

Secure transmission of visual image in the VTS system using fingerprinting

Bezpieczne przesyłanie obrazu wizyjnego w systemach VTS z wykorzystaniem fingerprintingu

Bartosz Czaplewski¹, Krzysztof Czaplewski²

¹ Gdansk University of Technology
Politechnika Gdańska
80-233 Gdańsk, ul. Narutowicza 11/12, e-mail: bartosz.czaplewski@eti.pg.gda.pl,

² Polish Naval Academy
Akademia Marynarki Wojennej
81-103 Gdynia, ul. Śmidowicza 69, e-mail: krzysztof@czaplewski.pl

Key words: fingercasting, encryption, JFD, video, piracy, VTS, navigation

Abstract

In order to maintain a high level of security in areas where vessel traffic services operate, there is a great use of industrial cameras, which transmit information to the VTS controllers. Two very important elements of the video image transmission are speed and confidentiality of the transmitted data. Currently, one of the major problems in front of such systems is to choose an appropriate method for safe image transmission that meets the requirements of speed and confidentiality. There is the possibility of using the tools of computer science known as fingerprints.

There are two ways to protect multimedia copyright. First way is to encrypt the data. Encryption provides that only privileged receivers with correct decryption keys will be able to decrypt information. Unfortunately, encryption isn't sufficient protection, because after decryption, privileged user may easily violate copyright by sharing decrypted data with public. Second way is called a digital fingerprinting, which rely on embedding of some additional binary sequence into image. This sequence, called a fingerprint, is unique for each receiver and it's unnoticeable for the human eye. With fingerprinting there is a possibility of further analysis of a intercepted copy which is suspected of being illegally shared.

The paper presents the theoretical basis for the use of fingerprinting methods to protect the media sent in VTS systems, as well as a description of the most promising methods, in terms of safety and cost, called Joint Fingerprinting and Decryption methods. Approach proposed by the authors can enhance the security of the video image transmission in VTS systems.

Słowa kluczowe: fingercasting, szyfrowanie, JFD, wideo, piractwo, VTS, nawigacja morska

Abstrakt

W celu utrzymania wysokiego poziomu bezpieczeństwa na obszarach objętych systemami nadzoru ruchu statków coraz częściej mają zastosowanie podsystemy kamer przemysłowych przesyłających informacje do kontrolerów systemów VTS. Jednym z bardzo ważnych elementów transmisji obrazu wizyjnego jest jego szybkość i poufność transmisji danych w trybie on-line. Aktualnie jednym z najważniejszych problemów determinujących powszechne wykorzystanie tego typu systemów jest dobranie odpowiedniego sposobu bezpiecznej transmisji obrazu wizyjnego, który spełni wymagania na szybkość i poufność. Można w tym celu wykorzystywać znane w naukach informatycznych narzędzia.

Istnieją dwie uzupełniające się metody ochrony zawartości multimediów oraz praw autorskich. Pierwszą metodą jest szyfrowanie, która zapewnia, że tylko zarejestrowani użytkownicy, posiadający odpowiednie klucze deszyfrujące, będą w stanie odszyfrować przesyłane treści multimedialne. Niestety, szyfrowanie nie jest wystarczającym zabezpieczeniem, gdyż po deszyfracji użytkownik mający dostęp do multimediów, może je ponownie udostępnić bez zgody autora, łamiąc tym samym prawa autorskie. Drugą metodą jest cyfrowy odcisk palca (ang. *digital fingerprinting*), polegający na osadzeniu dodatkowych, ukrytych sekwencji binarnych

w treści multimedialnej. Dane te, nazywane fingerprintami, jednoznacznie identyfikują stronę odbiorczą, a osadzone są w taki sposób, aby pozostały niezauważalne dla ludzkiego oka. Osadzone dodatkowe dane dają możliwość późniejszej analizy przechwyconej kopii podejrzanej o to, że jest bezprawnie udostępniona.

W referacie zostały przedstawione podstawy teoretyczne wykorzystania metody fingerprintingu do ochrony plików multimedialnych transmitowanych w systemach VTS, jak również opis najbardziej obiecujących metod pod względem bezpieczeństwa i kosztów, nazywanych metodami Łączonego Fingerprintingu i Deszyfracji (ang. *Joint Fingerprinting and Decryption*). Proponowane przez autorów podejście może zwiększyć poziom bezpieczeństwa transmisji obrazu wizyjnego w systemach VTS.

