Analysis and Strategies for Fighting Against Attacks on Watermark Systems

Jovan Spasik[1], Sofija Bogdanova[2] and Dimitar Taskovski[3]

Abstract - **In this paper we analyze the basic watermark characteristics and the strategies for fighting against attacks on watermark systems. This analysis shows that security of the watermark system depends on the watermark application and from information about watermarking system to which attacker can access.**

Keywords – **Analysis and strategies, watermarking system, watermarking applications, watermark characteristics.**

I. Introduction

Watermarking is a technology that helps the users to protect their intellectual and property rights via embedding invisible information directly into the carrier data. This information, which uniquely identifies true owners, is called watermark. Unfortunately, along with development of process of watermarking, there is an expansion of forgery and illegitimate use of digital data. The attackers represent the biggest problem because they are aiming to destroy the embedded watermark or to make it unreadable for watermark detector. That is the reason why there is a need for some kind of protection of intellectual property materials.

There are many papers in literature of signal processing which discuss about fighting against watermark attacks. The main problem is that these papers offer solutions for fighting only against specific attack or group of attacks. These solutions are not eligible against other attacks, so embedded watermarks remain unprotected.

Until now, nobody presented a complete strategy, which will be adequate for all possible attacks. These facts were incentive for writing this paper in which analyzes of known watermarking systems and watermark characteristic are made. This paper aim to show that robustness of watermarks depends from used watermarking application and its security. For fighting against different attacks according to different watermarking systems, different strategies should be recommended. Also, the paper should signify the negative characteristics and to note what the embedder must do in order to make more robust watermark.

The paper is organized in following way. Basic watermarking applications, watermark characteristic and comparative analysis are presented in the next chapter. In chapter 3 is presented classification of watermarking systems together with strategies for fighting against watermark attacks. This strategies are organized according to information to which attacker can access.

II. Watermarking Applications and Watermark Characteristic

One of the oldest applications of watermarking, or more precisely data hiding, is "secret communication". Today the watermarking can be used in variety of applications like "broadcast monitoring", "owner identification", "proof of ownership", "authentication", "fingerprinting" and "copy control"[2]. Common watermark characteristics like robustness, tamper resistance and computational cost are application dependent. In practice, it is probably impossible to design a watermarking system that satisfies all of these, so it is necessary to make tradeoffs between them that must be chosen with careful analysis of the application. Further, we present some basic watermark characteristic and strategies for choosing the right watermark in order to achieve right functionality.

"Robustness" is characteristic that demonstrate how resistant the watermark is against common signal processing applied on carrier data. In common, the watermark must be robust only in the period between processes of embedding and detection. On contrary, characteristic that shows how much watermark system is robust against hostile attacks is called "tamper resistance".

Table 1 represents sublimation of most common watermark applications and characteristics together with effects from different types of attacks on individual application.

The goal of **removal attacks** is to remove embedded watermark or to make it undetectable for watermark detector. **Collusion attacks** are important subclass of removal attacks.

The goal of collusion attacker is to create unmarked carrier data, by collecting many copies of the same carrier data marked with different watermarks [8].

Legend: C – Critical NC- not Critical R – robust F - Fragile	**Characteristics**			**Type of attacks**		
Application	Robustness	Speed	Quantity	Removal	Collusion	Forgery
1. Secret communication	F	slow	-	N	N	N
2. Broadcast monitoring	R	fast	+	C	N	N
3. Owner identification	R	slow	-	C	N	N
4. Proof of ownership	R	slow	-	C	N	C
5. Authentication	F	med	-	N	N	C
6. Fingerprinting	R	slow	-	C	C	N
7. Copy control	R	fast	+	C	N	N

Table 1: Analysis of watermarking applications, their characteristic and attacks influence

[1] Jovan Spasik, [2] Sofija Bogdanova and [3] Dimitar Taskovski are with the Faculty of Electrical Engineering and information Technologies, University "Ss. Cyril and Methodius", Skopje, R.Macedonia

In contrary, in **forgery attack** group, it is typical for the attacker not to remove the watermark, but to insert another valid watermark in order to fake the watermark detector.

“Computational cost” is characteristic which can be represented as speed of the watermarking system and used devices (embedders and detectors) needed for accomplishing the goal of application.

Conclusion to this chapter is that different set of standard should be used, according to the application for which watermark is created. If users act in accordance with this principle, they can achieve better watermark performances. Along with this, the attacks against watermark will have less success, or the overall performance of the watermarking system will be better.

III. CLASSIFICATION OF WATERMARKING SYSTEMS AND STRATEGIES AGAINST WATERMARK ATTACKS

Analyzing and focusing on the information to which attacker can access are one of the main strategies for designing and classifying security watermarking system [1]. According to this approach, we can assume that for example, attacker does or does not have access to the algorithms used for watermark embedding or that attacker is fair or un-fair player. Fair attacker is attacker that uses only publicly accessible information for his attacks. Unfair attacker can use observations for revealing the secret keys for embedding and detection or any information about functioning of the watermarking system.

