(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 October 2010 (07.10.2010)

PCT

(10) International Publication Number
WO 2010/113041 A2

(51) International Patent Classification:
H04L 9/18 (2006.01)

(21) International Application Number:
PCT/IB2010/001120

(22) International Filing Date:
15 March 2010 (15.03.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PCT/CN2009/071087 31 March 2009 (31.03.2009) CN

(71) Applicant *(for all designated States except US)*: **FRANCE TELECOM** [FR/FR]; 6, Place d'Alleray, F-75015 Paris (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LIAN, Shiguo** [CN/CN]; 10F, South Twr, Raycom Info Tech Park C, 2 Science Institute South RD, Haidian District, Pekin 100080 (CN). **DONG, Yuan** [CN/CN]; 10F, South Tower Raycom InfoTech Park C, 2 Science Institute South Road, Haidian District, Beijing 100080 (CN).

(74) Agent: **FRANCE TELECOM/FTR&D/PIV/BREVETS**; BORNIER Raynald, 38-40 Rue du Général Leclerc, F-92794 Issy les Moulineaux Cedex 9 (FR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

— *of inventorship (Rule 4.17(iv))*

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title**: FINGERPRINTING METHOD AND SYSTEM

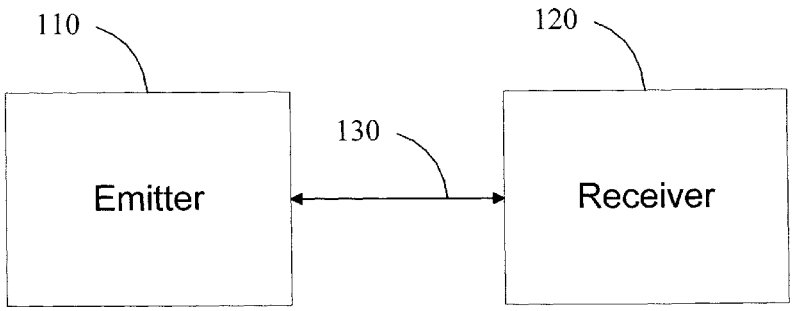


Figure 1

(57) **Abstract**: A method embedding a user code into a media content, said user code being selected from a plurality of available user codes, each available user code comprising elements of an alphabet of at least a first and second types, said first type being associated to a first subset of noise seeds taken from a set of distinct noise seeds, said method comprising the acts of receiving a media content comprising a set of noise sequences embedded therein; said noise sequences being generated using the set of distinct noise seeds and a given noise function; receiving a second subset of the noise seeds; removing a sub-set of noise sequences from the media content, said sub-set of noise sequences being generated from the received second sub-set of noise seeds and the given noise function; wherein the received second sub set of noise seeds corresponds to the noise seeds from the set of distinct noise seeds that are not associated to an element of the first type.

WO 2010/113041 A2

FINGERPRINTING METHOD AND SYSTEM

Field of the Invention

The present invention relates in general to digital right management and more specifically to media fingerprinting.

Background of the Invention

Fingerprinting is a suitable solution for media content distribution between a sender and a receiver or user. Indeed, for each user, a corresponding fingerprint or user code is embedded into his media content. In other words, a fingerprint corresponds to a unique user code, such as user ID (identification). Each user is provided with a copy of the media content, wherein the corresponding embedded fingerprint or user code allows further identifying said user, for example in the case of illegal redistribution by the user. Practically, for security enhancement, the media content may be encrypted in order to prevent their use by unauthorized customers for instance during transmission. In this case, the location where encrypting, decrypting and fingerprinting (i.e. embedding a fingerprint) of the media content takes place should thus be considered. Intuitively, a media content should be first fingerprinted then encrypted at a sender side then transmitted and decrypted at a receiver side. However, in this case, the sender should produce different copies for different users. This would be time and cost consuming and not efficient for multicasting (i.e. one sender transmits to a plurality of identified users). Balancing the transmission load between the sender and the users has been considered as a solution. In this case, the sender encrypts the media content for transmission and the receiver first decrypts then fingerprints the received encrypted media content. However, this may lead to potential hacking of the media content between decryption and fingerprinting. It is thus necessary to jointly decrypt and embed fingerprint (i.e. in a single step process) in order to avoid such problem. Today, a few existing joint decryption and fingerprinting solutions allow implementing decryption and fingerprinting at the same time. In these solutions, the media content is encrypted at the sender side (for example on a server in a communication network) and

decrypted into different copies at the receiver side (for example, a user terminal). Although these solutions may detect illegal distributors, they have some drawbacks.

For example, prior art document "Chamleon – A new kind of stream cipher", R. Anderson and C. Manifavas, Lecture Notes in Computer Science, Fast Software Encryption, Springer-Verlag, 1997, pp. 107-113 is based on a stream cipher, which allows encrypting a media content and embedding different fingerprints into the Least Significant Bit (LSB). This scheme is efficient, but is not robust to signal processing such as recompression or adding noise, which is a major drawback.

Furthermore, prior art document "Video fingerprinting and encryption principles for digital rights management", D. Kundur and K. Karthik, Proceedings of the IEEE, Vol. 92, No. 6, June 2004, pp. 918-932 is based on partial encryption. However, partial encryption brings confusion into sign bits of the Discrete Cosine Transform (DCT) coefficients in encryption and decrypts only part of the sign bits in decryption. This scheme is robust to signal processing, but the encrypted media is often still intelligible, and the decrypted media is greatly degraded.

Eventually, prior art document "Joint Fingerprint Embedding and Decryption for Video Distribution", S. Lian, Z. Liu, Z. Ren, H. Wang, 2007 IEEE International Conference on Multimedia and Expo (ICME2007) is based on motion vector encryption, which encrypts video data with motion vector encryption and embeds different fingerprint into the motion vectors. This scheme is suitable for video data, but inadequate for image, audio or text data where there is no motion vector.

Today there is a need for a fingerprinting solution that can be easily implemented on the existing communication infrastructures, overcoming the drawbacks of the prior art.

Summary of Invention

It is an object of the present system to overcome disadvantages and/or make improvement over the prior art.

To that extend, the invention proposes a method for identifying a media content authorized user, the authorized user being associated to a user code selected from a plurality of available user codes, each available user code comprising elements of an alphabet of at least a first and second types, said first type being associated to a subset of noise seeds taken from a set of distinct noise seeds, said method comprising the act of:

- correlating the media content with a set of noise sequences generated from the set of distinct noise seeds, to identify a subset of said set of noise sequences embedded into said media content prior to distribution of said media content,

- identifying elements of the first type that correspond to a subset of noise seeds that would generate the subset of embedded noise sequences,

- deriving a user code from the identified elements of the first type to retrieve the authorized user of the media content.

Drawbacks of the prior art are thus solved using as the identification of a user may be performed using a set of elements that are embedded into the media content and that may be identified by a checker who knows the set of distinct noise seeds or the corresponding noise sequences.

Furthermore, compare to existing solution, the present method involves fewer steps as only identified elements of the first type remain in the media content after fingerprinting.

The present invention also proposes a method for embedding a user code into a media content, said user code being selected from a plurality of available user codes, each available user code comprising elements of an alphabet of at least a first and second types, said first type being associated to a first subset of noise seeds taken from a set of distinct noise seeds, said method comprising the acts of:

- receiving a media content comprising a set of noise sequences embedded therein; said noise sequences being generated using the set of distinct noise seeds and a given noise function;

- receiving a second subset of the noise seeds,

- removing a sub-set of noise sequences from the media content, said sub-set of noise sequences being generated from the received second sub-set of noise seeds and the given noise function;

wherein the received second sub-set of noise seeds corresponds to the noise seeds from the set of distinct noise seeds that are not associated to an element of the first type.

The latter method may be performed by a receiver in a communication network that receives a media content previously embedded with noise sequences by an emitter.

The present invention allows in particular tracing people who illegally redistribute media contents such as e.g. video files, images, audio files etc...

An advantage of the present invention is that security is obtained by selecting noise strength that keeps the noised media content unintelligible.

Another advantage of the present invention is that the embedded fingerprint is robust to operations like adding noise.

Another advantage of the present invention is that the embedded fingerprint is easy to detect.

When encrypted, one advantage of the present invention allows securing a media content in perception, i.e. the encrypted media content cannot be read or understood.

The encrypted media content may also be decrypted into different copies comprising different user codes or fingerprints.

The invention also relates to a system for identifying a media content authorized user according to claim 7.

The invention also relates to a device for embedding a user code into a media content according to claim 8. The invention further relates to a device for authenticating a media content authorized user according to claim 11.

The invention also relates to a computer program for embedding a user code into a media content according to claim 14.

The invention also relates to a computer program for authenticating a media content authorized user according to claim 15.

