



US 20160004875A1

(19) **United States**  
(12) **Patent Application Publication**  
**Moskowitz**

(10) **Pub. No.: US 2016/0004875 A1**  
(43) **Pub. Date: Jan. 7, 2016**

(54) **DATA PROTECTION METHOD AND DEVICE**

Aug. 24, 2007, now Pat. No. 9,104,842, which is a division of application No. 10/602,777, filed on Jun. 25, 2003, now Pat. No. 7,664,263, which is a continuation of application No. 09/046,627, filed on Mar. 24, 1998, now Pat. No. 6,598,162.

(71) Applicant: **WISTARIA TRADING LTD**, Hamilton HM (BM)

(72) Inventor: **Scott A. Moskowitz**, Ft. Lauderdale, FL (US)

(73) Assignee: **WISTARIA TRADING LTD**, Hamilton HM (BM)

(21) Appl. No.: **14/855,281**

(22) Filed: **Sep. 15, 2015**

**Publication Classification**

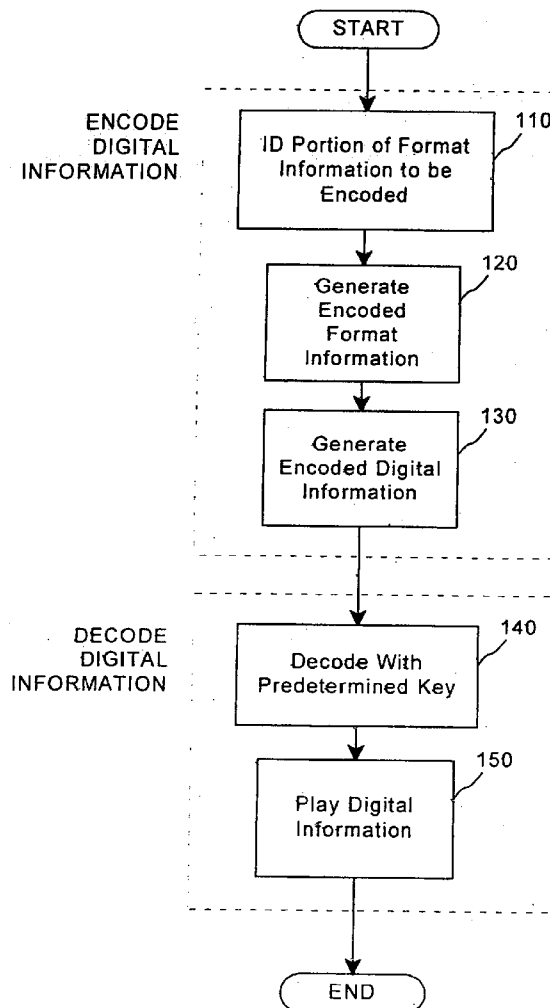
(51) **Int. Cl.**  
**G06F 21/60** (2006.01)  
**G06F 21/62** (2006.01)  
(52) **U.S. Cl.**  
CPC ..... **G06F 21/602** (2013.01); **G06F 21/6209** (2013.01)

**Related U.S. Application Data**

(60) Continuation of application No. 14/542,712, filed on Nov. 17, 2014, now Pat. No. 9,171,136, which is a continuation of application No. 13/556,420, filed on Jul. 24, 2012, now Pat. No. 8,930,719, which is a continuation of application No. 11/895,388, filed on

(57) **ABSTRACT**

An apparatus and method for encoding and decoding additional information into a digital information in an integral manner. More particularly, the invention relates to a method and device for data protection.



