

METHOD AND APPARATUS FOR IDENTIFYING IDENTITY INFORMATION

Technical Field

[01] Embodiments of this specification relate to the field of Internet technologies, and in particular, to a method and an apparatus for identifying identity information.

5 Background

[02] With the continuous development of Internet technologies, the problem of information security in a network has become increasingly prominent. At present, identity information of a user may be stolen by an illegal user. The illegal user may use the stolen identity information to register an account, and use the registered account to perform various illegal
10 operations, thereby bringing undesirable consequences to the user.

[03] Currently, whether identity information is used fraudulently by others can usually be determined according to a usage environment of an account. For example, if an account is authenticated on the same device as a large number of other accounts, the identity information corresponding to the account is likely used fraudulently by others. However, if the usage
15 environment of the account changes, this method cannot effectively identify the identity information that is used fraudulently. Therefore, there is an urgent need for an effective method for identifying identity information.

Summary of the Invention

[04] An objective of the embodiments of the present specification is to provide a method
20 and an apparatus for identifying identity information, which can improve the efficiency of identifying identity information.

[05] To achieve the above objective, some embodiments of this specification are implemented as follows.

[06] A method for identifying identity information, comprising:

25 [07] acquiring user data, the user data comprising identity information and account information of a user;

[08] establishing an association relationship between account information and identity information that are bound in the user data, and establishing an association relationship between two pieces of account information having common features in the user data; and

[09] determining a risk value of target identity information in the user data according to the established association relationships, and determining, according to the determined risk value, whether the target identity information has a risk of being used fraudulently.

[10] An apparatus for identifying identity information, comprising:

[11] a data acquisition unit configured to acquire user data, the user data comprising identity information and account information of a user;

[12] an association relationship establishing unit configured to establish an association relationship between account information and identity information that are bound in the user data, and establish an association relationship between two pieces of account information having common features in the user data; and

[13] an identifying unit configured to determine a risk value of target identity information in the user data according to the established association relationships, and determine, according to the determined risk value, whether the target identity information has a risk of being used fraudulently.

[14] An apparatus for identifying identity information comprises a memory and a processor. The memory is configured to store a computer program. When executed by the processor, the computer program implements the method for identifying identity information described above.

[15] As can be seen from the above, in one or more embodiments of this specification, after user data to be analyzed is acquired, association relationships between identity information and account information and between account information and account information in the user data may be established. The association between identity information and account information here can be reflected in that: the account information is registered or authenticated based on the

associated identity information; and the association between different pieces of account information can be reflected in that: two pieces of account information have common features. Further, the two pieces of account information having common features may include, for example, the two pieces of account information having an interaction behavior, the two pieces of account information being bound to the same object, and the two pieces of account information being logged-in on the same device. Generally speaking, among multiple pieces of account information associated with the same identity information, there are often two or more pieces of account information having common features. The reason is that when a user uses multiple accounts created by himself/herself, some of these accounts may be logged-in on the same device, and transferring money, assisting authentication, being bound to the same mobile phone, etc. may also occur in some accounts, thus making these accounts have common features. Once identity information of a certain user is fraudulently used, the account created by the fraudster based on the identity information of the user usually does not have or rarely has common features with other accounts of the user. In this way, by analyzing the associations between identity information and account information and between account information and account information in data of the user, a risk value of a certain piece of target identity information can be determined. Once the risk value is too high, it can be considered that there is a risk of fraudulent use of the target identity information. It can be seen that the technical solution provided by one or more of the embodiments of this specification analyzes the association between identity information and account information to determine whether the identity information is fraudulently used. Even if a fraudster changes an usage environment of an account, the method can still identify whether the account is in an abnormal state based on whether there are common features between the account and other accounts, thereby improving the efficiency of identifying identity information.

Brief Description of the Drawings

[16] In order to illustrate the technical solutions in one or more embodiments of this specification or the prior art more clearly, the accompanying drawings to be used in the description

of the embodiments or the prior art will be introduced briefly below. It is apparent that the accompanying drawings in the following description are only some of the embodiments described in this specification. For those of ordinary skill in the art, other accompanying drawings can further be obtained according to these accompanying drawings without any creative efforts.