Brief Description of the Drawings

Embodiments of the present invention will now be described solely by way of example and only with reference to the accompanying drawings, where like parts are provided with corresponding reference numerals, and in which:

Figure 1 schematically illustrates a system according to an embodiment of the present invention;

Figure 2 schematically illustrates a system according to an embodiment of the present invention;

Figure 3 schematically illustrates a method according to an embodiment of the present invention;

Figure 4A schematically illustrates an act of adding noise according to an embodiment of the present invention;

Figure 4B schematically illustrates random sequence generation according to an embodiment of the present invention;

Figure 5 schematically illustrates noise removing according to an embodiment of the present invention;

Figure 6 schematically illustrates user code detection according to an embodiment of the present invention;

Figure 7 schematically illustrates an example of software based implementation according to an embodiment of the present invention;

Figure 8 schematically illustrates an example of terminal based implementation according to an embodiment of the present invention;

Figure 9 schematically illustrates an example of set-top box based implementation according to an embodiment of the present invention.

Description of Preferred Embodiment

The following are descriptions of exemplary embodiments that when taken in conjunction with the drawings will demonstrate the above noted features and advantages, and introduce further ones.

In the following description, for purposes of explanation rather than limitation, specific details are set forth such as architecture, interfaces,

techniques, devices, etc..., for illustration. However, it will be apparent to those of ordinary skill in the art that other embodiments that depart from these details would still be understood to be within the scope of the appended claims.

Moreover, for the purpose of clarity, detailed descriptions of well-known devices, systems, and methods are omitted so as not to obscure the description of the present system. Furthermore, routers, servers, nodes, gateways or other entities in a telecommunication network are not detailed as their implementation is beyond the scope of the present system and method.

Unless specified otherwise, the exemplary embodiment will be described hereafter in its application to a user equipment and a server of a communication network. This is in no way limiting as the man skilled in the art may transpose the present teachings to another configuration.

In addition, it should be expressly understood that the drawings are included for illustrative purposes and do not represent the scope of the present system.

Figure 1 describes an illustrative embodiment of the system according to the invention. An emitter or sender 110 is connected to a receiver 120 through a link or a communication network 130.

The emitter 110 stores or is adapted to receive one or a plurality of original media content(s). An original media content may be for example a raw video, audio file, image, text etc... The emitter 110 may be for example a server in a communication network 130. The emitter is operable to embed noise sequences into the original media content according to the present invention and transmit the resulting media content to at least one receiver 120 over the communication network 130. For security purposes, the emitter 110 may also encrypt the resulting media content before transmission. The receiver 120 is operable to fingerprint the received resulting media content. When the resulting media content has been encrypted, the receiver 120 is also operable to jointly decrypt and fingerprint the received resulting media content. The receiver 120 may be,

for example, a user equipment such as e.g. a PC (Personal Computer) or a mobile telecommunication device (mobile phone, portable computer etc...).

In an additional embodiment according to the invention, the emitter 110 and the receiver 120 may be the same physical entity such as for instance a server.

Figure 2 describes an illustrative embodiment of the system according to the invention, detailing emitter 110 and receiver 120.

The emitter 110, or emitter unit, may comprise:

- an emitter user code unit 200,
- an emitter noise seed unit 210,
- an emitter media content unit 220,
- an encryption unit 230.

Whereas the receiver 120 comprises:

- a receiver user code unit 240,
- a receiver noise seed unit 250,
- a receiver media content unit 260,
- a decryption unit 270.

As described here above, emitter 100 and receiver 120 are connected via communication network 130.

The emitter user code unit 200 allows indexing and storing user codes that are further used for fingerprinting. Indeed, each user code or fingerprint corresponds to a user and provides identification through the associated user code. To each user of the receiver 120 corresponds a user code or fingerprinting that provides identification. A user code database and a user code table, both comprises in or accessible to the emitter user code unit 200, allow selecting a given user code. The emitter noise seed unit 210 allows storing and selecting noise seeds. A noise seed allows generating a noise sequence. The emitter media content unit 220 allows receiving, optionally storing, and modifying original media content. Modifications made to said original media content will be further described in reference to the method according to the invention. The encryption unit 230 is optional and allows encrypting the media content before transmitting it to the receiver 120 through the communication network 130.

The receiver user code unit 240 allows receiving and optionally storing user codes received from the communication network 130. The receiver noise seed unit 250 allows receiving and optionally storing noise seeds from the communication network 130. The receiver media content unit 260 allows receiving and optionally storing the media content received from the communication network 130. The decryption unit 270 is optional and allows decrypting the media content after having received it from the emitter 110 through the communication network 130. Decryption will be done in accordance with the encryption process that was performed on the emitter.

Figure 3 describes an illustrative embodiment of the method according to the invention which allows embedding in a receiver 120 a fingerprint or user code into a media content so that the user may further be identified through said user code. The description of the present figure is made in reference to Figure 2.

In act 300, a set of noise sequences are embedded into the original media content in order to obtain a noisy media content. These noise sequences are generated using a set of noise seeds. Noise sequences are typically sequences of random numbers comprised within a range [-L;L]. A noise seed is a number (for example a 32-bit integer) that allows generating a random sequence of random numbers within a given range of number. Such random sequence generation algorithms are known from the man skilled in the art through documents such as "P-adic chaos and random number generation", Christopher F. Woodcock, Nigel P. Smart, Experimental Mathematics, 1998, 7(4): 333-342 or "An Introduction to Cryptography", R. A. Mollin, CRC Press, 2006.

For example, for embedding n noise sequences, the original media content may be divided into d parts. Each noise sequence should in this case comprise d components that may be obtained through the random sequence generation. Each noise sequence will thus add its d-th noise component to the d-th part of the original media content. Beside, there may be as many noise seeds as noise sequences (n), even though it may be emphasized that noise sequences may be generated with fewer noise

seeds (using redundant ones). Noise seeds may be updated or changed from time to time to enhance security.

The n noise seeds may be different to assure they generate different sequences that may further be detected.

In act 310, for a given user code, a sub-set of noise seeds are selected using said given user code. It is emphasized that each user code ought to be unique for each user. The selection of noise seeds is performed depending on the user code.

An example of user code is illustrated as follows. A user code F_k of a k-th user is composed of components $f_{k,0},f_{k,1},\ldots,f_{k,n-1}$ ($f_{k,j}$=0 or -1, k=0,1,...,m-1, j=0,1,...,n-1, wherein m is the number of users and n the number of noise sequences). A typical example of user codes may be as follows:

$$
\begin{array}{lcccccc}
F_0: & 0 & -1 & -1 & -1 & -1 & -1 \ \ldots \\
F_1: & -1 & 0 & -1 & -1 & -1 & -1 \ \ldots \\
F_2: & -1 & -1 & 0 & 0 & -1 & -1 \ \ldots \\
F_3: & -1 & -1 & -1 & 0 & -1 & -1 \ \ldots \\
F_4: & -1 & 0 & 0 & -1 & 0 & -1 \ \ldots \\
F_5: & -1 & -1 & 0 & -1 & -1 & 0 \ \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \ \ldots
\end{array}
$$

In the case where there are n noise seeds, the selection may be performed for instance by selecting in the sequence of n noise seeds, the ones that corresponds to the components of F_k that are equal to -1. The components of F_k that are equal to 0 are not selected. Eventually, (n-t) noise seeds are selected.

Act 320 allows removing a sub-set of noise sequences from the media content, said sub-set of noise sequences being generated using said sub-set of noise seeds. In other words, (n-t) noise sequences generated by the selected noise seeds are removed from the noisy media content. The remaining noise sequences in media content are the ones that correspond to the zeros of the user code. As each zero is associated with a component in the sequence, this allows further reconstructing the user code. Detailed embodiments of acts 300, 320 and user code detection are further described here under, respectively in Figure 4A, Figure 5 and Figure 6.

In reference to Figure 2, act 300 and act 310 would typically be performed by emitter 110 whereas act 320 would typically be performed by receiver 120. The noisy media content obtained in act 300 would be typically transmitted to the receiver 120 through the communication network 130 along with the noise seeds selected in act 310 and optionally some user codes. Receiver 120 would then typically perform act 320 to remove the (n-t) noise sequences, providing thereby a fingerprinted media content. However, it should be emphasized that acts 300, 310 and 320 may also be performed by the same entity.

Furthermore, even though the method according to the invention may be carried out without encryption/decryption, it will be described hereunder using such an encryption/decryption scheme. Encryption is the process of transforming information (referred to as plaintext) using an algorithm (called cipher) to make it unreadable to anyone except those possessing special knowledge, usually referred to as a key. Decryption is the associated process for make said encrypted information readable using said key. Encryption corresponds to a particular way of scrambling information.

Figure 4A describes an illustrative additional embodiment of act 300 of the method according to the invention in reference to Figure 3.

The n noise sequences S_0, S_1, ..., S_{n-1} may be defined as flows:

* $S_i=s_{i,0}, s_{i,1},\ldots, s_{i,d-1}$,
* $-L \leq s_{i,j} < L$,
* i=0, 1, ... , n-1,
* j=0, 1, ... , d-1,
* [-L,L] being the range of numbers during random generation,
* d being the number of parts of the original media content.

