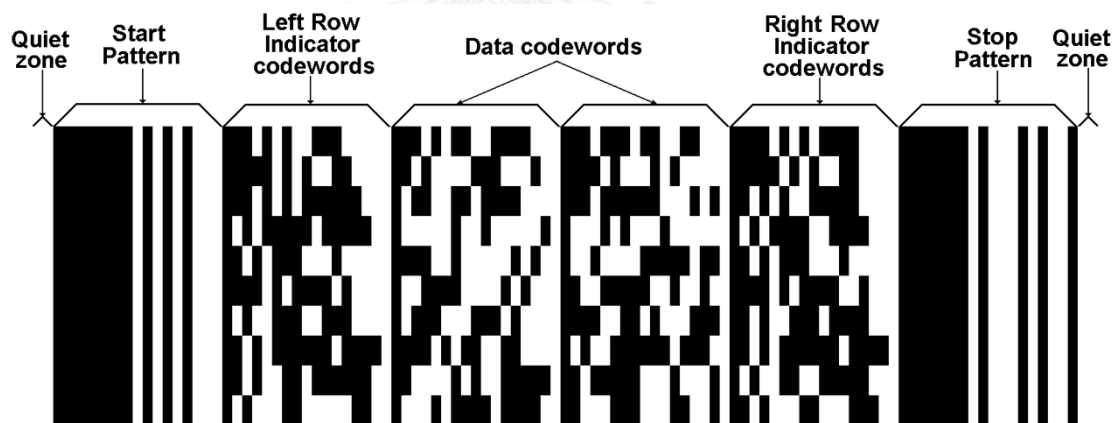


Figure 15. PDF 417 Symbol Structure

PDF 417 Symbol can contain 3 to 90 rows with 1 to 30 columns. The maximum possible number of code words per 1 symbol is 925. PDF 417 can be used to encode various formats of data, i.e., binary, text, and numeric data with three modes as follows:

- 1) Binary compaction mode: 5 codewords per 6 Byte
- 2) Text compaction mode: 1 codeword per 1-2 Characters
- 3) Numeric compaction mode: at most 15 codewords per 44 numeric digits.

### 3.3.2 QR Code

Another alternative 2D barcode that can contain binary data and handle some error is QR Code. The ISO standard, ISO 18004 - Information technology — Automatic identification and data capture techniques — QR Code 2005 bar code symbology specification [25], describes symbol definition and specification of all QR codes. Contrast to PDF 417 that has only 1 symbol structure, QR code have various types and structures. Each version of QR code has its own symbol structure. Examples of QR code symbol structures in version 6 and 7 are shown in Figure. 16.

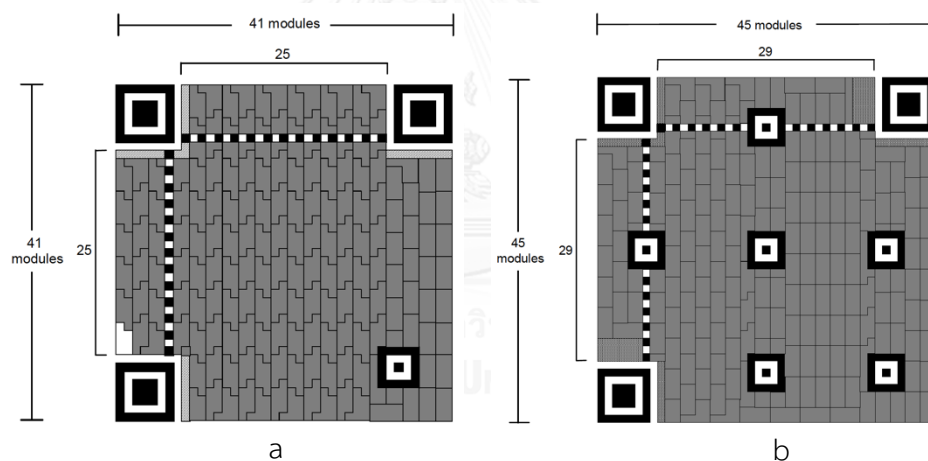


Figure 16. QR Code Symbol Structures (a) QR Code version 6 symbol structure (b) QR Code version 7 symbol structure

Different from PDF 417, number of codewords in QR code depends on version of QR code. In order to specify which version of QR code to be used depending on size of encoded data and preferred error correction level, Table 2 shows relationship between version, error correction level and data capacity of QR code

Table 2 Relationship of QR code version, error correction level and data capacity from version 3-

10

Version	Error Correction Level	Data Capacity			
		Numeric (digit)	Alphanumeric (character)	Byte	Kanji (character)
3	L	127	77	53	32
	M	101	61	42	26
	Q	77	47	32	20
	H	58	35	24	15
4	L	187	114	78	48
	M	149	90	62	38
	Q	111	67	46	28
	H	82	50	34	21
5	L	255	154	106	65
	M	202	122	84	52
	Q	144	87	60	37
	H	106	64	44	27
6	L	322	195	134	82
	M	255	154	106	65
	Q	178	108	74	45
	H	139	84	58	36
7	L	370	224	154	95
	M	293	178	122	75
	Q	207	125	86	53
	H	154	93	64	39
8	L	461	279	192	118
	M	365	221	152	93
	Q	259	157	108	66
	H	202	122	84	52
9	L	552	335	230	141
	M	432	262	180	111
	Q	312	189	130	80
	H	235	143	98	60
10	L	652	395	271	167
	M	513	311	213	131
	Q	364	221	151	93
	H	288	174	119	74



## Chapter 4. Proposed Methods

The proposed framework is described in two directions: creation and verification. Also, proposed image hashing algorithm based on three image transforms is shown afterwards.

### 4.1 Framework for Creating and Verifying Trustworthy Printout

Using the mentioned methodologies, this section aims to propose framework for creating and verifying the trustworthy printout.

Printout creation consists of following steps.

1. Define certified regions in document.
2. Compute image hash value of each region.
3. Digitally sign data from every region.
4. Create 2D barcode from data of selected region and digital signature from step 3.
5. Attach barcode to the document and print.

The document creation is summarized as presented in Figure. 17.

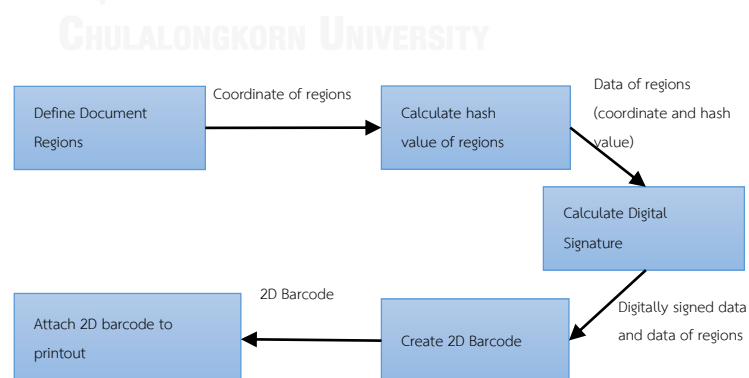


Figure 17. Printout creation process.

The printout verification mechanism of this framework can be accomplished via these steps.

1. Scan printout and extract barcode from scanned image.
2. Decode the 2D barcode to achieve digital signature value.
3. Verify validity and revocation of digital signature.
4. If digital signature is valid, then compute image hash value of each region.
5. Calculate hash distance from hash value calculated from each region and hash value specified in verified data.
6. For each region, if the hash distance from step 5 is less than a predefined classification threshold, then this framework can guarantee that “This region is valid and has not been modified”.

The process of printout verification is illustrated in Figure. 18.

