Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2424331 C 2012/01/03

(11)(21) 2 424 331

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2001/10/30
(87) Date publication PCT/PCT Publication Date: 2002/05/10
(45) Date de délivrance/Issue Date: 2012/01/03
(85) Entrée phase nationale/National Entry: 2003/03/31
(86) N° demande PCT/PCT Application No.: US 2001/048476
(87) N° publication PCT/PCT Publication No.: 2002/037489
(30) Priorité/Priority: 2000/10/31 (US09/703,343)

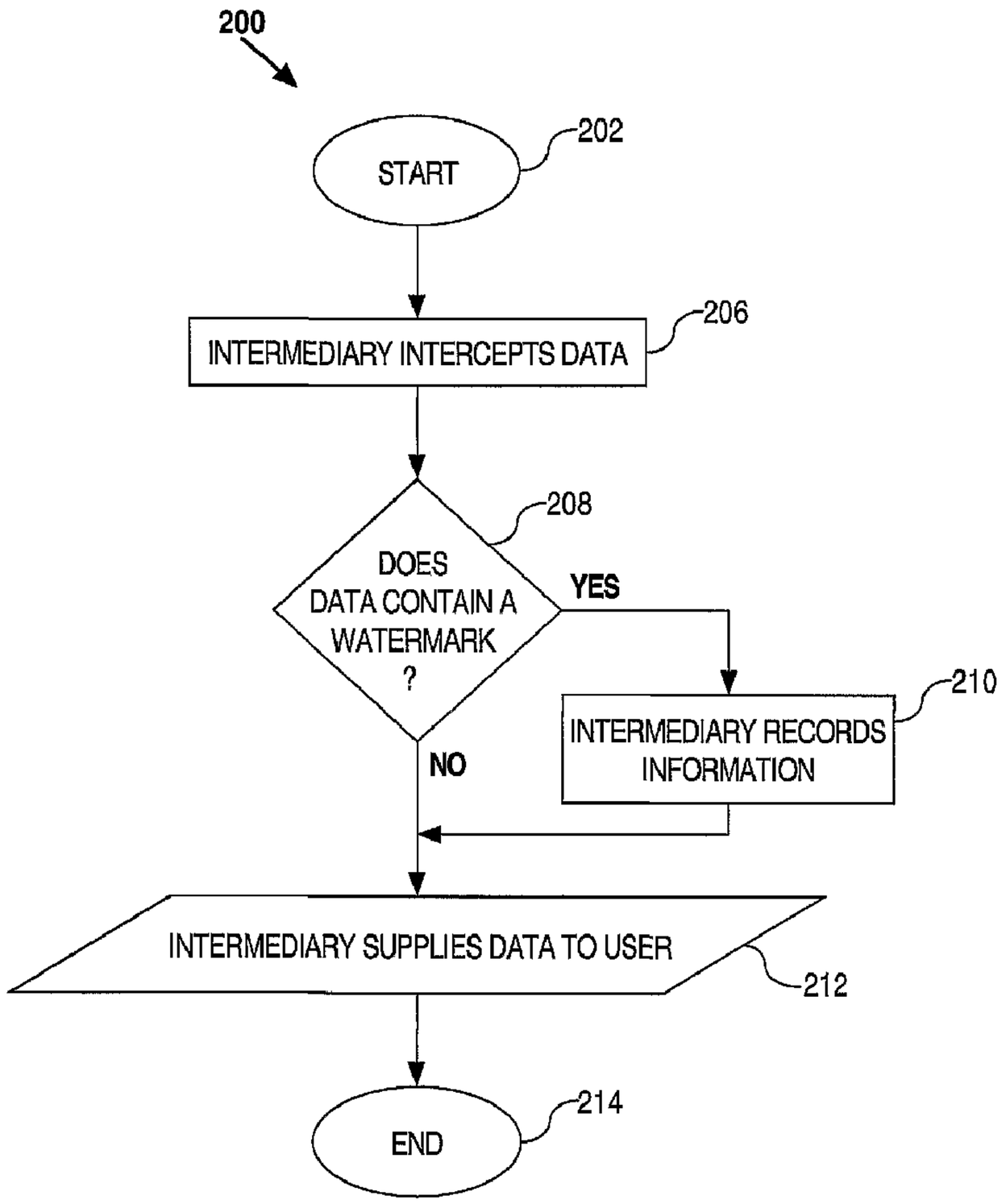
(51) Cl.Int./Int.Cl. *G11B 20/00* (2006.01), *G06F 1/00* (2006.01), *G06F 21/00* (2006.01), *G06T 1/00* (2006.01)

(72) Inventeur/Inventor:
LITTLEFIELD, ANDREW, US

(73) Propriétaire/Owner:
YAHOO! INC., US

(74) Agent: SMITHS IP

(54) Titre : METHODE DE LOCALISATION DE DONNEES
(54) Title: APPROACH FOR TRACKING DATA



(57) Abrégé/Abstract:
Tracking data in accordance with an embodiment of the invention generally involves determining whether data supplied by an intermediary includes rights data, such as a watermark, that indicates ownership rights in the data. If the data supplied by the

CA 2424331 C 2012/01/03

(11)(21) 2 424 331

(13) C

(57) Abrégé(suite)/Abstract(continued):
intermediary includes rights data, then a record is generated that indicates that the intermediary supplied data contains rights data. The intermediary may also record the source and destination of data passing through the intermediary. In addition, the intermediary may dynamically update rights data to indicate a source, destination or path of data passing through the intermediary. The intermediary may also generate statistical information about data passing through the intermediary.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 May 2002 (10.05.2002)

PCT

(10) International Publication Number
WO 02/37489 A2

(51) International Patent Classification[7]: G11B 20/00

(21) International Application Number: PCT/US01/48476

(22) International Filing Date: 30 October 2001 (30.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/703,343 31 October 2000 (31.10.2000) US

(71) Applicant: **INKTOMI CORPORATION** [US/US]; 4100 East Third Avenue, San Mateo, CA 94404 (US).

(72) Inventor: **LITTLEFIELD, Andrew**; 461 2nd Street, San Francisco, CA 94107 (US).

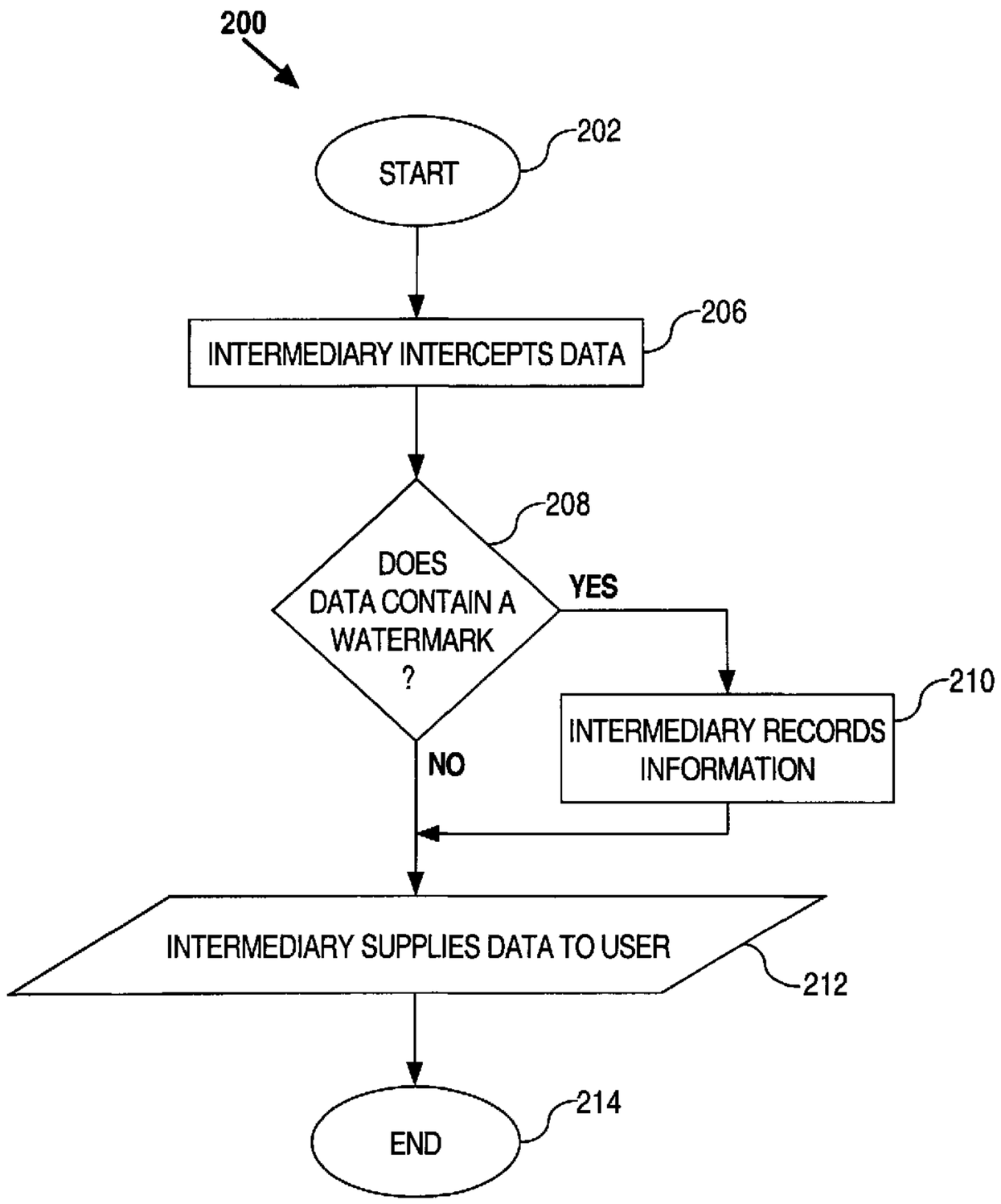
(74) Agents: **BECKER, Edward** et al.; Hickman Palermo Truong & Becker, LLP, 1600 Willow Street, San Jose, CA 95125 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,