The generation of the n noise sequences is performed with a random number generator under the control of noise seeds $R=r_0r_1r_2\ldots r_{n-1}$,wherein r_i is for example a 32-bit integer, i=0,1,...,n-1; $r_0, r_1, \ldots, r_{n-1}$ being the noise seeds for respectively S_0, S_1, ..., S_{n-1}. Random sequence generation is illustrated on Figure 4B wherein each noise seed $r_0, r_1, \ldots, r_{n-1}$ (respectively 471, 472, 473) allows generating S_0, S_1, ..., S_{n-1} noise sequences

(respectively 481, 482, 483) using random generation functions (respectively 491, 492, 493).

As described in Figure 4A, the original media content 410 (divided into d parts $P=p_0, p_1, \ldots, p_{d-1}$) is modified by embedding or adding in an act 460 the n noise sequences 440 previously generated in an act 430 (described here above in detail in Figure 4B) using the noise seeds $R=r_0r_1r_2\ldots r_{n-1}$ 420. This embedding is performed under the encryption key (or secret weight) $E= e_0, e_1, \ldots, e_{n-1}$ 450 (where $e_i=\pm1$, i=0,1,..., n-1 or more generally $e_i=\pm x$, x is a real number) to obtain an encrypted noisy media content $C=c_0, c_1, \ldots, c_{d-1}$ 470:

$$c_i = p_i + \sum_{j=0}^{n-1} e_j s_{j,i} \ (i = 0,1,...,d-1) \qquad \text{eq. (1)}$$

Figure 5 describes an illustrative embodiment of act 320 of the method according to the invention in reference to Figure 3.

After having selected (n-t) noise seeds $R_k=r_0, r_1, \ldots, r_{n-t-1}$ 500 in act 310 as previously described in Figure 3, (n-t) corresponding noise sequences $S_0, S_1, \ldots, S_{n-t-1}$ 520 are regenerated or reused in an act 510 using the scheme described in Figure 4B (random sequence generation from noise seeds). The function used to generate the corresponding (n-t) noise sequences on the receiver 120 is the same that the one used on the emitter 110 in act 300. The noise sequences 520 are then embedded or added in act 530 to the encrypted noisy media content $C=c_0, c_1, \ldots, c_{d-1}$ 470 under the decryption key (or secret weight) $E= e_0, e_1, \ldots, e_{n-1}$ 450 (corresponding to the one used for encryption) and in combination with the user code $F_k=f_{k,0}, f_{k,1}, \ldots, f_{k,n-1}$ 540 corresponding to the user k. Act 530 will allow removing (n-t) already embedded noise sequences (of the encrypted noisy media content) as the combination with the user code is made to add the negative value of the (n-t) noise sequences, through the exemplary choices of F_k. In other words, (n-t) selected noise sequences are substracted from the encrypted noisy media content. The selection of noise sequences in combination with user code is realized based on the selection of the corresponding (n-t) noise seeds that were also selected based upon

the user code. The resulting media content is the decrypted fingerprinted media content $P_k=p_{k,0}, p_{k,1}, \ldots, p_{k,n-1}$ 550. The decrypted fingerprinted media content 550 is decrypted using decryption key E that should be the same as the encryption key 450. The fingerprint is embedded by leaving t noise sequences in the media content. The remaining t noise sequences correspond to the zeros of the user code associated with the fingerprint. Indeed, as described here above, the -1 components of the user code will be added to the noisy media content in order to cancel the corresponding noise sequences, leaving thus only the zeros in identified parts of the media content, further allowing reconstructing the user code to identify the user of a given media content. An example of equations of noise removing process as detailed as follows:

- act 300 of noise addition (or embedding) allowed deriving equation (1) then as, described here above and shown in equations (2) and (3), act 320 allows noise removing:

$$p_{k,i} = c_i + \sum_{j=0}^{n-1} e_j f_{k,j} s_{j,i}$$

$$(i = 0,1,\ldots,d-1,\ k = 0,1,\ldots,m-1)$$

eq. (2),

$$p_{k,i} = p_i + \sum_{j=0}^{n-1} e_j s_{j,i} + \sum_{j=0}^{n-1} e_j f_{k,j} s_{j,i} = p_i + \sum_{j=0}^{n-1} e_j (f_{k,j} + 1) s_{j,i}$$

$$(i = 0,1,\ldots,d-1,\ k = 0,1,\ldots,m-1)$$

eq. (3),

wherein m is the number of users.

Equation (3) is derived by integrating equation (1) in equation (2).

As explained here above, as (n-t) of the $f_{k,j}$ are equal to -1, the last term of equation (3) containing ($f_{k,j}$ + 1) will be cancelled or equal to 0 for (n-t) values of the index j.

The receiver which performs the decrypting and fingerprinting should know the partition scheme in d parts of the received media content either by deriving it directly from the received media content or by knowing it in advance.

It is to be emphasized that the (n-t) selected noise seeds do not need to be send or received in any order (e.g. in the original order) but the receiver needs to be able to further use them in the correct order when removing noise in equation (2). In others words, as seen in equation (3), for a given user k, for a given part i (0<=i<=d-1), each component e_j of the secret weight E should be multiplied with the corresponding component $f_{k,j}$ of the user code F_k and the corresponding component $s_{j,i}$ of the noise sequence j for the part i (0<=j<=n-1). The term "corresponding" meaning that the same components of E and S multiplied at the emitter during the noise sequence embedding should also be multiplied at the receiver along with the component of the same rank of F_k as shown in equation (3).

Once the media content has been fingerprinted with a given user code, one (here under called checker) may want to detect said given user code to identify the associated user. For example, the owner of the rights of a fingerprinted media content may want to check who uploaded illegally on the Internet a copy of the media content.

Figure 6 describes an illustrative embodiment of user code detection according to the invention.

The method according to the invention allows identifying a media content authorized user, the authorized user being associated to a user code selected from a plurality of available user codes, each available user code comprising elements of an alphabet of at least a first and second types, said first type being associated to a subset of noise seeds taken from a set of distinct noise seeds, said method comprising the act of:

- correlating the media content with a set of noise sequences generated from the set of distinct noise seeds, to identify a subset of said set of noise sequences embedded into said media content prior to distribution of said media content,

- identifying elements of the first type that correspond to a subset of noise seeds that would generate the subset of embedded noise sequences,

- deriving a user code from the identified elements of the first type to retrieve the authorized user of the media content.

The user code detection, which involves a correlation process on each noise sequence as described here under in equation (4) may be performed in any order (i.e. each noise sequence independently from the others).

It is first then necessary for the checker to have all the noise seeds to create the corresponding noise sequences using the same function as the one used to generate the noise sequences at the emitter or at the receiver in respectively act 300 and 320. The checker may also used directly the same noise sequences which would have been stored in a database for example.

The decrypted fingerprinted media content $P_k=p_{k,0}, p_{k,1}, \ldots, p_{k,n-1}$ 550 is processed by the optional detection unit 610 with each of the noise sequences as follows:

$$< P_k, S_j >= \frac{\sum_{i=0}^{d-1} p_{k,i} s_{j,i}}{\sum_{i=0}^{d-1} s^2_{j,i}} = \frac{\sum_{i=0}^{d-1} [p_i + \sum_{t=0}^{n-1} e_t (f_{k,t} + 1) s_{t,i}] s_{j,i}}{\sum_{i=0}^{d-1} s^2_{j,i}} \quad \text{eq. (4)}$$

$$(j = 0,1,\ldots,n-1,\ k = 0,1,\ldots,m-1)$$

In the detection unit 610, each component $\{p_{k,i}\}$ of the decrypted fingerprinted media content $P_k=p_{k,0}, p_{k,1}, \ldots, p_{k,n-1}$ 550 is correlated with components $\{s_{j,i}\}$ of noise sequences to detect which component $\{s_{j,i}\}$ remains in the decrypted fingerprinted media content 550 after noise removing according to act 320. The correlation allows detecting in which parts of the media content noise is embedded, allowing thus knowing in which parts (of the d parts) of the media content zeros are associated with.

The remaining parts of the d parts correspond to -1. The user code 620 is hence reconstructed or determined.

The correlation algorithm is based for example in equation (4) on least mean squares, but any relevant correlation algorithm may be suitable to perform user code detection.

In equation (4), the least mean square method uses a threshold T to determine whether the correlated elements correspond or not.

The threshold T may have a reasonable range 0<T<1. In practice, it should be selected by analyzing the media content's properties. However, there are no general principles for various natural media content (images, audios, video, etc.). Generally, experiments should be done to get the statistical properties of the aimed media content. Experiments show that, for images, T=0.3 will get suitable detection results, while for audios, T=0.15 is more suitable.

In the example of equation (4), the correlation and determination is done as follows:

$$f_{k,j} = \begin{cases} 0, & \langle P_k, S_j \rangle \geq T \text{ or } \langle P_k, -S_j \rangle \geq T \\ -1, & \langle P_k, S_j \rangle < T \text{ and } \langle P_k, -S_j \rangle < T \end{cases} \quad \text{eq. (5)}$$

$$(j = 0,1,...,n-1,\ k = 0,1,...,m-1)$$

Noise seeds $r_0, r_1, ...,r_{n-1}$ or the corresponding noise sequences S_0, S_1, ..., S_{n-1} being distinct, their order or indexes or directly their values allows reconstructing the user code based on the detected elements of the first type (e.g. zeros).