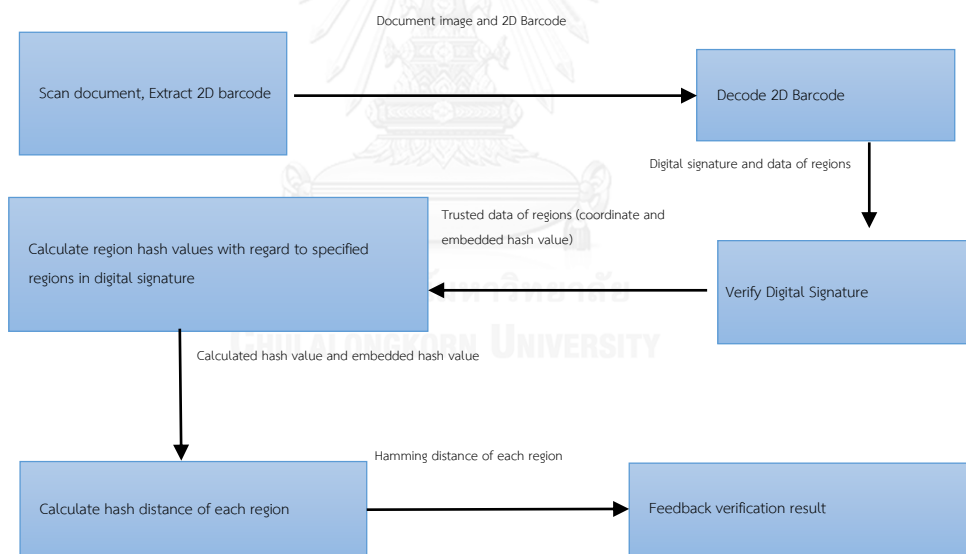


Figure 18. Printout verification process

## 4.2 Image Hashing Function Design

In applying image hash to text-based printout, ordinary image hashing functions based on DCT and Radon transforms are not sufficient to capture identity of printout. In this thesis, the modified algorithm based on the algorithm introduced by Wu, Zhou, and Niu [17], was designed and tested especially for print and scan scenario. However, there are some problems when applying to printout verification. The modified algorithm from [17] consists of three components. Radon transform is the first component. Then, Wavelet transform and Fourier transform are applied to improve algorithm accuracy. The process and output of each step are depicted in Figures. 19 and 20.

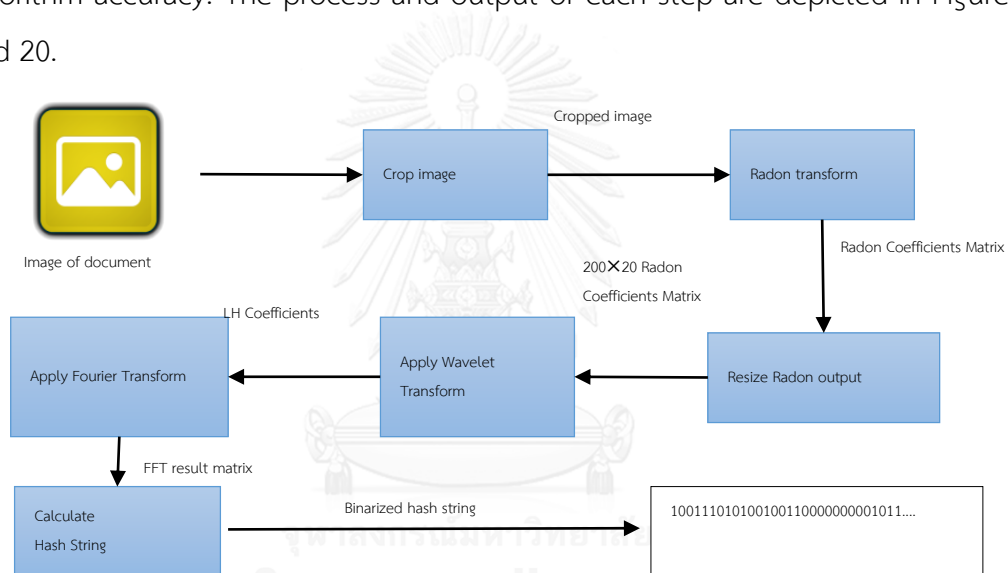


Figure 19. Image hash value computation process

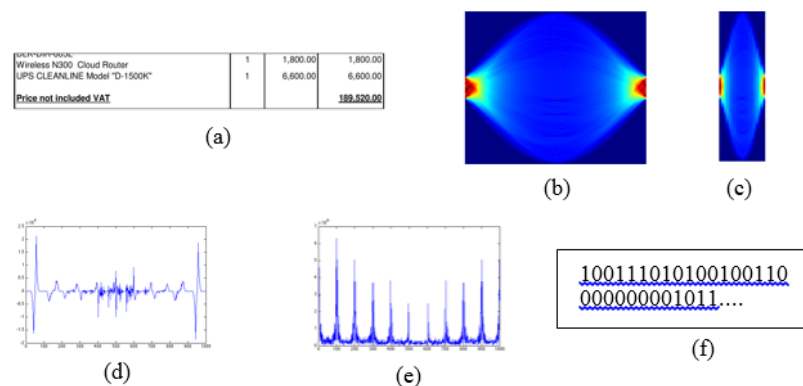


Figure 20. Output of each image hash value computation process. (a) Cropped Image, (b) Image from Radon transform, (c) Resized Image, (d) Image from Wavelet transform, (e) Image from Fourier transform, (f) Binary hash string.

In print and scan scenario, attacks from luminance change, shearing, rotation, noise and another small attack are unavoidable. These factors influence on the process of image hashing and cause low accuracy. Radon transform is a methodology widely used as a shape extraction method that is invariant to color change. This means feature extraction via Radon transform is possible to be performed under print and scan scenario smoothly.

So firstly, our proposed algorithm converts image region, which is cropped from a specified region of a printout, to grayscale and then applies radon transform to the region. This proposed method applies radon projection from -90 to 89 degrees with a sampling interval of 1 degree. Different from [17] processes of compensating the effect of rotating and shifting using cyclically shift radon transform output were removed. Since text-based image usually consists of a bunch of straight line in many angles, this makes anti-rotation and shift process with radon transform cannot perform accurately and decreases the accuracy of using image hashing function.

The output size from Radon transform depends on the image size. In order to get an appropriate hash length, the size of output image of radon transform is resized to 200×20.

The next step is to apply Wavelet transform. Since the vertical edge of the image from Radon transform is shown very clearly compared with the horizontal edge, this means that the LH sub-band from wavelet transform is more important than the other sub-bands. So, in this study, only LH sub-band was chosen instead of using all high frequencies [17]. This sub-band results in image of size 100×10.

To make a hash string of 1000 bits become accurate and discriminable between counterfeit and acceptable image, Fast Fourier Transform is applied and yields more different hash strings. Finally, the output from Fast Fourier transform is binarized into binary hash string using a mean value calculated from components as a threshold.

### 4.3 Printable Digital Signature Design

After image hash value is created, next issue is how to digitally sign this image hash value and embed digital signature to 2D barcode before print it. With existing 2D barcode, capacity of each announced barcodes are not enough to contain digital signature data with error correction codewords. This thesis proposed a new digital signature profile that extended from CAdES-BES digital signature [26]. This proposed digital signature profile is targeting on reducing size of digital signature as much as possible meanwhile the digital signature still trustworthy as CAdES-BES.

Extended from CAdES-BES definition this digital signature use RFC 3852 (Cryptographic Message Syntax) as a digital signature structure and specify appropriate value of digital signature regard to intention of use of digital signature. Structure of CMS digital signature can be illustrated as shown in Figure. 21.

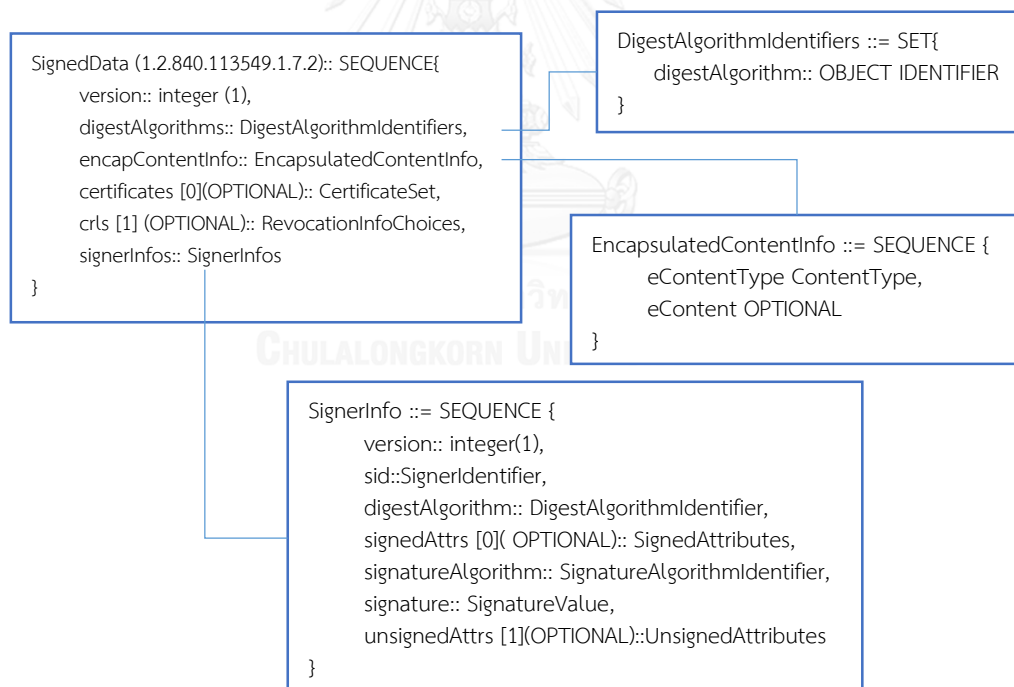


Figure 21. CMS digital signature data structure

In order to minimize size of digital signature, unused or optional value in CMS and CAdES-BES specifications have to be remove tables 3, 4, 5 and 6 show description and value of parameter used in proposed digital signature.

