Security Protocol Architecture for Website Authentications and Content Integrity

Belal Abuhaija, Nidal Shilbayeh, Mohammad Alwakeel
Faculty of Computes and information Technology
University of Tabuk
Tabuk, Kingdom of Saudi Arabia
{babuhaija, malwakeel}@ut.edu.sa, N_shilbayeh@yahoo.com

ABSTRACT: *With the spread of the Internet and networks, more people are relying on information sharing to research any subject. The number of Websites that are providing information on a single subject is huge. Some subjects and issues are somehow more sensitive to some people than others. However, there is no agreed mechanism for authenticating the content provider Web sites or preserving the integrity of its contents. One aim of this paper is to develop frame work architecture protocol for Web site content authentication to build a trusted Web sites directory called Islamic trust (ITRUST). We cannot treat all documents with the same measure of security. Another aim is to solely focus on the data integrity rather than intellectual property rights and to provide different documents with different security levels. In the proposed protocol we will employ public key infrastructure (PKI) and Digital Certificates (DC) for Website authentication and watermarking techniques for document integrity.*

Keywords: Watermarking, PKI, Security

Received: 29 April 2013, Revised 19 June 2013, Accepted 24 June 2013

1. Introduction

As the Internet usage has been growing at an exponential rate; many people using the Internet as their main source of information. Most of the time; this information can be in the form of multimedia technologies such as text documents, images video and audio. Most of the Web sites that offer such information do not need to verify their Web sites contents as they consider themselves as information holder and not information verifiers. However, some information needs to be verified for correctness and accuracy, especially when such information is considered to be sensitive and critical to someone faith and believes. In this research we are proposing a frame work for Web sites authentication and preserve content integrity and security.

The public access of the Internet and the spread of Web sites with various digital contents present a problem for users who would like to make sure that the information presented on Web sites is accurate. Religion and faith has emerged in recent years as the most contentious subject of them all. A lot of people are trying to explain Islam and Islam teachings according to their

special agendas. A huge number of Web sites about Islamic religion are spread on the Internet. Some Islamic teachings or the Sharia (i.e. the rules that guide Muslims) are split into several subjects. The core faith of Islam is the same; Such as how many prayers in a day or when is the month of Ramadan (the fasting month) starting or when to make HAJ (the holy pilgrimage to Mecca). All of these issues are agreed up on issues with very minor variations. However, when we are discussing the day to day practices and dealings with Muslims and non-Muslims (people of other Faiths) the issues are widely spread and disagreed up on issues. Such problem can lead to confusion about Islam and its principles for Muslims and Non-Muslims. To list a few examples; how do we as Muslims deal with Banks, or what are the rights of our non-Muslim neighbors, or even what are the rights of other Muslims, marriage between a Muslim and non-Muslim and the fruits of such marriage (i.e. the children).. Such issues are widely debatable and dis agreements arise from the lack of understanding the basic principles of Islam and the aim of our religion. We aim in this paper at providing framework architecture for Web sites that publishes Islamic teachings.

We are proposing a framework protocol that will authenticate the Web site itself and at the same time provide security and protection through watermarking techniques to the content regardless of the type of document (text, audio, video and images. Public Key Infrastructure (PKI) is a framework that enables integration of various services that are related to cryptography. The aim of PKI is to provide confidentiality, integrity, access control, authentication, and most importantly, non-repudiation. One important fact about PKI is that it offers the non-repudiation (non-rejection) which prevents parties from denying involvement in the online transaction, and also supports the digital signature and message encryption that further enhances the security elements within any network. Finally, PKI also supports cross-certification, and as such; identity can be enabled in integration among circles of trust [12].

There are three functionalities of any PKI, first, to able to register Websites and issue their public key certificates as well as to build trust relation between the websites in a public key infrastructure. Second, a policy to implement certificates revocation shall be put in place and a facility to publish this information to all concerned parties. The certificates duration should not be done indefinitely; they must be limited by a specific time. However; time is not the only reason that might revoke a certificate; compromising of private key is another reason among others. Third, encryption of messages and emails exchanged using public key and private key.

Digital watermarking has been suggested for use in many applications; Such as, preserving the intellectual property rights of owners [1], media fingerprinting system [2], broadcast and advertising monitoring [3] and covert channel communications [4].