[Continued on next page]

(54) Title: APPROACH FOR TRACKING DATA



(57) Abstract: Tracking data in accordance with an embodiment of the invention generally involves determining whether data supplied by an intermediary includes rights data, such as a watermark, that indicates ownership rights in the data. If the data supplied by the intermediary includes rights data, then a record is generated that indicates that the intermediary supplied data contains rights data. The intermediary may also record the source and destination of data passing through the intermediary. In addition, the intermediary may dynamically update rights data to indicate a source, destination or path of data passing through the intermediary. The intermediary may also generate statistical information about data passing through the intermediary.

WO 02/37489 A2

WO 02/37489 A2

CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

APPROACH FOR TRACKING DATA

FIELD OF THE INVENTION

The present invention relates generally to information management, and more specifically, to an approach for tracking the dissemination of data.

BACKGROUND OF THE INVENTION

The worldwide packet data communication network now commonly referred to as the "Internet" has experienced extraordinary growth and acceptance. The Internet provides access to many different types of information and data including electronic documents, music data files, e.g., MPEG Layer three (MP3) files, and video data files. With access to over several hundred million electronic documents, the Internet is arguably the largest single source of information in the world. As used herein, the term "electronic document" refers to any type of data or information in electronic form. Examples of electronic documents include, without limitation, text documents and web pages.

One problem created by the shear size and accessibility of the Internet is how to track the dissemination of data. Once a copy of data has been made available on the Internet, the copy is freely disseminated to potentially millions of users, with little or no control over to whom the data is disseminated. This problem is particularly acute for music data files, movie data files and image data files where illegal copies are readily made available for download to millions of users. More recently, music, video and image repositories have been created that store and make available for free, large amounts of pirated music, video and image data.

Various attempts have been made to control access to works of authorship to reduce the amount of pirating. One solution has been to encrypt data and provide the means to decrypt the encrypted data to only paying customers. Another solution is to provide works only through subscription Web sites so that only users who pay a subscription fee have access to the works. The primary shortcoming with both of these approaches is that once a clear, i.e., unencrypted, copy of the data has been obtained, the clear copy of the data may be freely distributed to millions of other users over the Internet.

Attempts have also been made to track the dissemination of works of authorship by encoding such works with data that indicates ownership, for example, an author field in a word processing document or a copyright notification contained in an executable file. One type of encoding, referred to as "watermarking," generally involves embedding a particular

digital signature, sequence or artifact referred to as a "watermark" in data so that the data (or copies thereof) can be definitively identified. The presence of a sufficiently unique watermark in a copy of data can provide a very high probability that the copy of data originated from a particular source. For example, suppose that a creator of an image wishes to sell copies of the image over the Internet to paying customers. The creator embeds in the image a watermark that identifies the creator. The watermark can be detected by an electronic device, such as a computer, but is not otherwise detectable by visual inspection. Any copy of the image that contains the watermark can then be identified as being created by the creator. Watermarking is well known in the industry and many sophisticated watermarking techniques have been developed. For example, a watermark may be created throughout an entire image so that cropping or reshaping the image does not remove the watermark.

Watermarking is very effective for determining whether particular data originates from a particular source and is therefore very helpful for determining whether copies of data have been illegally obtained. While watermarks can be used to determine whether a located copy of data is illegal, the primary limitation with watermarking is the difficulty of locating the illegal copies. The process of locating illegal copies of data is particularly difficult when the copies can reside anywhere on large communications networks such as the Internet. In the context of the Internet, a "Web Crawler" process is typically used to "crawl" the World Wide Web (the "Web") and locate illegal copies of data.

One problem is that the enormous size of the Web makes crawling the Web computationally expensive and time consuming. Another problem is that data changes very quickly on the Web, requiring repeated searches for illegal copies. Websites that serve as repositories for pirated copies of audio/video works can change their content and domains rapidly, making detection more difficult. Yet another problem is that many Websites that serve as repositories for music, video and image data are membership sites that require a user identification and password for access, which prevents Web Crawlers from retrieving Web pages (and the potentially illegal content contained therein) within the password-protected Web sites. Finally, some Web sites employ Web page structures known as "spider traps" that are designed to fool and confuse Web Crawlers.

Based upon the need to track access to data over communications networks such as the Internet and the limitations in prior approaches, an approach for tracking access to data over a communications network such as the Internet that does not suffer from limitations in prior approaches is highly desirable.

SUMMARY OF THE INVENTION

According to one aspect of the invention, a method is provided for tracking data. The method includes receiving, at an intermediary, data from a source and supplying the data from the intermediary to a user. The method also includes determining whether the data includes rights data that indicates ownership rights in the data and if the data includes the rights data, then recording that the data was supplied.

According to another aspect of the invention, a method is provided for tracking data disseminated over the Internet. The method includes the computer-implemented steps of receiving, at an intermediary over the Internet, data from a source and supplying the data from the intermediary to a user over the Internet. The method also includes the computer-implemented steps of determining whether the data contains a watermark that indicates ownership rights in the data and if the data contains a watermark, then recording that the data was supplied.

According to another aspect of the invention, a computer system is provided for tracking data. The computer system includes one or more processors and a memory communicatively coupled to the one or more processors. The memory contains one or more sequences of one or more instructions which, when executed by the one or more processors, cause the one or more processors to perform several steps. The steps include receiving, at an intermediary, data from a source and supplying the data from the intermediary to a user. The steps also include determining whether the data includes rights data that indicates ownership rights in the data and if the data includes the rights data, then recording that the data was supplied.

According to another aspect of the invention, an apparatus is provided. The apparatus includes an input/output mechanism configured to receive data from a source and supply the data to a user and a data tracking mechanism communicatively coupled to the input/output mechanism. The data tracking mechanism is configured to determine whether the data includes rights data that indicates ownership rights in the data, and if the data includes the rights data, then record that the data was supplied.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention are illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

FIG. 1 is block diagram of an arrangement for tracking data according to an embodiment of the invention;

FIG. 2 is a flow diagram of an approach for tracking data according to an embodiment of the invention;

FIG. 3 is a flow diagram of an approach for tracking data using dynamic watermarking according to an embodiment of the invention;

FIG. 4 is a block diagram of an arrangement upon which an approach for tracking data may be implemented in accordance with an embodiment of the invention; and

FIG. 5 is a block diagram of a computer system upon which embodiments of the invention may be implemented.