Introduction

In recent years, the popularity of services such as video on demand and Internet TV has greatly increased. Multimedia content is now transmitted in all possible ways, depending on their purpose, which currently range from entertainment to business applications. A similar phenomenon can be observed in traffic supervision centers, which are constantly monitoring the movement of vessels in coastal waters with the use of CCTV (Closed-Circuit Television) industrial cameras.

The creators and publishers of the multimedia content are copyright holders. This is a set of rules authorizing the owner to decide on the conditions of use of his intellectual property. Breaking these rules can result in loss of both financial and moral in case of author of multimedia content. In the case of disclosure or fabrication of the visual image, used by VTS systems, can result in life threatening. This means that there must be mechanism to ensure copyright protection for content available on the web. The growing interest of customers led to the creation of Digital Right Management, which is developing dynamically. Using the existing knowledge and methods for the DRM can be very convenient and efficient in VTS systems. Visual image transmitted in real time allows for immediate response by the marine traffic controller, but this image can be trusted only if it is accessible only by authorized service, which can be done physically, or in software. Physical security, which would involve permanent monitoring of the transmission line, would be expensive and limited to the wired network. Therefore, the approach proposed in the article could be very useful, because a network used for transmitting secured data can be both wired and wireless and maintenance costs would be much smaller.

Copyright protection is using two complementary methods of protection, such as encryption and digital fingerprinting [1]. Encryption provides data confidentiality, which means that it is not possible to capture the copies by third parties. Unfortunately, this is not sufficient, because the encryption does not protect against re-releasing a decrypted

copy by a dishonest user. In this case, additional protection is needed, which involves sending copies marked with digital fingerprints. Later analysis of the embedded fingerprint allows to identify the user who illegally shared a protected data. An identified pirate can be used if the evidence provided by the method are unquestionable or may be subjected to further observation, in order to obtain more information about him. These methods are called digital fingerprinting and it is easy to see that copyright protection is achieved not through preventing, but through counteracting the piracy.

It has to be assumed that the pirates are aware of the presence of fingerprints in their copies and they will attempt to remove or heavily damage the embedded fingerprints. Threats to the fingerprints can be divided into three groups: signal processing [2], which is usually unintentional way to remove or distortion of fingerprints, single-pirate attacks [2], such as geometric distortions and collusion attacks [1, 3, 4], which are the most dangerous. Signal processing, which is common in distribution of video images, can be a source of unintentional distortion or even removal of fingerprint. The problem may be a lossy compression, resampling, requantization, contrast and color enhancements. A single pirate can try to remove fingerprint in a pixel domain by using geometric distortions such as shrinking the image or cutting out pieces of the image. However, the greatest threat comes from the collusion attacks. In these attacks, organized groups of pirates analyze their marked copies of the same content and on the basis of this analysis they generate a pirated copy, which is free from fingerprint or contains damaged fingerprint, which doesn't identify real colluders. Methods must have a high resistance to collusion attacks and other attempts to remove the fingerprint.

Note that the fingerprinting system must operate in a network environment, where the number of customers can reach thousands. For suppliers of these services, it is important that the cost of transmitting a single copy is as low as possible. Thus, in order to maximize profits, suppliers have to handle the maximum number of clients with limited available bandwidth. Multicast transmission is perfectly



suitable for media distribution on a large scale because of its scalability. However, the implementation of fingerprinting system based on this type of transmission is problematic [5]. This is because the objective and purpose of fingerprinting and multicast are orthogonal to each other. Fingerprinting systems require that each user obtains a different, uniquely marked copy of the data, while the multicast transmission allows for efficient transfer of one, exactly the same, content to all users. For this reason, multicast cannot be applied directly to fingerprinting systems, because the uniqueness of each copy won't be preserved. This means that there is a need to design new fingerprinting methods matching characteristics of multicast transmission – a fingerprinting methods.

Of course, the use of the fingerprinting without encryption, or encryption without fingerprinting is pointless. Therefore, in any media distribution system, it is necessary to use both of these types of protection.

