Aspects of a Watermark Solution

Dominic BUCERZAN[1], Crina RATIU[2], Ioan DASCAL[2]
[1]Aurel Vlaicu University from Arad, Romania
[2]Vasile Goldis University from Arad, Romania
dominic@bbcomputer.ro, ratiu_anina@yahoo.com, nelu_dascal1@yahoo.com

Although watermarking is a relatively new technology, there are many ways of applying it on an electronic data set with the purpose of ensuring copyright integrity and authenticity of the electronic data. But, bearing in mind the evolution of information technology and of communication, a watermark may be the target of several attacks which aim at its robustness, its form and even at its removal. In order to reduce these threats, this paper proposes a solution - digital watermarking completed by a hash function which has an important role in the authenticity of the content of a message and in the security of the transmission of this message through computer networks which are the main support of collaborative systems.
Keywords: *Digital Watermarking, Hash Functions, Digital Information Security, Collaborative Systems*

1 Introduction

Digital Watermarking is associated with the old technique of hiding information, namely steganography. This kind of marking hides a secret of a personal message, in order to protect the right of property or to demonstrate authenticity. It is also possible to verify the originality of the message, check its content and data integrity in order to detect fraud. The marking of digital images, audio, video or electronic data in general, has as main target the establishing of the right of property and the checking of the original content.

Taking into consideration its features developed so far, we consider digital Watermarking a solution for ensuring the authenticity of shared information into collaborative activities like: collaborative e-Business, Web-Business and collaborative Education. We consider that the positive impact on securing this collaborative transaction is improved by our proposal to use an additional hash function.

The new digital technology for creating, processing and stocking multimedia products has been considered by its designers, producers, editors and clients to be extremely useful.

Simultaneously digital communication through computer networks has increased rapidly. Digital products may be easily copied, processed for various goals or publicly exposed.

Digital Watermarking consists of a signal inserted into an image, an audio signal or video signal inserted into any digital document. [1]

Watermark may be a series of numbers, the name of a company or someone's signature. The main feature of a watermark is that it cannot be seen or heard. So, if we see and listen to a multimedia object which contains a watermark, we cannot detect the watermark signal. However, the watermark may be detected and extracted with a certain algorithm.

The need to protect digital information against fraud and illegal manipulation has occurred in the same time with the communication through Internet. Electronic publishing and electronic commerce of digital information increase the danger of fraud and of intellectual theft [2].

As far as analogical signals are concerned, the problem solves itself, because by copying these kinds of data we get a lower quality than the original one audio and video file). On the other hand, digital information can be perfectly copied and it is difficult, if not impossible, to make the distinction between the original and copied version. Besides, there is not a mechanism that can help us

detect the illegal copying or the alteration of the content. In order to protect ourselves against copying and in order to protect the copyright for digital data, two complementary techniques have been developed: encrypting data and invisible marking. Cryptosystems may be used in order to protect digital data on its way from the sender to the receiver. The received and decrypted data is identical to the original one and it is no longer protected. The techniques of invisible watermarking may complete the cryptography by placing a secret undetectable signal, namely a transparent marking straight into the original data. The signal is inserted using a method that ensures its constant presence in the original data [5].

In this paper, we propose a new solution of digital watermarking which integrates a robust hash function to ensure the integrity of electronic data and to avoid possible attacks which aim the transmission of information through collaborative activities and more specific through computer networks.

2 Main Features of a Watermark

Perceptual invisibility

The changes brought by the usage of the watermark should not affect the quality of the data. However the differences between the original data and the data which contain the watermark may be detected if a comparison between the two products is made. As a consequence, these differences can remain hidden because the original data is available to its rightful owner. [1]

Complexity

Watermarks should be characterized by great complexity. This is mandatory in order to produce a great variety of watermarks different from one another.

Associated keys

Watermarks must be associated with an identification number named watermark key. The key is used to generate, detect and delete watermarks. At the same time the key must be individual and it must feature exclusively its rightful owner.

Trustworthy Detection

Watermarks must be sufficient proof of the copyright of the digital data. It clearly demonstrates the copyright of an image even if insignificant possibilities of error may occur.

Robustness

A digital image must undergo a large scale of more or less deliberate attacks (pirate attacks - intended or not, compression, filtering by cancellation of noise, resizing). All these attacks may or may not alter the inserted watermark. Obviously a watermark which is being used in the protection of copyright must be traceable within the limits in which the quality of the data remains pretty much the same.

It is practically impossible to create a system invulnerable to all kinds of attacks. New methods of overcoming marking systems will be invented and perfected in time. But basic knowledge about common attacks is required in order to find or improve existing security solutions.