In Fig.1 is presented “effort to deal with attacks” against “information to be kept secret” diagram. Here we can see that as “information to be kept secret” increases “the effort to deal with attacks” increases for un-fair attacks and decreases for fair attacks. Keeping in mind this notes, we can make analysis of watermark security with respect to public information ***P*** to which attacker can access, where a is detection algorithm, k_D is detection key and k_E is key used for embedding.

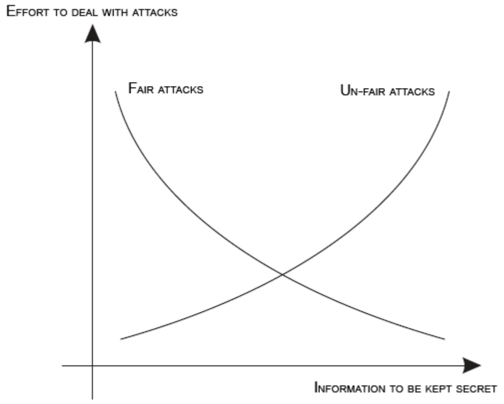

Figure 1: The diagram of “Effort to deal with attacks” versus “Information to be kept secret” in respect to fair and un-fair attacks

According to this information, the following classification can be made:

1. No public information is known: P = 0;
2. Embedding and detection algorithm are public P = {a};
3. Everything except embedding key is public P = {a, k_D};
4. Embedding and detection keys are public P = {a, k_D, k_E}

Scenario:	**Public information**			**Point of view:**	
	a	***k_E***	***k_D***	**Attacker**	**Owner**
1. Security-by-obscurity	no	no	no	Focused on un-fair attacks	To many secret information
2. Symmetric watermarking	yes	no	no	Balancing between fair and un-fair attacks	It is very hard to keep k_D in secret
3. Asymmetric watermarking	yes	no	yes	It is better to focus on fair attacks	Fair attacks are very dangerous
4. Playing with open cards	yes	yes	yes	Powerful fair attack exist	Nothing but hope

Table 2: Classification of watermarking systems in respect to information publicly accessible by attacker

3.1 Security by obscurity

This is the scenario where no public information about the watermarking system is known. In this scenario, designing of fair attack can be a very hard task. For that reason, the choice of attacker is to concentrate on un-fair attacks, because there could be a possibility some secret information about the watermarking system to outflow, thanks to the observations.

The watermarker’s point of view*:* This strategy was common in the beginning of watermarking’s history. The main concept was based on the fact that if all information about watermarking system is kept secret, then watermarking would definitely be secure. This was the main reason why research efforts were focused only at robustness requirements and watermarking security was neglected. Unfortunately, it is very hard to keep the watermarking algorithm secret and security-by-obscurity cannot fight against un-fair attacker.

In cryptography, the people are more aware about leaking of information. In 1883, Kerckof presented the basic rules in cryptography [4], where he stated that the cryptographic systems’ designers must be aware that the attacker knows their algorithm in details except the secret key, on which the security of cryptographic systems has to be based. The Kerckof’s basic rules are some kind of a warning about unsafeness of security-by-obscurity scenario. These rules are foundation of cryptography analysis and also foundation of security analysis of watermarking systems.

The attacker’s point of view: Because in this scenario there is no public information available, fair attacker tries to trick the watermark decoder with some transformation of contents of cover data. This is the case of direct confrontation of robustness of watermark and intentional signal processing. Main concern for watermarking systems represents geometrical distortion or geometrical attacks. Possible geometrical counterattacks are for example: embedding of template or extra signal used for synchronization of watermark embedder and detector [13]; embedding of the watermark signal in invariant domains [12]; introduction of redundancy in the watermark signal in order to reduce the space of potential delays; or using of self-registration of the image [10]. From day to day, robustness watermarking techniques become more powerful. If the fair attacker wants to remove the watermark signal, he must distort the cover signal to very low level. We may conclude that in the end the robustness will fight the fair attacker, at least in the context of given application, so in this scenario the only possible solution for attacker is to use un-fair strategies which will reveal algorithms and keys used for watermarking process. On the other side, the watermark owners are instructed from Kirckof's rules that security of watermarking system can not be protected only by obscurity. According to this, we can conclude that this scenario is not stable state in the table 2.

3.2 Symmetrical watermarking

Under presumption that only embedding and detection keys are secret, in this scenario the fair attacker tries to remove or to make the watermark undetectable based only on a priori information about embedding and detection algorithm. On contrary, un-fair attacker tries to detect the embedded watermark and based on this a posteriori information, to remove or to make watermark undetectable.