5 [17] FIG. 1 is a flowchart of a method for identifying identity information according to this specification;

 [18] FIG. 2 is a relationship diagram corresponding to a normal user according to this specification;

 [19] FIG. 3 is a relationship diagram including fraudulently used account information
10 according to this specification;

 [20] FIG. 4 is a schematic diagram of functional modules of an apparatus for identifying identity information according to this specification; and

 [21] FIG. 5 is a schematic structural diagram of an apparatus for identifying identity information according to this specification.

15 **Detailed Description**

 [22] A method and an apparatus for identifying identity information are provided in the embodiments of this specification.

 [23] In order to enable those skilled in the art to better understand the technical solutions in this specification, the technical solutions in the embodiments of this specification will be
20 described clearly and completely below with reference to the accompanying drawings in the embodiments of this specification. Obviously, the described embodiments are only a part of, not all, the embodiments of this specification. Based on the embodiments in this specification, all other embodiments obtained by a person of ordinary skill in the art without creative efforts shall fall within the protection scope of this specification.

25 [24] Referring to FIG. 1, a method for identifying identity information according to an embodiment of this specification may include the following steps.

[25] S1: User data is acquired, the user data including identity information and account information of a user.

[26] In this embodiment, the user data may be data to be analyzed, and the user data may be acquired from background servers of various websites. The user data may include account
5 information. For example, the account information may be an account name registered by the user, or a digital ID of the account name in a website. The user can fill in his/her own identity information when registering an account, or perform identity authentication after registering an account. In this way, registered account information and authenticated identity information can be bound in the background server of the website. The identity information may be information that can uniquely
10 represent the user's identity, such as ID card information, passport information, medical insurance card information, and driver's license information.

[27] S3: An association relationship between account information and identity information that are bound in the user data is established, and an association relationship between two pieces of account information having common features in the user data is established.

[28] In this embodiment, the bound account information and identity information in the user data may represent that the account information is authenticated by the identity information, and therefore, the bound account information and identity information may have an association. In addition, different pieces of account information may have common features during use. The common features can be embodied in many ways. For example, if there is an interaction behavior
20 between two pieces of account information, it can be considered that the two pieces of account information have common features. The interaction behavior may be transferring money to each other, sending chat information, or paying/collecting the same money together. For another example, if two pieces of account information are bound to the same object, it can also be considered that the two pieces of account information have common features. Being bound to the
25 same object may refer to that the two pieces of account information are bound to the same delivery address, bound to the same mobile phone number, bound to the same email address, etc. For

another example, if two pieces of account information are logged-in on the same device, it can also be considered that the two pieces of account information have common features. Of course, in practical applications, two pieces of account information having common features may also involve other more situations. The above-mentioned example situations in this embodiment are not exhaustive. Other situations involving two pieces of account information having common features and added by those skilled in the art upon understanding the essence of the technical solution in this embodiment shall also fall within the protection scope of this embodiment.

[29] In this embodiment, an association relationship between account information and identity information that are bound in the user data may be established, and an association relationship between two pieces of account information having common features in the user data may be established. In one embodiment, in order to visualize the association relationships, the association relationships between account information and identity information and between account information and account information may be indicated in the form of a relationship diagram. Please refer to FIG. 2. In the relationship diagram, account information and identity information may be represented by account nodes and an identity node, respectively, and there is a connection line between an account node and the identity node that are bound together, and there is also a connection line between two account nodes having common features. In this way, each piece of information in the user data is represented by using a node, and after adding connection lines between associated nodes, a relationship diagram corresponding to the user data can be established.

[30] S5: A risk value of target identity information in the user data is determined according to the established association relationships, and it is determined, according to the determined risk value, whether the target identity information has a risk of being used fraudulently.