Table 3. SignedData parameters description

Parameter	Description	Values
version	SignedData CMS version	1
digestAlgorithms	Collection of message digest algorithm identifier used by one or more signer	2.16.840.1.101.3.4.2.1 (SHA-256)
encapContentInfo	Signed Content (EncapsulatedContentInfo)	Specify in table 4
signerInfos	Collection of signer information (SignerInfo)	Specify in table 5

Table 4. EncapsulatedContentInfo parameters description

Parameter	Description	Values
contentType	Object identifier of this encapsulated content	1.2.840.113549.1.7.1 (data)
content	Octet string encoding of content data	Octet string encoding of an image hash data

Table 5. SignerInfo parameters description

Parameter	Description	Values
version	Version of signer info definition	1
sid		
IssuerAndSerial Number	Distunguish Name value of certificate Issuer and SerialNumber of certificate	“cn=ECDSA_SUB_CA, c=th”, 7011393726614152651
digestAlgorithm	Message digest algorithm identifier	2.16.840.1.101.3.4.2.1 (SHA-256)
signedAttrs[0]	Collection of attribute that are signed (SignedAttributes)	Specify in table 6
signatureAlgorithm	Signature Algorithm Identifier	1.2.840.10045.4.3.2 (ecdsa-with-SHA256)
Signature	Result of digital signature generation	34274169488719486431.....

Table 6. SignedAttributes parameters description

Parameter	Description	Values
contentType	Must be the same value as content type specified in EncapsulatedContentInfo – contentType	1.2.840.113549.1.7.1 (data)
signingTime	The time at which signer perform the signing process	2016-05-23 08:00:42 UTC
messageDigest	Message digest value of attribute encapContentInfo in digital signature (SignedData)	Message digest value of an image hash data
signingCertificate	ESS signing-certificate information of signer	ESS signing-certificate-v2.
unstructuredAddress	URL of X.509 signing-certificate and certificate chain	“http://certificate.localtion”

Another significant size reducing factor is removing signer certificate and certificate chain from signature by applying concept of JWS specifying signer certificate url into proposed digital signature profile instead of directly embedding X.509 certificate of signer and its chain within signature. An attribute unstructuredAddress in table 4 is added in order to specify online location of certificate and certificate chain of signer. Removal of signer certificate from digital signature brings security concerns about integrity and trustworthiness of signer certificate in location specified. However, the verifier can verify signer certificate against ESS-Signing Certificate information in an attribute signingCertificate. Figure 22. Illustrates structure of proposed digital signature.

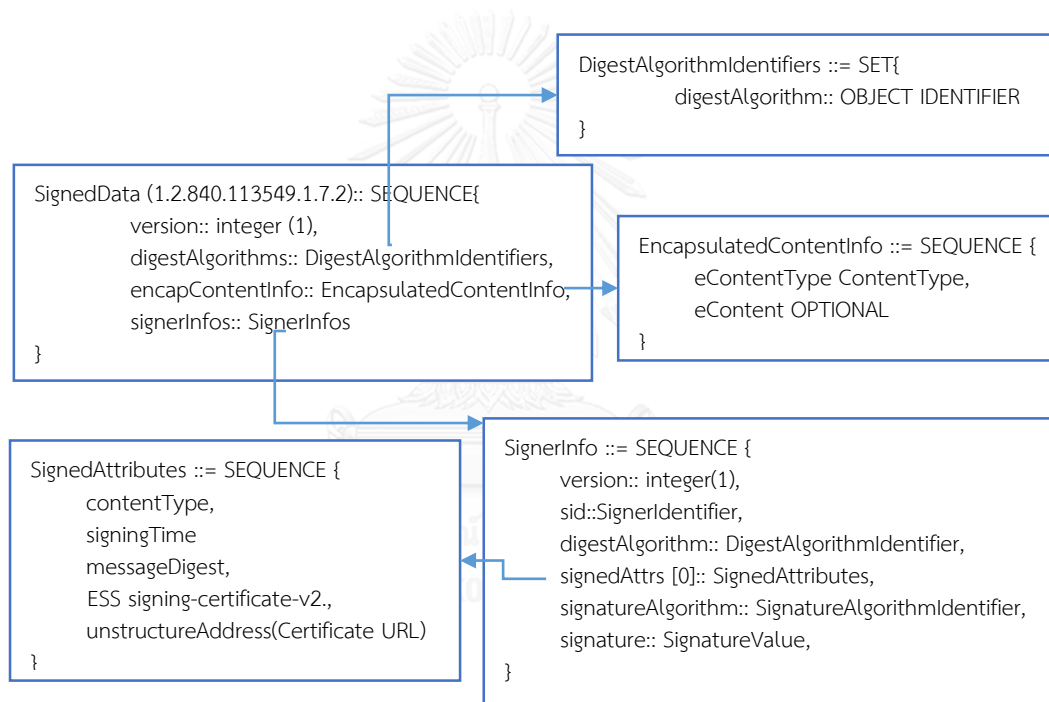


Figure 22. Proposed digital signature data structure

After image hash data are calculated and digital signature component is prepared, the final process is digital signing image with hash data, applying compression to digital signature, and embedding compressed digital signature to 2D barcode. Figure 23 shows process of 2D barcode creation

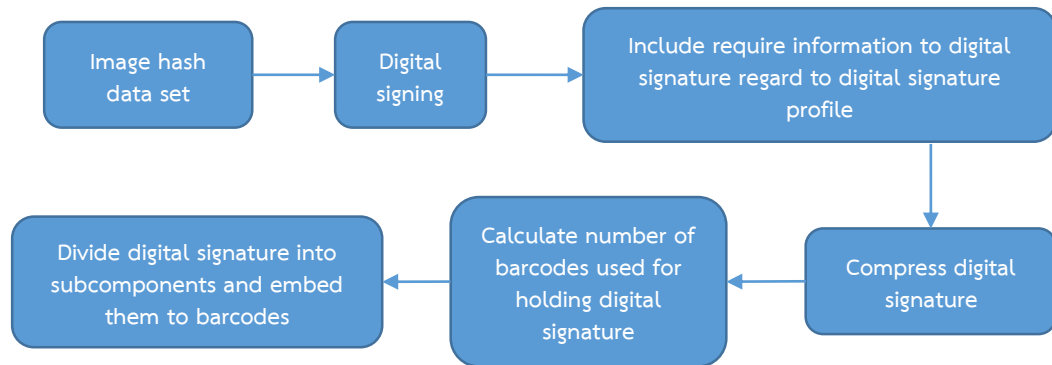


Figure 23. Process of barcode creation





## Chapter 5. Experimental Results

### 5.1 Image Hash Evaluation

For testing the proposed image hash algorithm, there were 100 original images obtained from 30 English documents, 30 Thai documents, 10 other language phrases, 10 Faces, 10 logos, and 10 signatures. Printing and scanning were performed via Epson AcuLaser CX17NF with 600 dpi for printing and 300 dpi for scanning. Figure 24 shows some examples of test data in this experiment.