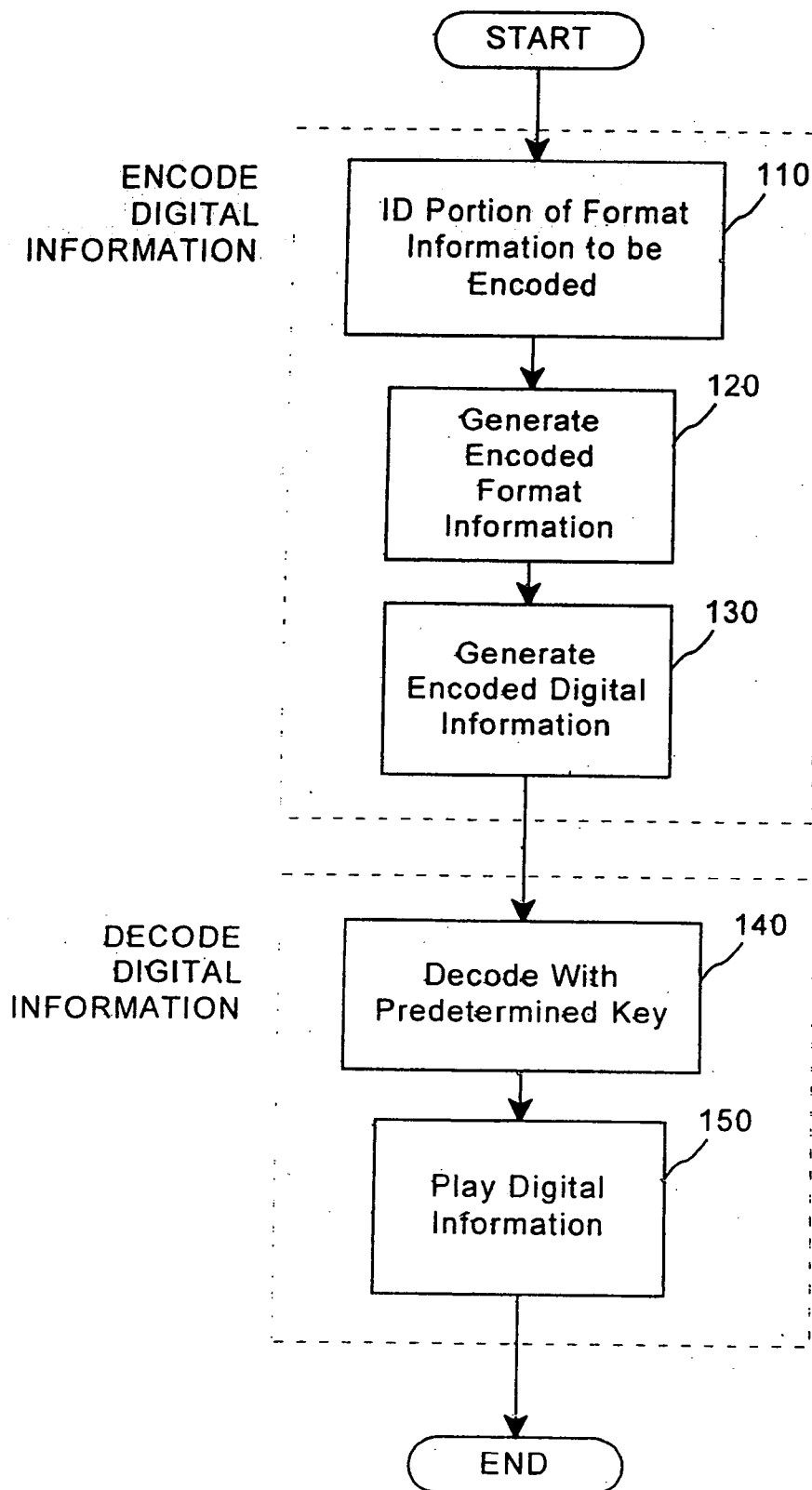


FIG. 1

## DATA PROTECTION METHOD AND DEVICE

### CROSS-REFERENCE TO RELATED APPLICATIONS

**[0001]** This application is a continuation of U.S. application Ser. No. 14/542,712, filed Nov. 17, 2014, which is a continuation of U.S. application Ser. No. 13/556,420, filed Jul. 24, 2012, issued Jan. 6, 2015 as U.S. Pat. No. 8,930,719, which is a continuation of U.S. application Ser. No. 11/895,388, filed Aug. 24, 2007, issued Aug. 11, 2015 as U.S. Pat. No. 9,104,842, which is a division of U.S. patent application Ser. No. 10/602,777, filed Jun. 25, 2003, issued Feb. 16, 2010 as U.S. Pat. No. 7,664,263, which is a continuation of U.S. patent application Ser. No. 09/046,627, filed Mar. 24, 1998, issued Jul. 22, 2003, as U.S. Pat. No. 6,598,162. The entire disclosure of U.S. application Ser. No. 14/542,712, filed Nov. 17, 2014, U.S. application Ser. No. 13/556,420, filed Jul. 24, 2012, U.S. application Ser. No. 11/895,388, filed Aug. 24, 2007, U.S. patent application Ser. No. 09/046,627, issued Jul. 22, 2003, as U.S. Pat. No. 6,598,162, are hereby incorporated by reference in their entireties.

### FIELD OF THE INVENTION

**[0002]** The invention relates to the protection of digital information. More particularly, the invention relates to a method and device for data protection.

**[0003]** With the advent of computer networks and digital multimedia, protection of intellectual property has become a prime concern for creators and publishers of digitized copies of copyrightable works, such as musical recordings, movies, video games, and computer software. One method of protecting copyrights in the digital domain is to use “digital watermarks.”

**[0004]** The prior art includes copy protection systems attempted at many stages in the development of the software industry. These may be various methods by which a software engineer can write the software in a clever manner to determine if it has been copied, and if so to deactivate itself. Also included are undocumented changes to the storage format of the content. Copy protection was generally abandoned by the software industry, since pirates were generally just as clever as the software engineers and figured out ways to modify the software and deactivate the protection. The cost of developing such protection was not justified considering the level of piracy which occurred despite the copy protection.

**[0005]** Other methods for protection of computer software include the requirement of entering certain numbers or facts that may be included in a packaged software’s manual, when prompted at start-up. These may be overcome if copies of the manual are distributed to unintended users, or by patching the code to bypass these measures. Other methods include requiring a user to contact the software vendor and to receive “keys” for unlocking software after registration attached to some payment scheme, such as credit card authorization. Further methods include network-based searches of a user’s hard drive and comparisons between what is registered to that user and what is actually installed on the user’s general computing device. Other proposals, by such parties as AT&T’s Bell Laboratories, use “kerning” or actual distance in pixels, in the rendering of text documents, rather than a varied number of ASCII characters. However, this approach can often be defeated by graphics processing analogous to sound process-

ing, which randomizes that information. All of these methods require outside determination and verification of the validity of the software license.

**[0006]** Digital watermarks can be used to mark each individual copy of a digitized work with information identifying the title, copyright holder, and even the licensed owner of a particular copy. When marked with licensing and ownership information, responsibility is created for individual copies where before there was none. Computer application programs can be watermarked by watermarking digital content resources used in conjunction with images or audio data. Digital watermarks can be encoded with random or pseudo-random keys, which act as secret maps for locating the watermarks. These keys make it impossible for a party to find the watermark without having the key. In addition, the encoding method can be enhanced to force a party to cause damage to a watermarked data stream when trying to erase a random-key watermark. Other information is disclosed in “Technology: Digital Commerce”, Denise Caruso, New York Times, Aug. 7, 1995; and “Copyrighting in the Information Age”, Harley Ungar, ONLINE MARKETPLACE, September 1995, Jupiter Communications.

**[0007]** Additionally, other methods for hiding information signals in content signals, are disclosed in U.S. Pat. No. 5,319,735—Preuss et al. and U.S. Pat. No. 5,379,345—Greenberg.

**[0008]** It is desirable to use a “stega-cipher” or watermarking process to hide the necessary parts or resources of the executable object code in the digitized sample resources. It is also desirable to further modify the underlying structure of an executable computer application such that it is more resistant to attempts at patching and analysis by memory capture. A computer application seeks to provide a user with certain utilities or tools, that is, users interact with a computer or similar device to accomplish various tasks and applications provide the relevant interface. Thus, a level of authentication can also be introduced into software, or “digital products,” that include digital content, such as audio, video, pictures or multimedia, with digital watermarks. Security is maximized because erasing this code watermark without a key results in the destruction of one or more essential parts of the underlying application, rendering the “program” useless to the unintended user who lacks the appropriate key. Further, if the key is linked to a license code by means of a mathematical function, a mechanism for identifying the licensed owner of an application is created.

**[0009]** It is also desirable to randomly reorganize program memory structure intermittently during program run time, to prevent attempts at memory capture or object code analysis aimed at eliminating licensing or ownership information, or otherwise modifying, in an unintended manner, the functioning of the application.

**[0010]** In this way, attempts to capture memory to determine underlying functionality or provide a “patch” to facilitate unauthorized use of the “application,” or computer program, without destroying the functionality and thus usefulness of a copyrightable computer program can be made difficult or impossible.