The detected user code may further be compared with a user code database or table (centralized or distributed in a network or on an entity (e.g. server, computer...)) containing all the users to identify the user which owns the rights on this media content.

The checker may be the media content distributor or anyone who may access the database wherein user codes and noise seeds are stored.

To avoid too many noise sequences in the decrypted fingerprinted media content 550 after noise removing (to avoid degrading the media content too much), it may be relevant to perform a limitation on this number of noise sequences. In others words, the number of zeros of each user code should be limited. Implementation testing showed that a number of three zeros is near an optimum as showed in equation (6) below, but this number is not, however, a limitation to the scope of the present invention:

$$\sum_{j=0}^{n-1} f_{k,j} \leq 3-n \quad (k=0,1,\ldots,m-1) \qquad \text{eq. (6)}$$

The method according to the invention has been described so far independently of the entity. The method according to the invention may be performed on a single entity, but may also be split between an emitter 110 and a receiver 120 linked through a network 130 as described in Figure 1.

In this case, emitter 110 would be likely performing act 300 and 310, while receiver 120 would likely be performing act 320.

In reference to Figures 2 and 3:

- the emitter user code unit 200 may be at least adapted to store user codes,
- the emitter noise seed unit 210 may be at least adapted to select noise seeds based upon a given user code according to act 310,
- the emitter media content unit 220 may be at least adapted to perform the noise embedding according to act 300 (optionally using encryption),
- the encryption unit 230 may be at least adapted to store the encryption key 450,
- the optional receiver user code unit 240 may comprise a detection unit 610 and/or store user codes and/or being adapted to connect to a user code database,
- the receiver noise seed unit 250 may be at least adapted to receive the selected noise seeds 500,
- the receiver media content unit 260 may be at least adapted to remove noise according to act 320,

- the decryption unit 270 may be at least adapted to store and/or receive the decryption key 450.

In an additional embodiment of the present invention, user code may be sent by and/or stored on the emitter 110 and also received on and/or stored on the receiver 120.

To enhance security, the selected noise seeds 500, the secret weight (encryption/decryption key) 450 and the user code (when transmitted) may be transmitted through a secret channel on the communication network 130, which depends on the implementations of the adding-noise process and removing-noise process. In the following, three cases of applications are described: software based implementation, terminal based implementation, and set-top box based implementation.

Figure 7 describes an illustrative embodiment of software based implementation of the method according to the invention.

In software based implementation, both the adding-noise act 300 and removing-noise act 320 may be implemented using software modules respectively 710 and 720. However, such a software implementation may prove to be easier for attackers to intrude into said software modules. Thus, all the noise seeds ($r_0r_1r_2...r_{n-t-1}$) 500, secret weight ($E= e_0e_1...e_{n-1}$) 450 and user code ($F_k=f_{k,0},f_{k,1},...,f_{k,n-1}$) should not be stored in the software modules permanently (or for a long time), but, alternatively, should be transmitted from the sender to the receiver in an encrypted form using a session key 730 shared between the emitter 110 and receiver 120. As shown in Figure 7, the selected noise seeds, the secret weight and the user code may be encrypted into $r'_0r'_1r'_2...r'_{n-t-1}$ 750, E' 760 and F'_k 740 using the session key 730 then transmitted to the receiver's software module 720, and decrypted using the session key 730.

Figure 8 describes an illustrative embodiment of terminal based implementation of the method according to the invention.

For example, a mobile phone is a terminal that may typically identify a user using in particular the Subscriber Identity Module (SIM) contained on a SIM card. In such a terminal based implementation, the user code ($F_k=f_{k,0},f_{k,1},\ldots,f_{k,n-1}$) 800 may be stored in the mobile or terminal 820 (receiver) in a secured manner on the SIM card. The selected noise seeds 500 and the secret weight 450 may be encrypted in $r'_0r'_1r'_2\ldots r'_{n-t-1}$ 830, E' 840 using a session key 850 and transmitted from a sender 810 in the same way as previously described in Figure 7. The user code ($F_k=f_{k,0},f_{k,1},\ldots,f_{k,n-1}$) 800 may be initialized for instance when the SIM card is activated.

Figure 9 describes an illustrative embodiment of set-top box based implementation of the method according to the invention.

A set-top box (i.e. for example a internet home box provided by a network operator) may identify the user. In this kind of implementation, all of the user code 800, secret weight 450 and selected noise seeds 500 may be stored both in the set-top box 920 and on sender side 910 securely. All parameters may be initialized for instance when the set-top box is activated.

The user code should not be known from the end user who buys the media content but only from the receiver that performs the fingerprinting. It is thus important to protect the user code and to restrict or prevent the access by the end user.

Claims

1. A method for embedding a user code into a media content, said user code being selected from a plurality of available user codes, each available user code comprising elements of an alphabet of at least a first and second types, said first type being associated to a first subset of noise seeds taken from a set of distinct noise seeds, said method comprising the acts of:
   - receiving a media content comprising a set of noise sequences embedded therein; said noise sequences being generated using the set of distinct noise seeds and a given noise function;
   - receiving a second subset of noise seeds,
   - removing a sub-set of noise sequences from the media content, said sub-set of noise sequences being generated from the received second sub-set of noise seeds and the given noise function;
   wherein the received second sub-set of noise seeds corresponds to the noise seeds from the set of distinct noise seeds that are not associated to an element of the first type in the selected user code.

2. A method according to claim 1, wherein the act of embedding the set of noise sequences into said media content is performed using a scrambling scheme to scramble said set of noise sequences, and the act of removing the sub-set of noise sequences from the media content is performed using the same scrambling scheme.

3. A method according to any of the preceding claims, wherein the act of removing a sub-set of noise sequences from the media content is performed according to:

$$p_{k,i} = p_i + \sum_{j=0}^{n-1} e_j s_{j,i} + \sum_{j=0}^{n-1} e_j f_{k,j} s_{j,i} = p_i + \sum_{j=0}^{n-1} e_j (f_{k,j} + 1) s_{j,i}$$

$$(i = 0,1,\ldots,d-1,\ k = 0,1,\ldots,m-1)$$

wherein p_i are the d parts of the media content, e_j is the t-th component among n of the encryption key E, $s_{j,i}$ is the i-th component among d of the j-th noise sequence among n and $f_{k,j}$ is the j-th component of the user code corresponding to user k.

4. A method for identifying a media content authorized user, the authorized user being associated to a user code selected from a plurality of available user codes, each available user code comprising elements of an alphabet of at least a first and second types, said first type being associated to a first subset of noise seeds taken from a set of distinct noise seeds, said method comprising the act of:
- correlating the media content with a set of noise sequences generated from the set of distinct noise seeds, to identify a subset of said set of noise sequences embedded into said media content prior to distribution of said media content,
- identifying elements of the first type that correspond to a subset of noise seeds that would generate the subset of embedded noise sequences,
- deriving a user code from the identified elements of the first type to retrieve the authorized user of the media content.

5. A method according to claim 4, wherein the subset of noise sequences embedded into said media content prior to distribution is embedded using a scrambling scheme to scramble said subset of noise sequences, and the act of correlating the media content comprises an act of correlating said media content with the set of noise sequences generated from the set of distinct noise seeds and scrambled with the same scrambling scheme.

6. A method according to claim 5, wherein the act of correlating the media content is performed using the correlation function:

$$< P_k, S_j >= \frac{\sum_{i=0}^{d-1} p_{k,i} s_{j,i}}{\sum_{i=0}^{d-1} s^2_{j,i}} = \frac{\sum_{i=0}^{d-1} [p_i + \sum_{t=0}^{n-1} e_t (f_{k,t} + 1) s_{t,i}] s_{j,i}}{\sum_{i=0}^{d-1} s^2_{j,i}}$$

$$(j = 0,1,...,n-1,\ k = 0,1,...,m-1)$$

with:

$$f_{k,j} = \begin{cases} 0, & \langle P_k, S_j \rangle \geq T \text{ or } \langle P_k, -S_j \rangle \geq T \\ -1, & \langle P_k, S_j \rangle < T \text{ and } \langle P_k, -S_j \rangle < T \end{cases}$$

$$(j = 0,1,...,n-1,\ k = 0,1,...,m-1)$$

wherein T is a pre-determined threshold allowing deciding whether P_k and S_j are correlated or not, k corresponds to user k, m corresponds to the number of users, S_j is the j-th noise sequence among n with components $s_{j,i}$, d corresponds to the number of part of the media content and e_t is the t-th component among n of the encryption key E and $f_{k,t}$ is the t-th component of the user code corresponding to user k.

.

