

GROUP FEATURE MINING METHOD AND DEVICE, EVENT IDENTIFICATION METHOD AND DEVICE, AND STORAGE MEDIUM

TECHNICAL FIELD

5 **[0001]** This specification relates to the field of Internet technologies, and in particular, to a group feature mining method and device, an event identification method and device, and a storage medium.

BACKGROUND

10 **[0002]** With development of Internet technologies, problems such as group crimes and network attacks that are performed through a network are increasingly serious, and a quantity of groups that need prevented and controlled on a prevention and control platform is also increasing. In an actual prevention and control process, the prevention and control platform can analyze related information of a large quantity of events based on experience of a prevention and control person, to identify a feature of a target group. Subsequently, the prevention and control platform can
15 prevent and control, based on the feature that is of the target group and that is identified by the prevention and control person, an event executed by the target group. However, the event has a large data amount and high complexity, and efficiency of performing manual analysis and identification by the prevention and control person is low. Consequently, it difficult for the prevention and control platform to implement prevention and control in a timely manner.

20 **[0003]** Content of the background part is merely information learned of by the inventor, and does not mean that the information has entered a public domain before an application date of this disclosure, and does not mean that the information can become a conventional technology of this disclosure.

SUMMARY

25 **[0004]** This specification provides a group feature mining method and device, an event

identification method and device, and a storage medium, to simultaneously mine a feature of a target group for a large amount of event data, improve efficiency of obtaining the feature of the target group, and prevent and control the target group in a timely manner.

[0005] According to a first aspect, this specification provides a group feature mining method, including: obtaining event information of a plurality of normal events and event information of a plurality of abnormal events, where event information of each event includes a plurality of attribute values for describing the event, and an operator of the plurality of abnormal events includes a target group; determining N attribute value combinations based on the event information of the plurality of normal events and the event information of the plurality of abnormal events, where N is an integer greater than or equal to 1, each attribute value combination includes at least one attribute value, an occurrence frequency of each attribute value combination in the plurality of abnormal events is greater than a first preset threshold, and an occurrence frequency of each attribute value combination in the plurality of normal events is 0; and generating a group feature corresponding to the target group based on the N attribute value combinations.

[0006] According to a second aspect, this specification further provides an event identification method, including: obtaining event information of a to-be-identified target event in response to receiving an event identification request, where the event information includes a plurality of attribute values for describing the target event; obtaining a group feature corresponding to a target group, where the group feature is mined in the method according to the first aspect, and is used to describe an attribute feature satisfied by an event executed by the target group; and matching the event information of the target event against the group feature, to determine whether the target event is an event executed by the target group.

[0007] According to a third aspect, this specification provides a group feature mining device, including at least one storage medium and at least one processor. The at least one storage medium stores at least one instruction set, to perform group feature mining; and the at least one processor is communicatively connected to the at least one storage medium. When running, the at least one processor reads the at least one instruction set, and performs the following operations based on an indication of the at least one instruction set: obtaining event information of a plurality of normal events and event information of a plurality of abnormal events, where event information of each event includes a plurality of attribute values for describing the event, and an operator of the plurality of abnormal events includes a target group; determining N attribute value combinations

based on the event information of the plurality of normal events and the event information of the plurality of abnormal events, where N is an integer greater than or equal to 1, each attribute value combination includes at least one attribute value, an occurrence frequency of each attribute value combination in the plurality of abnormal events is greater than a first preset threshold, and an occurrence frequency of each attribute value combination in the plurality of normal events is 0; and generating a group feature corresponding to the target group based on the N attribute value combinations.s

[0008] According to a fourth aspect, this specification provides an event identification device, including at least one storage medium and at least one processor. The at least one storage medium stores at least one instruction set, to perform event identification; and the at least one processor is communicatively connected to the at least one storage medium. When running, the at least one processor reads the at least one instruction set, and performs the following operations based on an indication of the at least one instruction set: obtaining event information of a to-be-identified target event in response to receiving an event identification request, where the event information includes a plurality of attribute values for describing the target event; obtaining a group feature corresponding to a target group, where the group feature is mined in the method according to the first aspect, and is used to describe an attribute feature satisfied by an event executed by the target group; and matching the event information of the target event against the group feature, to determine whether the target event is an event executed by the target group.

[0009] According to a fifth aspect, this specification provides a computer-readable non-transitory storage medium. The computer-readable non-transitory storage medium stores at least one instruction set, and when the at least one instruction set is executed by at least one processor, the group feature mining method according to the first aspect is implemented, or the event identification method according to the second aspect is implemented.

[0010] It can be learned from the foregoing technical solutions that, in the group feature mining method and device, and the storage medium provided in this specification, an attribute value combination can be determined based on obtained event information of a normal event and obtained event information of an abnormal event, and subsequently, a group feature that represents a target group is generated based on the attribute value combination. The attribute value combination can highly summarize the abnormal event and does not summarize the normal event. In other words, the attribute value combination frequently occurs in the event information of the

abnormal event, and does not occur in the event information of the normal event. It can be learned that in this solution, the event information of the normal event is considered in a process of generating the group feature of the target group, to avoid as far as possible a case in which the determined group feature once occurs in the event information of the normal event, so as to accurately prevent and control the target group, without affecting a normal user or group. In addition, in this solution, when the group feature of the target group is mined, a plurality of normal events and event information of the plurality of abnormal events can be simultaneously processed, and event information of a plurality of abnormal events can be generated through execution of a plurality of target groups. Therefore, in this solution, group feature mining efficiency can be improved, and prevention and control of the target group are performed in a more timely manner.