**[0011]** It is thus the goal of the present invention to provide a higher level of copyright security to object code on par with methods described in digital watermarking systems for digitized media content such as pictures, audio, video and multimedia content in its multifarious forms, as described in previous disclosures, “Steganographic Method and Device” Ser.

No. 08/489,172, filed Jun. 7, 1995, now U.S. Pat. No. 5,613,004, and “Human Assisted Random Key Generation and Application for Digital Watermark System”, Ser. No. 08/587,944, filed on Jan. 17, 1996, the disclosure of which is hereby incorporated by reference.

**[0012]** It is a further goal of the present invention to establish methods of copyright protection that can be combined with such schemes as software metering, network distribution of code and specialized protection of software that is designed to work over a network, such as that proposed by Sun Microsystems in their HotJava browser and Java programming language, and manipulation of application code in proposed distribution of documents that can be exchanged with resources or the look and feel of the document being preserved over a network. Such systems are currently being offered by companies including Adobe, with their Acrobat software. This latter goal is accomplished primarily by means of the watermarking of font, or typeface, resources included in applications or documents, which determine how a bitmap representation of the document is ultimately drawn on a presentation device.

**[0013]** The present invention includes an application of the technology of “digital watermarks.” As described in previous disclosures, “Steganographic Method and Device” and “Human Assisted Random Key Generation and Application for Digital Watermark System,” watermarks are particularly suitable to the identification, metering, distributing and authenticating digitized content such as pictures, audio, video and derivatives thereof under the description of “multimedia content.” Methods have been described for combining both cryptographic methods, and steganography, or hiding something in plain view. Discussions of these technologies can be found in Applied Cryptography by Bruce Schneier and The Code Breakers by David Kahn. For more information on prior art public-key cryptosystems see U.S. Pat. No. 4,200,770 Diffie-Hellman, U.S. Pat. No. 4,218,582 Hellman, U.S. Pat. No. 4,405,829 RSA, U.S. Pat. No. 4,424,414 Hellman Pohl. Computer code, or machine language instructions, which are not digitized and have zero tolerance for error, must be protected by derivative or alternative methods, such as those disclosed in this invention, which focuses on watermarking with “keys” derived from license codes or other ownership identification information, and using the watermarks encoded with such keys to hide an essential subset of the application code resources.

#### BACKGROUND OF THE INVENTION

**[0014]** Increasingly, commercially valuable information is being created and stored in “digital” form. For example, music, photographs and video can all be stored and transmitted as a series of numbers, such as 1’s and 0’s. Digital techniques let the original information be recreated in a very accurate manner. Unfortunately, digital techniques also let the information be easily copied without the information owner’s permission.

**[0015]** Because unauthorized copying is clearly a disincentive to the digital distribution of valuable information, it is important to establish responsibility for copies and derivative copies of such works. For example, if each authorized digital copy of a popular song is identified with a unique number, any unauthorized copy of the song would also contain the number. This would allow the owner of the information, such as a song publisher, to investigate who made the unauthorized copy. Unfortunately, it is possible that the unique number could be

erased or altered if it is simply tacked on at the beginning or end of the digital information.

**[0016]** As will be described, known digital “watermark” techniques give creators and publishers of digitized multimedia content localized, secured identification and authentication of that content. In considering the various forms of multimedia content, such as “master,” stereo, National Television Standards Committee (NTSC) video, audio tape or compact disc, tolerance of quality will vary with individuals and affect the underlying commercial and aesthetic value of the content. For example, if a digital version of a popular song sounds distorted, it will be less valuable to users. It is therefore desirable to embed copyright, ownership or purchaser information, or some combination of these and related data, into the content in a way that will damage the content if the watermark is removed without authorization.

**[0017]** To achieve these goals, digital watermark systems insert ownership information in a way that causes little or no noticeable effects, or “artifacts,” in the underlying content signal. For example, if a digital watermark is inserted into a digital version of a song, it is important that a listener not be bothered by the slight changes introduced by the watermark. It is also important for the watermark technique to maximize the encoding level and “location sensitivity” in the signal to force damage to the content signal when removal is attempted. Digital watermarks address many of these concerns, and research in the field has provided extremely robust and secure implementations.

**[0018]** What has been overlooked in many applications described in the art, however, are systems which closely mimic distribution of content as it occurs in the real world. For instance, many watermarking systems require the original un-watermarked content signal to enable detection or decode operations. These include highly publicized efforts by NEC, Digimarc and others. Such techniques are problematic because, in the real world, original master copies reside in a rights holders vaults and are not readily available to the public.

**[0019]** With much activity overly focused on watermark survivability, the security of a digital watermark is suspect. Any simple linear operation for encoding information into a signal may be used to erase the embedded signal by inverting the process. This is not a difficult task, especially when detection software is a plug-in freely available to the public, such as with Digimarc. In general, these systems seek to embed cryptographic information, not cryptographically embed information into target media content.

**[0020]** Other methods embed ownership information that is plainly visible in the media signal, such as the method described in U.S. Pat. No. 5,530,739 to Braudaway et al. The system described in Braudaway protects a digitized image by encoding a visible watermark to deter piracy. Such an implementation creates an immediate weakness in securing the embedded information because the watermark is plainly visible. Thus, no search for the embedded signal is necessary and the watermark can be more easily removed or altered. For example, while certainly useful to some rights owners, simply placing the symbol “C” in the digital information would only provide limited protection. Removal by adjusting the brightness of the pixels forming the “C” would not be difficult with respect to the computational resources required.

**[0021]** Other relevant prior art includes U.S. Pat. Nos. 4,979,210 and 5,073,925 to Nagata et al., which encodes information by modulating an audio signal in the amplitude/

time domain. The modulations introduced in the Nagata process carry a “copy/don’t copy” message, which is easily found and circumvented by one skilled in the art. The granularity of encoding is fixed by the amplitude and frequency modulation limits required to maintain inaudibility. These limits are relatively low, making it impractical to encode more information using the Nagata process.