7. A system for identifying a media content authorized user, the authorized user being associated to a user code selected from a plurality of available user codes, each available user code comprising elements of an alphabet of at least a first and second types, said first type being associated to a first subset of noise seeds taken from a set of distinct noise seeds, said system comprising:

an emitter unit for sending a media content to its authorized user over a communication network, said media content comprising a set of noise sequences embedded therein; said noise sequences being generated using the set of distinct noise seeds and a given noise function; said emitter being further operable to send a second subset of noise seeds corresponding to the noise seeds from the set of distinct noise seeds that are not associated to an element of the first type in the user code associated to the authorized user,

a device for embedding into the media content the user code of the authorized user, said device being operable to:

- receive the media content comprising the set of noise sequences embedded therein;
- receive the second subset of noise seeds;
- remove a sub-set of noise sequences from the media content, said sub-set of noise sequences being generated from the received second sub-set of noise seeds and the given noise function; thereby leaving embedded in said media content the subset of noise sequences corresponding to the first subset of noise seeds associated to an element of the first type in the user code;

a device for authenticating an authorized user in a media content distributed over a communication network, said device being operable to:

- correlate the media content with a set of noise sequences generated from the set of distinct noise seeds using the given noise function, to identify a subset of said set of noise sequences embedded into said media content prior to distribution of said media content,
- identify elements of the first type that correspond to a subset of noise seeds that would generate the subset of embedded noise sequences,
- derive a user code from the identified elements of the first type to retrieve the authorized user of the media content.

8. A device for embedding a user code into a media content, said user code being selected from a plurality of available user codes, each available user code comprising elements of an alphabet of at least a first and second types, said first type being associated to a first subset of noise seeds taken from a set of distinct noise seeds, said device being operable to:

- receive a media content comprising a set of noise sequences embedded therein; said noise sequences being generated using the set of distinct noise seeds and a given noise function;
- receive a second subset of the noise seeds,

- remove a sub-set of noise sequences from the media content, said sub-set of noise sequences being generated from the received second sub-set of noise seeds and the given noise function;

wherein the received second sub-set of noise seeds corresponds to the noise seeds from the set of distinct noise seeds that are not associated to an element of the first type.

9. A device according to claim 8, said device being further operable to embed the set of noise sequences into the media content using a scrambling scheme and remove the sub-set of noise sequences from the media content using the same scrambling scheme.

.

10. A device according to the preceding claim 9, said device being further operable to remove a sub-set of noise sequences from the media content according to:

$$p_{k,i} = p_i + \sum_{j=0}^{n-1} e_j s_{j,i} + \sum_{j=0}^{n-1} e_j f_{k,j} s_{j,i} = p_i + \sum_{j=0}^{n-1} e_j (f_{k,j} + 1) s_{j,i}$$

$$(i = 0,1,\ldots,d-1,\ k = 0,1,\ldots,m-1)$$

wherein p_i are the d parts of the media content, e_j is the t-th component among n of the encryption key E, $s_{j,i}$ is the i-th component among d of the j-th noise sequence among n and $f_{k,j}$ is the j-th component of the user code corresponding to user k.

11. A device for identifying a media content authorized user, the authorized user being associated to a user code selected from a plurality of available user codes, each available user code comprising elements of an alphabet of at least a first and second types, said first type being associated to a first subset of noise seeds taken from a set of distinct noise seeds, said device being operable to:

- correlate the media content with a set of noise sequences generated from the set of distinct noise seeds, to identify a subset of said set of

noise sequences embedded into said media content prior to distribution of said media content,

- identify elements of the first type that correspond to a subset of noise seeds that would generate the subset of embedded noise sequences,
- derive a user code from the identified elements of the first type to retrieve the authorized user of the media content.

12. A device according to claim 11, wherein the subset of noise sequences embedded into said media content prior to distribution is embedded using a scrambling scheme to scramble said subset of noise sequences, said device being operable to correlate said media content with the set of noise sequences generated from the set of distinct noise seeds and scrambled with the same scrambling scheme.

13. A device according to claim 12, said device being operable to correlate the media content i using the correlation function:

$$< P_k, S_j >= \frac{\sum_{i=0}^{d-1} p_{k,i} s_{j,i}}{\sum_{i=0}^{d-1} s^2_{j,i}} = \frac{\sum_{i=0}^{d-1} [p_i + \sum_{t=0}^{n-1} e_t (f_{k,t} + 1) s_{t,i}] s_{j,i}}{\sum_{i=0}^{d-1} s^2_{j,i}}$$

$$(j = 0,1,...,n-1,\ k = 0,1,...,m-1)$$

with:

$$f_{k,j} = \begin{cases} 0, & \langle P_k, S_j \rangle \geq T \text{ or } \langle P_k, -S_j \rangle \geq T \\ -1, & \langle P_k, S_j \rangle < T \text{ and } \langle P_k, -S_j \rangle < T \end{cases}$$

$$(j = 0,1,...,n-1,\ k = 0,1,...,m-1)$$

wherein T is a pre-determined threshold allowing deciding whether P_k and S_j are correlated or not, k corresponds to user k, m corresponds to the number of users, S_j is the j-th noise sequence among n with components $s_{j,i}$, d corresponds to the number of part of the media

content and e_t is the t-th component among n of the encryption key E and $f_{k,t}$ is the t-th component of the user code corresponding to user k.

14. A computer program providing computer executable instructions stored on a computer readable medium, which when loaded on to a data processor causes the data processor to perform a method for embedding a user code according to any of the claims 1 to 3.

15. A computer program providing computer executable instructions stored on a computer readable medium, which when loaded on to a data processor causes the data processor to perform a method for identifying a media content authorized user according to any of the claims 4 to 6.