DETAILED DESCRIPTION OF THE INVENTION

In the following description, for the purposes of explanation, specific details are set forth in order to provide a thorough understanding of the invention. However, it will be apparent that the invention may be practiced without these specific details. In other instances, well-known structures and devices are depicted in block diagram form in order to avoid unnecessarily obscuring the invention.

Various aspects of the invention are described in more detail hereinafter in the following sections: (1) functional overview; (2) record keeping and reporting; (3) watermarking; and (4) implementation mechanisms.

1. FUNCTIONAL OVERVIEW

Tracking data in accordance with an embodiment of the invention generally involves determining whether data, supplied through an intermediary, includes rights data that indicates ownership rights in the data. If the data supplied through the intermediary includes rights data, then a record is generated that indicates that the intermediary supplied data containing rights data. For purposes of explanation, rights data is described herein in the context of a watermark. The invention is not limited to the use of watermarks, however, and is applicable to any type of rights data.

FIG. 1 is a block diagram of an arrangement 100 for tracking data according to an embodiment of the invention. Arrangement 100 includes a data source 102, an intermediary 104 and a user 106. Data source 102 and intermediary 104 are communicatively coupled by a communications link 108. Intermediary 104 and user 106 are communicatively coupled by a communications link 110. Communications links 108,

110 may be implemented by any medium or mechanism that provides for the exchange of data between data source 102 and intermediary 104, and intermediary 104 and user 106, respectively. Examples of communications links 108, 110 include, without limitation, a network such as a Local Area Network (LAN), Wide Area Network (WAN), Ethernet or the Internet, or one or more terrestrial, satellite or wireless links.

Data source 102 may be any source of data. For example, in the context of the Internet, data source 102 may be one or more origin servers that store content. Intermediary 104 may be any medium or mechanism through which data passes when data is supplied from data source 102 to user 106. Examples of intermediary 104 include, without limitation, a network cache and a traffic server. Intermediary 104 may be configured with local storage, for example, a database. In operation, user 106 requests data from data source 102. The data is supplied by data source 102 to user 106 through intermediary 104. According to one embodiment of the invention, intermediary 104 is configured to determine whether the data contains a watermark. If the data contains a watermark, then intermediary 104 creates a record that intermediary 104 supplied the data.

FIG. 2 is a flow diagram 200 that illustrates an approach for tracking data according to an embodiment of the invention. After starting in step 202, in step 206, intermediary 104 intercepts data from data source 102. In step 208, intermediary 104 determines whether the data contains a watermark. If the data contains a watermark, then in step 210, intermediary 104 records that the data containing a watermark is being supplied. In step 212, intermediary 104 supplies the data (with or without a watermark) to user 106. The process is complete in step 214.

It should be noted that for purposes of explanation, the recording of information in step 212 is described as occurring before intermediary 104 supplies the data containing a watermark to user 106 in step 212. Intermediary 104 may alternatively provide the data containing a watermark to user 106 before recording the information.

This approach allows data supplied by intermediaries to be tracked. More specifically, this approach allows owners of data to track the dissemination of data, by intermediaries, which is particularly useful for policing unauthorized copies of data, such as copyrighted works of authorship.

2. RECORD KEEPING AND REPORTING

Intermediary 104 is configured to record various types of information when intermediary 104 supplies data containing a watermark. The type of information recorded by intermediary 104 depends upon the requirements of a particular application and the type of tracking desired. For example, intermediary 104 may record only that intermediary 104 supplied data containing a watermark, without specifying the origin or destination of the data. Intermediary 104 may also record that intermediary 104 supplied data containing a watermark to a particular destination, e.g., to user 106. For example, intermediary 104 may record an identification of user 106. As another example, intermediary 104 may record data that identifies a location associated with user 106, such as an Internet Protocol (IP) address, a Uniform Resource Locator (URL), or other location information.

According to another embodiment of the invention, intermediary 104 is configured to record time and date information, e.g., a timestamp, that indicates when intermediary 104 supplied data. Intermediary 104 may also be configured to record information about the source of the data containing a watermark that intermediary 104 supplied to user 106. This may include, for example, data that identifies data source 102, or a location associated with data source 102, such as an IP address or URL of data source 102. According to another embodiment of the invention, intermediary 104 is configured to extract and record information from a watermark, such as ownership information. Intermediary 104 may also be configured to record information about itself 104. For example, intermediary 104 may record that data supplied to user 106 passed through intermediary 104.

Intermediary 104 may also generate reporting information about data supplied to other entities. For example, intermediary 104 may record information that indicates the source and destination of data passing through intermediary 104. Intermediary 104 may also record time and date information associated with data passing through intermediary 104. Intermediary 104 may also record one or more attributes of data passing through intermediary 104. Reporting data generated by intermediary 104 may be made available to data source 102 or other entities (not illustrated).

According to another embodiment of the invention, intermediary 104 is configured to determine whether data source 102 is associated with an owner of rights in data passing through intermediary 104. If data source 102 is not associated with the owner of rights in the data, then intermediary 104 is configured to generate and send a

message to the owner of the rights in the data indicating that the data, in which the owner owns rights, was received from a source, i.e., data source 102, not associated with the owner. According to another embodiment of the invention, intermediary 104 is configured to refuse to provide the data to user 106 when data source 102 is not associated with the owner of rights in the data. In this capacity, intermediary 104 acts as a security gatekeeper for unauthorized copies of data.

3. WATERMARKING

Many different types of watermarks may be used with the present approach for tracking data and the invention is not limited to any particular type of watermark or watermarking technique. For example, watermarks may be implemented by encoding a portion or an entire set of data. A watermark may be implemented in a manner so as to be detectable without the use of machine processing. For example, image data may be encoded with a watermark that is visible to the human eye when the image data is displayed. This technique is useful as a deterrent to illegal copying of data. Alternatively, a watermark may be invisible to the human eye and only detected by machine processing. For example, image or video data may be encoded in a manner such that a watermark is detectable by a computer, but invisible to the human eye. As another example, sound data may be encoded in a manner such that a watermark is audible or inaudible to the human ear.

The type of information associated with a watermark may vary from implementation to implementation. For example, a watermark may indicate the creator of data and/or the owner of rights to data. A watermark may include, for example, a copyright notice. Watermarks may also identify the legitimate owner of a copy of the data. Watermarks may also include a location of, for example, a physical address, an IP address or URL of an owner.

A single watermark may include numerous items of information. For example, a watermark in a particular copy of an image may identify company A as the legitimate owner of that copy, company B as the owner of the copyrights in the image, the duration and start date of a license agreement between companies A and B, the IP addresses of companies A and B, etc.

According to one embodiment of the invention, intermediary 104 updates watermarks according to an approach referred to herein as "dynamic watermarking." According to the dynamic watermarking approach, intermediary 104 receives data from

data source 102 and determines whether the data contains a watermark. If the data contains a watermark, then intermediary 104 dynamically updates a watermark and supplies the data with the dynamically-updated watermark to user 106.

Dynamic watermarking may involve deleting existing information from a watermark, adding new information to a watermark, or both. For example, intermediary 104 may update a watermark to include additional information about the source of the data, such as the location of data source 102. As another example, intermediary 104 may update the watermark to include time and date information, e.g., a timestamp, that indicates the time and date that the data with the dynamically-updated watermark is supplied by intermediary 104 to user 106. Intermediary 104 may also update a watermark to include information about user 106 or a location associated with user 106. For example, intermediary 104 may dynamically update a watermark to specify an IP address or URL of user 106. Intermediary 104 may also update a watermark to include information about itself, such as an IP address or URL of intermediary 104.

FIG. 3 is a flow diagram 300 that illustrates an approach for tracking data using dynamic watermarking according to an embodiment of the invention. After starting in step 302, in step 306, intermediary 104 intercepts data from data source 102. In step 308, intermediary 104 determines whether the data contains a watermark. If the data contains a watermark, then in step 310, intermediary 104 dynamically updates the watermark to add or remove information, as previously described. In step 312, intermediary 312 records that the data containing a watermark is being supplied, or other information, such as the origin or destination of the data. In step 314, intermediary 104 supplies the data (with or without a watermark) to user 106. The process is complete in step 316.