[0011] Further, in the event identification method and device, and the storage medium provided in this specification, a target event can be identified based on the group feature generated in the mining method. After it is identified that the target event is an event executed by the target group, execution of the target event can be impeded by using a means, to prevent and control the target group. According to the event identification solution, prevention and control timeliness and accuracy can be improved.

[0012] Other functions of the group feature mining method and device, the event identification method and device, and the storage medium provided in this specification are listed in the following descriptions. The creative aspects of the group feature mining method and device, the event identification method and device, and the storage medium provided in this specification can be fully explained by practice or by using the method, device, and combination described in the following detailed example.

BRIEF DESCRIPTION OF DRAWINGS

[0013] To describe the technical solutions in the embodiments of this specification more clearly, the accompanying drawings required for describing the embodiments are described briefly below. Clearly, the accompanying drawings in the following descriptions show some embodiments of this specification, and a person of ordinary skill in the art can still derive other accompanying drawings from these accompanying drawings without creative efforts.

[0014] FIG. 1 is a schematic diagram of an application scenario of group feature mining and

event identification according to an embodiment of this specification;

[0015] FIG. 2 is a diagram of a hardware structure of an electronic device according to an embodiment of this specification;

5 [0016] FIG. 3 is a schematic flowchart of a group feature mining method according to an embodiment of this specification;

[0017] FIG. 4 is a schematic flowchart of determining N attribute value combinations in Manner 1;

[0018] FIG. 5 is a schematic diagram of a header table established in Manner 1;

[0019] FIG. 6 is a schematic diagram of an attribute value tree established in Manner 1;

10 [0020] FIG. 7 is a schematic flowchart of determining N attribute value combinations in Manner 2;

[0021] FIG. 8 is a schematic diagram of a header table established in Manner 2; and

[0022] FIG. 9 is a schematic flowchart of an event identification method according to an embodiment of this specification.

15 DESCRIPTION OF EMBODIMENTS

[0023] Specific application scenarios and requirements of this specification are provided in the following descriptions, so that a person skilled in the art can manufacture and use content of this specification. For the person skilled in the art, various local modifications to the disclosed embodiments are apparent, and the general principles defined herein can be applied to other
20 embodiments and applications without departing from the spirit and scope of this specification. Therefore, this specification is not limited to the shown embodiments, but has a widest scope consistent with that of the claims.

[0024] The terms used herein are only intended to describe a particular example embodiment, but do not impose a limitation. For example, unless otherwise specified in the context, the singular
25 forms "one", "a", and "the" used herein can also include a plural form. When used in this specification, the terms "include", "comprise", and/or "have" indicate/indicates existence of an associated integer, step, operation, element, and/or component, but does not exclude existence of one or more other features, integers, steps, operations, elements, components, and/or groups, or another feature, integer, step, operation, element, component, and/or group can be added to the

system/method.

[0025] In consideration of the following descriptions, these and other features of this specification, operations and functions of related components of the structure, and economy of combination and manufacturing of components can be significantly improved. With reference to the accompanying drawings, all of these form a part of this specification. However, it should be clearly understood that the accompanying drawings are merely used for the purpose of illustration and description, and are not intended to limit the scope of this specification. It should be further understood that the accompanying drawings are not drawn to scale.

[0026] The flowchart used in this specification illustrates operations implemented by a system in some embodiments of this specification. It should be clearly understood that operations in the flowchart can be implemented out of sequence. On the contrary, operations can be implemented reversely or simultaneously. In addition, one or more other operations can be added to the flowchart. One or more operations can be removed from the flowchart.

[0027] For ease of description, this specification provides the following explanations about the terms that appear in the following descriptions.

[0028] Target group: In a field in which prevention and control is required, the target group can be understood as a prevented and controlled organization or group. In another field, the target group can be understood as an organization or a group that needs to be targeted in the field. One target group can include one or more objects. The object can be represented by using a user account, a user name, a device identifier (ID), a card number, etc., or the object can be represented by using a combination of two or more of the foregoing items.

[0029] A group feature is an attribute feature satisfied by an event executed by a group.

[0030] An attribute value includes an attribute and a value. In other words, the attribute value includes information about two dimensions: the attribute and the value. For example, the attribute value is a card number 1111. The card number is an attribute, and 1111 is a value.

[0031] A normal event is an event executed by an object or a group other than the foregoing target group. For ease of description below, the object or the group other than the target group is referred to as a normal group. Usually, an attribute value for describing the normal event is a normal value, without an exception, or some attribute values are slightly abnormal, but do not affect characterization of an event.

[0032] An abnormal event is an event executed by the foregoing target group, or some events

executed by a normal group. An attribute value for describing the some events is abnormal, or a combination of attribute values for describing the some events is abnormal.

[0033] As described in the background, in a process of preventing and controlling a target group, a prevention and control person can manually analyze event information, to mine a group feature corresponding to the target group. The foregoing analysis and mining process relies heavily on professional experience accumulated by the prevention and control person. Analysis and mining accuracy varies based on different professional degrees of the prevention and control person. In addition, one analysis and mining process can be specific to only one target group. Consequently, efficiency is relatively low, and it is difficult to implement prevention and control in a timely and effective manner.

[0034] In view of this, this specification provides a group feature mining method, to simultaneously process a large amount of event information, and improve group feature mining efficiency. In addition, in the mining method, a normal event and an abnormal event can be distinguished, and the normal event is used as a reference to obtain a group feature that highly represents an abnormal event executed by a target group, to improve group feature mining accuracy.