[31] In this embodiment, after association relationships between various pieces of information are established in the user data, the established association relationships can be analyzed to obtain a difference between a normal user and an identity fraudster, and identity

information that may be at risk of being used fraudulently is identified based on the difference obtained by the analysis. Specifically, taking the relationship diagram shown in FIG. 2 as an example, a relationship diagram of identity information and account information of a normal user is shown in FIG. 2. In FIG. 2, the identity information of the user is represented by an identity node, and the user authenticates account information A and account information B through his/her own identity information. Therefore, connection lines are established between both nodes of account information A and account information B and the identity node. In addition, there has been a transfer behavior between account information A and account information C, and account information A and account information D have been logged-in on the same device. In addition, account information C and account information D are bound to the same delivery address. Therefore, according to the above common features, a connection line can be established between the node of account information A and the node of account information C, a connection line can be established between the node of account information A and the node of account information D, and a connection line can be established between the node of account information C and the node of account information D. Similarly, connection lines between other account nodes can be established based on common features. It can be seen from FIG. 2 that although the user created two different accounts using his/her own identity information, these two accounts likely have common features with another identical account because these two accounts are used by the user. In the relationship diagram, it can be expressed that there is also an association relationship between respective account groups associated with the two accounts.

[32] However, please refer to FIG. 3, assuming that the identity information of the user in FIG. 2 is used fraudulently, the fraudster creates account information H using the identity information of the user, and uses account information H to transfer money with two other pieces of account information I and J, respectively. Then, because the fraudster and the user are two different subjects, there is probably no intersection between relationship networks of the two different subjects, which can be reflected in the relationship diagram as follows: there is usually

no association between an account group where the account information used by the fraudster is located and an account group where the account information used by the user.

[33] As can be seen from the above, in the relationship diagram, there is an obvious difference between account information created by a normal user and account information created by a fraudster based on identity information of the normal user. To quantify this difference, in this embodiment, a concept of the shortest path between different account nodes can be introduced. Specifically, among paths connecting a first account node and a second account node in the relationship diagram, a path including the fewest connection lines may be used as the shortest path between the first account node and the second account node. For example, various paths connecting account node C and account node B in FIG. 3 may include at least a plurality of paths such as a path from C to B, a path from C to D to A to an identity node to B, and a path from C to A to the identity node to B. Among these paths, the path having the fewest connection lines is the path from C to B, which contains only one connection. Therefore, the shortest path for connection between account node C and account node B is the path from C to B directly. For another example, the shortest path from account node A to account node I should be the path from A to the identity node to H to I, which contains 3 connection lines. It can be seen that the identity node may be included in some shortest paths (such as the shortest path between A and I), or may not be included in some shortest paths (such as the shortest path between C and B). It can be found from the relationship diagram that, for multiple pieces of account information created by the user himself/herself, in an account group associated with the account information, the shortest paths between different account nodes may probably not include the identity node of the user. For example, in FIG. 3, all the shortest paths from node C to node B, from node C to node F, from node G to node A, and from node E to node A do not include the identity node. However, in an account group associated with the account information created by the fraudster, if an account node in the account group wants to generate a shortest path with an account node in the account group associated with the user, the shortest path generated often includes the identity node of the user. For example, the

shortest path from account node H, I, or J to any one of account nodes A to G in FIG. 3 respectively includes the identity node. It can be seen that if the identity information of the user is not used fraudulently, the identity node corresponding to the identity information of the user will not be included in a large number of shortest paths. However, if the identity information of the user is used fraudulently, the identity node corresponding to the identity information of the user will be included in a large number of shortest paths.

[34] Based on the above difference, this embodiment can calculate a risk value of a target identity node in the relationship diagram. The target identity node may be an identity node currently to be analyzed. When the risk value of the target identity node is calculated, in the established relationship diagram, the number of the shortest paths where a target identity node representing the target identity information is located may be counted, and the total number of the shortest paths in the relationship diagram may be counted. Then, a ratio of the number of the shortest paths where the target identity node is located to the total number of the shortest paths may be determined as the risk value of the target identity information. It can be seen that a larger number of the shortest paths where the target identity node is located corresponds to a higher risk value. In this way, it can be determined based on the determined risk value whether the target identity information has a risk of being used fraudulently. Specifically, the determined risk value may be compared with a specified threshold, and if the determined risk value is greater than the specified threshold, it is determined that the target identity information has a risk of being used fraudulently. The specified threshold may be a value obtained through calculation for relationship diagrams of a large number of normal users. For example, a ratio of the number of the shortest paths where an identity node is located to the total number calculated based on a relationship diagram of a normal user is usually not higher than 0.4, then 0.4 or 0.4 added with a certain redundancy value can be used as the specified threshold. A value of the specified threshold may be set flexibly according to an actual application scenario, which is not limited in this specification.