Digital watermarking has many forms depending on the purpose of the mark and on the application using the mark. One technique used is to keep the watermark content intact in the case of tampering with the watermark, this is called robust technique. Another technique which is rendered the watermark unusable and destroyed if tampered with is called fragile watermark.

Embedding copyrights information in the document usually uses invisible watermark technique which is used in most digital form of watermark. Invisible watermarks are transparent when the content is viewed, listened to or heard [16]. However, visible watermarks are embedded in such a way that is visible when the content is viewed. In such context, text documents mostly uses visible watermark, while other forms of documents mostly uses invisible watermark.

From the above, it is clear that the main goal of Digital watermarking techniques is to preserve the intellectual rights of the author of the documents (i.e. text, audio, video and image). However, in our approach we are concerned with the integrity of the content and the authentication of the provider rather than the copyrights of the author as in most cases the authors are providing the information free of charge.

There are several objectives to this research; first; developing a framework for Web site authentication. Second, use appropriate watermark techniques in order to preserve the integrity of the documents. Third, provide a monitoring data integrity and websites security. As far as we know this is the first time that a research is concerned with the integrity of the published content rather than the copyrights of the authors.

This paper is organized as follows. Section II gives some literature review; then section III introduces the main entities of the proposed protocol; section IV is discussion and implementation procedures and we conclude with section V.

2. Literature Review

Hundreds of websites on the Internet are providing information about Islam and Islamic teachings. The fundamentals of Islam

or the five pillars of Islam are mostly agreed up on issues among the main stream Muslims. However, the interpretations of some teachings of Islam might be at times disagreed upon issues. For example what does a specific verse in the Holy book (Quran) means? Or other issues that govern the relationships between Muslims and Non-Muslims, what are their rights? Or even when it comes to marriage; which is widely spread these days or even dealings in the banks, can we get a loan with interest or not? All such issues and more are mostly debatable issues among Muslim scholars. Such issues needs to be articulated well and answered by a well-recognized Muslim Scholars (Olama). Many websites are providing information about such issues without an authenticated reference from the Muslim Scholars. Others are presenting the material to mean other than the intention of the question or the issue at hand. One other very important issue that often arises is to which doctrine this Fatwa belongs to. Muslims have four main stream doctrines or ideologies. The masses of Muslims can follow any one of these doctrines. In this research we are aiming at providing streamline architecture to provide correct and reliable information to Muslims and Non-Muslims about Islam and any other issues that may arise. We aim at ensuring that the websites are secure and authenticated through PKI and the documents they publish are presented well and secure through digital watermarking techniques. Digital watermarking is a technique for embedding information into a document, the document might be of any type, for example text, audio video or image. At a later stage this information might be extracted or detected for the purpose of authentication or identifications. It has never been the objective to check for the integrity of the content for the purpose of preserving the information. This might be the case only as part of the main objective of preserving the intellectual rights.

In [5], the authors presented a framework to monitor media broadcasts utilizing Public Key Infrastructure (PKI) and Digital Certificates (DC). They proposed an independent monitoring agency to operate the framework. The authors concentrate on IPTV as the new way of delivering audio and video contents across IP networks. The authors deploy a PKI infrastructure to establish a trust relationship between the three entities involved, the broadcaster, the monitoring agency and the rights entity. In our approach we are utilizing PKI principles to establish the trust relations with different entities at two levels as well be explained later.

In [6], the authors acknowledge that it is difficult to protect against copying. However, the requirements for copyright protection schemes have been presented. Such as, there should be a mechanism to uniquely identify the owner of the multimedia object. Copyright token should be difficult to detect or remove from the stream; as well as the ability of the author to trace the unauthorized copies of the stream. The authors argue that watermarking alone is not sufficient to resolve the rightful ownership of digital data. They proposed the use of protocol relying on public key infrastructure is necessary as well. We are using similar approach to gain confidence into the published material as we are concerned with the integrity of the product rather than the rightful ownership. We might be concerned with ownership in case somebody challenges the content of the document. In [7], the authors investigated the problem of tracing unauthorized distribution of sensitive intelligent documents by proposing a watermark-based distribution protocol which complements the traditional cryptography based access control scheme.