[0035] Further, this specification further provides an event identification method, to identify a target event based on the group feature mined in the mining method, in other words, identify whether the target event is an event executed by the target group. If the target event is an event executed by the target group, execution of the target event is prevented from being completed, to prevent and control the target group. According to the event identification method, timeliness and accuracy of preventing and controlling the target group can be improved.

[0036] The following describes, with reference to FIG. 1, an application scenario of group feature mining and event identification solutions provided in this specification.

[0037] FIG. 1 is a schematic diagram of an application scenario of group feature mining and event identification according to an embodiment of this specification. As shown in FIG. 1, a scenario 001 can include a target group 100, a terminal device 200, a server 300, a network 400, a mining device 500, a normal group 600, and a terminal device 700.

[0038] The target group 100 can execute an abnormal event by using the terminal device 200, and the normal group can execute a normal event by using the terminal device 700. A person skilled in the art can understand that the abnormal event and the normal event are different in different service scenarios. For example, in a payment scenario, the normal event can be a payment

event executed by an authorized user by using a card/account of the authorized user, and the abnormal event can be a payment event executed by an unauthorized user by embezzling a card/account of the authorized user.

[0039] The terminal device 200 or 700 can display an interaction interface to a user, so that the user executes some events such as a payment event, a transfer event, etc. in the interaction interface. The terminal device 200 or 700 can include a mobile device, a tablet, a laptop computer, a device built in a motor vehicle or similar content, or any combination thereof. In some embodiments, the mobile device can include a smart home device, a smart mobile device, a virtual reality device, an augmented reality device or a similar device, or any combination thereof. In some embodiments, the smart home apparatus can include a smart television, a desktop computer, or any combination. In some embodiments, the smart mobile device can include a smartphone, a personal digital assistant, a game device, a navigation device, or any combination thereof. In some embodiments, the virtual reality device or the augmented reality device may include a virtual reality helmet, virtual reality glasses, a virtual reality patch, an augmented reality helmet, augmented reality glasses, an augmented reality patch or similar content, or any combination thereof. For example, the virtual reality device or the augmented reality device may include Google glasses, a head-mounted display, and VR. In some embodiments, the apparatus built in the motor vehicle can include a vehicle-mounted computer, a vehicle-mounted television, etc. In some embodiments, the terminal device 200 or 700 can be a device with a location technology, and is configured to locate the terminal device 200 or 700.

[0040] In some embodiments, one or more applications (app) can be installed in the terminal device 200 or 700. The app can provide the user with a capability of executing some events (for example, a payment event, or a transfer event). The app includes but is not limited to a web browser app program, a search app program, a chat app program, a shopping app program, a video app program, a financial app program, etc.

[0041] As shown in FIG. 1, the terminal device 200 and the terminal device 700 can be communicatively connected to the server 300. In some embodiments, the server 300 can be further communicatively connected to another terminal device, and receive data sent by the terminal device. In some embodiments, the terminal device 200 and the terminal device 700 can interact with the server 300 through the network 400, to receive or send a message, etc. The server 300 can be a server that provides various services, for example, a background server that provides support

for an overseas shopping platform installed in the terminal device 200. In this case, the server 300 can store event information of an abnormal transaction event executed by the target group 100 and event information of a normal transaction event executed by the normal group 600.

[0042] The network 400 is configured to provide a medium for a communication connection between the terminal device and the server 300, or is configured to provide a medium for a communication connection between the mining device 500 and the server 300. The network 400 can facilitate exchange of information or data. In some embodiments, the network 400 can be any type of wired or wireless network, or a combination thereof. For example, the network 400 can include a cable network, a wired network, an optical fiber network, a telecommunications network, an intranet, the Internet, a local area network (LAN), a wide area network (WAN), a wireless local area network (WLAN), a metropolitan area network (MAN), a public switched telephone network (PSTN), a Bluetooth network, a ZigBee network, a near field communication (NFC) network, or a similar network. In some embodiments, the network 400 can include one or more network access points. For example, the network 400 can include a wired or wireless network access point, for example, a base station or an Internet switching point. Through the access point, one or more components in the terminal device 200, the server 300, and the mining device 500 can be connected to the network 400 to exchange data or information.

[0043] The mining device 500 is an electronic device with a certain data processing capability. The mining device 500 can execute data and instructions of a group feature mining method described in this specification, and can execute or be configured to execute the data and instructions. In some embodiments, the mining device 500 can include a hardware device with a data information processing function and a necessary program required for driving the hardware device to work. The mining device 500 can communicate with the server 300, and obtain event information of the normal event and event information of the abnormal event from the server 300. For example, the mining device 500 can obtain the event information from the server 300 through the network 400, or a person skilled in the art can manually import the event information in the server 300 into the mining device 500.

[0044] In some embodiments, after the mining device 500 obtains a group feature in the group feature mining method provided in this specification, the group feature is transmitted to the server 300 through the network 400. The server 300 can implement, based on the group feature, an event identification method provided in this specification. The server 300 can execute data and

instructions of the event identification method described in this specification, and can execute or be configured to execute the data and instructions. In some other embodiments, after the mining device 500 obtains a group feature in the group feature mining method provided in this specification, the group feature is transmitted to the terminal device 200 or 700 through the network 400, or the group feature is transmitted to the server 300 through the network 400 and then distributed by the server 300 to the terminal device 200 or 700. The terminal device 200 or 700 can implement, based on the group feature, an event identification method provided in this specification. The terminal device 200 or 700 can execute data and instructions of the event identification method described in this specification, and can execute or be configured to execute the data and instructions.