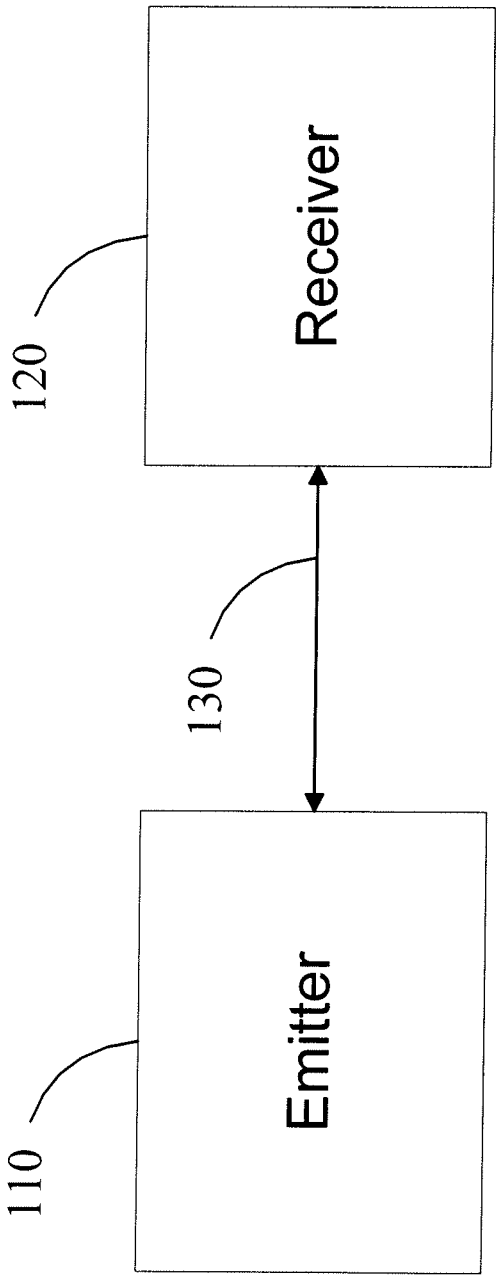


Figure 1

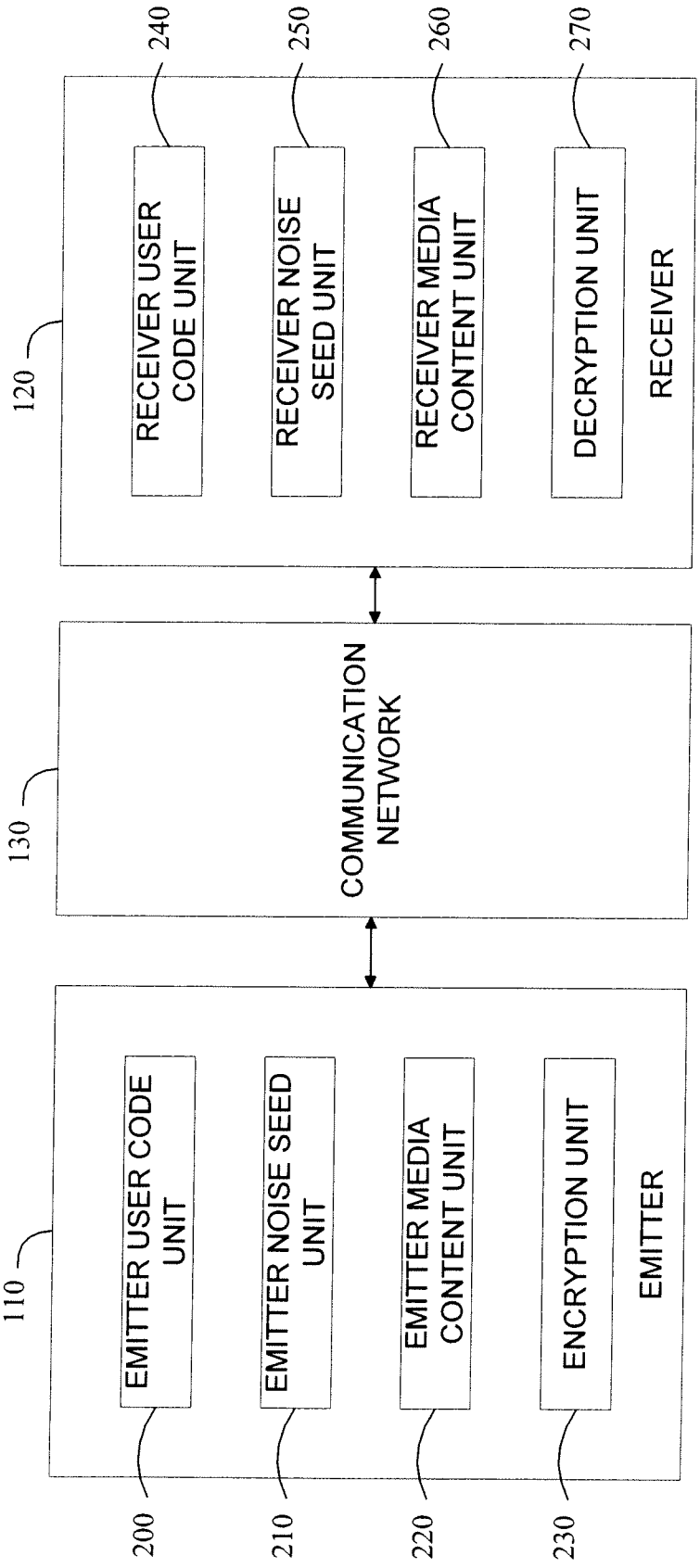

110
EMITTER
EMITTER USER CODE UNIT
EMITTER NOISE SEED UNIT
EMITTER MEDIA CONTENT UNIT
ENCRYPTION UNIT
200
210
220
230
130
COMMUNICATION NETWORK
120
RECEIVER
RECEIVER USER CODE UNIT
RECEIVER NOISE SEED UNIT
RECEIVER MEDIA CONTENT UNIT
DECRYPTION UNIT
240
250
260
270

Figure 2

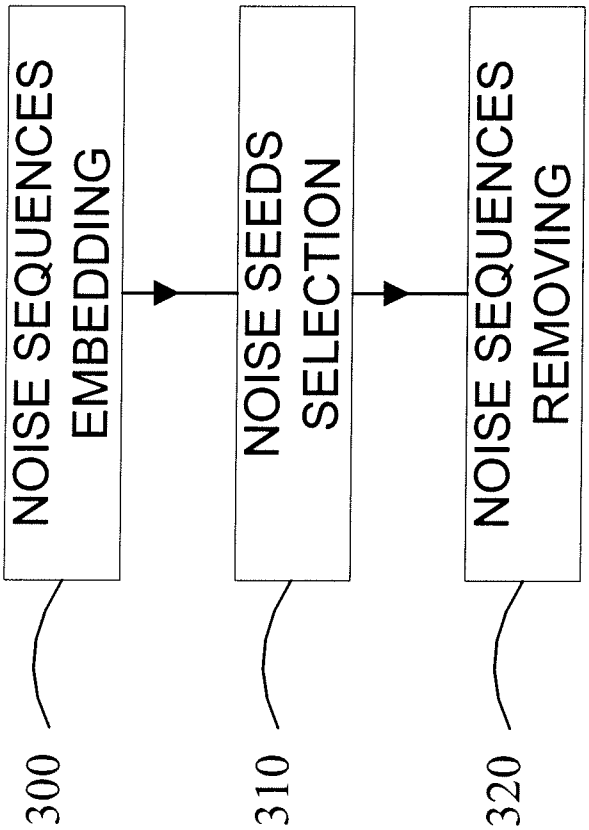

NOISE SEQUENCES EMBEDDING
NOISE SEEDS SELECTION
NOISE SEQUENCES REMOVING
300
310
320

Figure 3

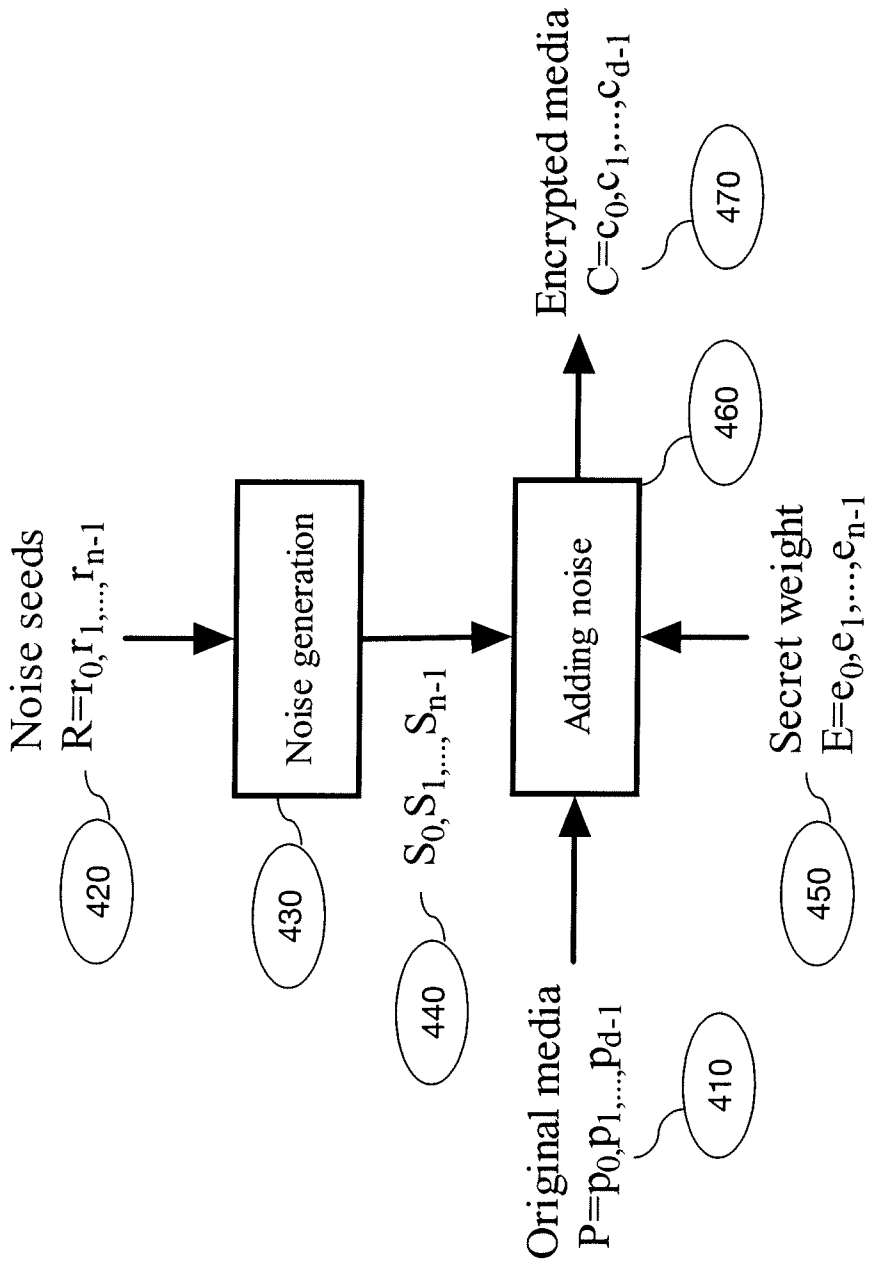

Noise seeds
R=r0,r1,...,rn-1
420
Noise generation
430
S0,S1,...,Sn-1
440
Original media
P=p0,p1,...,pd-1
410
Adding noise
460
Secret weight
E=e0,e1,...,en-1
450
Encrypted media
C=c0,c1,...,cd-1
470