A category of attacks consists in the unintended modification of the marked image (namely the modification of the image supposing that there is no watermark). This kind of threat should be recognized by any fragile marking, but it is mentioned because it is the most frequent way of attack. Variations of the attack consist, for example, in removing one face of a person and replacing it with another one [1].

Another type of attack is the one in which the image is modified without affecting the marking or without creating a new marking and the detector may consider it to be authentic. There are fragile makings which quickly discover certain modification of the image, but which may not detect elaborated modifications. Such an example is a marking inserted in the most unimportant bit of the image. Trying to change the image without realizing that the watermark can be found in the least significant bit will modify the marking and will be detected; if someone tries to alter the image without changing any of the least significant bits or without replacing them with a new set of less significant bits, the detector will consider the image to be authentic [5].

Somebody may be interested in completely destroying the marking without living any trace of his intervention. In order to succeed, one may insert a random noise into the image, he may use special programs created to destroy the marking or he may use statistical analysis in order to estimate the original image.

3 Solution for Watermark

It is common knowledge how difficult it is to protect an image against theft and against being used without the consent of its author. One of the solutions for this problem is the usage of a transparent watermark, which allows he image to be seen entirely and make the image useless for whoever may want to use it without authorization. In the following lines the procedure of applying a watermark in the Adobe program will be described.

A transparent layer 1280x6800 in Photoshop will be opened. After the new document has been created go to Layer / New/Layer. Now that we have a new layer, choose the palette Horizontal Type Tool or just press T key in order to open the writing text tool. After we have finished writing in the new layer of the document, another palette will be used to create a digital stamp, namely Custom Shape Tool (see Figure 1).

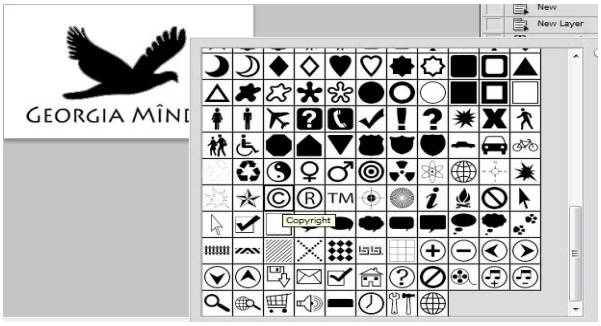

Fig. 1. Custom Shape Tool window - used to create a digital stamp

In the pallet Custom Shape Tool (U) several types of drawings and shapes can be found, from which we can choose a bird and the sign for Copyright. After the lineament of the text and images in order to fit the page press CTRL key and simultaneously click left to choose the two layers. After the layers have been selected click right on a layer and press Merge Layers. After having made all these steps we must go to Filters menu - Stylize and Boss and give the following measures (see Figure 2).

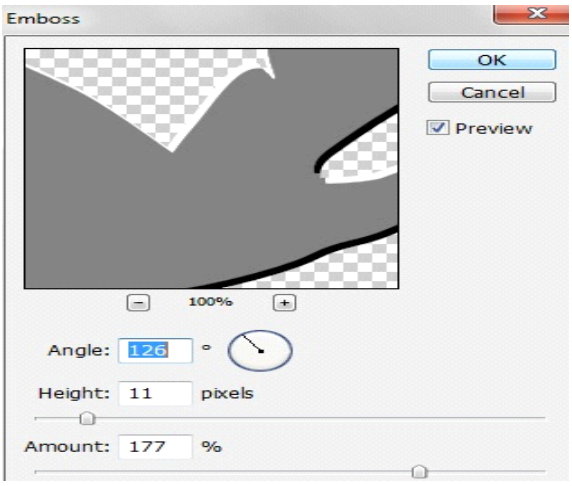

Fig. 2. Stylize and Boss window – see measures

After having finished all these steps we must go to layer / shape 3/ and above we can see an option which changes from Soft to Soft Light.

4 Authentication of the Image

As we can conclude so far, the application of a watermark is no longer an issue nowadays. There are several applications which facilitate Watermarking. However taking into consideration the evolution of technology of changing an image, watermarks can be removed. In order to avoid this kind of situation it is advised to use a "hash" function which has an important role in the identification of the content of a message sent through computer networks. [6]

The "hash" functions (category of functions that in specialized articles are also named dispersion functions or digest functions) are the spring of nowadays security data base. These functions change a series of symbols of arbitrary lengths such as a password of 8 characters or a 1000 pages document, in a series of symbols with relatively short of fixed dimensions. The main idea is that when a difference occurs in the case of the inserted series the result must change as well.