FIG. 4 is a block diagram of an arrangement 400 upon which an approach for tracking data may be implemented in accordance with an embodiment of the invention. Arrangement 400 includes data sources 402, 404, 406, users 408, 410, 412, communications links 414, 416, 418, 420, 422, 424 and a communications network 426. Communications links 414, 416, 418, 420, 422, 424 may be implemented by any medium or mechanism that provides for the exchange of data between their respective endpoints. Examples of communications links 414, 416, 418, 420, 422, 424 include, without limitation, networks such as Local Area Networks (LANs), Wide Area Networks (WANs), Ethernets or the Internet, or one or more terrestrial, satellite or wireless links. Communications network 426 may be any type of communications network, for example

a Local Area Network (LAN), a Wide Area Network (WAN), or a packet-based network such as an Ethernet or the Internet.

Users 408, 410, 412 request data from data sources 402, 404, 406. The data is provided to users 408, 410, 412 over communications links 414, 416, 418, 420, 422, 424 and communications network 426. Data being provided to users 408, 410, 412 is passed through and stored on an intermediary, which in the present example is a cache 428. For example, suppose user 408 requests a data item from data source 406. The data item is supplied by data source 406 over communications link 418 and communications network 426 and stored in cache 428. The data item is then supplied by cache 428 to user 408. Intermediaries such as cache 428 may be used to store frequently-requested versions of data to reduce the amount of time required to supply the frequently-requested versions of data to users 408, 410, 412.

According to one embodiment of the invention, cache 428 is configured to determine whether data being supplied by cache 428 includes a watermark. If the data being supplied by cache 428 includes a watermark, then cache 428 records information about the data being supplied. As previously described, cache 428 may record a variety of different types of information. For example, cache 428 may record only that data containing a watermark has been supplied by cache 428. Cache 428 may also record that data containing a watermark has been supplied by cache 428 to a particular user, which in the present example, is user 408. This may include, for example, an identity of user 408 or a location associated with user 408, such as an IP address or URL of user 408. Cache 428 may also record information about the source of the data that includes the watermark and that is being supplied by cache 428. For example, cache 428 may record an identity of data source 406. As another example, cache 428 may record a location associated with data source 406, such as an IP address or URL of data source 406.

Cache 428 may be further configured to record time and date information, e.g., a timestamp, that indicates the time at which the data containing the watermark was supplied by cache 428 to user 408. Cache 428 may also extract and record information contained in the watermark of the data being supplied to user 408. For example, suppose that a particular watermark identifies an owner of the data or copyright information. Cache 428 may also record the owner of the data or the copyright information. Cache 428 may also be configured to generate statistical information about data containing watermarks that is supplied by cache 428. Cache 428 may be further configured to report

statistical information to various entities, for example to data sources 402, 404, 406, or the owners of supplied data.

In the context of a communications network, such as the Internet, cache 428 is typically implemented as storage, such as one or more databases, with one or more software processes that manage storing and retrieving data from the storage. In this context, the inspection of data supplied by cache 428 may be performed by a software process executing at cache 428, e.g., integrated functionality, or may be implemented as a disparate process that interacts with cache 428. The invention is not limited to any particular implementation of cache 428. Furthermore, in the present example, cache 428 is illustrated in FIG. 4 as being contained in communications network 426, but the invention is not limited to this context and is also applicable to applications where cache 428 resides outside of communications network 426.

In the present example, configuring cache 428 in the manner described allows data containing watermarks to be tracked. This is particularly useful for tracking the dissemination, e.g., the downloading, of unauthorized copies of data. Cache 428 may also be configured to generate reporting data for data supplied by cache 428. For example, cache 428 may be configured to generate reporting data that specifies the source and destination of data provided from data sources 402, 404, 406 to users 408, 410, 412 to allow data to be tracked.

4. IMPLEMENTATION MECHANISMS

The approach for tracking data may be implemented in a wide variety of applications and contexts. For example, any type of intermediary, such as a network cache, may be configured to track data in accordance with the approach described herein. The approach may also be implemented as a stand-alone mechanism that interacts with a intermediary. The approach may be implemented in hardware circuitry, in computer software, or a combination of hardware circuitry and computer software and the invention is not limited to a particular hardware or software implementation. For example, the approach may be implemented as a process executing in conjunction with a network cache manager.

Figure 5 is a block diagram that illustrates a computer system 500 upon which an embodiment of the invention may be implemented. Computer system 500 includes a bus 502 or other communication mechanism for communicating information, and a processor 504 coupled with bus 502 for processing information. Computer system 500 also includes

a main memory 506, such as a random access memory (RAM) or other dynamic storage device, coupled to bus 502 for storing information and instructions to be executed by processor 504. Main memory 506 also may be used for storing temporary variables or other intermediate information during execution of instructions to be executed by processor 504. Computer system 500 further includes a read only memory (ROM) 508 or other static storage device coupled to bus 502 for storing static information and instructions for processor 504. A storage device 510, such as a magnetic disk or optical disk, is provided and coupled to bus 502 for storing information and instructions.

Computer system 500 may be coupled via bus 502 to a display 512, such as a cathode ray tube (CRT), for displaying information to a computer user. An input device 514, including alphanumeric and other keys, is coupled to bus 502 for communicating information and command selections to processor 504. Another type of user input device is cursor control 516, such as a mouse, a trackball, or cursor direction keys for communicating direction information and command selections to processor 504 and for controlling cursor movement on display 512. This input device typically has two degrees of freedom in two axes, a first axis (e.g., x) and a second axis (e.g., y), that allows the device to specify positions in a plane.

The invention is related to the use of computer system 500 for tracking data. According to one embodiment of the invention, the tracking of data is provided by computer system 500 in response to processor 504 executing one or more sequences of one or more instructions contained in main memory 506. Such instructions may be read into main memory 506 from another computer-readable medium, such as storage device 510. Execution of the sequences of instructions contained in main memory 506 causes processor 504 to perform the process steps described herein. One or more processors in a multi-processing arrangement may also be employed to execute the sequences of instructions contained in main memory 506. In alternative embodiments, hard-wired circuitry may be used in place of or in combination with software instructions to implement the invention. Thus, embodiments of the invention are not limited to any specific combination of hardware circuitry and software.

The term "computer-readable medium" as used herein refers to any medium that participates in providing instructions to processor 504 for execution. Such a medium may take many forms, including but not limited to, non-volatile media, volatile media, and transmission media. Non-volatile media includes, for example, optical or magnetic disks, such as storage device 510. Volatile media includes dynamic memory, such as main

memory 506. Transmission media includes coaxial cables, copper wire and fiber optics, including the wires that comprise bus 502. Transmission media can also take the form of acoustic or light waves, such as those generated during radio wave and infrared data communications.

Common forms of computer-readable media include, for example, a floppy disk, a flexible disk, hard disk, magnetic tape, or any other magnetic medium, a CD-ROM, any other optical medium, punch cards, paper tape, any other physical medium with patterns of holes, a RAM, a PROM, and EPROM, a FLASH-EPROM, any other memory chip or cartridge, a carrier wave as described hereinafter, or any other medium from which a computer can read.

Various forms of computer readable media may be involved in carrying one or more sequences of one or more instructions to processor 504 for execution. For example, the instructions may initially be carried on a magnetic disk of a remote computer. The remote computer can load the instructions into its dynamic memory and send the instructions over a telephone line using a modem. A modem local to computer system 500 can receive the data on the telephone line and use an infrared transmitter to convert the data to an infrared signal. An infrared detector coupled to bus 502 can receive the data carried in the infrared signal and place the data on bus 502. Bus 502 carries the data to main memory 506, from which processor 504 retrieves and executes the instructions. The instructions received by main memory 506 may optionally be stored on storage device 510 either before or after execution by processor 504.