Figure 4A

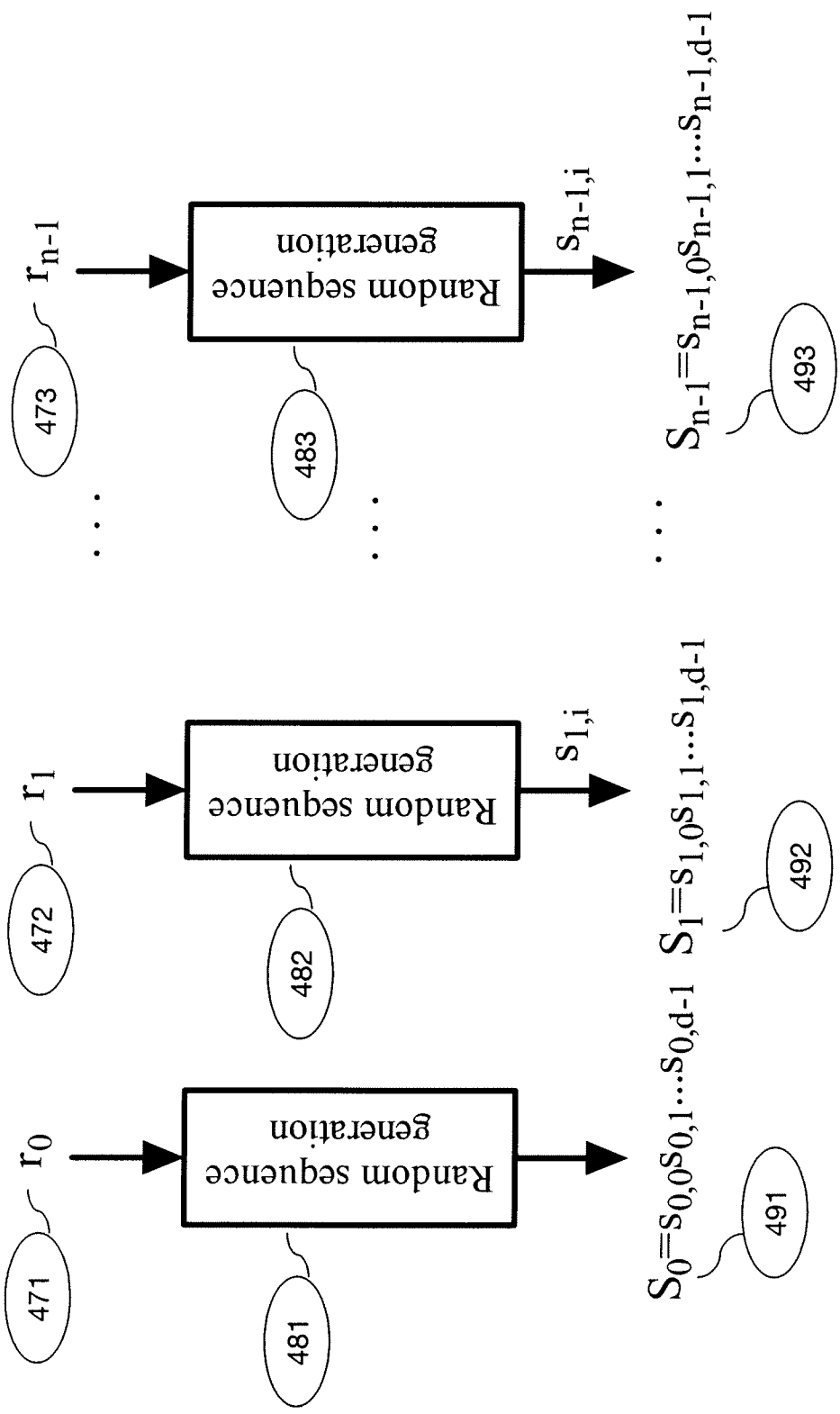

Random sequence generation
Random sequence generation
Random sequence generation
r_0
r_1
r_{n-1}
$s_{1,i}$
$s_{n-1,i}$
$S_0=S_{0,0}S_{0,1}...S_{0,d-1}$
$S_1=S_{1,0}S_{1,1}...S_{1,d-1}$
$S_{n-1}=S_{n-1,0}S_{n-1,1}...S_{n-1,d-1}$
471
472
473
481
482
483
491
492
493

Figure 4B

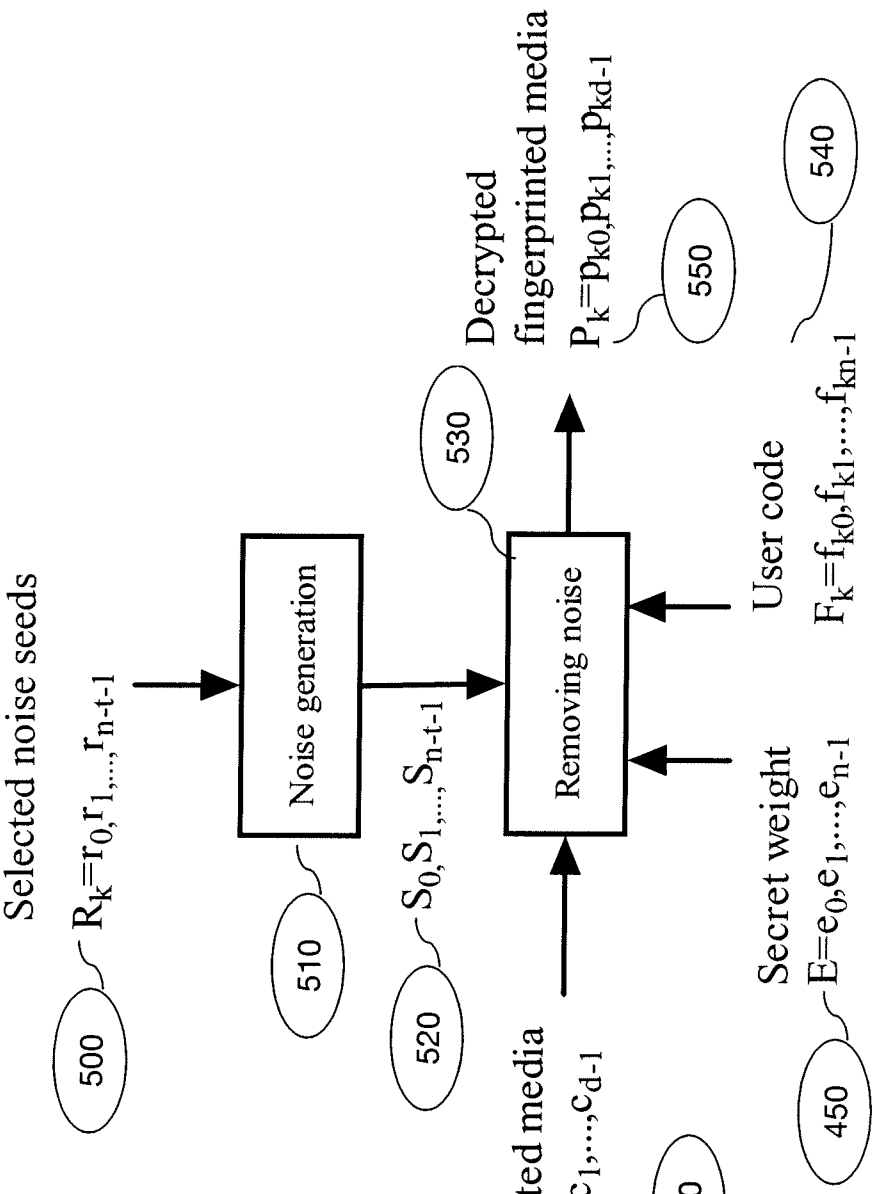

Selected noise seeds
$R_k=r_0,r_1,...,r_{n-t-1}$
500
Noise generation
510
$S_0,S_1,...,S_{n-t-1}$
520
530
Decrypted fingerprinted media
$P_k=p_{k0},p_{k1},...,p_{kd-1}$
550
Encrypted media
$C=c_0,c_1,...,c_{d-1}$
470
Removing noise
User code
$F_k=f_{k0},f_{k1},...,f_{kn-1}$
540
Secret weight
$E=e_0,e_1,...,e_{n-1}$
450