[0045] It should be understood that a quantity of terminal devices, a quantity of servers 300, a quantity of networks 400, and a quantity of mining devices 500 in FIG. 1 are merely an example. There can be any quantity of terminal devices, any quantity of servers 300, any quantity of networks 400, and any quantity of mining devices 500 based on an implementation requirement.

[0046] It should be noted that the group feature mining method can be completely performed on the mining device 500, or can be completely performed on the server 300, or can be partially performed on the mining device 500 and partially performed on the server 300. An example in which the group feature mining method is performed on the mining device 500 is used for description in the following descriptions. The event identification method can be completely performed on the terminal device 200, or can be completely performed on the server 300, or can be partially performed on the terminal device 200 and partially performed on the server 300. An example in which the event identification method is performed on the server 300 is used for description in the following descriptions.

[0047] FIG. 2 is a diagram of a hardware structure of an electronic device 800 according to an embodiment of this specification. The electronic device 800 can be used as the mining device 500 in FIG. 1, or can be used as the server 300 in FIG. 1, or can be used as the terminal device 200 or 700 in FIG. 1.

[0048] As shown in FIG. 2, the electronic device 800 can include at least one storage medium 830 and at least one processor 820. In some embodiments, the electronic device 800 can further include a communication port 850 and an internal communication bus 810. In addition, the electronic device 800 can further include an I/O component 860.

[0049] The internal communication bus 810 can be connected to different system components, including a storage medium 830, a processor 820, and a communication port 850.

[0050] The I/O component 860 supports an input/output between the electronic device 800 and another component.

5 **[0051]** The communication port 850 is used for data communication between the electronic device 800 and the outside. For example, the communication port 850 can be used for data communication between the electronic device 800 and a network 400. The communication port 850 can be a wired communication port, or can be a wireless communication port.

[0052] The storage medium 830 can include a data storage apparatus. The data storage
10 apparatus can be a non-transitory storage medium, or can be a transitory storage medium. For example, the data storage apparatus can include one or more of a disk 832, a read-only storage medium (ROM) 834, or a random access storage medium (RAM) 836. The storage medium 830 further includes at least one instruction set stored in the data storage apparatus. The instruction set can be computer program code. When the electronic device 800 is the mining device 500, the
15 computer program code can include a program, a routine, an object, a component, a data structure, a process, a module, etc. for performing the group feature mining method provided in this specification. When the electronic device 800 serves as the server 300 or the terminal device 200 or 700, the computer program code can include a program, a routine, an object, a component, a data structure, a process, a module, etc. for performing the event identification method provided
20 in this specification.

[0053] The processor 820 can be communicatively connected to the storage medium 830 and the communication port 850 through the internal communication bus 810. The processor 820 is configured to execute the at least one instruction set. When the electronic device 800 runs as the mining device 500, the processor 820 reads the at least one instruction set and performs, based on
25 an indication of the at least one instruction set, the group feature mining method provided in this specification. When the electronic device 800 runs as the server 300 or the terminal device 200 or 700, the processor 820 reads the at least one instruction set, and performs, based on an indication of the at least one instruction set, the event identification method provided in this specification.

[0054] The processor 820 can be in a form of one or more processors. In some embodiments,
30 the processor 820 can include one or more hardware processors, for example, a microcontroller, a microprocessor, a reduced instruction set computer (RISC), an application-specific integrated

circuit (ASIC), an application-specific instruction set processor (ASIP), a central processing unit (CPU), a graphics processing unit (GPU), a physical processing unit (PPU), a microcontroller unit, a digital signal processor (DSP), a field programmable gate array (FPGA), an advanced RISC machine (ARM), a programmable logic device (PLD), any circuit or processor that can perform one or more functions, or any combination thereof.

[0055] To merely describe a problem, only one processor 820 is shown in the electronic device 800 in the accompanying drawings. However, it should be noted that the electronic device 800 in this specification can further include a plurality of processors. Therefore, operations and/or method steps disclosed in this specification can be performed by one processor or can be performed jointly by a plurality of processors, as described in this specification. For example, if the processor 820 in the electronic device 800 in this specification performs step A and step B, it should be understood as that step A and step B can be jointly or separately performed by two different processors 820 (for example, a first processor performs step A, and a second processor performs step B, or a first processor and a second processor jointly perform step A and step B).

[0056] FIG. 3 is a flowchart of a group feature mining method P100 according to an embodiment of this specification. As described above, a mining device 500 can perform the mining method P100 in this specification. Specifically, a processor 820 in the mining device 500 can read an instruction set stored in a local storage medium of the mining device 500, and then perform, based on a stipulation of the instruction set, the mining method P100 described in this specification. The method P100 can be applied to feature mining of a card stolen group in the transaction field, feature mining of a network attack group in the network security field, feature mining of a fraud group in the financial field, etc. To describe the method P100 clearly and comprehensively, feature mining of a card stolen group (a target group) of a shopping platform is used as an example for description in below, and application of the method P100 to another field can be obtained by analogy.

[0057] As shown in FIG. 3, the group feature mining method P100 can include the following steps.

[0058] S110: Obtain event information of a plurality of normal events and event information of a plurality of abnormal events, where event information of each normal event includes a plurality of attribute values for describing the normal event, event information of each abnormal event includes a plurality of attribute values for describing the abnormal event, and an operator of

the plurality of abnormal events includes the target group.