The scheme uses the embedded identity of the users in a distribution list that are allowed access to the intelligent document. The distribution protocol consists of generating the watermark and intelligence user certificates, then the acquisition of the watermarked intelligence document and finally resolving the policy violation if it happened. To maintain watermark secrecy, the document provider is not allowed access or provided by the watermarks. The authors concluded that the protocol provides a concrete support for non-repudiation in the document distribution process as it identifies each user who made copies of the intelligent document and make them accountable for the producing additional copies. In our protocol we use the similar concept as we embed the website information in the document to identify the provider. This shall help in case of discrepancies to go back and check the original document as published by the website. Therefore, the website is held accountable for the integrity and reliability of the information. In [8], the authors presented a distribution protocol to address the management of documents in large enterprises. The protocol uses registration certificates to distribute the end user identity information. The problem of binding the end user identity to copies of sensitive documents whether the document is hard of soft format is a hard challenge in enterprise settings. To support non-repudiation, document policies management reinforcement is essential to identify the end user who owns the copy. The authors differentiated the end user role from the document provider role in their protocol. The proposed protocol distinguished copies made by the document provider from copies made by the end user. The authors presented four role entities in the document distribution; Document provider, enterprise registration authority, end user and policy enforcer. The authors claim that the proposed protocol provides strong support for non-repudiations by identifying the end user. The protocol addresses the issues of maintaining watermark secrecy and provides protection from Trojan horse attacks. In our protocol we are using the same approach by identifying the website provider in case that somebody challenges the document. The reason for this is to monitor the provider for document contents integrity. The website is a public domain and since the documents are available to everybody who seeks information it will not be practical to monitor the end user. However,

such information can be used in sending alerts to content providers to be on the lookup for any security issues with their website contents. In other words, the change in the document if any is occurring after the download and the website contents is still intact.

One of the main issues that we are addressing in our proposed architecture is that the security measure for the documents is not the same for all website contents. Thus, we categorize and classify documents based on their importance and sensitivity. For example, Quran reciting should have the highest priority, while lectures fall far behind. However, Fatwa (scholar opinion on current issues) needs to be tagged by the doctrine of the scholar as this is a very important issue since the follower of one doctrine can keep with the same ideology.

3. Watermarking Framwork Protocol

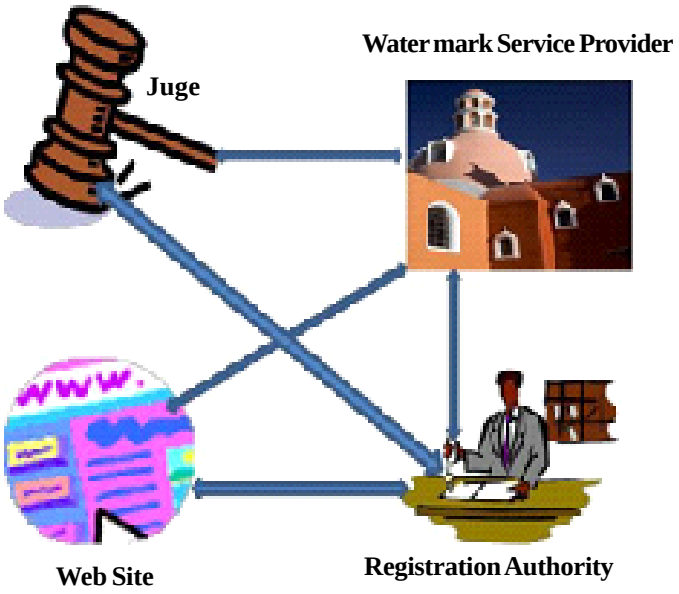

Figure 1. Protocol Entities

Figure 1 illustrates the proposed protocol which has four main entities. First, the content provider is the website who publishes the documents (text, audio, video and images). The content provider has a direct relation with the registration authority. Second, registration authority (RA) which registers, authenticate and monitor the websites or content providers. Third, watermark service provider (WSP) which provides different watermarking techniques to documents based on the nature of the document. Fourth, judge authority (JA) which approves and ensures the safety of the contents of the documents. The relationships between entities vary based on each entity's functionality. Below is a description of the functionality of each entity.

2.1 Functionality of Content Provider (CP)

• The Content Provider registers with the RA.

• The registration uses Public Key Infrastructure (PKI) as shall be explained in section IV.

• CP applies for Digital Certificate from RA (may use X.509 V3 standard).

• After receiving the Digital certificate; it may communicate with WSP.