Computer system 500 also includes a communication interface 518 coupled to bus 502. Communication interface 518 provides a two-way data communication coupling to a network link 520 that is connected to a local network 522. For example, communication interface 518 may be an integrated services digital network (ISDN) card or a modem to provide a data communication connection to a corresponding type of telephone line. As another example, communication interface 518 may be a local area network (LAN) card to provide a data communication connection to a compatible LAN. Wireless links may also be implemented. In any such implementation, communication interface 518 sends and receives electrical, electromagnetic or optical signals that carry digital data streams representing various types of information.

Network link 520 typically provides data communication through one or more networks to other data devices. For example, network link 520 may provide a connection through local network 522 to a host computer 524 or to data equipment operated by an

Internet Service Provider (ISP) 526. ISP 526 in turn provides data communication services through the worldwide packet data communication network now commonly referred to as the "Internet" 528. Local network 522 and Internet 528 both use electrical, electromagnetic or optical signals that carry digital data streams. The signals through the various networks and the signals on network link 520 and through communication interface 518, which carry the digital data to and from computer system 500, are exemplary forms of carrier waves transporting the information.

Computer system 500 can send messages and receive data, including program code, through the network(s), network link 520 and communication interface 518. In the Internet example, a server 530 might transmit a requested code for an application program through Internet 528, ISP 526, local network 522 and communication interface 518. In accordance with the invention, one such downloaded application provides for the tracking of data as described herein.

The received code may be executed by processor 504 as it is received, and/or stored in storage device 510, or other non-volatile storage for later execution. In this manner, computer system 500 may obtain application code in the form of a carrier wave.

The novel approach described herein for tracking data provides several advantages over prior approaches. First, the approach provides for the tracking of data in a manner that does not require repeated "crawling" of a communications network. More particularly, the approach provides for the tracking of the usage of data, for example, source, destination, frequency and timestamping, which is particularly useful for tracking the pirating of data. The approach may also be implemented transparent to users 408, 410, 412, making it a "low friction", i.e., a less intrusive or disruptive, approach.

In the foregoing specification, the invention has been described with reference to specific embodiments thereof. However, various modifications and changes may be made thereto without departing from the broader spirit and scope of the invention. The specification and drawings are, accordingly, to be regarded in an illustrative sense rather than a restrictive sense.

CLAIMS

What is claimed is:

1. A method for tracking data at an intermediary device connected via a network to a source and to a user device, the method comprising the computer-implemented steps of:

receiving from the source, at the intermediary device, data to be supplied to the user device;

determining, at the intermediary device, that the data includes digital rights data that indicates an owner of rights in the data; and

in response to determining that the data includes the digital rights data, then:

determining, at the intermediary device, whether the source is associated with the owner of rights in the data;

if the source is associated with the owner of rights in the data, then:

the intermediary device supplying the data to the user device; and

the intermediary device recording that the data was supplied.

2. The method as recited in Claim 1, further comprising recording the source from which the intermediary receives the data.

3. The method as recited in Claim 1, further comprising recording the user to which the data was supplied.

4. The method as recited in Claim 3, further comprising recording that the data was supplied to a location associated with the user.

5. The method as recited in Claim 1, further comprising recording a time at which the data was supplied.

6. The method as recited in Claim 5, wherein the step of recording a time at which the data was supplied includes updating the digital rights data to indicate a time at which the data was supplied.

7. The method as recited in Claim 5, further comprising informing the source that the

data was supplied at the time.

8. The method as recited in Claim 1, further comprising the intermediary device dynamically updating the digital rights data to indicate a user to which the data was supplied.

9. The method as recited in Claim 1, further comprising the intermediary device dynamically updating the digital rights data to indicate that the intermediary device supplied the data.

10. The method as recited in Claim 1, wherein the digital rights data is a watermark.

11. The method as recited in Claim 1, wherein the data is music data.

12. The method as recited in Claim 1, wherein the data is image data.

13. The method as recited in Claim 1, wherein the data is video data.

14. The method as recited in Claim 1, wherein the data is digital data.

15. The method as recited in Claim 1, wherein the intermediary device is a network cache.

16. The method as recited in Claim 1, wherein the method further comprises:
if the source is not associated with the owner of rights in the data, then generating and transmitting to the owner of rights in the data, report data that indicates that the data was received from the source not associated with the owner of rights in the data.

17. The method as recited in Claim 1, wherein the method further comprises:
if the source is not associated with the owner of rights in the data, then the intermediary device not allowing the data to be supplied to the user.

18. A method for tracking data disseminated over the Internet at an intermediary

device connected via the Internet to a source and to a user device, the method comprising the computer-implemented steps of:

receiving from the source, at the intermediary device, data to be supplied to the user device;

the intermediary device supplying the data from the intermediary device to the user device over the Internet;

the intermediary device determining whether the data contains digital rights data that indicates an owner of rights in the data; and

if the data contains the digital rights data, then the intermediary device recording that the data was supplied.

19. A computer system for tracking data, said computer system being an intermediary device connected via a network to a source and to a user device, said computer system comprising:

one or more processors; and

a memory communicatively coupled to the one or more processors and containing one or more sequences of one or more instructions which, when executed by the one or more processors, cause the one or more processors to perform the steps of:

receiving from the source, at the intermediary device, data to be supplied to the user device;

determining, at the intermediary device, that the data includes digital rights data that indicates an owner of rights in the data, and

in response to determining that the data includes the digital rights data, then:

determining, at the intermediary device, whether the source is associated with the owner of rights in the data;

if the source is associated with the owner of rights in the data, then:

the intermediary device supplying the data to the user device; and

the intermediary device recording that the data was supplied.

20. A network cache deployed as an intermediary device between a source device and

a user device, the network cache comprising:

an input/output means configured to receive data from the source device and supply the data to the user device; and

a data tracking means communicatively coupled to the input/output means and configured to

determine whether the data received from the source device includes digital rights data that indicates ownership rights in the data, and

if the data includes the digital rights data, then record that the data was supplied.

21. A computer-readable medium carrying one or more sequences of one or more instructions for tracking data at an intermediary device connected via a network to a source and to a user device, the one or more sequences of one or more instructions including instructions which, when executed by the one or more processors, cause the one or more processors to perform the steps of:

receiving from the source, at the intermediary device, data to be supplied to the user device;

determining, at the intermediary device, that the data includes digital rights data that indicates an owner of rights in the data; and

in response to determining that the data includes the digital rights data, then:

determining, at the intermediary device, whether the source is associated with the owner or rights in the data;

if the source is associated with the owner of rights in the data, then:

the intermediary device supplying the data to the user device; and

the intermediary device recording that the data was supplied.

22. The computer-readable medium as recited in Claim 21, further comprising one or more sequences of one or more instructions including instructions which, when executed by one or more processors, cause the one or more processors to perform the step of: recording the source from which the intermediary device receives the data.

23. The computer-readable medium as recited in Claim 21, further comprising one or more sequences of one or more instructions including instructions which, when

executed by one or more processors, cause the one or more processors to perform the step of: recording the user to which the data was supplied.

24. The computer-readable medium as recited in Claim 23, further comprising one or more sequences of one or more instructions including instructions which, when executed by one or more processors, cause the one or more processors to perform the step of: recording that the data was supplied to a location associated with the user.

25. The computer-readable medium as recited in Claim 21, further comprising one or more sequences of one or more instructions including instructions which, when executed by one or more processors, cause the one or more processors to perform the step of: recording a time at which the data was supplied.

26. The computer-readable medium as recited in Claim 25, wherein the step of recording a time at which the data was supplied includes updating the digital rights data to indicate a time at which the data was supplied.

27. The computer-readable medium as recited in Claim 25, further comprising one or more sequences of one or more instructions including instructions which, when executed by one or more processors, cause the one or more processors to perform the step of: informing the source that the data was supplied at the time.

28. The computer-readable medium as recited in Claim 21, further comprising one or more sequences of one or more instructions including instructions which, when executed by one or more processors, cause the one or more processors to perform the step of: the intermediary device dynamically updating the digital rights data to indicate a user to which the data was supplied.

29. The computer-readable medium as recited in Claim 21, further comprising one or more sequences of one or more instructions including instructions which, when executed by one or more processors, cause the one or more processors to perform the step of: the intermediary device dynamically updating the digital rights data to indicate that the intermediary device supplied the data.