[0059] In a group feature mining scenario of the shopping platform, an event can be understood as a transaction, and one event is one transaction. The mining device 500 can obtain transaction data of a completed transaction such as transaction information of a normal transaction and transaction information of an abnormal transaction from a server of the shopping platform. Usually, after a transaction is completed, the shopping platform identifies, by using identification information, whether the transaction is a normal transaction or an abnormal transaction.

[0060] Event information of an event can include a plurality of attribute values for describing the event. Specifically, the event can have a plurality of attributes and a value corresponding to each attribute. In this specification, each attribute and a value corresponding to the attribute are referred to as an attribute value. For example, it is assumed that an event x is a payment event, and event information of the event x can be represented in the following form:

{[attribute 1, value 1], [attribute 2, value 2], [attribute 3, value 3], ...}.

[0061] Alternatively, event information of the event x can be represented in a form shown in Table 1.

Table 1

Event identifier	Attribute 1	Attribute 2	Attribute 3	...
Event x	Value 1	Value 2	Value 3	...

[0062] An attribute value for describing each transaction in transaction information of the transaction include a basic attribute value and a statistical attribute value. The basic attribute value is an attribute value displayed in transaction details after the transaction is completed, for example, a shopping account number, a payment card number, a shopping amount, a shopping time, a delivery address, a shopping device ID, or interaction between a purchaser and a seller. The statistical attribute value is an attribute value obtained by collecting statistics about transactions in a recent period of time of a transaction after the transaction is completed, or an attribute value obtained by processing a basic attribute value of the transaction. For example, a transaction time of the transaction is January 1, and the statistical attribute value can be a total quantity of transactions completed by the shopping account on January 1, a transaction amount, etc.

[0063] In addition, a plurality of obtained abnormal transactions are not necessarily executed

by the target group, and some abnormal transactions may be generated by a normal group by performing a wrong operation. Herein, there are two cases. In Case 1, the plurality of obtained abnormal transactions are all executed by the target group. In Case 2, some abnormal transactions in the plurality of obtained abnormal transactions are executed by the target group, and the other
5 abnormal transactions are generated by the normal group.

[0064] In Case 1, after initially obtaining transaction information of abnormal transactions, the mining device 500 can obtain an abnormality label of each abnormal transaction. The abnormality label can represent a degree of abnormality of the abnormal transaction. Therefore, the mining device 500 can select the plurality of abnormal transactions from the initially obtained abnormal
10 transactions based on the abnormality label of each abnormal transaction. It can be understood that, after a transaction is completed, the shopping platform sets an abnormality label for each transaction, or the shopping platform sets an abnormality label for a transaction whose degree of abnormality reaches a certain degree. The abnormality label can be understood as the identification information for distinguishing whether the transaction is a normal transaction or an abnormal
15 transaction, and a value corresponding to the identification information can be quantized. For example, if the value corresponding to the identification information of the transaction is greater than or equal to a threshold #1, it indicates that the transaction is an abnormal transaction. If the value corresponding to the identification information of the transaction is greater than or equal to a threshold #2, it indicates that the transaction is an abnormal transaction, and the transaction is a
20 transaction executed by the target group. Alternatively, the abnormality label and the identification information are two pieces of independent information. As an abnormal transaction, the transaction not only includes identification information for identifying that the transaction is an abnormal transaction, but also includes abnormality label information for identifying whether the transaction is a transaction executed by the target group.

[0065] In Case 2, after initially obtaining transaction information of abnormal transactions, the mining device 500 does not distinguish whether the abnormal transactions are executed by the target group, and determines that the initially obtained abnormal transactions are all the plurality of abnormal transactions. For example, if the mining device 500 initially obtains an abnormal transaction set A, the plurality of abnormal transactions obtained by the mining device 500 in S110
30 are all abnormal transactions in the abnormal transaction set A.

[0066] For the two different cases, implementation of subsequent steps can be different.

Specific differences are described below.

[0067] S120: Determine N attribute value combinations based on the event information of the plurality of normal events and the event information of the plurality of abnormal events, where N is an integer greater than or equal to 1, each attribute value combination includes at least one attribute value, an occurrence frequency of each attribute value combination in the plurality of abnormal events is greater than a first preset threshold, and an occurrence frequency of each attribute value combination in the plurality of normal events is 0.

[0068] The mining device 500 can determine N attribute value combinations based on transaction information of a plurality of normal transactions and transaction information of a plurality of abnormal transactions in two manners.

[0069] Manner 1

[0070] The mining device 500 determines M candidate attribute value combinations based on the transaction information of the plurality of abnormal transactions. An occurrence frequency of each candidate attribute value combination in the plurality of abnormal transactions is greater than a preset threshold #3, and each candidate attribute value combination includes one or more attribute values.

[0071] The M candidate attribute value combinations include a candidate attribute value combination #1, and the candidate attribute value combination #1 includes an attribute value a, an attribute value b, and an attribute value c. If transaction information of one abnormal transaction includes an entirety of the attribute value a, the attribute value b, and the attribute value c, it is considered that the candidate attribute value combination #1 occurs for one time in the plurality of abnormal transactions. If the attribute value a, the attribute value b, and the attribute value c occur as a whole for m times in the plurality of abnormal transactions, it is considered that the candidate attribute value combination #1 occurs for m times in the plurality of abnormal transactions. Herein, m is an integer greater than 1.

[0072] The occurrence frequency can be a quantity of times, and the preset threshold #3 is also a quantity of times. Alternatively, the occurrence frequency is a percentage of a quantity of times of occurrence to a quantity of abnormal transactions, and the preset threshold #3 is also a percentage.