2.2 Functionality of Registration Authority (RA)

• RA checks the credentials of CP; if approved.

• Step 2: RA generates public key and private key upon the request from CP.

• Step 3: The PK and PrK are generated based on the IP address, URL and other information supplied by CP.

• Step 4: RA store a copy of the public key and private key in a secure database.

• Step 5: RA transmit a copy of private key CP in a secure manner and publish the public key of the CP.

• Step 6: RA issues digital certificate to CP to establish CP credentials bind the CP to public key.

• Step 7: RA store the digital certificate and it keeps a digital certificate revocation list.

2.3 Functionality of Watermark Service Provider (WSP)

• WSP shall authenticate RA (same procedure as suggested between CP and RA) to add another layer of security to the system.

• Step 2: WSP checks the credentials of the CP using the digital certificate generated by RA. If approved;

• WSP forwards any document received from CP to JA.

• WSP request approval on documents (text, audio and video) from JA.

• Provides watermarking (visible or invisible) to the content provider and may use finger print in the case of audio and video to any document that is approved by JA.

• WSP forwards the approved documents to RA and CP.

2.4 Functionality of Judge Authority (JA)

• JA registers with RA. Step 2: Repeat steps 2-4 as in CP.

• WSP checks the credentials of JA through its digital certificate

• JA receives documents from WSP. Step 5: JA checks the content of the document.

• If the document is approved then the JA uses its digital signature on the document. This is to distinguish different judges inside this entity.

• WSP sends the signed document to RA.

3. Framework Implimentations

The above procedure deals with two issues; first, establishing a trust relationships between the content providers or Web sites, the judges and the watermark service provider by using PKI and by issuing digital certificates for the CP and the judges. Key management and distribution is an essential part of this process. Secure socket layer (SSL) protocol shall be used to authenticate the content providers by using several cryptographic algorithms. The registration authority shall generate a file that contains all the public keys which is corresponding to each entity; after all it is the entity that is required to check the identity of each request coming from the CP or the JA.

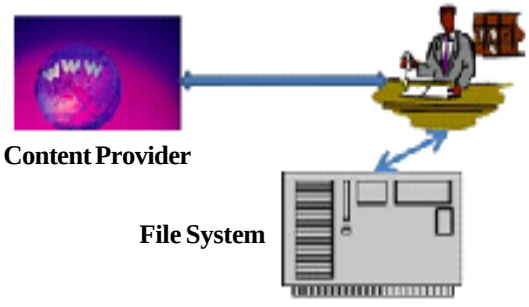

Figure 2. Key Establishment and Management

The website can use its private key combination to encrypt any text documents that are sent to RA, the RA has access to the private key/ public key of the website. Certain information is required to issue the websites with key pairs. Such as URL, IP addresses, location and contact person.

The RA entity has an additional responsibility to register, authenticate and issue digital certificates. This same procedure can be used within the Judge Authority (JA) to register, authenticate and provide judges with digital certificates. In the proposed protocol, the main role for the RA is to verify the submitted public keys and then manage the issued certificates along with all

associated process. Therefore, the entire processes of authentication and authorization services are supported and enhanced. The RA shall ask the WSP to generate visible logo for the Website [11]. This visible logo should be used on all text or image documents that are published by the Website for identifications.

Inside the judge authority there are many judges to approve the documents. Each judge needs to have public key and private key along with digital certificates [9-10] to identify the person who is reviewing the document. In such case before the judge can be approved for reviewing he/she must register with the RA. The RA verifies the judge identity and issues digital certificate to the judge. WSP can use such information to verify that the judge is registered and has a valid digital certificate from the RA. The importance of the PKI in enabling an appropriate security during transmission cannot be minimized in any proposed protocol. It has been cited in many references that information security is as good as the requirements [17]. The number of Websites that can utilize such protocol are humongous. Therefore; transmission security is an essential part of the network infrastructure.

WSP is concerned with the watermarking services to preserve and protect the integrity of the documents. Once the JA communicates with the WSP that a certain document is approved; WSP generates the appropriate watermarking information based on the type of document. In case of an image; water marking as stated earlier can be of two types fragile and robust. The latter is concerned with the copy right protection as in ownership proof and identifications or copy prevention and control. Such issues are related to intellectual rights which are not of a concern to us in the proposed architecture. We focus on fragile watermarking which is useful in two applications. First, authentication of the media source and second; content integrity verification which is more of concern to us in our proposed architecture [19]. Content integrity watermarking must have the ability to protect against cut and paste manipulations of digital images to present and use such images in different settings. The watermark embedding technique shall be used to prevent the modification of part of the image or cut part of the image to produce a new totally different image as illustrated in Figure 3 below.