The watermarker's point of view: The first concern of the watermarker is choosing a watermark technique that has to be robust against common transformation of watermark content, which attacker can use in current application. Theoretically, it is possible to find watermarking technique, which is robust against common transformation allowed for particular application. On the other side, the attacker can use attacks that are more sophisticated because watermark extraction functions are public and he can act in the embedding domain. More sophisticated attacks are based on noise filtering or more specially Wiener filtering which can be use to separate the watermark from cover signal. Possible countermeasure proposed by Su is satisfaction of Power Spectrum Condition [7] which states that Spectral power density of watermark should be shaped according to Spectral power density of features vector. Another possibility proposed by Le Guelvouit is the owner first to embed watermark and then to attack watermarked signal with Wiener filter [9]. In this way, the efficiency of filters for removing the watermark is diminished. As second, the watermarker must be sure that watermarking systems implemented in public electronic devices can not be hacked. Security followed by rules from this section is possible only if there is a method for securing the keys.

The attacker's point of view: In this scenario attacker assumes that watermarker did a great job and that he has to deal with robust watermark. Because of that, the attacker strategy is to reject the rules of the game. Revealing of the secret keys helps the attacker to forgery watermarked content with very low level of distortion. The keys are used for decoding the secret messages. When the attacker has the keys, the data which has to be embedded and the watermark content, the synthesis of the watermark signal and its subtraction from the watermarked content are considerably easier. It should be noted that this simple attack is not working with models with quantization index, because in that case the embedded signal depends on the cover signal. Similar algorithm for modulation schemes with quantization index is presented in [3]. In this case the carrier signal is more degraded.

3.3. Asymmetric watermarking

In this scenario, both the algorithm **a** and detection key $\mathbf{k}_\mathbf{D}$ are public information. The security of embedding key is essential for security of the watermarking system. This concept may sound strange but it is reality in cryptography where, the digital signs are based on asymmetry. The key used for embedding and related verifying key are different, but with verifying key we can check the data which was signed with private key.

If we like to use this concept in watermarking, the processes of embedding and detection should be asymmetrical and they must be based on different keys. This is the reason why the public detection watermarking is usually referred as asymmetrical watermarking.

The attacker's point of view: The asymmetry of embedding and detection keys is not sufficient to offer security in watermarking system. However, asymmetry does not imply robust public key watermarking. So far, all known asymmetrical systems where detection key $\boldsymbol{k}_\mathrm{D}$ is public, are hacked. In paper [11] is presented an attack which is valid for almost any known asymmetrical system. As conclusion, we can say that even if asymmetry is helpful, it is far from being adequate to ensure security in a public detection system.

The watermarker's point of view: The knowledge of detection algorithm and key for detection involve knowledge of boundaries of detection region. That is the main reason why asymmetrical systems are not safe methods for detection with public key. In scenarios where attacker knows the boundaries of detection regions, the "closes point" attacks presents deadlock. These closest-point attacks can be prevented only by using sufficiently sophisticated detectors, irrelevant from the methods used for embedding [5] or function that indicates detection region but not reveals boundaries of this region. The function that conveys this principle is fractal function. Unfortunately, until now there is no way to build watermark system that satisfies this function, but this is a good starting point that yields to good characteristic which can be used in public key detection watermarking. The attacker must test every point of the space in order to reveal the boundary of detection region and that process can last extremely long. As addition, the watermarker must verify that there is no other way for revealing the boundary and to be sure, that knowledge of detection keys will not bring any a priori information about embedding key.

3.4 Playing with open cards

The principle presented in previous scenario reduces the amount of information that owner should keep in secret. That yields toward reducing the effort needed for watermarker to resist against un-fair attacks and against stronger fair attacks. At the end, this process will bring the situation in which all information is public. In that situation, the attacker can access to information about detection and embedding keys used by watermarking system.

The watermarker's point of view: It is clear that playing with open cards scenario is most appropriate for un-fair attackers. In this scenario, the watermarker does not care about protection of information because all information is public. The only question is: Is robustness possible against fair attacker?

In fair version of the "closest point" attack, the attacker knows the boundary of detection region and according to this, watermark can be made unreadable by simply moving the host data in closest point of un-detectable region.

Un-fair version of "closest point" attack differs from its fair version only in determination of detection region. The un-fair attackers estimate boundaries of detection regions through error and trial procedures.

The following chapter presents possibility asymmetric watermarking schemes to give solution for fighting the "closest point" attack.