[0073] It should be understood that, that the attribute value a, the attribute value b, and the attribute value c occur as a whole in transaction information of one abnormal transaction does not

mean that the attribute value a, the attribute value b, and the attribute value c are adjacent attribute values. Once transaction information of one abnormal transaction includes the attribute value a, the attribute value b, and the attribute value c, it is considered that the entirety of the attribute value a, the attribute value b, and the attribute value c occurs in the abnormal transaction.

5 [0074] In some embodiments, as shown in FIG. 4, the mining device 500 can perform de-duplication on attribute values in the transaction information of the plurality of abnormal transactions (n1 abnormal transactions in the example in FIG. 4, where n1 is greater than 1), to obtain a plurality of target attribute values. Then, for each target attribute value in the plurality of target attribute values, the mining device 500 can determine a degree of support for the target
10 attribute value based on an occurrence frequency of the target attribute value in the plurality of abnormal transactions. Further, the mining device 500 determines the M candidate attribute value combinations based on the transaction information of the plurality of abnormal transactions and degrees of support for the plurality of target attribute values. Subsequently, some candidate attribute value combinations are excluded from the M candidate attribute value combinations
15 based on the transaction information of the plurality of abnormal transactions (n2 normal transactions in the example in FIG. 4, where n2 is greater than 1), to obtain the N attribute value combinations.

[0075] For example, the plurality of abnormal transactions are 10 abnormal transactions. Transaction information of an abnormal transaction 1 includes attribute values a, b, c, e, f, and o,
20 transaction information of an abnormal transaction 2 includes attribute values a, c, and g, transaction information of an abnormal transaction 3 includes attribute values e and i, transaction information of an abnormal transaction 4 includes attribute values a, c, d, e, and g, transaction information of an abnormal transaction 5 includes attribute values a, c, e, g, and l, transaction information of an abnormal transaction 6 includes attribute values e and j, transaction information
25 of an abnormal transaction 7 includes attribute values a, b, c, e, f, and p, transaction information of an abnormal transaction 8 includes attribute values a, c, and d, transaction information of an abnormal transaction 9 includes attribute values a, c, e, g, and m, and transaction information of an abnormal transaction 10 includes attribute values a, c, e, g, and n. De-duplication is performed on the attribute values in the transaction information of the 10 abnormal transactions, to obtain
30 target attribute values a, b, c, d, e, f, g, i, j, l, m, n, o, and p. De-duplication can be understood as that the same attribute values in different abnormal transactions are used as one target attribute

value. For example, if the attribute value a exists in the plurality of abnormal transactions, the target attribute values include only one attribute value a, and de-duplication is performed on the other attribute values a. After the plurality of target attribute values are obtained, a degree of support is calculated for each target attribute value. For example, Degree of support=Occurrence frequency of a target attribute value/Quantity of abnormal transactions, or Degree of support=Occurrence frequency of a target attribute value. For the attribute value a, the attribute value a occurs for eight times in the abnormal transactions, and the quantity of abnormal transactions is 10. Therefore, Degree of support for the attribute value a=8/10=80%, or a degree of support for the attribute value a is 8. A degree of support for another target attribute value is determined by analogy. Details are not described again. The following describes the solution by using an example in which Degree of support=Occurrence frequency of a target attribute value.

[0076] In a manner of determining the M candidate attribute value combinations, the mining device 500 establishes a header table based on values of the degrees of support for the target attribute values, and sorts target attribute values of each abnormal transaction based on the values of the degrees of support for the target attribute values. Further, an attribute value tree is established based on the header table and sorted abnormal transactions. After the attribute value tree is established, mining is sequentially performed upward from a bottom of the header table. In a mining process, an attribute value subtree corresponding to each target attribute value in the header table is found in the attribute value tree, and the M candidate attribute value combinations are mined based on the attribute value subtree corresponding to each target attribute value.

[0077] For example, the degree of support for each target attribute value in a, b, c, d, e, f, g, i, j, l, m, n, o, and p is determined, and a target attribute value whose degree of support is less than a threshold #4 is deleted. For example, degrees of support for the target attribute values i, j, l, m, n, o, and p are all less than the threshold #4. Therefore, the i, j, l, m, n, o, and p do not occur in the header table, and the header table is established in descending order of values of degrees of support for the remaining target attribute values a, b, c, d, e, f, and g. The established header table is shown in FIG. 5. As shown in FIG. 5, a target attribute value whose degree of support is less than the threshold #4 is deleted from transaction information of each abnormal transaction, and for each abnormal transaction, target attribute values included in each abnormal transaction are sorted based on values of degrees of support, to form transaction information of sorted abnormal transactions.

[0078] Then, the attribute value tree is established based on the header table and the transaction

information of the sorted abnormal transactions. In other words, the transaction information of the sorted abnormal transactions is sequentially inserted into the tree. The attribute value tree established based on the header table and the transaction information of the sorted abnormal transactions is shown in FIG. 6. A number in FIG. 6 represents a total quantity of times that a target attribute value is repeated in a process of inserting the transaction information of the abnormal transaction into the tree. An arrow in FIG. 6 represents a link from a target attribute value in the header table to a node in the tree.

[0079] Then, mining is sequentially performed upward from the bottom of the header table based on the attribute value tree, and one time of mining is performed for one target attribute value.