[35] In one embodiment, all identity nodes appearing in a relationship diagram can be processed in the above manner, so as to screen out identity information that has a risk of being used fraudulently. For the identity information with a risk of being used fraudulently, various pieces of account information bound to the identity information with a risk of being used fraudulently can be extracted from the user data. Then, the various pieces of account information may be analyzed one by one, so as to screen out suspicious account information that may be created by a fraudster. When screening suspicious account information, it can be determined whether current account information is simultaneously authenticated on the same device as a large amount of other account information. If so, the current account information may be suspicious account information created by the fraudster. Specifically, for the current account information in the account information bound to the target identity information, an authentication device corresponding to the current account information may be determined according to a historical operation record of the current account information. Then, historical authentication information generated by the authentication device can be searched, and the quantity of accounts authenticated in the authentication device can be counted from the historical authentication information. If the counted quantity is greater than or equal to a specified quantity threshold, it can be determined that the current account information is suspicious account information. The specified quantity threshold can be set flexibly according to an actual situation. For example, an average quantity of authentications of normal authentication devices can be counted, and then the average quantity of authentications or the average quantity of authentications added with a certain redundancy value is used as the specified quantity threshold.

[36] Referring to FIG. 4, an apparatus for identifying identity information is further provided in an embodiment of this specification, including:

[37] a data acquisition unit configured to acquire user data, the user data including identity information and account information of a user;

[38] an association relationship establishing unit configured to establish an association relationship between account information and identity information that are bound in the user data, and establish an association relationship between two pieces of account information having common features in the user data; and

5 [39] an identifying unit configured to determine a risk value of target identity information in the user data according to the established association relationships, and determine, according to the determined risk value, whether the target identity information has a risk of being used fraudulently.

[40] In one embodiment, the association relationship establishing unit includes:

10 [41] a relationship diagram establishing module configured to establish a relationship diagram corresponding to the user data, wherein the relationship diagram includes an identity node for representing identity information in the user data and account nodes for representing account information in the user data, wherein there is a connection line between an account node and the identity node that are bound, and there is a connection line between two account nodes having
15 common features.

[42] In one embodiment, among paths connecting a first account node and a second account node in the relationship diagram, a path including the fewest connection lines is used as the shortest path between the first account node and the second account node; and correspondingly, the identifying unit includes:

20 [43] a shortest path counting module configured to, in the established relationship diagram, count the number of the shortest paths where a target identity node representing the target identity information is located, and count the total number of the shortest paths in the relationship diagram; and

[44] a ratio calculation module configured to determine a ratio of the number of the
25 shortest paths where the target identity node is located to the total number of the shortest paths as the risk value of the target identity information.

[45] In one embodiment, the identifying unit includes:

[46] a threshold comparison module configured to compare the determined risk value with a specified threshold, and if the determined risk value is greater than the specified threshold, determine that the target identity information has a risk of being used fraudulently.

5 [47] In one embodiment, the apparatus further includes:

[48] a suspicious account screening unit configured to extract account information bound to the target identity information from the user data if the target identity information has a risk of being used fraudulently, and screen suspicious account information out from the account information bound to the target identity information.

10 [49] In one embodiment, the suspicious account screening unit includes:

[50] an authentication device determining module configured to, for current account information in the account information bound to the target identity information, determine an authentication device corresponding to the current account information; and

[51] an authentication quantity counting module configured to count the quantity of
15 accounts authenticated in the authentication device, and if the counted quantity is greater than or equal to a specified quantity threshold, determine the current account information as suspicious account information.