Figure 5

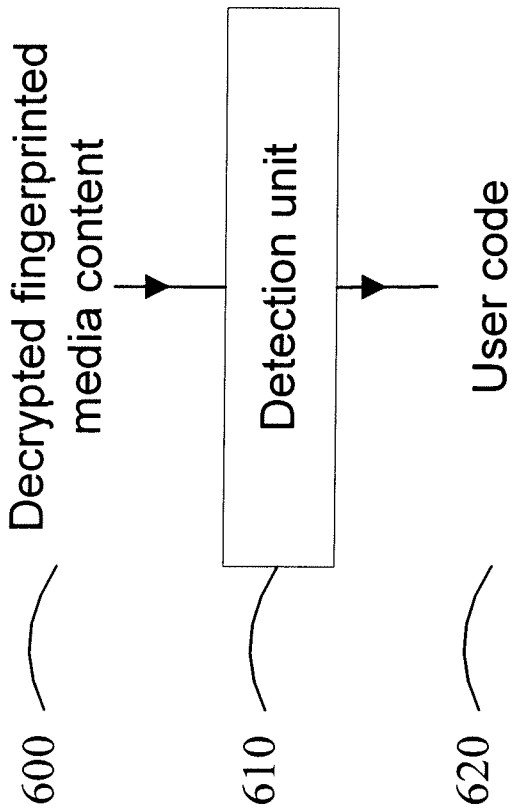

Decrypted fingerprinted media content
Detection unit
User code
600
610
620

Figure 6

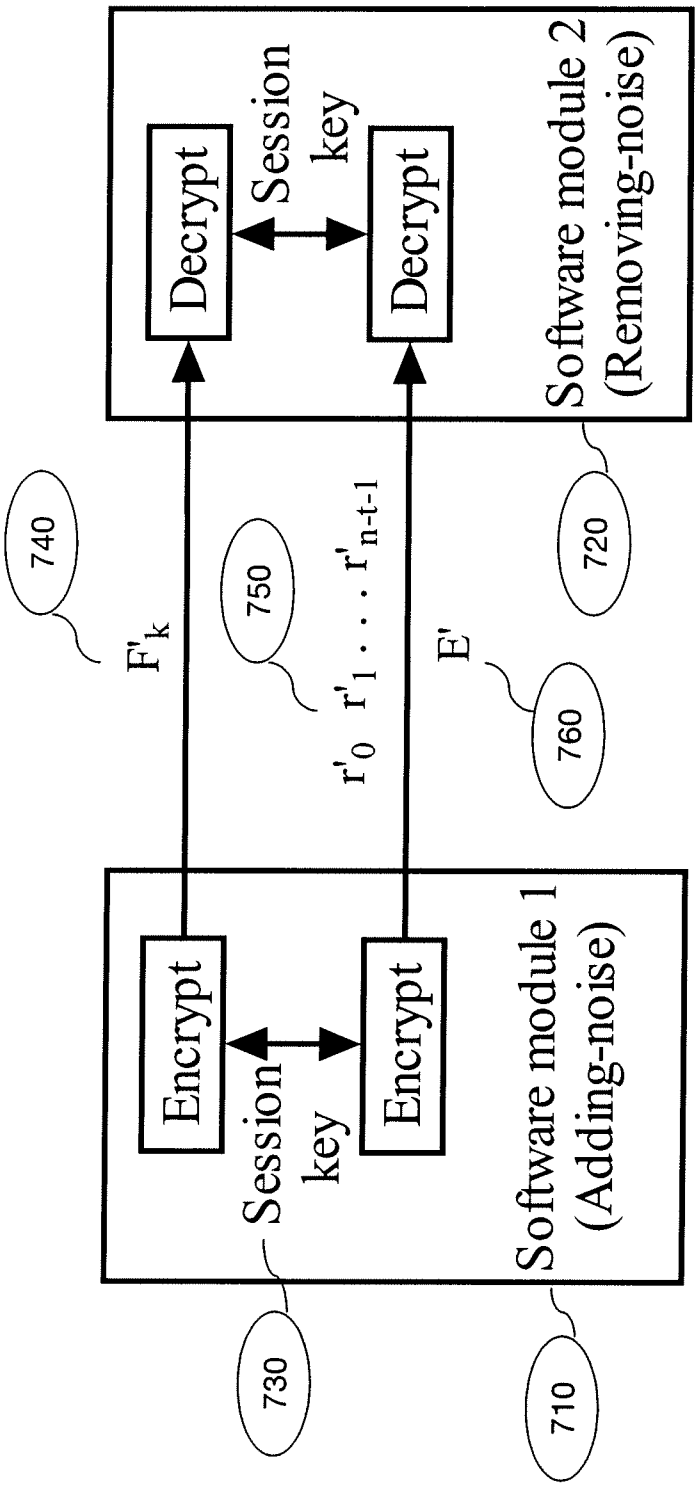

Software module 1
(Adding-noise)
Encrypt
Session
key
Encrypt
Software module 2
(Removing-noise)
Decrypt
Session
key
Decrypt
F'_k
r'_0 r'_1 . . . r'_{n-t-1}
E'
730
710
740
750
760
720

Figure 7

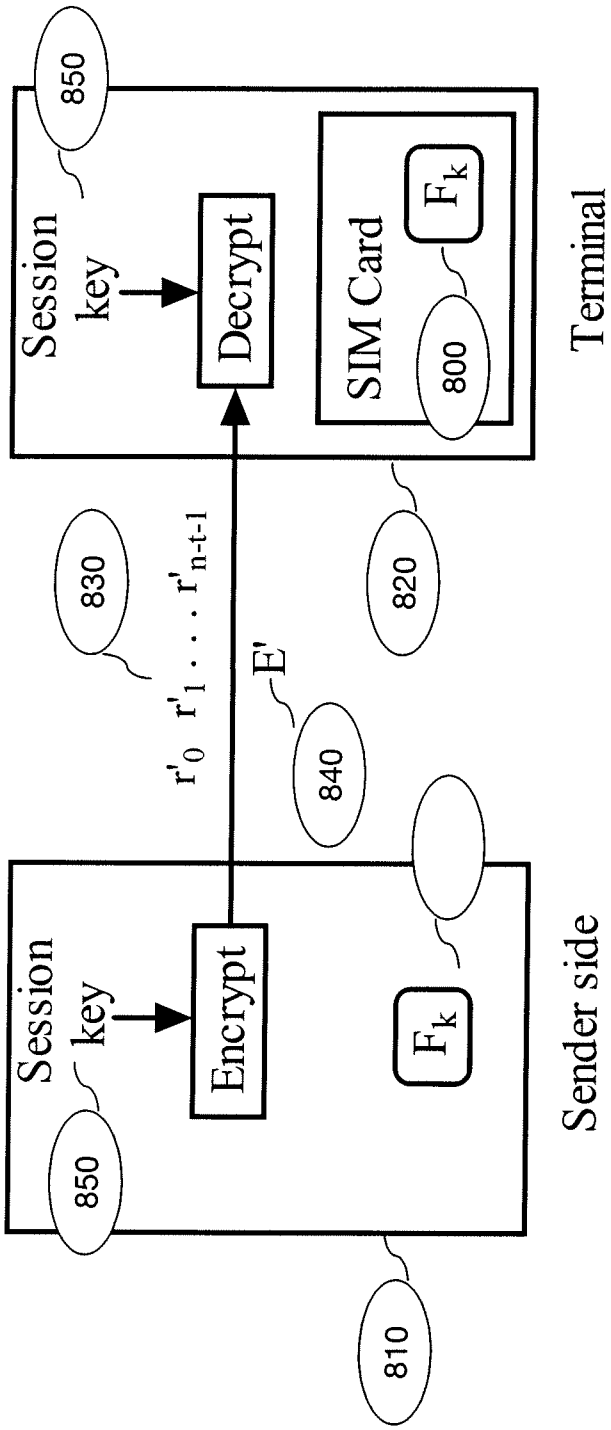
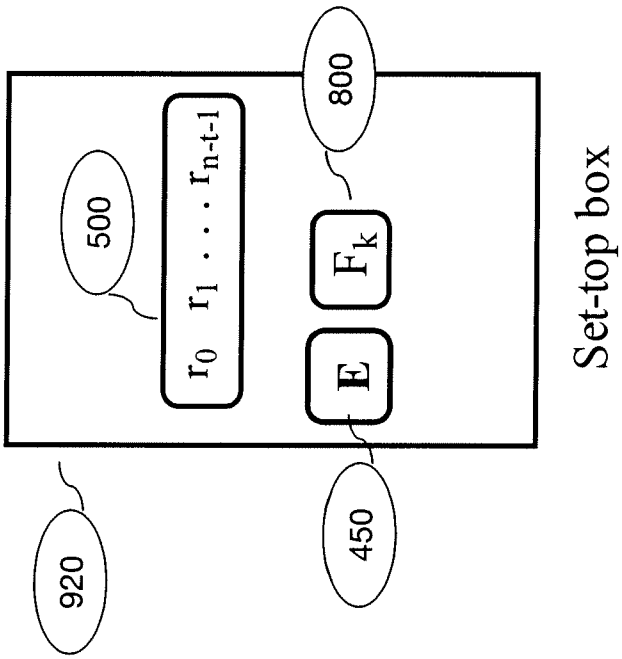

Session key
850
Encrypt
F_k
Sender side
810
r'_0 $r'_1 \ldots r'_{n-t-1}$
830
E'
840
820
Session key
850
Decrypt
SIM Card
F_k
800
Terminal

Figure 8


920
500
r_0 r_1 . . . r_{n-t-1}
E
F_k
800
450
Set-top box

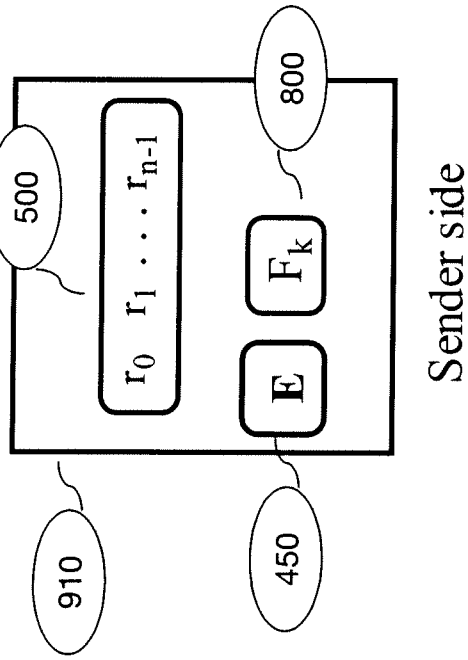

910
500
r_0 r_1 . . . r_{n-1}
E
F_k
800
450
Sender side

Figure 9