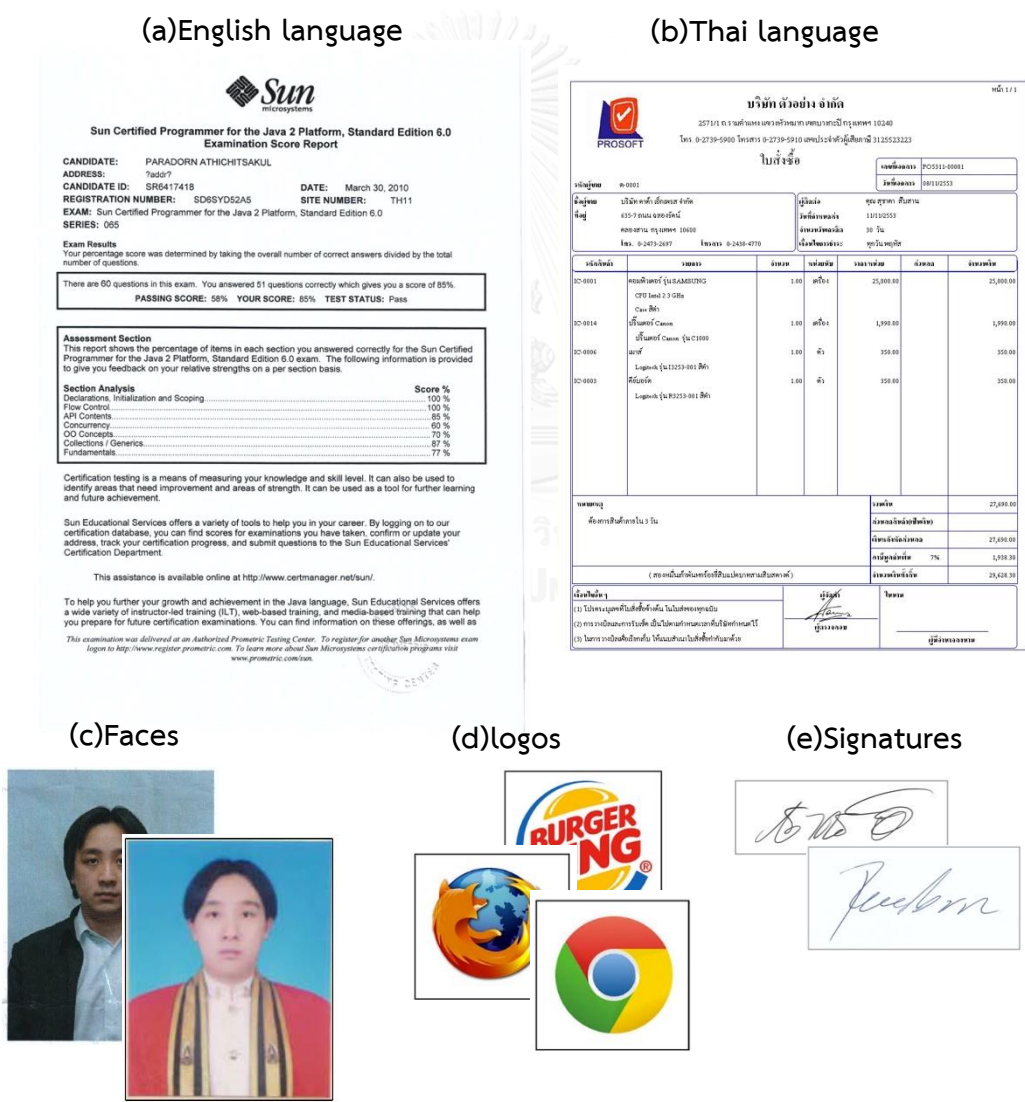


Figure 24. Example test data in this experiment

### 5.1.1 Image Hash Algorithm Performance

In order to prove that the image hash algorithm can be implemented with fingerprinting printout which has some possible changes from printing and scanning process such as small amount of noise, change of brightness contrast and color, folding, etc. Two types of attack were defined and applied to dataset as follows.

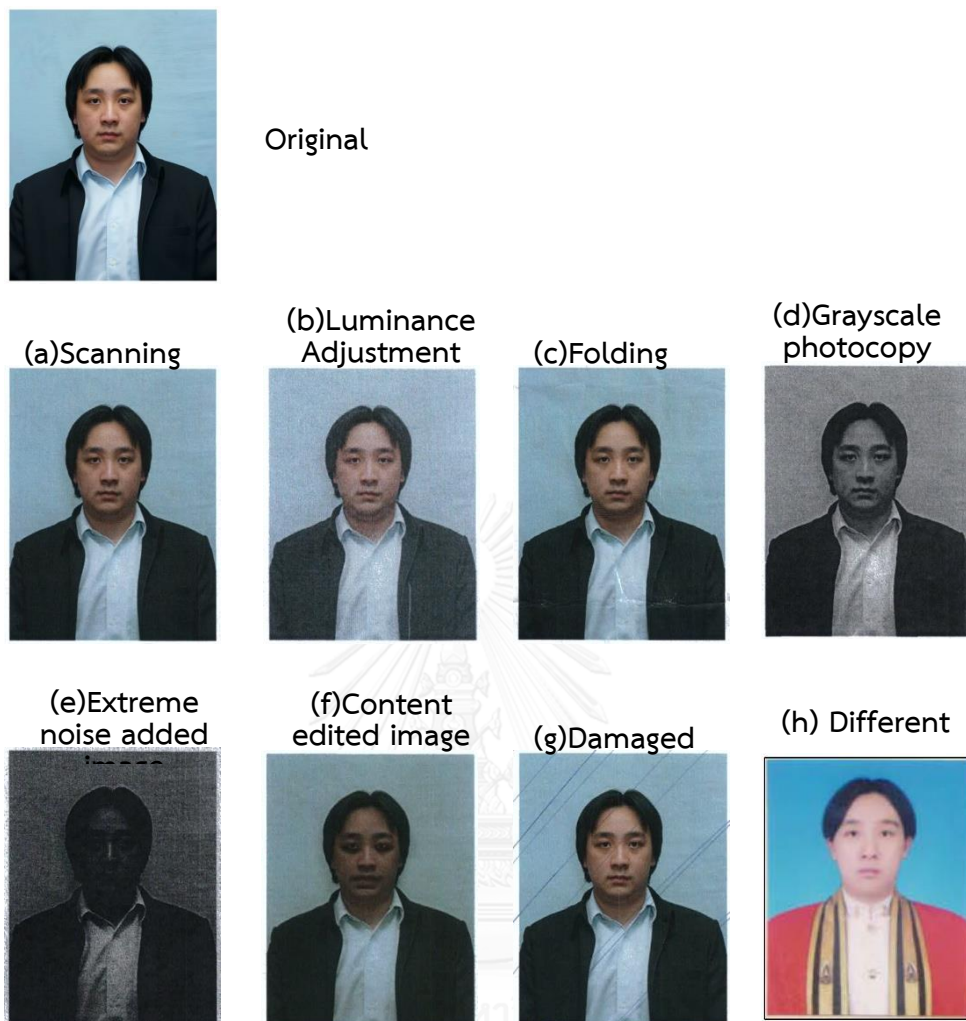
#### **Acceptable attack**

1. Scanning: Just only scan original image.
2. Luminance adjustment: Luminance of original image is adjusted, before printing and scanning.
3. Folding: Original image is folded before scanning.
4. Grayscale photocopying: Use photocopy machine to perform grayscale image before scanning

#### **Unacceptable attack**

1. Extreme noise adding: Use photocopy machine to perform grayscale image with maximum brightness or darkness before scanning.
2. Document content editing: Digitally edit partial word or number from scanned image, then print and scan the image again.
3. Damaged document: This attack performs physical damage via adding ink directly to image before scanning.
4. Different document: This attack is performed to compare the original image with other images.

Figure 25 illustrates attacks considered in this thesis.



CHULALONGKORN UNIVERSITY

Figure 25. Example images of each attacks

Table 7. Performance of Hashing under Attacks.

Attack	Hamming Distance (%)			
	Mean	Std. Deviation	Max	Min
Acceptable Attack				
Scanning	5.15	2.93	21.20	0.80
Grayscale Photocopying	7.98	4.20	19.20	0.80
Folding	5.76	3.32	18.00	1.20
Luminance Adjustment	6.58	3.15	21.20	1.20
Unacceptable Attack				
Document Content Editing	10.24	5.03	22.40	2.20
Damaged Document	19.21	8.72	37.00	1.80
Extreme Noise Adding	21.88	8.31	40.60	4.80
Different document	25.38	6.52	37.20	5.40

The experimental result of each attack is described in Table 7. This table shows that acceptable attack such as color change and small amount of noise does not make a significant distance with this algorithm. In contrast, unacceptable attack such as damaged document, large noise adding and editing document provides a large distance value via the proposed algorithm. With an ordinary process, user just prints a document and rescans it to verify integrity. This proposed algorithm rarely rejects the document as a different document. The next part shows algorithm accuracy and comparison of our algorithm and the other existing image hash algorithms.