[52] Referring to FIG. 5, an apparatus for identifying identity information comprising a memory and a processor is further provided in an embodiment of this specification. The memory
20 is configured to store a computer program. When executed by the processor, the computer program implements the method for identifying identity information described above. Specifically, at the hardware level, the apparatus for identifying may include a processor, an internal bus, and a memory as shown in FIG. 5. The memory may include a memory and a non-volatile memory. The processor reads a corresponding computer program from the non-volatile memory into the memory
25 and runs the computer program. Those of ordinary skill in the art can understand that the structure shown in FIG. 5 is merely illustrative, which does not limit a structure of the above apparatus for

identifying. For example, the apparatus for identifying may further include more or fewer components than those shown in FIG. 5. For example, it may also include other processing hardware, such as a Graphics Processing Unit (GPU), or have a configuration different from that shown in FIG. 5. Definitely, in addition to a software implementation, this application does not
5 exclude other implementations, such as a logic device or a combination of software and hardware.

[53] In the embodiment of the present specification, the processor may include a central processing unit (CPU) or a graphics processor (GPU), and definitely may also include other single-chip microcomputers, logic gate circuits, integrated circuits, etc. with logic processing capabilities, or an appropriate combination thereof. The memory described in the embodiments of the present
10 application may be a memory device for saving information. In a digital system, a device that can save binary data may be a memory; in an integrated circuit, a circuit that has a storage function but does not have a physical form may also be a memory, such as a RAM and a FIFO; in the system, a storage device in a physical form may also be referred to as a memory, etc. During implementation, the memory may also be implemented in the form of a cloud memory. The specific
15 implementation is not limited in this specification.

[54] It should be noted that the apparatus described in the embodiments of this specification may also include other implementations according to the description of the related method embodiments. For a specific implementation, reference may be made to the description of the method embodiments, and details are not repeated herein.

20 [55] As can be seen from the above, in one or more embodiments of this specification, after user data to be analyzed is acquired, association relationships between identity information and account information and between account information and account information in the user data may be established. The association between identity information and account information here can be reflected in that: the account information is registered or authenticated based on the
25 associated identity information; and the association between different pieces of account information can be reflected in that: two pieces of account information have common features.

Further, the two pieces of account information having common features may include, for example, the two pieces of account information having an interaction behavior, the two pieces of account information being bound to the same object, and the two pieces of account information being logged-in on the same device. Generally speaking, among multiple pieces of account information associated with the same identity information, there are often two or more pieces of account information having common features. The reason is that when a user uses multiple accounts created by himself/herself, some of these accounts may be logged-in on the same device, and transferring money, assisting authentication, being bound to the same mobile phone, etc. may also occur in some accounts, thus making these accounts have common features. Once identity information of a certain user is fraudulently used, the account created by the fraudster based on the identity information of the user usually does not have or rarely has common features with other accounts of the user. In this way, by analyzing the associations between identity information and account information and between account information and account information in data of the user, a risk value of a certain piece of target identity information can be determined. Once the risk value is too high, it can be considered that there is a risk of fraudulent use of the target identity information. It can be seen that the technical solution provided by one or more of the embodiments of this specification analyzes the association between identity information and account information to determine whether the identity information is fraudulently used. Even if a fraudster changes an usage environment of an account, the method can still identify whether the account is in an abnormal state based on whether there are common features between the account and other accounts, thereby improving the efficiency of identifying identity information.

[56] The foregoing describes specific embodiments of this specification. Other embodiments are within the scope of the appended claims. In some cases, the actions or steps recorded in the claims can be performed in an order different from the orders in the embodiments and can still achieve desired results. In addition, the processes depicted in the accompanying drawings do not necessarily require the particular order or sequential order shown to achieve the

desired results. In certain implementations, multitasking and parallel processing are also possible or may be advantageous.

[57] Although the present application provides method operation steps as described in the embodiments or flowcharts, more or fewer operation steps may be included based on conventional or non-creative efforts. The sequence of steps listed in the embodiments is only one way in many execution sequences of steps, and does not represent a unique execution sequence. When an actual apparatus or client terminal product is executed, it may be executed sequentially or in parallel according to the method shown in the embodiments or the accompanying drawings (for example, a parallel processor or a multi-thread processing environment).