30. The computer-readable medium as recited in Claim 21, wherein the digital rights data is a watermark.

31. The computer-readable medium as recited in Claim 21, wherein the data is music data.

32. The computer-readable medium as recited in Claim 21, wherein the data is image data.

33. The computer-readable medium as recited in Claim 21, wherein the data is video data.

34. The computer-readable medium as recited in Claim 21, wherein the data is digital data.

35. The computer-readable medium as recited in Claim 21, wherein the intermediary device is a network cache.

36. The computer-readable medium as recited in Claim 21, wherein the computer-readable medium further includes instructions which, when executed by the one or more processors, cause the one or more processors to perform the steps of:

if the source is not associated with the owner of rights in the data, then generating and transmitting to the owner of rights in the data, report data that indicates that the data was received from the source not associated with the owner of rights in the data.

37. The computer-readable medium as recited in Claim 21, wherein the computer-readable medium further includes instructions which, when executed by the one or more processors, cause the one or more processors to perform the steps of:

if the source is not associated with the owner of rights in the data, then the intermediary device not allowing the data to be supplied to the user.

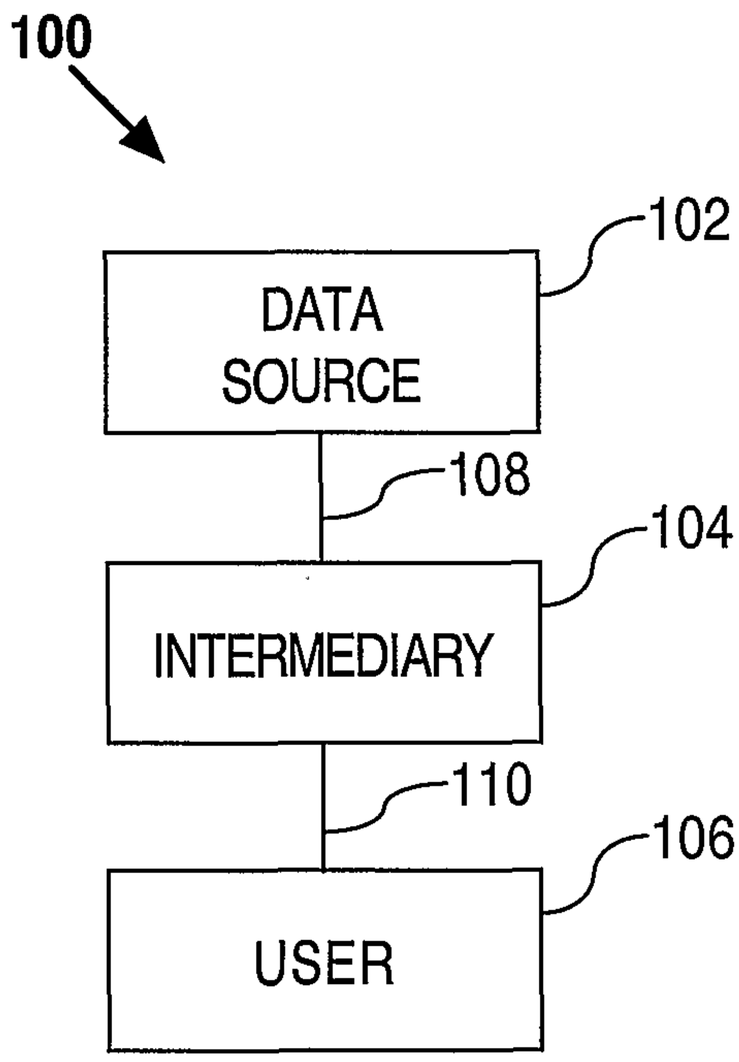


FIG. 1

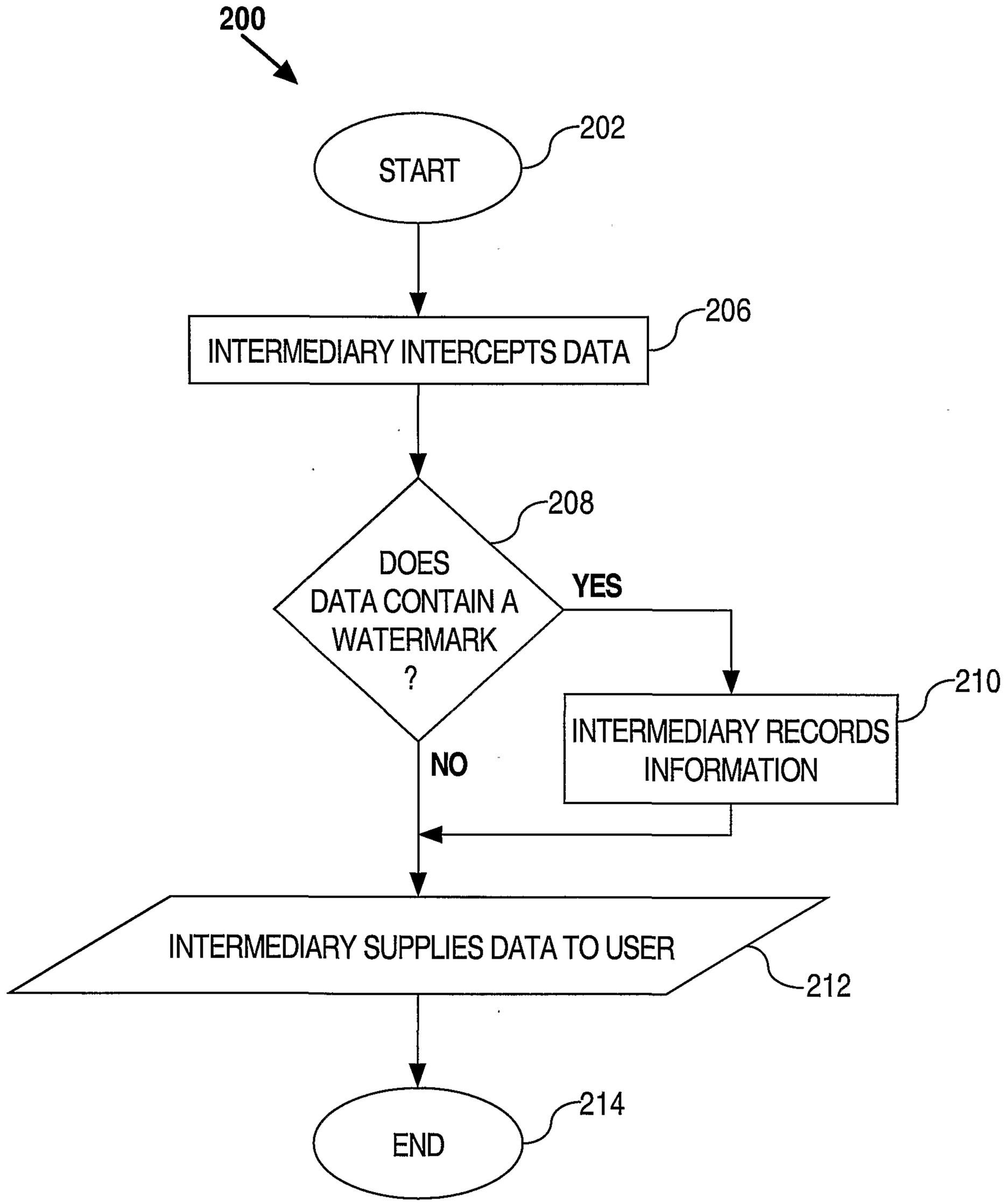

200
START
202
INTERMEDIARY INTERCEPTS DATA
206
DOES DATA CONTAIN A WATERMARK ?
208
YES
NO
INTERMEDIARY RECORDS INFORMATION
210
INTERMEDIARY SUPPLIES DATA TO USER
212
END
214