**[0022]** Although U.S. Pat. No. 5,661,018 to Leighton describes a means to prevent collusion attacks in digital watermarks, the disclosed method may not actually provide the security described. For-example, in cases where the watermarking technique is linear, the “insertion envelope” or “watermarking space” is well-defined and thus susceptible to attacks less sophisticated than collusion by unauthorized parties. Over-encoding at the watermarking encoding level is but one simple attack in such linear implementations. Another consideration not made by Leighton is that commercially-valuable content may already exist in a un-watermarked form somewhere, easily accessible to potential pirates, gutting the need for any type of collusive activity. Digitally signing the embedded signal with preprocessing of watermark data is more likely to prevent successful collusion. Furthermore, a “baseline” watermark as disclosed is quite subjective. It is simply described elsewhere in the art as the “perceptually significant” regions of a signal. Making a watermarking function less linear or inverting the insertion of watermarks would seem to provide the same benefit without the additional work required to create a “baseline” watermark. Indeed, watermarking algorithms should already be capable of defining a target insertion envelope or region without additional steps. What is evident is the Leighton patent does not allow for initial prevention of attacks on an embedded watermark as the content is visibly or audibly unchanged.

**[0023]** It is also important that any method for providing security also function with broadcasting media over networks such as the Internet, which is also referred to as “streaming.” Commercial “plug-in” products such as RealAudio and RealVideo, as well as applications by vendors VDONet and Xtreme, are common in such network environments. Most digital watermark implementations focus on common file base signals and fail to anticipate the security of streamed signals. It is desirable that any protection scheme be able to function with a plug-in player without advanced knowledge of the encoded media stream.

**[0024]** Other technologies focus solely on file-based security. These technologies illustrate the varying applications for security that must be evaluated for different media and distribution environments. Use of cryptolopes or cryptographic containers, as proposed by IBM in its Cryptolope product, and InterTrust, as described in U.S. Pat. Nos. 4,827,508, 4,977,594, 5,050,213 and 5,410,598, may discourage certain forms of piracy. Cryptographic containers, however, require a user to subscribe to particular decryption software to decrypt data. IBM’s InfoMarket and InterTrust’s DigiBox, among other implementations, provide a generalized model and need proprietary architecture to function. Every user must have a subscription or registration with the party which encrypts the data. Again, as a form of general encryption, the data is scrambled or encrypted without regard to the media and its formatting. Finally, control over copyrights or other neighboring rights is left with the implementing party, in this case, IBM, InterTrust or a similar provider.

**[0025]** Methods similar to these “trusted systems” exist, and Cerberus Central Limited and Liquid Audio, among a

number of companies, offer systems which may functionally be thought of as subsets of IBM and InterTrust’s more generalized security offerings. Both Cerberus and Liquid Audio propose proprietary player software which is registered to the user and “locked” in a manner parallel to the locking of content that is distributed via a cryptographic container. The economic trade-off in this model is that users are required to use each respective companies’ proprietary player to play or otherwise manipulate content that is downloaded. If, as is the case presently, most music or other media is not available via these proprietary players and more companies propose non-compatible player formats, the proliferation of players will continue. Cerberus and Liquid Audio also by way of extension of their architectures provide for “near-CD quality” but proprietary compression. This requirement stems from the necessity not to allow content that has near-identical data make-up to an existing consumer electronic standard, in Cerberus and Liquid Audio’s case the so-called Red Book audio CD standard of 16 bit 44.1 kHz, so that comparisons with the proprietary file may not yield how the player is secured. Knowledge of the player’s file format renders its security ineffective as a file may be replicated and played on any common player, not the intended proprietary player of the provider of previously secured and uniquely formatted content. This is the parallel weakness to public key crypto-systems which have gutted security if enough plain text and cipher text comparisons enable a pirate to determine the user’s private key.

**[0026]** Many approaches to digital watermarking leave detection and decoding control with the implementing party of the digital watermark, not the creator of the work to be protected. A set of secure digital watermark implementations address this fundamental control issue forming the basis of key-based approaches. These are covered by the following patents and pending applications, the entire disclosures of which are hereby incorporated by reference: U.S. Pat. No. 5,613,004 entitled “Steganographic Method and Device” and its derivative U.S. patent application Ser. No. 08/775,216 (which issued Nov. 11, 1997, as U.S. Pat. No. 5,687,236), U.S. patent application Ser. No. 08/587,944 entitled “Human Assisted Random Key Generation and Application for Digital Watermark System” (which issued Oct. 13, 1998, as U.S. Pat. No. 5,822,432), U.S. patent application Ser. No. 08/587,943 entitled “Method for Stega-Cipher Protection of Computer Code” (which issued Apr. 28, 1998, as U.S. Pat. No. 5,748,569), U.S. patent application Ser. No. 08/677,435 entitled “Optimization Methods for the Insertion, Protection, and Detection of Digital Watermarks in Digitized Data” (which issued Mar. 30, 1999, as U.S. Pat. No. 5,889,868) and U.S. patent application Ser. No. 08/772,222 entitled “Z-Transform Implementation of Digital Watermarks” (which issued Jun. 20, 2000, as U.S. Pat. No. 6,078,664). Public key crypto-systems are described in U.S. Pat. Nos. 4,200,770, 4,218,582, 4,405,829 and 4,424,414, the entire disclosures of which are also hereby incorporated by reference.

**[0027]** In particular, an improved protection scheme is described in “Method for Stega-Cipher Protection of Computer Code,” U.S. patent application Ser. No. 08/587,943 (which issued Apr. 28, 1998, as U.S. Pat. No. 5,748,569). This technique uses the key-based insertion of binary executable computer code within a content signal that is subsequently, and necessarily, used to play or otherwise manipulate the signal in which it is encoded. With this system, however, certain computational requirements, such as one digital

player per digital copy of content, may be necessitated. For instance, a consumer may download many copies of watermarked content. With this technique, the user would also be downloading as many copies of the digital player program. While this form of security may be desirable for some applications, it is not appropriate in many circumstances.

**[0028]** Finally, even when digital information is distributed in encoded form, it may be desirable to allow unauthorized users to play the information with a digital player, perhaps with a reduced level of quality. For example, a popular song may be encoded and freely distributed in encoded form to the public. The public, perhaps using commonly available plug-in digital players, could play the encoded content and hear the music in some degraded form. The music may sound choppy, or fuzzy or be degraded in some other way. This lets the public decide, based on the available lower quality version of the song, if they want to purchase a key from the publisher to decode, or “clean-up,” the content. Similar approaches could be used to distribute blurry pictures or low quality video. Or even “degraded” text, in the sense that only authenticated portions of the text can be determined with the predetermined key or a validated digital signature for the intended message.

**[0029]** In view of the foregoing, it can be appreciated that a substantial need exists for a method allowing encoded content to be played, with degraded quality, by a plug-in digital player, and solving the other problems discussed above.