[58] In the 1990s, it can be obviously distinguished whether an improvement to a technology is an improvement in hardware (e.g., improvement to circuit configuration, such as a diode, a transistor, or a switch) or an improvement in software (improvement to a process). However, with the development of technologies, improvements to many method procedures at present can be considered as direct improvements to hardware circuit configurations. Almost all designers program improved processes into hardware circuits to obtain corresponding hardware circuit configurations. Therefore, it cannot be said that improvement to one process cannot be implemented by a hardware entity module. For example, a Programmable Logic Device (PLD) (e.g., a Field Programmable Gate Array (FPGA)) is such an integrated circuit, and its logic functions are determined by a user through programming the device. Designers "integrate" a digital system onto a piece of PLD by independent programming without asking a chip manufacturer to design and manufacture a dedicated integrated circuit chip. Moreover, at present, instead of manually making an integrated circuit chip, this programming is also implemented mostly using "logic compiler" software, which is similar to a software compiler used for program development and compilation. However, the original code before compilation also has to be compiled using a specific programming language, which is known as a Hardware Description Language (HDL). There is not only one, but many kinds of HDLs, such as ABEL (Advanced Boolean Expression

Language), AHDL (Altera Hardware Description Language), Confluence, CUPL (Cornell University Programming Language), HDCal, JHDL (Java Hardware Description Language), Lava, Lola, MyHDL, PALASM, and RHDL (Ruby Hardware Description Language). At present, VHDL (Very-High-Speed Integrated Circuit Hardware Description Language) and Verilog are the most
5 commonly used. Those skilled in the art should also know that a hardware circuit for implementing a logic method procedure can be easily obtained by slightly logically programming the method procedure using the above several hardware description languages and programming it into an integrated circuit.

[59] The apparatus, modules, or units illustrated in the foregoing embodiments can
10 specifically be implemented by a computer chip or an entity, or implemented by a product having a certain function. A typical implementation device is a computer. Specifically, for example, the computer can be a personal computer, a laptop computer, a cellular phone, a camera phone, a smart phone, a personal digital assistant, a media player, a navigation device, an email device, a game console, a tablet computer, a wearable device, or a combination of any of these devices.

[60] For ease of description, the apparatus is divided into various units based on
15 functions, and the units are described separately. Of course, when implementing this specification, the functions of all units can be implemented in one or more pieces of software and/or hardware.

[61] Those skilled in the art should understand that the embodiments of the present
20 invention can be provided as a method, an apparatus, or a computer program product. Therefore, the present invention may be in the form of a complete hardware embodiment, a complete software embodiment, or an embodiment combining software and hardware. Moreover, the present invention can be in the form of a computer program product implemented on one or more computer-usable storage media (including, but not limited to, a magnetic disk memory, a CD-ROM, an optical memory, and the like) including computer usable program code.

[62] The present invention is described with reference to flowcharts and/or block
25 diagrams of the method, the device (apparatus), and the computer program product according to

the embodiments of the present invention. It should be understood that computer program instructions may be used to implement each process and/or block in the flowcharts and/or block diagrams and combinations of processes and/or blocks in the flowcharts and/or block diagrams. The computer program instructions may be provided to a general-purpose computer, a special-
5 purpose computer, an embedded processor, or a processor of another programmable data processing device to generate a machine, such that instructions executed by the computer or the processor of another programmable data processing device generate an apparatus configured to implement functions designated in one or more processes in a flowchart and/or one or more blocks in a block diagram.

10 [63] The computer program instructions may also be stored in a computer-readable memory that can guide a computer or another programmable data processing device to work in a specific manner, such that the instructions stored in the computer-readable memory generate an article of manufacture including an instruction apparatus, and the instruction apparatus implements functions designated in one or more processes in a flowchart and/or one or more blocks in a block
15 diagram.