Classical Approach

Figure 1 shows a schematic of classic fingerprinting method in order to detect the users who violate copyrights. Operation of the fingerprinting system can be divided into three stages: fingerprint embedding, survival of an attack, identification of pirates.

In the first stage, distribution side needs to create a set of labels that are unique to each user and gather them in a database. These labels are called fingerprints, because they identify individual receivers, and their presence in the pirates copy will be evidence in case of attack. Then fingerprints are embedded in the multimedia data in such a way as to be difficult to remove and not to cause large changes in the data.

Multimedia content ordered by k users is represented by vector x . In order to create a marked copy of the content, distribution side needs a set of unique fingerprints $\{f_i\}$, where $i = 1, 2, \dots, k$. These

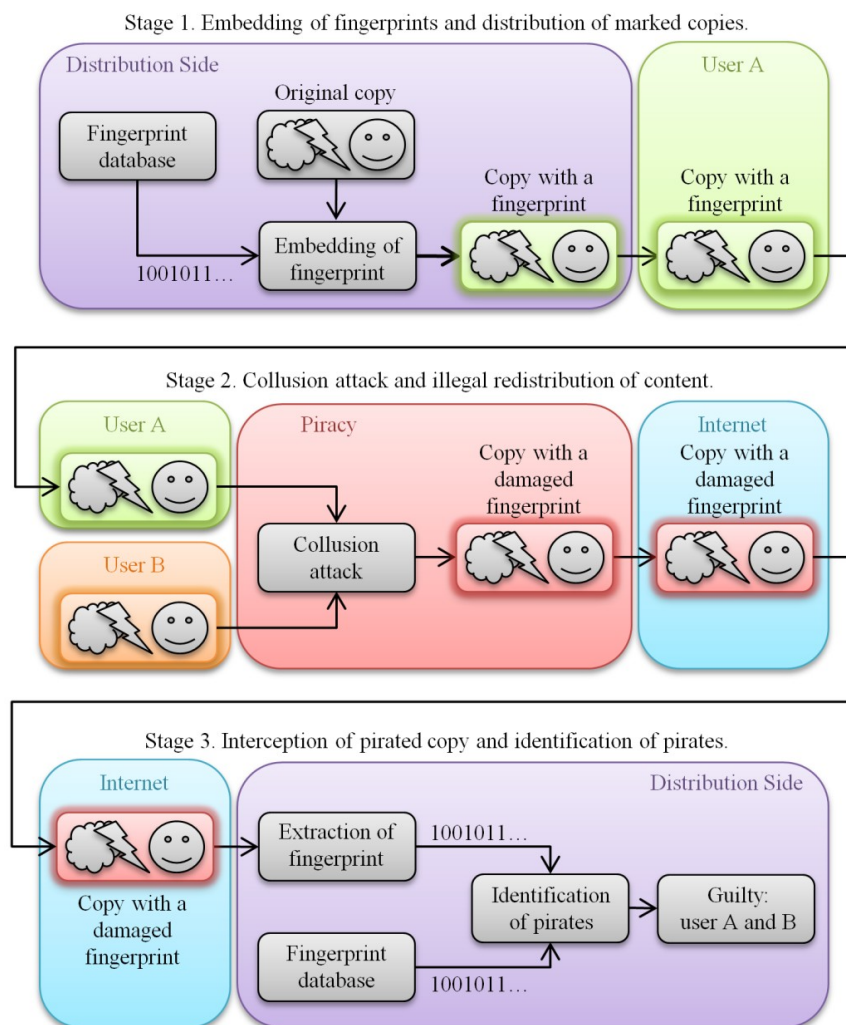


Fig. 1. General scheme of the fingerprinting system [own study]

Rys. 1. Ogólny schemat systemu fingerprintingu [opracowanie własne]

fingerprints have to be assigned to individual users. The greatest efficiency can be achieved through the use of fingerprints, which are orthogonal to each other. Unfortunately, the number of orthogonal sequences of fixed length are limited, which means the use of orthogonal fingerprints is possible only for a relatively small group of users. The use of non-orthogonal fingerprints lowers the efficiency of the system, but allows to provide service for a much larger number of users.