#### SUMMARY OF THE INVENTION

**[0030]** The disadvantages of the art are alleviated to a great extent by a method for combining transfer functions with predetermined key creation. In one embodiment, digital information, including a digital sample and format information, is protected by identifying and encoding a portion of the format information. Encoded digital information, including the digital sample and the encoded format information, is generated to protect the original digital information.

**[0031]** In another embodiment, a digital signal, including digital samples in a file format having an inherent granularity, is protected by creating a predetermined key. The predetermined key is comprised of a transfer function-based mask set to manipulate data at the inherent granularity of the file format of the underlying digitized samples.

**[0032]** It is thus a goal of the present invention, to provide a level of security for executable code on similar grounds as that which can be provided for digitized samples. Furthermore, the present invention differs from the prior art in that it does not attempt to stop copying, but rather, determines responsibility for a copy by ensuring that licensing information must be preserved in descendant copies from an original. Without the correct license information, the copy cannot function.

**[0033]** An improvement over the art is disclosed in the present invention, in that the software itself is a set of commands, compiled by software engineer, which can be configured in such a manner as to tie underlying functionality to the license or authorization of the copy in possession by the user. Without such verification, the functions sought out by the user in the form of software cease to properly work. Attempts to tamper or “patch” substitute code resources can be made highly difficult by randomizing the location of said resources in memory on an intermittent basis to resist most attacks at disabling the system.

**[0034]** With these and other advantages and features of the invention that will become hereinafter apparent, the nature of

the invention may be more clearly understood by reference to the following detailed description of the invention, the appended claims and to the several drawings attached herein.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**[0035]** FIG. 1 is a block flow diagram of a method for copy protection or authentication of digital information according to an embodiment of the present invention.

#### DETAILED DESCRIPTION

**[0036]** In accordance with an embodiment of the present invention, a method combines transfer functions with predetermined key creation. Increased security is achieved in the method by combining elements of “public-key steganography” with cryptographic protocols, which keep in-transit data secure by scrambling the data with “keys” in a manner that is not apparent to those with access to the content to be distributed. Because different forms of randomness are combined to offer robust, distributed security, the present invention addresses an architectural “gray space” between two important areas of security: digital watermarks, a subset of the more general art of steganography, and cryptography. One form of randomness exists in the mask sets that are randomly created to map watermark data into an otherwise unrelated digital signal. The second form of randomness is the random permutations of data formats used with digital players to manipulate the content with the predetermined keys. These forms can be thought of as the transfer function versus the mapping function inherent to digital watermarking processes.

**[0037]** According to an embodiment of the present invention, a predetermined, or randomly generated, key is used to scramble digital information in a way that is unlike known “digital watermark” techniques and public key cryptosystems. As used herein, a key is also referred to as a “mask set” which includes one or more random or pseudo-random series of bits. Prior to encoding, a mask can be generated by any cryptographically secure random generation process. A block cipher, such as a Data Encryption Standard (DES) algorithm, in combination with a sufficiently random seed value, such as one created using a Message Digest 5 (MD5) algorithm, emulates a cryptographically secure random bit generator. The keys are saved in a database, along with information matching them to the digital signal, for use in descrambling and subsequent viewing or playback. Additional file format or transfer property information is prepared and made available to the encoder, in a bit addressable manner. As well, any authenticating function can be combined, such as Digital Signature Standard (DSS) or Secure Hash Algorithm (SHA).

**[0038]** Using the predetermined key comprised of a transfer function-based mask set, the data representing the original content is manipulated at the inherent granularity of the file format of the underlying digitized samples. Instead of providing, or otherwise distributing, watermarked content that is not noticeably altered, a partially “scrambled” copy of the content is distributed. The key is necessary both to register the sought-after content and to descramble the content into its original form.

**[0039]** The present invention uses methods disclosed in “Method for Stega-Cipher Protection of Computer Code,” U.S. patent application Ser. No. 08/587,943 (which issued Apr. 28, 1998, as U.S. Pat. No. 5,748,569), with respect to transfer functions related to the common file formats, such as PICT, TIFF, AIFF, WAV, etc. Additionally, in cases where the

content has not been altered beyond being encoded with such functional data, it is possible for a digital player to still play the content because the file format has not been altered. Thus, the encoded content could still be played by a plug-in digital player as discrete, digitally sampled signals, watermarked or not. That is, the structure of the file can remain basically unchanged by the watermarking process, letting common file format based players work with the “scrambled” content.

**[0040]** For example, the Compact Disc-Digital Audio (CD-DA) format stores audio information as a series of frames. Each frame contains a number of digital samples representing, for example, music, and a header that contains file format information. As shown in FIG. 1, according to an embodiment of the present invention some of the header information can be identified and “scrambled” using the predetermined key at steps 110 to 130. The music samples can remain unchanged. Using this technique, a traditional CD-DA player will be able to play a distorted version of the music in the sample. The amount of distortion will depend on the way, and extent, that the header, or file format, information has been scrambled. It would also be possible to instead scramble some of the digital samples while leaving the header information alone. In general, the digital signal would be protected by manipulating data at the inherent granularity, or “frames,” of the CD-DA file format. To decode the information, a predetermined key is used before playing the digital information at steps 140 and 150.

**[0041]** A key-based decoder can act as a “plug-in” digital player of broadcast signal streams without foreknowledge of the encoded media stream. Moreover, the data format orientation is used to partially scramble data in transit to prevent unauthorized descrambled access by decoders that lack authorized keys. A distributed key can be used to unscramble the scrambled content because a decoder would understand how to process the key. Similar to on-the-fly decryption operations, the benefits inherent in this embodiment include the fact that the combination of watermarked content security, which is key-based, and the descrambling of the data, can be performed by the same key which can be a plurality of mask sets. The mask sets may include primary, convolution and message delimiter masks with file format data included.

**[0042]** The creation of an optimized “envelope” for insertion of watermarks provides the basis of much watermark security, but is also a complementary goal of the present invention. The predetermined or random key that is generated is not only an essential map to access the hidden information signal, but is also the descrambler of the previously scrambled signal’s format for playback or viewing.

**[0043]** In a system requiring keys for watermarking content and validating the distribution of the content, different keys may be used to encode different information while secure one way hash functions or one-time pads may be incorporated to secure the embedded signal. The same keys can be used to later validate the embedded digital signature, or even fully decode the digital watermark if desired. Publishers can easily stipulate that content not only be digitally watermarked but that distributors must check the validity of the watermarks by performing digital signature-checks with keys that lack any other functionality. The system can extend to simple authentication of text in other embodiments.