[64] The computer program instructions may also be loaded to a computer or another programmable data processing device, such that a series of operation steps are executed on the computer or another programmable device to generate computer-implemented processing, and therefore, the instructions executed on the computer or another programmable device provide steps
20 for implementing functions designated in one or more processes in a flowchart and/or one or more blocks in a block diagram.

[65] In a typical configuration, the computing device includes one or more central processing units (CPUs), an input/output interface, a network interface, and a memory.

[66] The memory can include a volatile memory, a Random Access Memory (RAM),
25 and/or non-volatile memory in computer-readable media, e.g., a Read-Only Memory (ROM) or a flash RAM. The memory is an example of a computer-readable medium.

[67] The computer-readable medium includes non-volatile and volatile media as well as movable and non-movable media, and can implement information storage by means of any method or technology. The information can be a computer-readable instruction, a data structure, a module of a program, or other data. Examples of the storage medium of a computer include, but are not limited to, a phase change memory (PRAM), a static random access memory (SRAM), a dynamic random access memory (DRAM), other types of RAMs, a ROM, an electrically erasable programmable read-only memory (EEPROM), a flash memory or other memory technologies, a compact disk read-only memory (CD-ROM), a digital versatile disc (DVD) or other optical storages, a cassette tape, a magnetic tape/magnetic disk storage, graphene storage, or other magnetic storage devices, or any other non-transmission medium, and can be used to store information accessible to a computing device. According to the definition in this text, the computer-readable medium does not include transitory media, such as a modulated data signal and a carrier.

[68] It should be further noted that the terms "include," "comprise," or any other variations thereof are intended to cover non-exclusive inclusion, so that a process, method, commodity, or device including a series of elements not only includes the elements, but also includes other elements not explicitly listed, or further includes elements inherent to the process, method, commodity, or device. In the absence of more limitations, an element defined by "including a/an..." does not exclude that the process, method, commodity, or device including the element further has other identical elements.

[69] Those skilled in the art should understand that the embodiments of this specification can be provided as a method, an apparatus, or a computer program product. Therefore, this specification may be in the form of a complete hardware embodiment, a complete software embodiment, or an embodiment combining software and hardware. Moreover, this specification can be in the form of a computer program product implemented on one or more computer-usable

storage media (including, but not limited to, a magnetic disk memory, a CD-ROM, an optical memory, and the like) including computer-usable program code.

[70] This specification can be described in general context of a computer-executable instruction executed by a computer, for example, a program module. Generally, the program module includes a routine, a program, an object, an assembly, a data structure, and the like used for executing a specific task or implementing a specific abstract data type. This specification can also be implemented in distributed computing environments. In the distributed computing environments, a task is executed by using a remote processing device connected via a communication network. In the distributed computing environments, the program module may be located in local and remote computer storage media including a storage device.

[71] All the embodiments in this specification are described in a progressive way, the same and similar parts among the embodiments can be referred to by each other, and each embodiment emphasizes a part different from other embodiments. In particular, the apparatus and server embodiments are described relatively simply since they are basically similar to the method embodiment, and the description of part of the method embodiment may be referenced for relevant parts.

[72] The foregoing description is merely embodiments of this specification, and is not used to limit this specification. For those skilled in the art, this specification can have various alterations and changes. Any modification, equivalent replacement, improvement, etc. made within the spirit and principle of the specification should be included in the scope of protection of the claims.

CLAIMS

1. A method for identifying identity information, comprising:

acquiring user data, the user data comprising identity information and account information of a user;

5 establishing an association relationship between account information and identity information that are bound in the user data, and establishing an association relationship between two pieces of account information having common features in the user data; and

determining a risk value of target identity information in the user data according to the established association relationships, and determining, according to the determined risk value,

10 whether the target identity information has a risk of being used fraudulently,

wherein the establishing an association relationship between account information and identity information that are bound in the user data, and establishing an association relationship between two pieces of account information having common features in the user data comprises:

establishing a relationship diagram corresponding to the user data, the relationship diagram
15 comprising an identity node for representing identity information in the user data and account nodes for representing account information in the user data, wherein there is a connection line between an account node and the identity node that are bound, and there is a connection line between two account nodes having common features; and

wherein among paths connecting a first account node and a second account node in the
20 relationship diagram, a path comprising the fewest connection lines is used as the shortest path between the first account node and the second account node; and accordingly, the determining a risk value of target identity information in the user data comprises:

counting, in the established relationship diagram, the number of the shortest paths where a target identity node representing the target identity information is located, and counting the total
25 number of the shortest paths in the relationship diagram, and

determining a ratio of the number of the shortest paths where the target identity node is located to the total number of the shortest paths as the risk value of the target identity information.