For each target attribute value, a conditional pattern base corresponding to the target attribute value is found, in other words, a subtree in which the target attribute value is used as a leaf node is found. After the subtree is obtained, a count of each attribute value node in the subtree is set to a count of a leaf node, and a node whose count is less than a degree of support is deleted. After mining is completed for all target attribute values in the header table, the M candidate attribute value combinations can be obtained. For example, starting from the target attribute value f at the bottom of the header table, a conditional pattern base of f is first searched for. Because f has only one node in the tree, there is only one candidate path, namely, {a:8, c:8, e:6, b:2, f:2}. Then, a count of another node on the path is set to a count of the node of f. Therefore, a subtree corresponding to the node of f becomes {a:2, c:2, e:2, b:2, f:2}. A combination of two, three, ..., or more target attribute values is obtained based on the subtree corresponding to the node of f. By analogy, after mining is performed for each target attribute value in the header table, a plurality of combinations of target attribute values are obtained, and are used as the M candidate attribute value combinations.

[0080] In another manner of determining the M candidate attribute value combinations, after determining degrees of support for all target attribute values, the mining device 500 deletes a target attribute value whose degree of support is less than a threshold #4. In other words, when the candidate attribute value combination is determined subsequently, a deleted target attribute value is not considered. For target attribute values that remain after deletion, a degree of support for a combination of any two target attribute values, a degree of support for a combination of any three target attribute values, and a degree of support for a combination of any four target attribute values are sequentially determined, until a degree of support for a combination of a maximum quantity of target attribute values is determined. The maximum quantity is a quantity of target attribute

values that remain after deletion.

[0081] For example, when the degree of support for the combination of two target attribute values is determined, Degree of support for a combination of target attribute values=Occurrence frequency of the combination of the target attribute values in abnormal transactions/Quantity of abnormal transactions. A combination ace of target attribute values occurs for six times in abnormal transactions, and a quantity of abnormal transactions is 10. Therefore, Degree of support for a combination ace of target attribute values=6/10=60%, or Degree of support for a combination ace of target attribute values=6. A degree of support for another combination of a plurality of target attribute values is determined by analogy. Details are not described again.

[0082] After degrees of support for different combinations of target attribute values are determined, for any combination of target attribute values, whether the combination of target attribute values is a candidate attribute value combination is determined based on a degree of support for the combination of target attribute values and the threshold #3. For example, if the degree of support for the combination of target attribute values is less than or equal to the threshold #3, the combination of target attribute values is not a candidate attribute value combination; or if the degree of support for the combination of target attribute values is greater than the threshold #3, the combination of target attribute values is a candidate attribute value combination.

[0083] The following describes a specific manner of determining the N attribute value combinations from the M candidate attribute value combinations based on the transaction information of the plurality of normal transactions. Herein, M is an integer greater than or equal to N.

[0084] In some embodiments, the mining device 500 obtains the N attribute value combinations by excluding a first attribute value combination from the M candidate attribute value combinations based on the transaction information of the plurality of normal transactions. The first attribute value combination has the same attribute values as transaction information of one of the plurality of normal transactions, and a quantity of same attribute values is greater than or equal to a threshold #5.

[0085] For example, the M candidate attribute value combinations are candidate attribute value combinations aceg, ace, acg, aeg, and ac. When the threshold #5 is a quantity of attribute values in the candidate attribute value combination, if a candidate attribute value combination occurs in transaction information of a normal transaction, it is considered that the candidate attribute value

combination is the first attribute value combination. For the candidate attribute value combination
ace, whether the attribute value combination ace occurs in transaction information of each normal
transaction is searched for. If transaction information of one normal transaction includes the
attribute value combination ace, it is considered that the attribute value combination ace is the first
5 attribute value combination.

[0086] For another example, when the threshold #5 is a product of a preset percentage and a
quantity of attribute values in the candidate attribute value combination, if the preset percentage is
60%, and more than 60% of the attribute values in the candidate attribute value combination occur
in transaction information of a normal transaction, it is considered that the candidate attribute value
10 combination is the first attribute value combination. For the candidate attribute value combination
ace, whether two attribute values or three attribute values in the attribute value combination ace
occur in transaction information of each normal transaction is searched for. If transaction
information of a normal transaction includes two attribute values or three attribute values in the
attribute value combination ace, it is considered that the attribute value combination ace is the first
15 attribute value combination.

[0087] In a manner in the foregoing example, whether the candidate attribute value
combination is the first attribute value combination is sequentially determined, and the first
attribute value combination is excluded from the M candidate attribute value combinations, to
obtain the N attribute value combinations. In this manner, the determined N attribute value
20 combinations can be prevented from occurring in the transaction information of the normal
transaction. Therefore, when a transaction executed by the target group is subsequently identified
based on the N attribute value combinations, a transaction of the normal group can be prevented
from being affected, and identification accuracy can be improved.

[0088] Manner 2

25 **[0089]** The mining device 500 determines N candidate attribute value combinations based on
the transaction information of the plurality of abnormal transactions and the transaction
information of the plurality of normal transactions.

[0090] In some embodiments, as shown in FIG. 7, the mining device 500 can perform de-
duplication on attribute values in the transaction information of the plurality of abnormal
30 transactions in a de-duplication manner in Manner 1, to obtain a plurality of target attribute values.
After obtaining the plurality of target attribute values, the mining device 500 determines a degree

of support for each target attribute value. A manner used herein is different from the manner of determining the degree of support in Manner 1. In Manner 2, when determining the degree of support for the target attribute value, the mining device 500 further needs to consider an occurrence frequency of the target attribute value in the normal transactions in addition to an occurrence frequency of the target attribute value in the abnormal transactions. Subsequently, the N attribute value combinations are determined based on the degree of support for the target attribute value, the transaction information of the plurality of abnormal transactions (n1 abnormal transactions in the example shown in FIG. 7, where n1 is greater than 1), and the transaction information of the plurality of normal transactions (n2 normal transactions in the example shown in FIG. 7, where n2 is greater than 1).