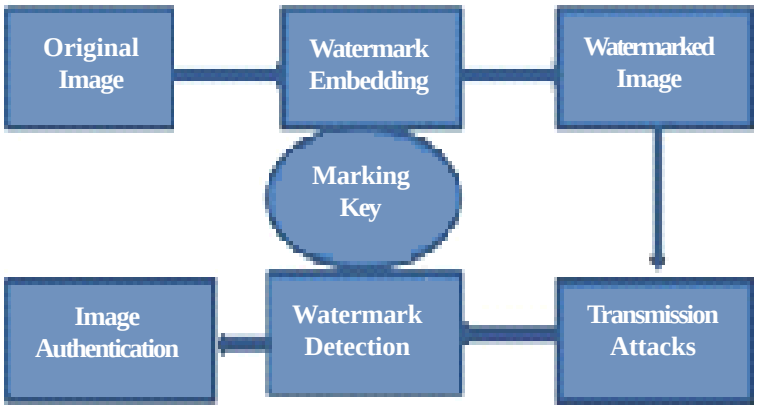

Figure 3. Image Authentication block diagram

The watermark shall be usable for black and white images as well as colored images. Such watermarks should be invisible to the viewer. However, one other image that shall be posted at the websites as a logo of authentication shall be visible to all people accessing the site. For documents with audio contents; the audio marking key is unique to the documents with audio contents. However, audio can be of several types such as the Quran recitations or typical lectures. In this case the watermarking key shall be different and suitable for the application. It is important when using the marking key (i.e. inserting information) that the content of the Quran is not altered or modified; this can be ensured by using fragile watermark. The insertion of digital watermarks should not affect the quality of audio as well as it should not be noticeable by the user [18]. The same can be said about video. However, audio watermark in general can be of two types; the time domain and frequency domain. A technique that is used in the time domain is Low-bit encoding. While in the frequency domain; techniques like Discrete Fourier Transform and Discrete Cosine Transform can be used.

As far as text documents are concerned; visible Logos and other techniques can be employed. WSP requirements and algorithms shall be flexible and scalable to accommodate a wide variety of documents. Watermark attacks can be of several types depending on the document type; however, the basic principal of the attack is that the attacker can detect the watermarking that has been used and then the attacker can read and modify the hidden message and ultimately remove the hidden information. The new document approval requests that are sent by the CP's; the WSP entity can verify visible logos of the websites. These visible

watermarking logos on the website pages [11] are produced up on registrations with the RA. Information provided by the Websites that are used in the private/public key generations can be used by WSP as well to provide digital watermarks. WSP can use many different algorithms to secure a document depending on the nature of the document i.e. text, audio, video or image [13-14]. However, the protocol can do much more than that. The websites should also publish the registration authority that approves the website. This is to ensure that that incase of any disputed documents the end user can communicate directly with the RA to check the integrity of the document. The registration authority can periodically check the contents of the CP's that is under its control to make sure that an acceptable level of security is available as well as no tampering with the documents took place.

4. Conclusion

In this work we are proposing the first step in a tenacious subject in which its effects can go a long way. The above framework protocol is designed to ensure that the published information is accurate and correct. It is also a starting step in building a trust relation between the CP and the end user. The protocol has several layers of security build into it; while at the same time flexible enough to fit a wide range of opinions which are all valid and main stream.

Website integrity will be maintained through separation between the integrity of the website and the integrity of the contents.

The watermarking procedure is kept with one entity, in this case WSP such that it prevents the revealing of techniques to any other entity and keeps it secret.

Future work will deal with detailed implementations of such protocol and jump start the process of document integrity and trusts. WSP algorithms and procedure will be paid a special attention due to the fact that text, audio, video and image documents have different watermarking requirements.

5. Acknowledgment

The Authors Gratefully Acknowledges the Support for this work from it Research center for the holy Quran and its Sciences (Noor), Taibah University, Madinah, Saudi Arabia.