After generating a set of fingerprints $\{f_i\}$, it is necessary to scale them so when added to the original signal x , they will be imperceptible to the human eye. For this purpose, fingerprints are multiplied by α coefficient, which is the strength of embedding and it is associated with a just noticeable difference for human perception. Coefficient α is particularly important, since the appropriate selection of its value will bring the maximum resistance of embedded fingerprint while maintaining invisibility. Marking a copy is performed by adding a scaled fingerprint to the original signal:

$$x_i = x + \alpha \cdot f_i \quad (1)$$

where x_i is a marked copy sent to i -th user, x is a original signal, f_i is a fingerprint of i -th user and α is the strength of embedding.

It may happen that one or more users are going to violate the copyright and join together in an attack designed to generate a pirated copy, which is free of fingerprint or contains a damaged fingerprint. This is the second stage of the fingerprinting system, which takes place in a hostile environment of rogue users called pirate. At this stage it is important that embedded fingerprint survives any manipulation and will be present in the pirated copy. Marked copy is exposed to a number of intentional or unintentional operations, sometimes related to the attacks. Changes caused by these operations can be described as additive noise, and then a damaged copy can be described by the formula:

$$x'_i = x + \alpha \cdot f_i + z \quad (2)$$

where x'_i is a marked copy received by i -th user, x is a original signal, $\alpha \cdot f_i$ is a scaled fingerprint of i -th user and z is the cumulative changes that have occurred in the image.

The third stage of the system begins when the pirated copy is intercepted. The first step is to extract the fingerprint, which survived the attacks of pirate. Then, the extracted fingerprint is compared to the database, which contains all users fingerprints. An appropriate analysis should allow the identification of users who have violated copyright.

Detectors can be different depending on the fingerprinting method used, but there are two basic types of detection [1]: blind detection, otherwise known as non-coherent detection and non-blind detection, otherwise known as coherent detection.

In the case of non-blind detection, knowledge of the original signal is used. Deduction the original data from a tested copy makes detection much easier because it increases FNR fingerprint-to-noise ratio, which is defined by the formula [6]:

$$\text{FNR} = 10 \log \frac{P_F}{P_Z} \quad (3)$$

where P_F is the average power of fingerprint and P_Z is the average power of interference z . However, in order to exploit the original data, its storage is necessary. Maintaining a database containing all unmarked data requires significant resources and it is expensive. However, it is not a problem if the distribution side performs a detection of pirates.

In the case of blind detection, the original data is not known during detection. This kind of detection is difficult, because original signal is treated as a interference, which has a much higher level than the embedded fingerprint. For blind detection, FNR fingerprint-to-noise ratio is defined by the formula [6]:

$$\text{FNR} = 10 \log \frac{P_F}{P_Z + P_X} \quad (4)$$

where P_F is the average power of fingerprint, P_Z is the average power of interference z and P_X is the average power of the original signal. There is no need to store the unmarked data, which means that the detection can be performed by a trusted third party.

Unfortunately, the classical approach of fingerprinting is not free from defects. First of all, the operation of embedding fingerprints in the multimedia data is performed for each user individually. Distribution side creates as many marked copies as there are registered users. Thus, the number of calculations that would make the distribution side would grow linearly with the number of users, which leads to poor scalability of the system.

Secondly, each of these copies are sent separately to each user. If k customers ordered the same content, k independent streams would be sent over the network. This means that the total bandwidth required to provide the service to all receivers would be multiple of bandwidth required to provide the service to one receiver, which also leads to poor scalability of the system. It should be noted that in this case, there would be multiple data streams sent over the network, and these streams differ very

little because fingerprints are imperceptible. This will lead to a waste of bandwidth.

Thirdly, legally sent copies have to be available only to authorized persons. Thus, it is necessary to use additional encryption algorithm because fingerprinting method doesn't ensure this. Additional algorithm leads to an even greater number of operations performed by a distribution side for each receiver.

It has been tried to overcome these disadvantages by adopting a classical approach to the use of multicast transmission. Many solutions have been proposed, which are very different from each other and based on completely different ideas. In some solution, such as General Fingerprint Multicast Scheme [7], fingerprint embedding is realized in the DCT transform. It uses the fact that in order to maintain the invisibility of fingerprints, not all the DCT coefficients can be used to fingerprint embedding. This means that the coefficients, which can not be used are identical for each marked copy, therefore, they may be sent to all users through multicast transmission. Meanwhile, coefficients with embedded fingerprint are sent to users through unicast transmission. Thus, the required bandwidth is much smaller. However, this method results in much higher computational complexity for both the transmitter and receiver, because both sides must not only encrypt or decrypt data, but also to split or combine two streams: unicast with fingerprint and common multicast. Further problems may arise in case of loss of synchronization between the component streams.