The hash functions are vital for the

cryptographic systems which protect communication systems. Susan Landau, in her article "Find me a hash", names this kind of function "tape function"- the duct tape of cryptography, because it has so many uses: to prove a message is genuine, to ensure the integrity of the software, to create passwords which you can only use once, to create digital signatures and it has also the function of enabling many internet communication protocols. [7]

The characteristics of a hash cryptographic function

- it is a one way function which means it is difficult to reverse it.
- it can easily face collisions/ collision resistant (it is difficult to find messages which generate the same type of hash)

The hash function may be regarded as a surjective function (because there is a possibility, however remote, to find n messages for the same hash). It is not a bijective function because this would mean that:

- it could be reversed, which is completely wrong
- there is a correspondence 1 to 1.

Not all the hash functions are proper to cryptography. An example of a hash function which cannot be regarded as a cryptographic function is a function which gives you the sum of a series of numbers. This function is not cryptographic, first of all because we can easily find two numbers for which the function may give us the same result (for ex. 23 and 50) and secondly because it has no fixed length [8].

An example of daily used hash function is the algorithm used to determine the Control number of the Personal Numeric Code.

The resume function (digest function) is mainly used for checking the integrity of a file. For example, if you have downloaded software from a website, how can you check if you have downloaded it without having been altered? One method would be to download the file once again and to compare the number of bits: if the number of bits is the same that means that the file has been downloaded correctly, but, if they differ, which binary figures are the right ones? In order to find this out a third download may be made just for comparison. However, this method is not very efficient. But if the site published the hash values of the download you would be able to check them yourself.

The goal of hash functions is not to ensure the top secret character of transmissions, but to create a value h=H (M), called also resume (digest), extremely hard to forge. The role of hash function is not to secure the data, but to create a value h=H(M) called digest, very difficult to forge. All of these functions have an important role in the digital signature. The hash function which has been mostly used till recently is MD5.

MD5 developed by Ron Rivest has been designed because specialists have considered that previous MD4 algorithm has been on top of cryptographic attacks. MD5 goes a little back in time, giving up a little bit of speed (comparing to MD4) in favor of a greater security.

MD5 algorithm has been made public because of possible adjustments and for a possible acceptance as a standard procedure. This algorithm is mainly designed for the application of digital signature, where a file must be safely compressed before being encrypted using a secret key through a cryptographic system with a public key.

5 Application

In order to obtain a considerable degree of security of the transferred data through an uncertain environment such as computer networks, especially through internet, we have developed a program which applies a watermark to a file (data set), in this case an image. Then for the obtained image a new digest will be calculated by applying the MD 5 algorithm.

In order to reach the above mentioned goals JAVA has been chosen as the environment programming language, because it has been designed for a complex environment such as the internet. Java can work on distributed and heterogeneous platforms, it is neutral from an architectural point of view, it is compatible with network environment, the

compiler and the run system make the programs run faster, java is a dynamic programming language,.

As it has been mentioned in the previous sentences there are several methods to erase or replace a watermark from a digital image with other types of watermarks. In order to increase the security of electronic data against deleting or replacing the data, we have designed an application which marks an image through a watermark backed up by a digest function which uses the MD5 algorithm.

The first step consists of loading the image. Next the image is marked with a watermark. The user can select witch watermark to use from the main menu of the program. In the pictures below, it is presented a transparent watermark and a white t-shirt .Then the hash code is processed from the new image. One can see that for a different watermark a different digest code is obtained. The algorithm of the digest function may be applied to the unmarked image as well.

69b1c8a4daf207de3a673266babc54b5

Fig. 3. The loaded image – unmarked and the digest code obtained after having applied the digest function

bdefec04fdace43361f8dc298b30c8f0

93e343b57757265dc2c34bb7e91d9555

Fig. 4. The loaded image – marked with two different types of watermark and the digest code obtained after having applied the digest function

In the following lines a part of the main code is presented in order to visualize the application

of the watermark and the implementation of the MD5 algorithm. The following code is used for watermark:

```
  //Se pastraza imaginea initiala in tabloul pixeli si se foloseste in calcule
tabloul pix
                for(i=0;i<latime*inaltime;i++)
                   pix[i]=pixeli[i];
                //Se descompune tabloul pixeli in culorile componente
                for(i=0;i<inaltime;i++)
                      for(int j=0;j<latime;j++){
                         int alfa=(pix[i*latime+j]>>24) & 0xff;
                         valoareAlfa[i*latime+j]=alfa;
                            int rosu=(pix[i*latime+j]>>16) & 0xff;
                            culoareRosie[i*latime+j]=rosu;
                            int verde=(pix[i*latime+j]>>8) & 0xff;
                            culoareVerde[i*latime+j]=verde;
                            int albastru=pix[i*latime+j] & 0xff;
                            culoareAlbastra[i*latime+j]=albastru; }
                imagineNoua=imagine;
```