**[0044]** Before such a market is economically feasible, there are other methods for deploying key-based watermarking coupled with transfer functions to partially scramble the content to be distributed without performing full public key

encryption, i.e., a key pair is not necessarily generated, simply, a predetermined key’s function is created to re-map the data of the content file in a lossless process. Moreover, the scrambling performed by the present invention may be more dependent on the file in question. Dissimilarly, encryption is not specific to any particular media but is performed on data. The file format remains unchanged, rendering the file useable by any conventional viewer/player, but the signal quality can be intentionally degraded in the absence of the proper player and key. Public-key encryption seeks to completely obscure the sensitive “plaintext” to prevent comparisons with the “ciphertext” to determine a user’s private keys. Centralized encryption only differs in the utilization of a single key for both encryption and decryption making the key even more highly vulnerable to attacks to defeat the encryption process. With the present invention, a highly sought after photograph may be hazy to the viewer using any number of commonly available, nonproprietary software or hardware, without the authorized key. Similarly, a commercially valuable song may sound poor.

**[0045]** The benefit of some form of cryptography is not lost in the present invention. In fact, some piracy can be deterred when the target signal may be known but is clearly being protected through scrambling. What is not anticipated by known techniques, is an ala carte method to change various aspects of file formatting to enable various “scrambled states” for content to be subsequently distributed. An image may lack all red pixels or may not have any of the most significant bits activated. An audio sample can similarly be scrambled to render it less-than-commercially viable.

**[0046]** The present invention also provides improvements over known network-based methods, such as those used for the streaming of media data over the Internet. By manipulating file formats, the broadcast media, which has been altered to “fit” within electronic distribution parameters, such as bandwidth availability and error correction considerations, can be more effectively utilized to restrict the subsequent use of the content while in transit as well as real-time viewing or playing.

**[0047]** The mask set providing the transfer function can be read on a per-use basis by issuing an authorized or authenticating “key” for descrambling the signal that is apparent to a viewer or a player or possessor of the authenticating key. The mask set can be read on a per-computer basis by issuing the authorized key that is more generalized for the computer that receives the broadcast signals. Metering and subscription models become viable advantages over known digital watermark systems which assist in designating the ownership of a copy of digitized media content, but do not prevent or restrict the copying or manipulation of the sampled signal in question. For broadcast or streamed media, this is especially the case. Message authentication is also possible, though not guaranteeing the same security as an encrypted file as with general crypto systems.

**[0048]** The present invention thus benefits from the proprietary player model without relying on proprietary players. No new players will be necessary and existing multimedia file formats can be altered to exact a measure of security which is further increased when coupled with digital watermarks. As with most consumer markets for media content, predominant file formats exist, de facto, and corresponding formats for computers likewise exist. For a commercial compact disc quality audio recording, or 16 bit 44.1 kHz, corresponding file formats include: Audio Interchange File Format (AIFF),

Microsoft WAV, Sound Designer II, Sun's .au, Apple's Quicktime, etc. For still image media, formats are similarly abundant: TIFF, PICT, JPEG, GIF, etc. Requiring the use of additional proprietary players, and their complementary file formats, for limited benefits in security is wasteful. Moreover, almost all computers today are multimedia-capable, and this is increasingly so with the popularity of Intel's MMX chip architecture and the PowerPC line of microchips. Because file formatting is fundamental in the playback of the underlying data, the predetermined key can act both as a map, for information to be encoded as watermark data regarding ownership, and a descrambler of the file that has been distributed. Limitations will only exist in how large the key must be retrofitted for a given application, but any manipulation of file format information is not likely to exceed the size of data required versus that for an entire proprietary player.

**[0049]** As with previous disclosures by the inventor on digital watermarking techniques, the present invention may be implemented with a variety of cryptographic protocols to increase both confidence and security in the underlying system. A predetermined key is described as a set of masks. These masks may include primary, convolution and message delimiter mask. In previous disclosures, the functionality of these masks is defined solely for mapping. The present invention includes a mask set which is also controlled by the distributing party of a copy of a given media signal. This mask set is a transfer function which is limited only by the parameters of the file format in question. To increase the uniqueness or security of each key used to scramble a given media file copy, a secure one way hash function can be used subsequent to transfer properties that are initiated to prevent the forging of a particular key. Public and private keys may be used as key pairs to further increase the unlikeliness that a key may be compromised.

**[0050]** These same cryptographic protocols can be combined with the embodiments of the present invention in administering streamed content that requires authorized keys to correctly display or play the streamed content in an unscrambled manner. As with digital watermarking, symmetric or asymmetric public key pairs may be used in a variety of implementations. Additionally, the need for certification authorities to maintain authentic key-pairs becomes a consideration for greater security beyond symmetric key implementations. The cryptographic protocols makes possible, as well, a message of text to be authenticated by a message authenticating function in a general computing device that is able to ensure secure message exchanges between authorizing parties.

**[0051]** An executable computer program is variously referred to as an application, from the point of view of a user, or executable object code from the point of view of the engineer. A collection of smaller, atomic (or indivisible) chunks of object code typically comprise the complete executable object code or application which may also require the presence of certain data resources. These indivisible portions of object code correspond with the programmers' function or procedure implementations in higher level languages, such as C or Pascal. In creating an application, a programmer writes "code" in a higher level language, which is then compiled down into "machine language," or, the executable object code, which can actually be run by a computer, general purpose or otherwise. Each function, or procedure, written in the programming language, represents a self-contained portion of the larger program, and implements, typically, a very small

piece of its functionality. The order in which the programmer types the code for the various functions or procedures, and the distribution of and arrangement of these implementations in various files which hold them is unimportant. Within a function or procedure, however, the order of individual language constructs, which correspond to particular machine instructions is important, and so functions or procedures are considered indivisible for purposes of this discussion. That is, once a function or procedure is compiled, the order of the machine instructions which comprise the executable object code of the function is important and their order in the computer memory is of vital importance. Note that many "compilers" perform "optimizations" within functions or procedures, which determine, on a limited scale, if there is a better arrangement for executable instructions which is more efficient than that constructed by the programmer, but does not change the result of the function or procedure. Once these optimizations are performed, however, making random changes to the order of instructions is very likely to "break" the function. When a program is compiled, then, it consists of a collection of these sub-objects, whose exact order or arrangement in memory is not important, so long as any sub-object which uses another sub-object knows where in memory it can be found.