FIG. 2

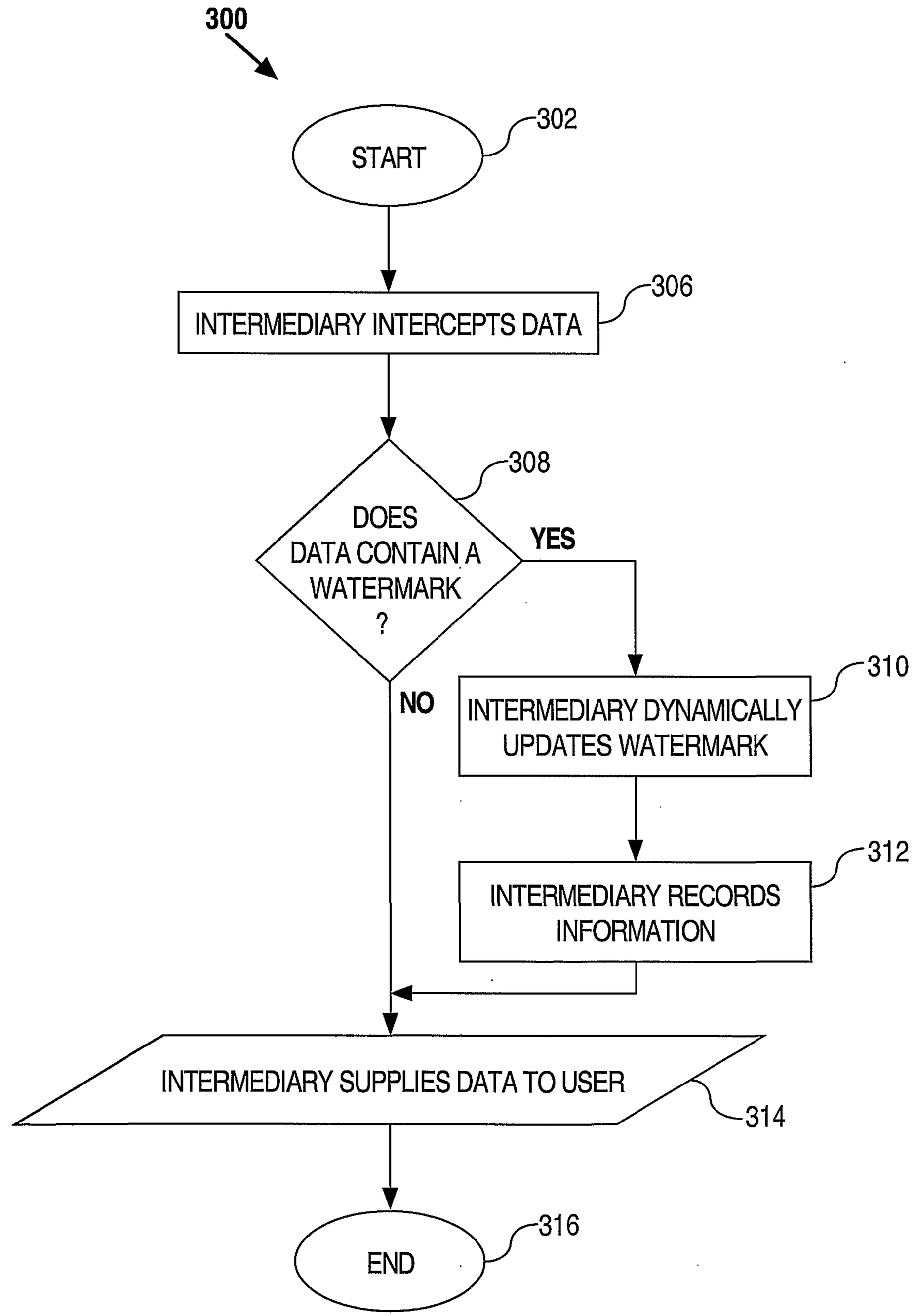

300
START
302
INTERMEDIARY INTERCEPTS DATA
306
DOES DATA CONTAIN A WATERMARK ?
308
YES
NO
INTERMEDIARY DYNAMICALLY UPDATES WATERMARK
310
INTERMEDIARY RECORDS INFORMATION
312
INTERMEDIARY SUPPLIES DATA TO USER
314
END
316

FIG. 3

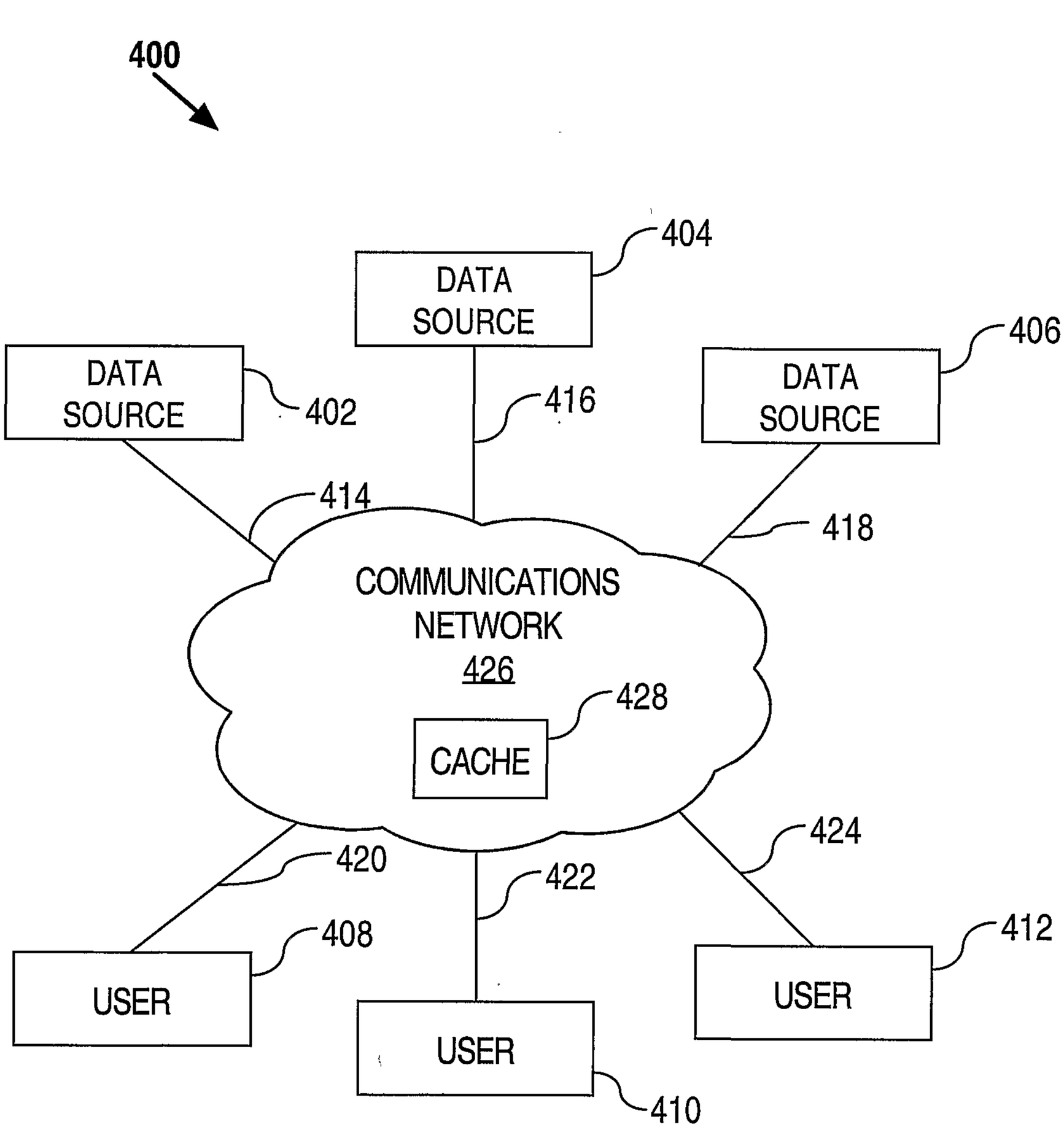

400
DATA SOURCE
404
DATA SOURCE
402
DATA SOURCE
406
416
414
418
COMMUNICATIONS NETWORK
426
428
CACHE
424
420
422
408
USER
USER
USER
412
410