There are other methods, such as Watermarking Multicast with a Hierarchy of Intermediaries [8], based on the concept of fingerprints reflecting the actual location of the user in the network. This method uses a specialized network devices that are able to create and embed fingerprints in the transmitted multimedia data streams. Fragments of fingerprint are progressively embedded in the transmitted data, as the stream is going through further nodes in network. The method ensures high scalability and security, but it has a very big disadvantage. It is required that all network elements have the specific functionality for fingerprinting, which is not fulfilled for most network devices in existing infrastructure. This implies a very large additional costs of implementation.

There is also a solution as in [9], in which two differently marked copies are sent through multicast transmission to all users, but each packet or video frame is encrypted with another key. Users have unique sets of keys for decryption. These sets are created. In such a way that the i -th key decrypts

only one packet or frame of the i -th pair of two received. In this case, the user's fingerprint is a unique combination of decrypted packets or frames of two marked copies. In this method, there is a great scalability, because it uses only two multicast transmissions, but there are two major drawbacks: such methods are very vulnerable to collusion attacks and the main problem is enormous number of keys.

The most promising solutions are methods called Joint Fingerprinting and Decryption. These methods are very different from the classical approach, and thus are free from its disadvantages. The following section is devoted to the JFD methods.

Joint Fingerprinting and Decryption

JFD methods are cryptographic methods, which are the solution to the problem of scalability in fingerprinting systems [6, 10, 11, 12, 13]. These methods not only allow to take full advantage of the multicast channel, but also to move operation of embedding on the user side, thus reducing the number of computations performed by the distribution side, which previously had to perform these operations for each user separately. In addition, JFD method is also an encryption algorithm, which in the classical approach had to be added.

The most characteristic feature of the Joint Fingerprinting and Decryption methods is that fingerprint embedding is performed by the receiver. The general idea is to encrypt multimedia content and send it to all users who put their fingerprints during decryption process. Over the network, only one copy is sent, so JFD methods are suitable for multicast transmission. As a result, the bandwidth requirements are substantially reduced. Also, the computational requirements are much smaller, because only one copy is encrypted and nothing is fingerprinted on the distribution side.

Figure 2 show a general scheme of the JFD methods. Firstly, distribution side encrypts the original data by using an encryption key K_E , which is not known to any user. The obtained result is a ciphertext that is sent over a network using multicast. Then, the ciphertext is decrypted using a unique decryption key in joint fingerprinting and decryption process. Each user has a unique decryption key K_D , which is different from the encryption key K_E . K_D keys are generated by the distribution side based on K_E and fingerprints associated to each user.

Differences between K_E and K_D cause slight differences between the decrypted copy and original copy of the multimedia content. These differences



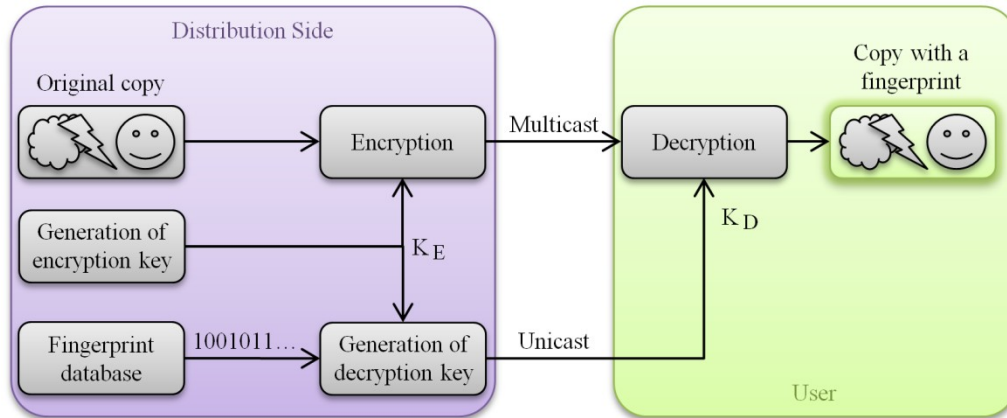


Fig. 2. Fingerprint embedding on the user side using JFD method [own study]

Rys. 2. Osadzanie fingerprintów po stronie użytkownika, korzystając z metody JFD [opracowanie własne]

should be imperceptible to the human eye and unique across all users. Therefore, these changes are user's fingerprint. A user don't have access to the decrypted content without fingerprint because it is embedded during decryption process.