**[0052]** The memory address of the first instruction in one of these sub-objects is called the "entry point" of the function or procedure. The rest of the instructions comprising that sub-object immediately follow from the entry point. Some systems may prefix information to the entry point which describes calling and return conventions for the code which follows, an example is the Apple Macintosh Operating System (MacOS). These sub-objects can be packaged into what are referred to in certain systems as "code resources," which may be stored separately from the application, or shared with other applications, although not necessarily. Within an application there are also data objects, which consist of some data to be operated on by the executable code. These data objects are not executable. That is, they do not consist of executable instructions. The data objects can be referred to in certain systems as "resources."

**[0053]** When a user purchases or acquires a computer program, she seeks a computer program that "functions" in a desired manner. Simply, computer software is overwhelmingly purchased for its underlying functionality. In contrast, persons who copy multimedia content, such as pictures, audio and video, do so for the entertainment or commercial value of the content. The difference between the two types of products is that multimedia content is not generally interactive, but is instead passive, and its commercial value relates more on passive not interactive or utility features, such as those required in packaged software, set-top boxes, cellular phones, VCRs, PDAs, and the like. Interactive digital products which include computer code may be mostly interactive but can also contain content to add to the interactive experience of the user or make the underlying utility of the software more aesthetically pleasing. It is a common concern of both of these creators, both of interactive and passive multimedia products, that "digital products" can be easily and perfectly copied and made into unpaid or unauthorized copies. This concern is especially heightened when the underlying product is copyright protected and intended for commercial use.

**[0054]** The first method of the present invention described involves hiding necessary "parts" or code "resources" in digitized sample resources using a "digital watermarking" process, such as that described in the "Steganographic Method



and Device” patent application. The basic premise for this scheme is that there are a certain sub-set of executable code resources, that comprise an application and that are “essential” to the proper function of the application. In general, any code resource can be considered “essential” in that if the program proceeds to a point where it must “call” the code resource and the code resource is not present in memory, or cannot be loaded, then the program fails. However, the present invention uses a definition of “essential” which is more narrow. This is because, those skilled in the art or those with programming experience, may create a derivative program, not unlike the utility provided by the original program, by writing additional or substituted code to work around unavailable resources. This is particularly true with programs that incorporate an optional “plug-in architecture,” where several code resources may be made optionally available at run-time. The present invention is also concerned with concentrated efforts by technically skilled people who can analyze executable object code and “patch” it to ignore or bypass certain code resources. Thus, for the present embodiment’s purposes, “essential” means that the function which distinguishes this application from any other application depends upon the presence and use of the code resource in question. The best candidates for this type of code resources are NOT optional, or plug-in types, unless special care is taken to prevent work-arounds.

**[0055]** Given that there are one or more of these essential resources, what is needed to realize the present invention is the presence of certain data resources of a type which are amenable to the “stega-cipher” process described in the “Steganographic Method and Device” patent U.S. Pat. No. 5,613,004. Data which consists of image or audio samples is particularly useful. Because this data consists of digital samples, digital watermarks can be introduced into the samples. What is further meant is that certain applications include image and audio samples which are important to the look and feel of the program or are essential to the processing of the application’s functionality when used by the user. These computer programs are familiar to users of computers but also less obvious to users of other devices that run applications that are equivalent in some measure of functionality to general purpose computers including, but not limited to, set-top boxes, cellular phones, “smart televisions,” PDAs and the like. However, programs still comprise the underlying “operating systems” of these devices and are becoming more complex with increases in functionality.

**[0056]** One method of the present invention is now discussed. When code and data resources are compiled and assembled into a precursor of an executable program the next step is to use a utility application for final assembly of the executable application. The programmer marks several essential code resources in a list displayed by the utility. The utility will choose one or several essential code resources, and encode them into one or several data resources using the stegacipher process. The end result will be that these essential code resources are not stored in their own partition, but rather stored as encoded information in data resources. They are not accessible at run-time without the key. Basically, the essential code resources that provide functionality in the final end-product, an executable application or computer program, are no longer easily and recognizably available for manipulation by those seeking to remove the underlying copyright or license, or its equivalent information, or those with skill to substitute alternative code resources to “force” the applica-

tion program to run as an unauthorized copy. For the encoding of the essential code resources, a “key” is needed. Such a key is similar to those described in U.S. Pat. No. 5,613,004, the “Steganographic Method and Device” patent. The purpose of this scheme is to make a particular licensed copy of an application distinguishable from any other. It is not necessary to distinguish every instance of an application, merely every instance of a license. A licensed user may then wish to install multiple copies of an application, legally or with authorization. This method, then, is to choose the key so that it corresponds, is equal to, or is a function of, a license code or license descriptive information, not just a text file, audio clip or identifying piece of information as desired in digital watermarking schemes extant and typically useful to stand-alone, digitally sampled content. The key is necessary to access the underlying code, i.e., what the user understands to be the application program.

**[0057]** The assembly utility can be supplied with a key generated from a license code generated for the license in question. Alternatively, the key, possibly random, can be stored as a data resource and encrypted with a derivative of the license code. Given the key, it encodes one or several essential resources into one or several data resources. Exactly which code resources are encoded into which data resources may be determined in a random or pseudo random manner. Note further that the application contains a code resource which performs the function of decoding an encoded code resource from a data resource. The application must also contain a data resource which specifies in which data resource a particular code resource is encoded. This data resource is created and added at assembly time by the assembly utility. The application can then operate as follows:

**[0058]** 1) when it is run for the first time, after installation, it asks the user for personalization information, which includes the license code. This can include a particular computer configuration;

**[0059]** 2) it stores this information in a personalization data resource;

**[0060]** 3) Once it has the license code, it can then generate the proper decoding key to access the essential code resources.

**[0061]** Note that the application can be copied in an uninhibited manner, but must contain the license code issued to the licensed owner, to access its essential code resources. The goal of the invention, copyright protection of computer code and establishment of responsibility for copies, is thus accomplished.