## 5.1.2 Image hash Algorithm Accuracy

### 5.1.2.1 Accuracy of the Proposed Algorithm

Table 8 shows means of hamming distance on acceptable and unacceptable attacks, False Acceptance Rate (FAR) and False Rejection Rate (FRR) on scanning process of the proposed algorithm. At the intersection point between FAR and FRR, FAR can be reduced to 17.5% with 18.5 % of FRR and 9% of ordinary scanning rejection rate. The accuracy graph of this proposed algorithm is shown in Figure. 26.

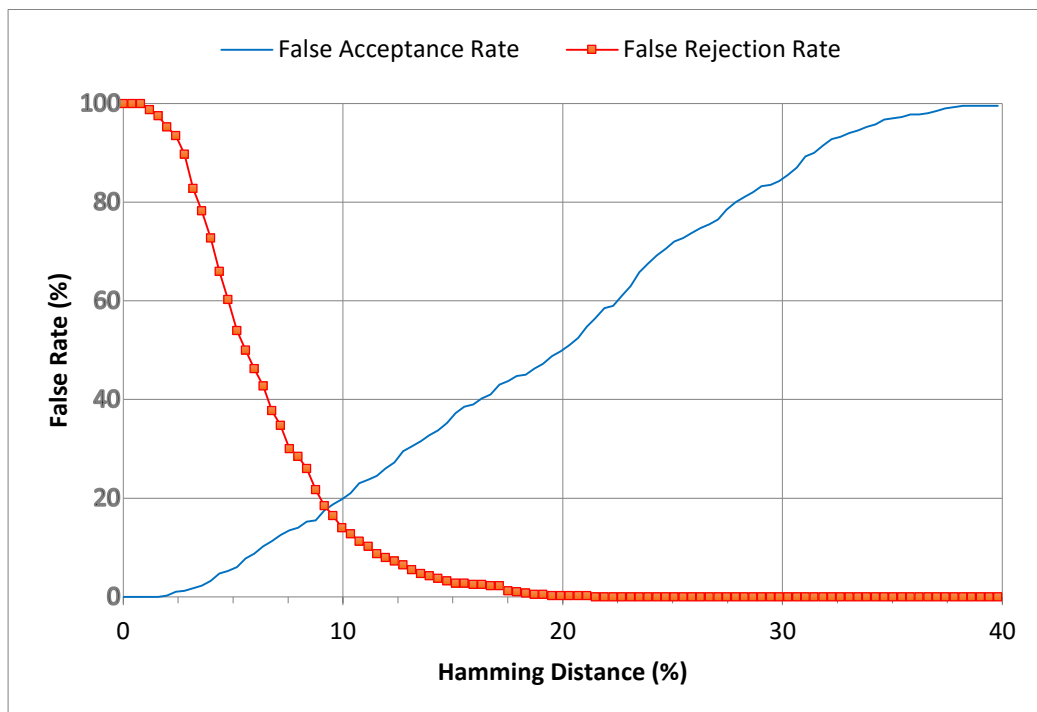


Figure 26. Accuracy of the proposed algorithm.

Table 8. Performance and Accuracy of the Proposed Algorithm.

Mean value of Hamming Distance on Attacks		Error Rate		
Acceptable Attacks	Unacceptable Attacks	False Acceptance	False Rejection	
			Overall	Scanning
6.37%	19.18%	17.5%	18.5%	9%

### 5.1.2.2 Accuracy of Wu, Zhou, and Niu's Algorithm

Table 9 presents the results of method based on [17]. With the optimal threshold, FAR is decreased to 22.75% with 23.25 % of FRR and 19% of ordinary scanning rejection. The accuracy graph of Radon-Based algorithm is illustrated in Figure. 27.

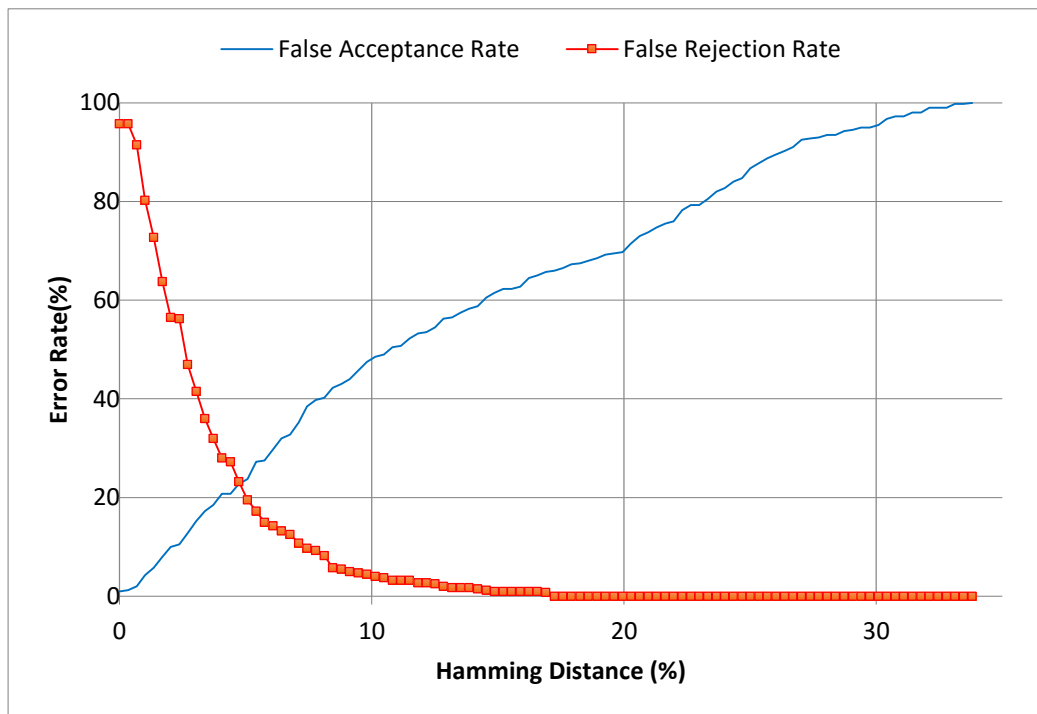


Figure 27. Accuracy of Wu, Zhou, and Niu's algorithm

Table 9. Performance and Accuracy of the Approach Proposed by [17]

Mean value of Hamming Distance on Attacks		Error Rate		
Acceptable Attacks	Unacceptable Attacks	False Acceptance	False Rejection	
			Overall	Scanning
3.36%	13.06%	22.75%	23.25%	19%

### 5.1.2.3 Accuracy of DCT-Based Algorithm

Table 10 shows accuracy of DCT-Based algorithm. With appropriate threshold, FAR can be reduced to 21.75% with 33 % of FRR and 19% of ordinary scanning rejection. The curves of FAR and FRR from DCT-Based algorithm is shown in Figure. 28.