The code for MD5 algorithm is [10]:

```
public String md5suma(int[] p,int lat,int in) {//metoda care foloseste algoritmul
MD5
        //Se transforma tabloul de numere intregi p[] in sir de caractere (String)
       String sir=Integer.toString(p[0]);
       for(int j=1;j<lat*in;j++)
           sir=sir+Integer.toString(p[j]);
         byte[] md5hash=new byte[32];
       try{
           MessageDigest md=MessageDigest.getInstance("MD5","SUN");
            //se creaza o instanta a clasei MessageDigest pentru algoritmul MD5 al
firmei SUN
           md.update(sir.getBytes("iso-8859-1"),0,sir.length());
            //se calculeaza rezumatul sirului sir
           md5hash=md.digest();
       }catch(NoSuchProviderException nspe){
       }catch(NoSuchAlgorithmException nsae){
       }catch(UnsupportedEncodingException uee){
       }
        //se converteste sirul de bytes la sir de caractere(String) cu metoda
conversieHex()
       return conversieHex(md5hash);
    }
    public String conversieHex(byte[] data){
       StringBuffer buf=new StringBuffer();
       for(int i=0;i<data.length;i++){
           int jumByte=(data[i]>>4) & 0x0F;
           int dJum=0;
           do{
              if((0<=jumByte) && (jumByte<=9))
                  buf.append((char)('0'+jumByte));
              else
                  buf.append((char)('a'+(jumByte-10)));
              jumByte=data[i] & 0x0F;
           }while(dJum++<1);
       }
       return buf.toString();}}
```

6 Conclusions

Watermarking has received a great deal of attention from the specialists and many applications have been developed. The watermarking techniques may be classified by their role. There are watermarks used for identifying the author, the fingerprint watermarking used to identify a buyer of a certain multimedia product or to identify the validity concerning the integrity of that

product.
In the domain of watermarking a lot of research is needed in the future, reviewed analysis and publications so that as many people as possible may become familiarized with this concept.
The development of a marking technique takes into consideration several aspects. Many of the existent techniques borrowed concepts from other research fields.
The watermarking technique must take into consideration a series of elements, such as cryptography, authentication, steganography, human perception, data fusion, and communication through unsecure channels. The choice of these elements for a marking method or a given application is still in a blur because of certain needs. Finally a compromise must be reached between imperceptibility, robustness, and complexity. We propose a new solution of digital watermarking which integrates a robust hash function for the improvement of the integrity of electronic data and of the transmission security of digital data through collaborative activities.
The proposed solution is still in work and we must focus on the form of our digital watermarking, and on the optimization of the speed of the hash function algorithm.

References

[1] M. Arnold, M. Schmucker and S. Wolthusen, *Techniques and Applications of Digital Watermarking and Content Protection*, Artech House, INC., 2003.
[2] A. H. Paquet and R. K. Ward, "Wavelet-Based Digital Watermarking for Image Authentication," Accepted for Publication at the *IEEE Canadian Conference on Acoustics Speech and Signal Processing (ICASSP)*, 2002.
[3] C. Nafornita, *Digital Watermarking in the Wavelet Domain*, Editura Politehnica, 2005.
[4] J. B. Knudsen, *Java Cryptography*, O'Reilly, 1998.
[5] I. J. Cox , M. L. Miller, J.A. Bloom, J. Fridrich and T. Kalker, *Digital Watermarking and Steganography*.
[6] http://ro.wikipedia.org/wiki/Func%C8%9Bie_hash
[7] http://www.securizare.ro/content/view/291/38/
[8] http://hackpedia.info/viewtopic.php?f=20&t=6805
[9] http://cs.unitbv.ro/~costel/secupdfs/3_3_0_Valori%20Hash%20si%20IntegritateaDatelor
[10]http://download.oracle.com/docs/cd/E17476_01/javase/1.4.2/docs/guide/security/jce/JCERefGuide.html

Dominic BUCERZAN (b. May 17, 1956) received his M. Sc. in Information Technology from "Aurel Vlaicu" University of Arad, Romania and a PhD in Economic Cybernetics from the "Bucharest Academy of Economic Studies" (2005), with a paper in the field of Information Security. Currently he works as a lecturer in informatics at the Department of Mathematics-Informatics, Faculty of Exact Sciences, "Aurel Vlaicu" University of Arad, Romania. His current research interests include aspects of IT Security and Cryptography. He is author or co-author of 4 books and more than 45 papers and participated in 35 conferences and workshops.

Crina RATIU (b. October 23, 1983) received her Master of Science in Information Technology (2008) from "Aurel Vlaicu" University of Arad, Romania. At present she is a candidate for a PhD in Business Informatics at "Babes-Bolyai" University of Cluj-Napoca , Romania. Her current research is focused on IT Security in Business Informatics Systems. She published articles in the above mentioned field of interest.