**[0062]** This invention represents a significant improvement over prior art because of the inherent difference in use of purely informational watermarks versus watermarks which contain executable object code. If the executable object code in a watermark is essential to an application which accesses the data which contains the watermark, this creates an all-or-none situation. Either the user must have the extracted watermark, or the application cannot be used, and hence the user cannot gain full access to the presentation of the information in the watermark bearing data. In order to extract a digital watermark, the user must have a key. The key, in turn, is a function of the license information for the copy of the software in question. The key is fixed prior to final assembly of the application files, and so cannot be changed at the option of the user. That, in turn, means the license information in the software copy must remain fixed, so that the correct key is available to the software. The key and the license information

are, in fact, interchangeable. One is merely more readable than the other. In U.S. Pat. No. 5,613,004, the “Steganographic Method and Device, patent”, the possibility of randomization erasure attacks on digital watermarks was discussed. Simply, it is always possible to erase a digital watermark, depending on how much damage you are willing to do to the watermark-bearing content stream. The present invention has the significant advantage that you must have the watermark to be able to use the code it contains. If you erase the watermark you have lost a key piece of the functionality of the application, or even the means to access the data which bear the watermark.

**[0063]** A preferred embodiment would be implemented in an embedded system, with a minimal operating system and memory. No media playing “applets,” or smaller sized applications as proposed in new operating environments envisioned by Sun Microsystems and the advent of Sun’s Java operating system, would be permanently stored in the system, only the bare necessities to operate the device, download information, decode watermarks and execute the applets contained in them. When an applet is finished executing, it is erased from memory. Such a system would guarantee that content which did not contain readable watermarks could not be used. This is a powerful control mechanism for ensuring that content to be distributed through such a system contains valid watermarks. Thus, in such networks as the Internet or set-top box controlled cable systems, distribution and exchange of content would be made more secure from unauthorized copying to the benefit of copyright holders and other related parties. The system would be enabled to invalidate, by default, any content which has had its watermark(s) erased, since the watermark conveys, in addition to copyright information, the means to fully access, play, record or otherwise manipulate, the content.

**[0064]** A second method according to the present invention is to randomly re-organize program memory structure to prevent attempts at memory capture or object code analysis. The object of this method is to make it extremely difficult to perform memory capture-based analysis of an executable computer program. This analysis is the basis for a method of attack to defeat the system envisioned by the present invention.

**[0065]** Once the code resources of a program are loaded into memory, they typically remain in a fixed position, unless the computer operating system finds it necessary to rearrange certain portions of memory during “system time,” when the operating system code, not application code, is running. Typically, this is done in low memory systems, to maintain optimal memory utilization. The MacOS for example, uses Handles, which are double-indirect pointers to memory locations, in order to allow the operating system to rearrange memory transparently, underneath a running program. If a computer program contains countermeasures against unlicensed copying, a skilled technician can often take a snapshot of the code in memory, analyze it, determine which instructions comprise the countermeasures, and disable them in the stored application file, by means of a “patch.” Other applications for designing code that moves to prevent scanning-tunnelling microscopes, and similar high sensitive hardware for analysis of electronic structure of microchips running code, have been proposed by such parties as Wave Systems. Designs of Wave Systems’ microchip are intended for preventing attempts by hackers to “photograph” or otherwise determine “burn in” to microchips for attempts at reverse

engineering. The present invention seeks to prevent attempts at understanding the code and its organization for the purpose of patching it. Unlike systems such as Wave Systems’, the present invention seeks to move code around in such a manner as to complicate attempts by software engineers to reengineer a means to disable the methods for creating licensed copies on any device that lacks “trusted hardware.” Moreover, the present invention concerns itself with any application software that may be used in general computing devices, not chipsets that are used in addition to an underlying computer to perform encryption. Wave Systems’ approach to security of software, if interpreted similarly to the present invention, would dictate separate microchip sets for each piece of application software that would be tamperproof. This is not consistent with the economics of software and its distribution.

**[0066]** Under the present invention, the application contains a special code resource which knows about all the other code resources in memory. During execution time, this special code resource, called a “memory scheduler,” can be called periodically, or at random or pseudo random intervals, at which time it intentionally shuffles the other code resources randomly in memory, so that someone trying to analyze snapshots of memory at various intervals cannot be sure if they are looking at the same code or organization from one “break” to the next. This adds significant complexity to their job. The scheduler also randomly relocates itself when it is finished. In order to do this, the scheduler would have to first copy itself to a new location, and then specifically modify the program counter and stack frame, so that it could then jump into the new copy of the scheduler, but return to the correct calling frame. Finally, the scheduler would need to maintain a list of all memory addresses which contain the address of the scheduler, and change them to reflect its new location.

**[0067]** The methods described above accomplish the purposes of the invention--to make it hard to analyze captured memory containing application executable code in order to create an identifiable computer program or application that is different from other copies and is less susceptible to unauthorized use by those attempting to disable the underlying copyright protection system. Simply, each copy has particular identifying information making that copy different from all other copies.

**[0068]** Although various embodiments are specifically illustrated and described herein, it will be appreciated that modifications and variations of the present invention are covered by the above teachings and within the purview of the appended claims without departing from the spirit and intended scope of the invention.

1.-57. (canceled)

**58.** A system for copy protection of digital information, comprising:

a computer comprising a processor, a device for receiving digital information, and memory;

wherein said computer is programmed to receive said digital information including a digital sample, said digital sample comprising at least one of audio, video, multimedia, and format information, said format information comprising a first portion and a second portion,

wherein said computer:

identifies, in said digital sample, said format information;

generates, using a key and an encoding algorithm, encoded first portion format information from a first portion of said format information; and

generates encoded digital information, comprising said at least one of audio, video, and multimedia, and said encoded first portion of said format information and said second portion of said format information.

59. A computer based system for generating an executable program, comprising:

a computer system comprising a processor and a memory, wherein said computer system is programmed to:

store data in a non-transient computer readable medium;

store in said memory (1) a plurality of data resources, wherein said plurality of data resources includes a specifier data resource, which specifier data resource contains data specifying one or more of said plurality of data resources; and (2) a plurality of code resources, wherein said plurality of code resources includes a decode code resource, wherein said decode code resource functions to decode code resource encoded in specified data resources;

encode, using an encode key, a first one of said plurality of code resource in one or more data resources speci-

fied by said specifier data resource, so that said first one of said plurality of code resources is stored as encoded information in said one or more data resources;

assemble an executable program containing said plurality of data resources and said plurality of code resources;

wherein the said executable program is configured to prompt for a license code;

wherein said executable program is configured to use said decode code resource and either said license code or a decode key derived from said license code, on data resources specified by said specifier data resource, to access said encoded code resources;

wherein said executable program cannot access said encoded code resources without said license code; and

wherein said executable program is configured to store personalization information including said license code in a personalization data resource.

\* \* \* \* \*