In order to ensure confidentiality of the transmitted data and invisibility of the embedded fingerprints simultaneously, it is necessary to prepare a set of decryption keys S_K . Each user has a decryption key K_D belonging to the set S_K and should be able to decrypt the data sent to him without loss of perceptual quality. However, the use of a false key will result in a distortion of such a level that the media will be unusable. It is described by the following conditions, which must be met with high probability [11]:

$$\text{if } K_{\text{used}} \in S_K \text{ then } p(X_D, X) < \delta_1 \quad (5)$$

$$\text{if } K_{\text{used}} \notin S_K \text{ then } p(X_D, X) > \delta_2 \quad (6)$$

where K_{used} is the key used to decrypt data, S_K is the set of proper decryption keys, $p(X_D, X)$ is a measure of perceptual difference between the decrypted data X_D and original data X , δ_1 is the global masking threshold, which is equal or smaller than just noticeable difference for the human, δ_2 is the threshold of acceptable interference in the content.

JFD methods are perfectly suited for multicast transmissions. Fingerprinting is performed on the receiver side. Therefore, computational requirements for the distribution side are minimal. Only one copy is encrypted and sent using multicast transmission to all registered users. Therefore, bandwidth required for the operation of the system is independent of the number of users. In this methods, the receiver side is not burdened with any additional operations, because fingerprinting is performed during the decryption process. These properties provide high scalability of JFD methods,

which is extremely important in case of a large number of receiver points.

The first time when JFD term was used, it was method based on scrambling of DCT (Discrete Cosine Transform) coefficients [11]. It combines a the Chameleon method [10] and selective video encryption [14]. In this method, the entire video frame is coded using DCT transform. Later, the encoded frame is divided into two sets of elements called essential features and nonessential features. Essential features are the low-frequency components, because they contain most of the signal energy. Encrypting the entire frame would require a large amount of computation. Therefore, in order to reduce computational requirements, only essential features are encrypted and nonessential features remain unchanged. Encryption is based on the fact that set of essential features is divided into n subsets and signs of the coefficients in these subsets are reversed depending on the used key. Each receiver has a unique set of keys, which allow him to decrypt only p of n subsets. The remaining $k = n - p$ subsets are still encrypted and the combination of DCT coefficients, which remain encrypted is the fingerprint.

Unfortunately, there are visible distortions in a fingerprinted image, which reduce image quality and is unacceptable. Also, the encrypted image is not pseudo-random signal and it is possible to see properties of the original data. In addition, this method is not resistant to collusion attacks.

Another method, which is continually being developed, is called the Hillcast and its description can be found in [6, 12, 13]. Hillcast method uses block cipher, based on simple matrix multiplication for encryption, similar to the Hill cipher. A randomly picked fragment of image is encrypted with the group key and sent to all users through multicast transmission. There are two variants of this

method: encryption of pixel values of DCT coefficients. There are also unique decryption keys for each user, which are sent to them through unicast transmission. Differences between the group key and the decryption key cause a fingerprint embedding during the decryption process. The study [6] showed that the embedded fingerprints remain unnoticed to the human eye and the method is resistant to collusion attacks done by up to 30 pirates.

However, in case of Hillcast, it is necessary to provide an additional mechanism to ensure the safety of used historical cipher. The literature indicates the existence of two different approaches to the problem of insecurity of the Hill cipher. The first approach involves the invention of nonlinear version of the Hill cipher [15], while the second approach is to design secure cyclic mechanism of key exchange [12, 16].

Conclusions

A large number of industrial cameras, which are constantly monitoring the vessel traffic, are used in order to ensure the safety of coastal waters. These systems are designed to not only collect the necessary information for vessel traffic control, but also to protect traffic participants from the raids of sea pirates. VTS systems will effectively serve their purpose only if traffic controllers have constant access to the transmitted video images. Furthermore, the images transmitted from the cameras have to be available only for traffic controllers at designated terminals. It can be achieved through the physical security of the transmission although it can be very expensive and it limits choice of medium to wired solutions. Therefore, it is suggested to use software solutions.