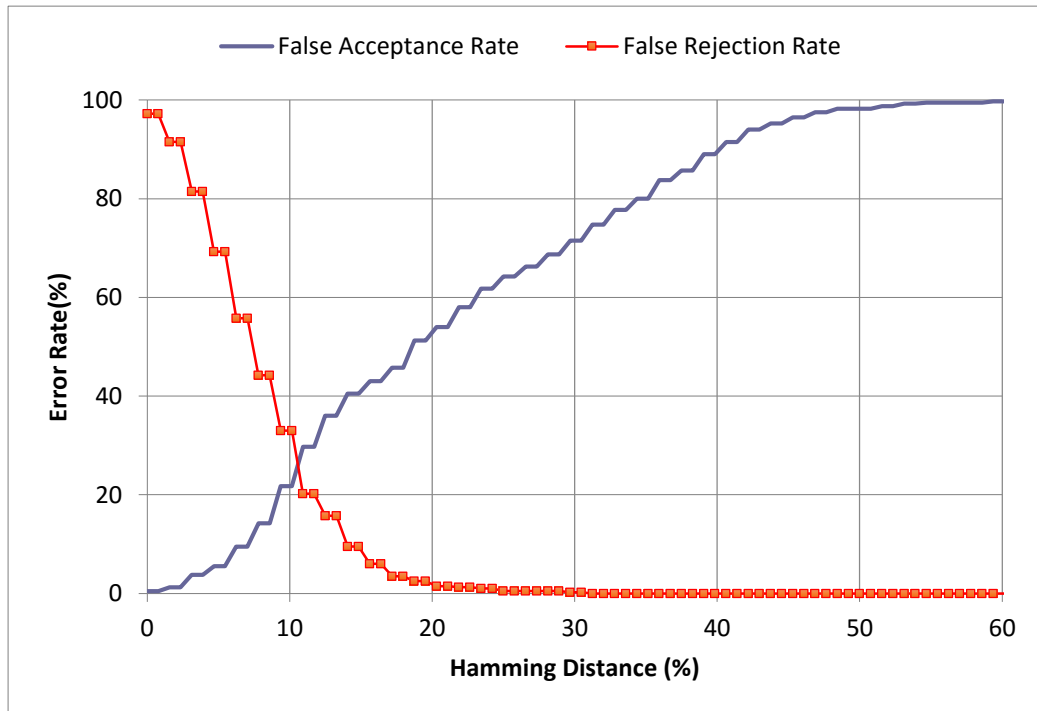


Figure 28. Accuracy of DCT-based algorithm.

Table 10. Performance and Accuracy of DCT-Based Algorithm

Mean value of Hamming Distance on Attacks		Error Rate		
Acceptable Attacks	Unacceptable Attacks	False Acceptance	False Rejection	
			Overall	Scanning
8.37%	21.79%	21.75%	33%	19%

#### 5.1.2.4 Accuracy of pHash's DCT-Based Algorithm

Results of DCT-Based method from [27] shown in table 11. With the optimal situation, FAR is 19.75% with 26 % of FRR and 13% of ordinary scanning rejection at the intersection point of FAR and FRR curves. Figure. 29 demonstrates the error rate of DCT-Based Image Hashing algorithm.

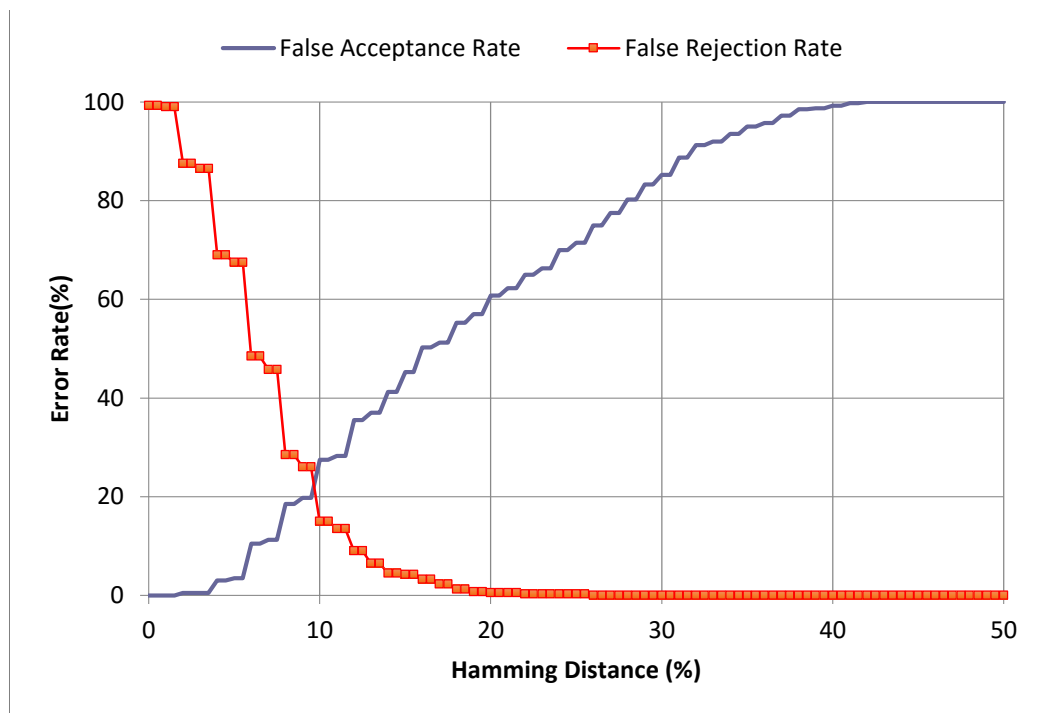


Figure 29. Accuracy of pHash's DCT-based algorithm.

Table 11. Performance and Accuracy of pHash's DCT-Based Algorithm [27]

Mean value of Hamming Distance on Attacks		Error Rate		
Acceptable Attacks	Unacceptable Attacks	False Acceptance	False Rejection	
			Overall	Scanning
11.25%	29.02%	19.75%	26%	13%



Accuracy comparison of all above algorithms is illustrated in Figure. 30. The figure clearly shows that the intersection point between FAR and FRR or critical point of the proposed algorithm is lower than those of the other existing algorithms. In the other words, this proposed algorithm yields better performance of capturing identity of printout than other existing algorithms.

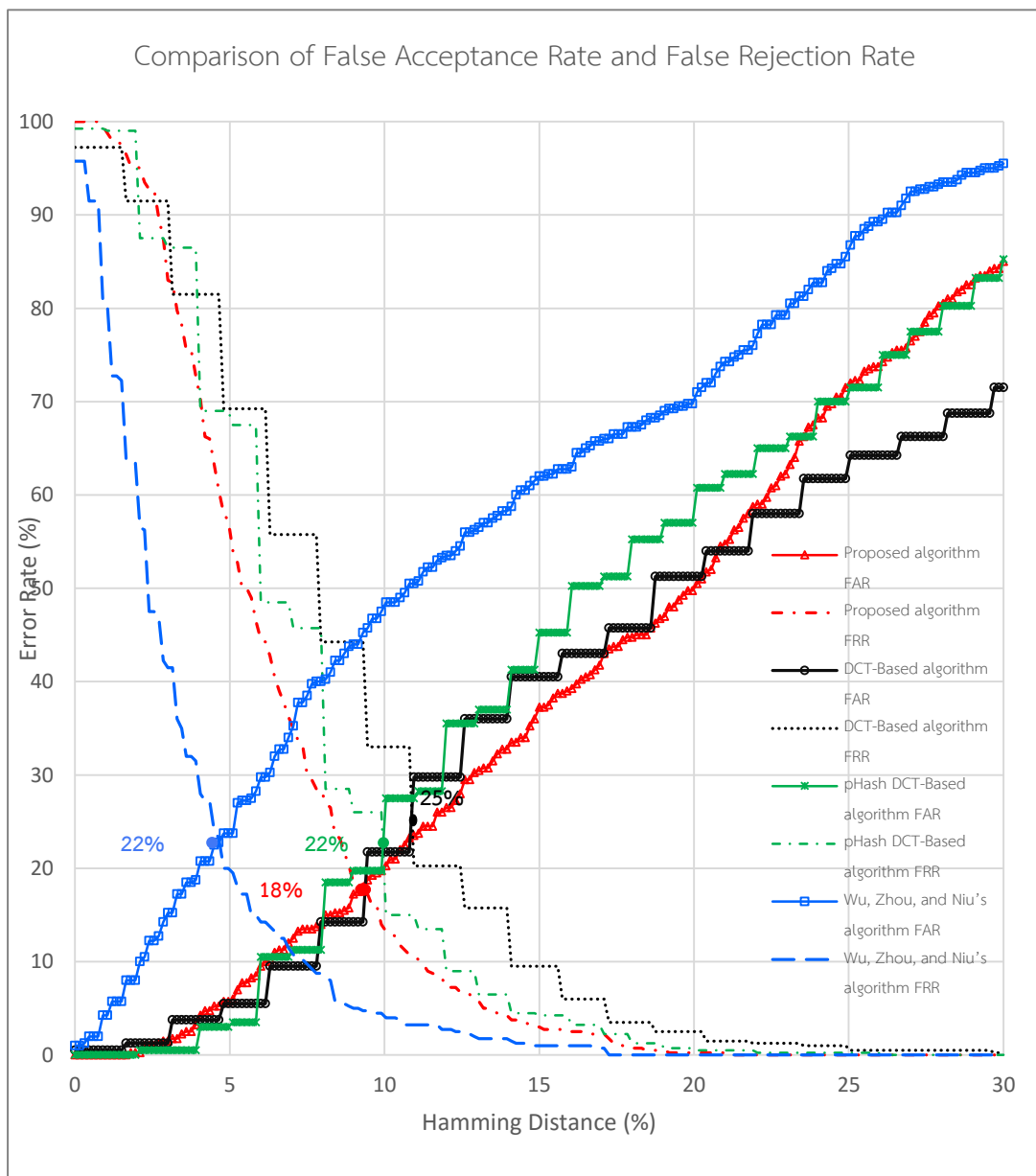


Figure 30. Comparison of our algorithm and the other existing image hash algorithms in terms of FAR and FRR.