The knowing of boundaries of detection region is not helpful for the attacker, if detection region is defined in such complicated manner, that moving the point inside and outside from it with respect to visual distortion, would be computational impossible process. Unfortunately, the usage of this complicate detection region makes watermark embedding process to be very complex. Here, asymmetrical watermarking can help the watermarker, by offering him a simple description for detection region. According to this scenario, the needed asymmetry between embedding and detection of watermark (which represent entering and exiting the region) can be accomplish by offering the different description of detection region (different key) to the watermarker and to the attacker as shown in Figure2.

If it is possible to design detection region for which it is very simple to move point inside and very hard to move point outside the region, with acceptable visible distortions, then there is no need to make different keys for embedding and detection and there is no need for keeping them secret too. With other words, the watermarker can play with open cards. It is obvious that designing such detection region is almost impossible, but until we get explicit answer for that, the scenario for playing with open card can not be ignored.

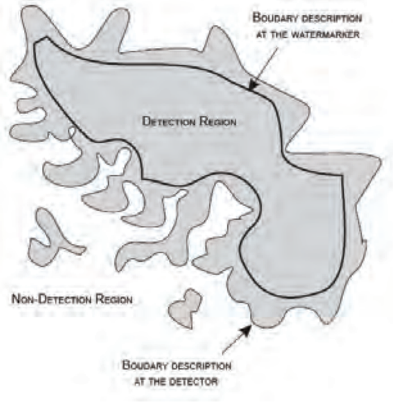
Boudary description at the watermarker
Detection Region
Non-Detection Region
Boudary description at the detector

Figure 2: Asymmetrical watermarking; Comparison of boundaries descriptions offered to watermarker (simple) and attacker (complex)

The attacker's point of view: Scenario for playing with open cards is favorable for attacker's point of view because in this case attackers can easily perform lots of different attacks. However, if someone succeeds to design a basic asymmetrical region with features described in previous chapter, the efforts of attacker will become hopeless, because in such situation he will not have possibility to use un-fair attacks.

IV. CONCLUSION:

No matter how robust the watermark can be against one attack, it can be very fragile against another. We demonstrate that for successful fight against attacks on embedded watermarks, it is extremely important to understand the process of watermarking. The lack of understanding the functionality of used watermarking technique is common mistake, which produces weak watermarking systems and un-robust watermarks. This paper clarifies that the watermarker should make good strategy, which includes securing the watermarking system, used techniques and watermarks itself. Security of the watermark system depends on the watermark application in which watermark should be used and from information about watermarking system to which attacker can access. As conclusion, the embedder must have knowledge of all the previous mentioned, in order to design more robust watermarking system.

REFERENCES

[1] M. Barni, F. Bartolini, T. Furon, "A general framework for robust watermarking security", Signal Processing 83, p.2069–2084, 2003.

[2] I. J. Cox, M. L. Miller and J. A. Bloom, "Watermarking applications and their properties", Published in the Int. Conf. on Information Technology'2000, Las Vegas, 2000.

[3] J. Eggers, B. Girod, "Informed Watermarking", Kluwer Academic Publishers, Dordrecht, 2002.

[4] A. Kerckhoffs, "La cryptographie militaire", J. Sci. Militaires 9, p.5–38, 1883.

[5] M. L. Miller, "Is asymmetric watermarking necessary or sufficient?", in: Proceedings of the European Signal Processing Conference - EUSIPCO 2002

[7] J. Su, J. Eggers, B. Girod, "Analysis of digital watermarks subjected to optimum linear filtering and additive noise", Signal Process 81, p.1141–1175, 2001.

[8] I. Cox, J. Kilian, F. T. Leighton and T. Shamoon, "Secure spread spectrum watermarking for multimedia", IEEE Trans. on Image Processing, 6(12):1673–1687, 1997.

[9] S. Pateux, G. Le Guelvouit, C. Guillemot, "Perceptual watermarking of non i.i.d. signals based on wide spread spectrum using side information", Proceedings of the IEEE International Conference on Image Processing, Vol. 3, Rochester, NY, USA, 2002.

[10] P. Bas, B. Macq, "A new video-object watermarking scheme robust to object manipulation", in: Proceedings of the International Conference on Image Processing, IEEE, Thessaloniki, Greece, 2001.

[11] T. Furon, I. Venturini, P. Duhamel, "Unified approach of asymmetric watermarking schemes", in: P.W. Wong, E. Delp (Eds.), Security and Watermarking of Multimedia Contents III, SPIE, San Jose, CA, USA, 2001.

[12] J. O'Ruanaidh, T. Pun, "Rotation, scale and translation invariant spread spectrum digital image watermarking, Special issue on copyright protection and control", Signal Processing 66(3): 303–317, 1998.

[13] S. Pereira, T. Pun, "Fast robust template matching for affine resistant image watermarks", in: A. Pfitzmann (Ed.), Proceedings of the Third International Workshop on Information Hiding, Springer, Dresden, Germany, pp. 199–210, 1999.