FIG. 4

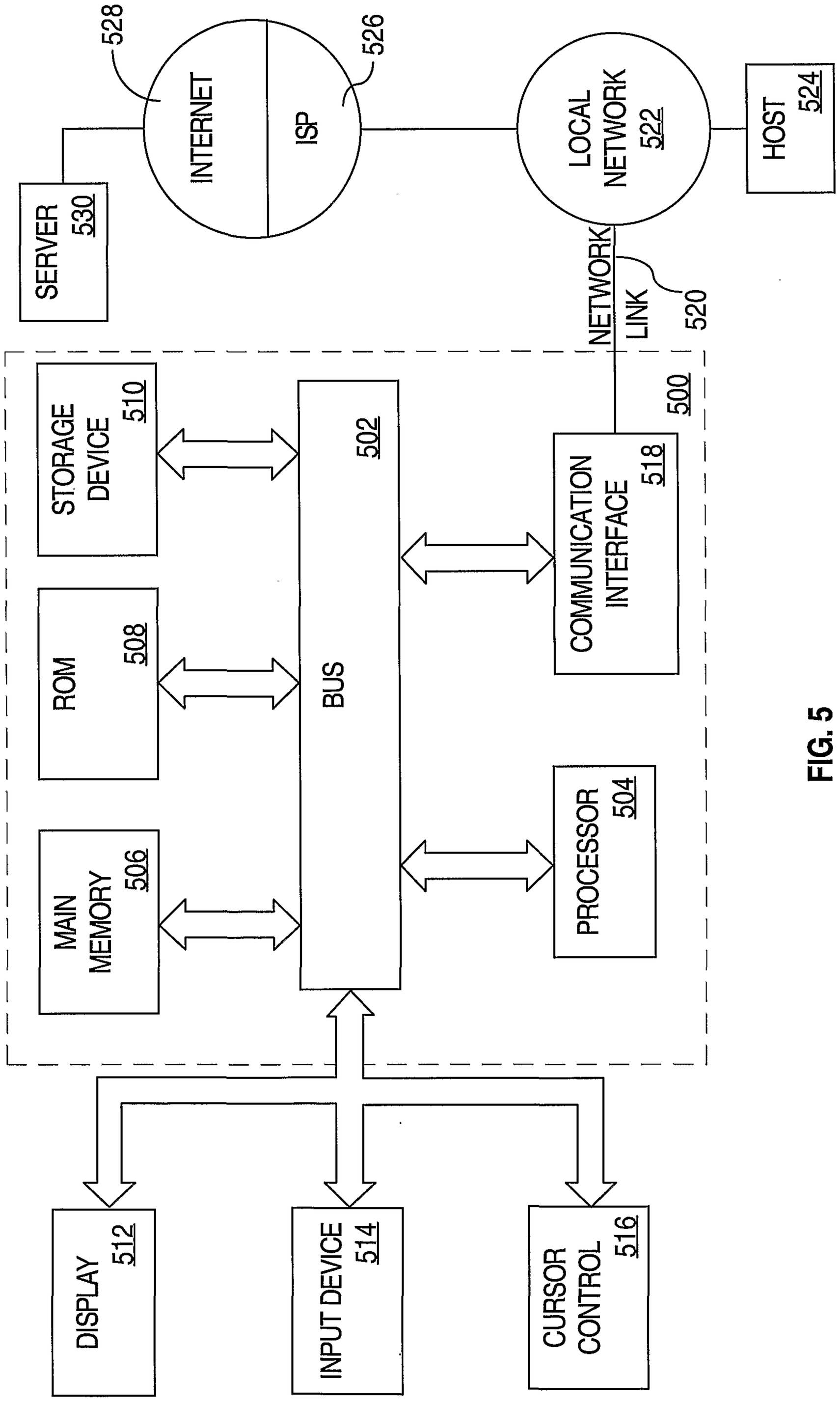
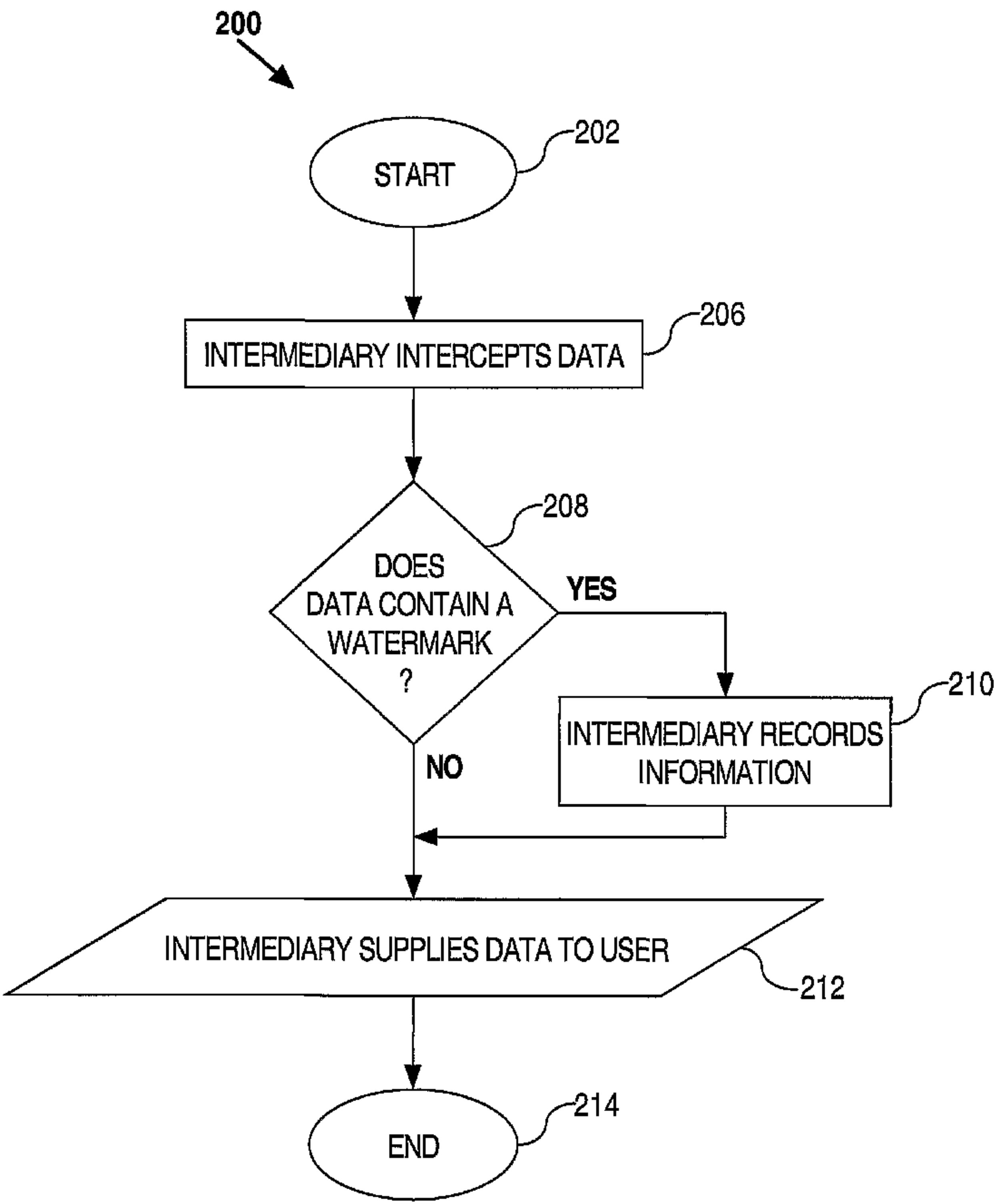

528
INTERNET
ISP
526
LOCAL NETWORK 522
HOST 524
SERVER 530
NETWORK LINK 520
500
STORAGE DEVICE 510
502
COMMUNICATION INTERFACE 518
ROM 508
BUS
MAIN MEMORY 506
PROCESSOR 504
DISPLAY 512
INPUT DEVICE 514
CURSOR CONTROL 516

FIG. 5


200
START
202
INTERMEDIARY INTERCEPTS DATA
206
DOES DATA CONTAIN A WATERMARK ?
208
YES
NO
INTERMEDIARY RECORDS INFORMATION
210
INTERMEDIARY SUPPLIES DATA TO USER
212
END
214