2. The method according to claim 1, wherein the two pieces of account information having common features comprises at least one of the following:

- 5 the two pieces of account information having an interaction behavior;
- the two pieces of account information being bound to the same object; and
- the two pieces of account information being logged-in on the same device.

3. The method according to claim 1, wherein the determining whether the target identity information has a risk of being used fraudulently comprises:

- 10 comparing the determined risk value with a specified threshold, and if the determined risk value is greater than the specified threshold, determining that the target identity information has a risk of being used fraudulently.

4. The method according to claim 1, wherein after determining whether the target identity information has a risk of being used fraudulently, the method further comprises:

- 15 extracting account information bound to the target identity information from the user data if the target identity information has a risk of being used fraudulently, and screening suspicious account information out from the account information bound to the target identity information.

5. The method according to claim 4, wherein the screening suspicious account information out from the account information bound to the target identity information comprises:

- 20 determining, for current account information in the account information bound to the target identity information, an authentication device corresponding to the current account information; and

 counting the quantity of accounts authenticated in the authentication device, and if the counted quantity is greater than or equal to a specified quantity threshold, determining the current account

- 25 information as suspicious account information.

6. An apparatus for identifying identity information, comprising:

a data acquisition unit configured to acquire user data, the user data comprising identity information and account information of a user;

an association relationship establishing unit configured to establish an association relationship between account information and identity information that are bound in the user data, and establish
5 an association relationship between two pieces of account information having common features in the user data; and

an identifying unit configured to determine a risk value of target identity information in the user data according to the established association relationships, and determine, according to the determined risk value, whether the target identity information has a risk of being used fraudulently,

10 wherein the association relationship establishing unit comprises:

a relationship diagram establishing module configured to establish a relationship diagram corresponding to the user data, wherein the relationship diagram includes an identity node for representing identity information in the user data and account nodes for representing account information in the user data, wherein there is a connection line between an account node and the
15 identity node that are bound, and there is a connection line between two account nodes having common features, and

wherein among paths connecting a first account node and a second account node in the relationship diagram, a path comprising the fewest connection lines is used as the shortest path between the first account node and the second account node; and correspondingly, the identifying
20 unit comprises:

a shortest path counting module configured to, in the established relationship diagram, count the number of the shortest paths where a target identity node representing the target identity information is located, and count the total number of the shortest paths in the relationship diagram;
and

a ratio calculation module configured to determine a ratio of the number of the shortest paths where the target identity node is located to the total number of the shortest paths as the risk value of the target identity information.

7. The apparatus according to claim 6, wherein the identifying unit comprises:

5 a threshold comparison module configured to compare the determined risk value with a specified threshold, and if the determined risk value is greater than the specified threshold, determine that the target identity information has a risk of being used fraudulently.

8. The apparatus according to claim 6, wherein the apparatus further comprises:

a suspicious account screening unit configured to extract account information bound to the
10 target identity information from the user data if the target identity information has a risk of being used fraudulently, and screen suspicious account information out from the account information bound to the target identity information.

9. The apparatus according to claim 8, wherein the suspicious account screening unit comprises:

15 an authentication device determining module configured to, for current account information in the account information bound to the target identity information, determine an authentication device corresponding to the current account information; and

an authentication quantity counting module configured to count the quantity of accounts authenticated in the authentication device, and if the counted quantity is greater than or equal to a
20 specified quantity threshold, determine the current account information as suspicious account information.

10. An apparatus for identifying identity information comprising a memory and a processor, wherein the memory is configured to store a computer program, and when executed by the processor, the computer program implements the method according to any of claims 1 to 5.