### 5.2.1 Digital Signature Size

Tables 12 and 13 show sizes of digital signature with and without content data obtained from signing process. These two tables clearly show that the proposed digital signature size is smaller than the other defined digital signatures. To evaluate size of signature without content data, PKCS#7 is only available format for comparison. Note that the proposed digital signature outperforms PKCS#7. Additionally, compressed size of proposed digital signature without content data is 337 bytes which is sufficient to fill in one 2D-barcode.

*Table 12. Size comparison of digital signature with content data*

Digital signature structure	File size (bytes)	Compressed file size (bytes)
PKCS#7 (CMS)	7385	2276
CAdES-BES	6737	1949
CAdES-LT	10194	3661
JWS	8067	1784
<b><u>Proposed Digital Signature</u></b>	<b><u>6251</u></b>	<b><u>1530</u></b>

*Table 13. Size comparison of digital signature without content data*

Digital signature structure	File size (bytes)	Compressed file size (bytes)
PKCS#7 (CMS)	694	616
<b><u>Proposed Digital Signature</u></b>	<b><u>455</u></b>	<b><u>377</u></b>

### 5.3 Printout and Printout Verification

This part reveals process of embedding digital signature to printout and verifying digital signature. Figure. 33 is the demonstration of result printout. Using proposed digital signature to digitally signed image hash value resulting in digital signature size of 1530 bytes. Also, three PDF-417 barcodes are used to contain this digital signature as shown in Figure. 33.

**Sun**  
microsystems

**Sun Certified Programmer for the Java 2 Platform, Standard Edition 6.0  
Examination Score Report**

**CANDIDATE:** PARADORN ATHICHITSAKUL  
**ADDRESS:** ?addr?  
**CANDIDATE ID:** SR6417418      **DATE:** March 30, 2010  
**REGISTRATION NUMBER:** SD6SYD52A5      **SITE NUMBER:** TH11  
**EXAM:** Sun Certified Programmer for the Java 2 Platform, Standard Edition 6.0  
**SERIES:** 065

**Exam Results**  
Your percentage score was determined by taking the overall number of correct answers divided by the total number of questions.

There are 60 questions in this exam. You answered 51 questions correctly which gives you a score of 85%.  
**PASSING SCORE: 58% YOUR SCORE: 85% TEST STATUS: Pass**

**Assessment Section**  
This report shows the percentage of items in each section you answered correctly for the Sun Certified Programmer for the Java 2 Platform, Standard Edition 6.0 exam. The following information is provided to give you feedback on your relative strengths on a per section basis.

Section Analysis	Score %
Declarations, Initialization and Scoping.....	100 %
Flow Control.....	100 %
API Contents.....	85 %
Concurrency.....	60 %
OO Concepts.....	70 %
Collections / Generics.....	87 %
Fundamentals.....	77 %

Certification testing is a means of measuring your knowledge and skill level. It can also be used to identify areas that need improvement and areas of strength. It can be used as a tool for further learning and future achievement.

Sun Educational Services offers a variety of tools to help you in your career. By logging on to our certification database, you can find scores for examinations you have taken, confirm or update your address, track your certification progress, and submit questions to the Sun Educational Services' Certification Department.

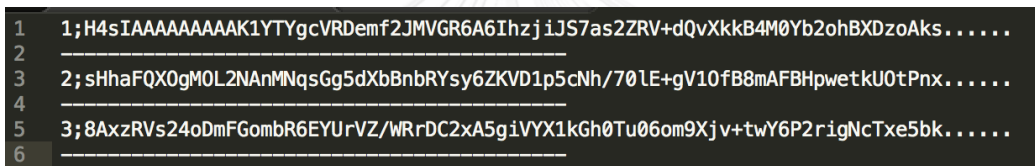
This assistance is available online at <http://www.certmanager.net/sun/>.

To help you further your growth and achievement in the Java language, Sun Educational Services offers a wide variety of instructor-led training (ILT), web-based training, and media-based training that can help you prepare for future certification examinations. You can find information on these offerings, as well as

This examination was delivered at an Authorized Prometric Testing Center. To register for another Sun Microsystems exam logon to <http://www.register.prometric.com>. To learn more about Sun Microsystems certification programs visit [www.prometric.com/sun](http://www.prometric.com/sun).

Figure 33. Digital signature embedded printout

Scanning barcodes in Figure. 32 contains three strings shown in Figure. 33. These strings can be combined to be the embedded digital signature. After scanning and obtaining embedded digital signature, trustworthiness of digital signature is verified. Then, the remaining task is calculating image hash from scanned image and compare it with image hash data set from digital signature. In this comparison, the threshold is set to be 9% of distance since 9% of hamming distance is an intersection point of false acceptance and false rejection of proposed algorithm that represented in Figure 25. It means that if distance from image hash comparison is more than 9% then compared image is altered and cannot be fully trusted. Table 14 demonstrates distance of each region and clearly shows that in normal situation this proposed system can verify and prove trustworthiness of printout correctly.



```

1 1;H4sIAAAAAAAAAAK1YTYgcVRDemf2JMVGR6A6Ihzj iJS7as2ZRV+dQvXkkB4M0Yb2ohBXDzoAks.....
2 -----
3 2;sHhaFQX0gMOL2NAnMNqsGg5dXbBnbRYsy6ZKVD1p5cNh/701E+gV10fB8mAFBHpwetkU0tPnx.....
4 -----
5 3;8AxzRVs24oDmFGombR6EYUrVZ/WRrDC2xA5giVYX1kGh0Tu06om9Xjv+twY6P2rigNcTxe5bk.....
6 -----

```

Figure 34. Scanned barcode

Table 14. Distance between specified regions obtained from original image and scanned image.

Specified Region Obtained from Original Image	Specified Region Obtained from Scanned Image	Distance
<b>CANDIDATE:</b> PARADORN ATHICHITSAKUL <b>ADDRESS:</b> ?addr? <b>CANDIDATE ID:</b> SR6417418	<b>CANDIDATE:</b> PARADORN ATHICHITSAKUL <b>ADDRESS:</b> ?addr? <b>CANDIDATE ID:</b> SR6417418	5.4%
On March 30, 2010	On March 30, 2010	6.6%
API Contents: ..... 85 % Concurrency: ..... 60 %	API Contents: ..... 85 % Concurrency: ..... 60 %	3%
There are 60 questions in this exam. You answered 51 questions correctly which gives you a score of 85%. <b>PASSING SCORE:</b> 58% <b>YOUR SCORE:</b> 85% <b>TEST STATUS:</b> Pass	There are 60 questions in this exam. You answered 51 questions correctly which gives you a score of 85%. <b>PASSING SCORE:</b> 58% <b>YOUR SCORE:</b> 85% <b>TEST STATUS:</b> Pass	2.8%

## Chapter 6. Analysis and Discussion

In theory, the image hashing algorithm presented in [17] performs well, however, its accuracy is practically lower than expectation because shearing, rotation, noise and other similar factors influence on radon transform output. To overcome this limitation, some steps of [17] were modified to obtain to the new algorithm with more robustness and flexibility. In Table 15, the FRR is varied depending on the fixed FAR at 15.25%.

Note that user will face 25.75% of FRR when applying this proposed framework with 15.25% of FAR as shown in table 15. Although the framework outperforms the other methods, but this FRR can be further improved by integrating an alternative validation mechanism as the post-processing stage.