References

[1] O'Ruanaidh, J. J. K., Dowling, W. J., Boland, F. M. (1996). Watermarking digital images for copyright protection, Vision, *Image and Signal Processing, IEE Proceedings*, Aug.

[2] Kirovski, D., Malvar, H., Yacobi, Y. (2004). A dual watermark-fingerprint system, *IEEE MultiMedia Journal*, July-Sept.

[3] Jian Zhao. (1997). Applying Digital Watermarking Techniques to Online Multimedia Commerce, *In*: Proc. of the International Conference on Imaging Science, Systems, and Applications (CISSA97), June 30-July 3, Las Vegas, USA.

[4] Mazurczyk, W., Kotulski, Z. (2008). Covert Channel for Improving VoIP Security, *Advances in Information Processing and Protection*.

[5] Dominik Birk, Sean Gaines, Christoph Wegener. (2008). A Framework for Digital Watermarking Next Generation Media Broadcasts, *IAENG International Journal of Computer Science*.

[6] Katzenbeisser, S. (2001). On the design of copyright protection protocols for multimedia distribution using symmetric and public key watermarking, *In*: Proc. 12th Int. Workshop Database and Expert Systems Applicat., Sept., p. 815–819.

[7] Cheung, S. -C., Chiu, D. K. W., Ho, C. (2008). The use of digital watermarking fo intelligence multimedia document distribution, *J. Theor. Appl. Electron. Commer. Res.*, 3 (3) 103–118.

[8] Cheung, S. C., Chiu, D. K. W. (2003). A watermark infrastructure for enterprise document management, *In*: Proc. 36th Hawaii Int. Conf. System Sciences (HICSS'03), Hilton Waikoloa Village, HI.

[9] Johnson, M., Ishwar, P., Prabhakan, V., Schonberg, D., Ramchandran, K. (2004). On compressing encrypted data, *IEEE Trans. Signal Process.*, 52 (10) 2992–3006, Oct.

[10] K. Kuroda, M. Nishigaki, M. Soga, A. Takubo, and I. Nakamura, A digital watermark using public-key cryptography for open algorithm, *In*: Proc. ICITA 2002, Also Available [Online]: http://charybdis.mit.csu.edu.au/~mantolov/CD/ICITA2002/ p. 13 1-21.pdf.

[11] Moulin, P., Wang, Y. (2004). New results on steganographic capacity, *In*: Proc. Conf. Information Sciences and Systems (CISS 2004). Princeton, *NJ*, Mar., p. 813–818.

[12] Ray Bird, Inder Gopal, Amir Herzberg, Philippe Janson, Shay Kutten, Refik Molva, Moti Yung. (1995). The KryptoKnight Family of Light-Weight Protocols for Authentication andKey Distribution, *IEEE/ACM Transactions on Networking*, 3 (1), February.

[13] Yaxun Zhou, Wei Jin. (2011). A novel image zero-watermarking scheme based on DWT-SVD; *Multimedia Technology* (ICMT), 2011 International Conference, Date: Aug, p. 2873-2876.

[14] Yan Yang, Rong Huang, Mintao Xu. (2009). A Novel Audio Watermarking Algorithm for Copyright Protection Based on DCT, *Electronic Commerce and Security*, *Second International Symposium*; Date: Oct., p. 184-188.

[15] Jungyeop Kim, Sungmin Won, Wenjun Zeng, Soohong Park. (2011). Copyright protection of vector map using digital watermarking in the spatial domain, Digital Content, Multimedia Technology and its Applications (IDCTA), 2011 7th International Conference, Date: Dec., p.154-159.

[16] Pei-YuLin. (2012). ImperceptibleVisibleWatermarking Scheme Using Color Distribution Modulation, Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC), 2012 9th International Conference on, p. 4-7 Sept.

[17] Xiao Le Li et al. (2012). Design and Verification of Security Protocol for Information Transmission in Digital Campus, *Advanced Materials Research*, June.

[18] Xiumei Wen, Xuejun Ding, Jianhua Li. LitingGao, Haoyue Sun. (2009). An Audio Watermarking Algorithm Based on Fast Fourier Transform. *International Conference on Information Management, Innovation Management and Industrial Engineering*.

[19] Hasan, M. H., Gilani, S. A. M. (2006). A Fragile Watermarking Scheme for Color Image Authentication, *World Academy of Science, Engineering and Technology*, 19, p. 39-43.