In order to provide confidentiality of transmitted information, it is necessary to use an encryption algorithm. However, there is a need for a mechanism that will allow for detect the leaks of confidential information from the system, and then track down those involved in the illegal sharing of recordings. The paper presents an approach to the problem of piracy known from such services as video on demand or Internet TV. Fingerprinting methods allow to achieve a satisfactory level of security using the existing network infrastructure without having to build a new one. As can be seen in the examples of listed various methods, it is possible to minimize the cost of fingerprinting by using the appropriate method, adapted to multicast transmission.

Fingerprinting methods, especially solutions based on approach of Joint Fingerprinting and

Decryption, can greatly contribute to the safety of coastal waters. Therefore, fingerprinting should be taken into consideration when designing new or improving existing VTS systems.

References

1. LIU K.J.R., TRAPPE W., WANG Z.J., WU M., ZHAO H.: Multimedia fingerprinting forensics for traitor tracing. EURASIP Book Series on Signal Processing and Communications, vol. 4, Hindawi Publishing Corporation, 2005.
2. COX I.J., KILIAN J., LEIGHTON F.T., SHAMMOON T.G.: Secure spread spectrum watermarking for multimedia. IEEE Trans. Image Processing, vol. 6, 12, 1997, 1673–1687.
3. WU M., TRAPPE W., WANG Z.J., LIU K.J.R.: Collusion-resistant fingerprinting for multimedia. IEEE Signal Processing Mag., vol. 21, 2004, 15–27.
4. ZHAO H., WU M., WANG Z.J., LIU K.J.R.: Forensic analysis of nonlinear collusion attacks for multimedia fingerprinting. IEEE Trans. Image Processing, vol. 14, 5, 2005, 646–661.
5. BARCZ M.: Review and analysis of fingerprinting methods for multicast distribution of video signals. Gdansk University of Technology, Faculty of Electronics, Telecommunications and Informatics, master thesis, 2009.
6. CZAPLEWSKI B.: Implementation and research of fingerprinting method based on generalized Hill cipher. Gdansk University of Technology, Faculty of Electronics, Telecommunications and Informatics, master thesis, 2011.
7. ZHAO H., LIU K.J.R.: Bandwidth efficient fingerprint multicast for video streaming. Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, 2004, 849–852.
8. JUDGE P., AMMAR M.: WHIM: Watermarking multicast video with a hierarchy of intermediaries. Proc. 10th International Workshop on Network and Operating System Support for Digital Audio and Video, 2000.
9. PARNES R., PARVIAINEN R.: Large scale distributed watermarking of multicast media through encryption, Proc. IFIP Int. Conf. Communications and Multimedia Security Issues of the New Century, 2001, 17.
10. ANDERSON R., MANIFAVAS C.: Chameleon – A new kind of stream cipher. Lecture Notes in Computer Science, Fast Software Encryption, Springer, vol. 1267, 1997, 107–113.
11. KUNDUR D., KARTHIK K.: Video fingerprinting and encryption for digital rights management. Proc. IEEE, vol. 92, 6, 2004, 918–932.
12. RYKACZEWSKI R.: Hillcast – A method of joint decryption and fingerprinting for multicast distribution of multimedia data. Gdansk University of Technology, Faculty of Electronics, Telecommunications and Informatics Annals, 8, series: Information Technology, 2010.
13. CZAPLEWSKI B., RYKACZEWSKI R.: Joint fingerprinting and cryptographic protection of data using the Hill cipher. Gdansk University of Technology, Faculty of Electronics, Telecommunications and Informatics Annals, 9, series: ICT Young, 2011.
14. CHENG H., LI X.: Partial encryption of compressed images and videos. IEEE Trans. Signal Processing, vol. 48, 2000, 2439–2451.
15. TOORANI M., FALAHATI A.: A secure variant of the Hill cipher. IEEE Symposium on Computers and Communications, 2009.
16. SAEEDNIA S.: How to make the Hill cipher secure. Cryptologia, vol. 24, 2000, 353–360.