*Table 15. FRR of each algorithm at 15.25% FAR*

Algorithm	False Acceptance	False Rejection
Proposed Method	15.25%	25.75%
Radon-Based	15.25%	41.5%
DCT-Based	15.25%	36.23%
pHash DCT-Based	15.25%	42.75%

Although the experimental result show that performance of proposed algorithm is great and better than the existing algorithm, but in very detail image or very small editing, performance of proposed algorithm slightly drops and not be able to discriminate edit content out from original one Figure. 35 shows example of large detail that this proposed algorithm cannot identify counterfeiting activity, hamming distance of image hash value via proposed algorithm is only 7.6% which less than the specify threshold.

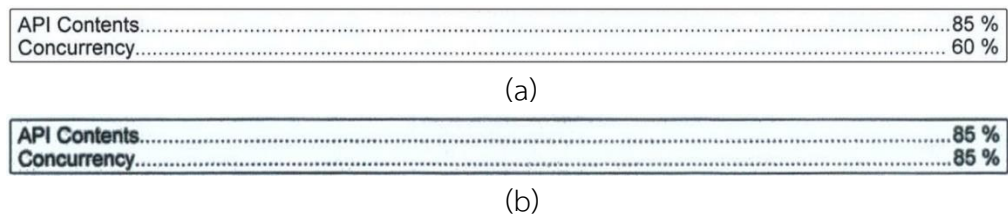


Figure 35. Large detail image (a) Original image (b) Edit image

Another interesting issue on prove and implementing this proposed framework is sizing of embedded words in 2D barcode, since 2D barcode was not design to held a very large size word or binary like whole digital signature, using 2D barcode at maximum capacity such as 900 code words going to lead a problem on a decoding process. At maximum capacity of 2D barcode almost pixel use to represent a data, so noise, damage or any flaw in printing and scanning process will influence and bring unrecoverable problem to decoding process.

## Chapter 7. Conclusion and Future Research

Applying a digital signature to printout to increase the trustworthiness of printing document is a good idea. However, the problem lies in selecting what should be used as a data representation of printout since cryptographic hash cannot be applied to printout directly. Consequently, this thesis proposes a printout verification process using image hashing algorithm based on Radon, Wavelet, and Fast Fourier transforms to capture fingerprint of printout. The result of proves that the proposed image hashing algorithm is possible to be employed as a fingerprinting mechanism for printout. For the future work, in order to implement this framework in real world situation, FAR and FRR should become less. With this reason, pre-processing of printout should be considered together with applying some alternative validation mechanism as a post-processing and some steps of algorithm may be further analyzed and adapted.



## REFERENCES

- [1] W. Shanks, "Building and Managing a PKI Solution for Small and Medium Size Business," Available: <https://www.sans.org/reading-room/whitepapers/certificates/building-managing-pki-solution-small-medium-size-business-34445>. (Accessed, 24 July 2016).
- [2] NIST, "Digital Signature Standard (DSS)," FIPS Publication 186-3, 2009.
- [3] "Electronic Signature," ELECTRONIC TRANSACTIONS ACT B.E. 2544, chapter 2, section 26, pp. 1-2, 2001.
- [4] M. Warasart and P. Kuacharoen, "Paper-based Document Authentication using Digital Signature and QR Code," in Proceedings of the 2012 4TH International Conference on Computer Engineering and Technology, Bangkok, Thailand, vol. 40, pp. 94-98, 2012.
- [5] L. Weng and B. Preneel, "A Secure Perceptual Hash Algorithm for Image Content Authentication," *Communications and Multimedia Security, Lecture Notes in Computer Science*, vol. 7025, pp. 108-121, 2011.
- [6] L. Yu and S. Sun, "Image Robust Hashing Based on DCT Sign," in Proceedings of the International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Pasadena, CA, USA, pp. 131-134, 2006.
- [7] J. Dittmann, A. Steinmetz, and R. Steinmetz, "Content-based digital signature for motion pictures authentication and content-fragile watermarking," in Proceedings of the 1999 IEEE International Conference on Multimedia Computing and Systems, Florence, Italy, vol. 2, pp. 209-213, 1999.
- [8] M. Schneider and C. Shih-Fu, "A robust content based digital signature for image authentication," in Proceedings of the International Conference on Image Processing, Lausanne, Switzerland, vol. 3, pp. 227-230, 1996.
- [9] L. Chun-Shien and H. Y. M. Liao, "Structural digital signature for image authentication: an incidental distortion resistant scheme," *IEEE Transactions on Multimedia*, vol. 5, no. 2, pp. 161-173, 2003.

- [10] C. Winter, M. Steinebach, and Y. Yannikos, "Fast indexing strategies for robust image hashes," *Digital Investigation*, vol. 11, no. 1, pp. S27-S35, 2014.
- [11] J. Fridrich and M. Goljan, "Robust hash functions for digital watermarking," in Proceedings of the *International Conference on Information Technology: Coding and Computing*, Vegas, NV, USA, pp. 178-183, 2000.
- [12] IETF, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile," RFC 5280, 2008.
- [13] Z. Jie, "A Novel Block-DCT and PCA Based Image Perceptual hashing Algorithm," *International Journal of Computer Science Issues*, vol. 10, no. 1, pp. 399-403, 2013.
- [14] V. Monga and B. L. Evans, "Perceptual Image Hashing Via Feature Points: Performance Evaluation and Tradeoffs," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3452-3465, 2006.
- [15] Z. Hanling, X. Caiqiong, and G. Guangzhi, "Content Based Image Hashing Robust to Geometric Transformations," in Proceedings of the *2009 Second International Symposium on Electronic Commerce and Security*, Nanchang, China, vol. 2, pp. 105-108, 2009.
- [16] D. V. Jadhav and R. S. Holambe, "Feature extraction using Radon and wavelet transforms with application to face recognition," *Neurocomputing*, vol. 72, no. 7-9, pp. 1951-1959, 2009.
- [17] D. Wu, X. Zhou, and X. Niu, "A novel image hash algorithm resistant to print-scan," *Signal Processing*, vol. 89, no. 12, pp. 2415-2424, 2009.
- [18] G. Y. Chen and B. Kegl, "Feature extraction using Radon, wavelet and fourier transform," in Proceedings of the *IEEE International Conference on Systems, Man and Cybernetics*, Montral, Que, Canada, pp. 1020-1025, 2007.
- [19] E. Magli, G. Olmo, and L. L. Presti, "Pattern recognition by means of the Radon transform and the continuous wavelet transform," *Signal Processing*, vol. 73, no. 3, pp. 277-289, 1999.
- [20] ETSI, "CMS Advanced Electronic Signatures (CAAdES)," *Electronic Signatures and Infrastructures (ESI)*, 2011.
- [21] IETF, "JSON Web Signature (JWS)," RFC 7515, 2015.

- [22] H. Wang, L. Wang, and C. Bai, "The application of digital signature Technology and fingerprint identification in 2D barcode person identity," in Proceedings of the *World Automation Congress (WAC)*, Puerto Vallarta, Mexico, pp. 1-4, 2012.
- [23] D. J. S. Geldenhuys and A. J. Hoffman, "A Digital Signature Issuing and Verification System for Auto Identification Tokens," in Proceedings of the *2012 8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, pp. 1-7, 2012.
- [24] ISO, "PDF417 bar code symbology specification," Information technology -- Automatic identification and data capture techniques, 2006.
- [25] ISO, "QR Code 2005 bar code symbology specification," Information technology — Automatic identification and data capture techniques, 2006.
- [26] ETSI, "CAAdES Basic Electronic Signature (CAAdES-BES)," Electronic Signatures and Infrastructures (ESI); CMS Advanced Electronic Signatures (CAAdES), pp. 16-18, 2011.
- [27] C. Zauner, "*Implementation and Benchmarking of Perceptual Image Hash Functions*," Available: [http://www.phash.org/docs/pubs/thesis\\_zauber.pdf](http://www.phash.org/docs/pubs/thesis_zauber.pdf). (Accessed, 25 July 2016).

## VITA

Name: Paradorn Athichitsakul

Affiliation: Advanced Virtual and Intelligent Computing (AVIC) Center,  
Department of Mathematics and Computer Science, Faculty of Science,  
Chulalongkorn University, Thailand

Work: I work for Electronic Transaction Development Agency (Public  
Organization) (ETDA) in Research and Development Division as a Software Engineer.  
My responsibility is studying standardization and implement it in order to create a  
software or service that can improve Thailand electronic transaction.

Education: I graduated Bachelor degree in Information Engineering at King  
Mongkut's Institute of Technology Ladkrabang in 2009.