[0091] For example, the plurality of abnormal transactions are the 10 abnormal transactions, and two normal transactions exist, and respectively have attribute values a, c, q, y, w and attribute values a, c, h, u, and v. For each target attribute value in the target attribute values a, b, c, d, e, f, g, i, j, l, m, n, o, and p, the mining device 500 obtains a first occurrence frequency of the target attribute value in the plurality of abnormal transactions, obtains a second occurrence frequency of the target attribute value in the plurality of normal transactions, and determines the degree of support for the target attribute value based on the first occurrence frequency and the second occurrence frequency.

[0092] For example, for each target attribute value, the mining device 500 can determine a difference between the first occurrence frequency and the second occurrence frequency, and determine the degree of support for the target attribute value based on the difference. The first occurrence frequency is a quantity of times of occurrence of the target attribute value in the abnormal transactions, and the second occurrence frequency is a quantity of times of occurrence of the target attribute value in the normal transactions. Degree of support for the target attribute value=(First occurrence frequency–Second occurrence frequency); or Degree of support for the target attribute value=(First occurrence frequency–Second occurrence frequency)/(Quantity of normal transactions+Quantity of abnormal transactions); or Degree of support for the target attribute value=(First occurrence frequency–Second occurrence frequency)/Quantity of abnormal transactions; or Degree of support for the target attribute value=(First occurrence frequency/Quantity of abnormal transactions)–(Second occurrence frequency/Quantity of normal transactions); Alternatively, the degree of support for the target attribute value is determined in a

manner similar to the foregoing formulas. Determining manners are not listed one by one herein. In this specification, "/" in a formula indicates division. Degree of support for the target attribute value=(First occurrence frequency–Second occurrence frequency) is used as an example for description. If a first occurrence frequency of the target attribute value a is 8, and a second occurrence frequency of the target attribute value a is 2, Degree of support for the target attribute value a=8–2=6. A degree of support for another target attribute value is obtained by analogy.

5 [0093] In a manner of determining the N attribute value combinations, the mining device 500 establishes a header table based on values of the degrees of support for the target attribute values, and sorts target attribute values of each abnormal transaction based on the values of the degrees of support for the target attribute values. Further, an attribute value tree is established based on the header table and sorted abnormal transactions. After the attribute value tree is established, mining is sequentially performed upward from a bottom of the header table. In a mining process, an attribute value subtree corresponding to each target attribute value in the header table is found in the attribute value tree, and the M candidate attribute value combinations are mined based on the attribute value subtree corresponding to each target attribute value.

10 [0094] For example, the degree of support for each target attribute value in a, b, c, d, e, f, g, i, j, l, m, n, o, and p is determined, and a target attribute value whose degree of support is less than a threshold #4 is deleted. For example, degrees of support for the target attribute values i, j, l, m, n, o, and p are all less than the threshold #4. Therefore, the i, j, l, m, n, o, and p do not occur in the header table, and the header table is established in descending order of values of degrees of support for the remaining target attribute values a, b, c, d, e, f, and g. The established header table is shown in FIG. 8. As shown in FIG. 8, a target attribute value whose degree of support is less than the threshold #4 is deleted from transaction information of each abnormal transaction, and for each abnormal transaction, target attribute values included in each abnormal transaction are sorted based on values of degrees of support, to form transaction information of sorted abnormal transactions.

25 [0095] Subsequently, a process in which the attribute value tree is established based on the header table and the transaction information of the sorted abnormal transactions and the N attribute value combinations are sequentially mined upward from the bottom of the header table based on the attribute value tree is similar to a process of mining the M candidate attribute value combinations in Manner 1. For details, references can be made to the foregoing descriptions. Details are not described again.

[0096] In another manner of determining the N attribute value combinations, after determining degrees of support for all target attribute values, the mining device 500 deletes a target attribute value whose degree of support is less than a threshold #4. In other words, when the candidate attribute value combination is determined subsequently, a deleted target attribute value is not considered. For target attribute values that remain after deletion, a degree of support for a combination of any two target attribute values, a degree of support for a combination of any three target attribute values, and a degree of support for a combination of any four target attribute values are sequentially determined, until a degree of support for a combination of a maximum quantity of target attribute values is determined. The maximum quantity is a quantity of target attribute values that remain after deletion.

[0097] For example, Degree of support for a combination of target attribute values=Occurrence frequency of the combination of target attribute values in abnormal transactions; or Degree of support for a combination of target attribute values=Occurrence frequency of the combination of target attribute values in abnormal transactions/Quantity of abnormal transactions; or Degree of support for a combination of target attribute values=Occurrence frequency of the combination of target attribute values in abnormal transactions/(Quantity of abnormal transactions+Quantity of normal transactions). If a combination of target attribute values occurs in the normal transaction in addition to the abnormal transaction, Degree of support for the combination of target attribute values=Occurrence frequency of the combination of target attribute values in abnormal transactions–Occurrence frequency of the combination of target attribute values in normal transactions; or Degree of support for the combination of target attribute values=(Occurrence frequency of the combination of target attribute values in abnormal transactions–Occurrence frequency of the combination of target attribute values in normal transactions)/Quantity of abnormal transactions; or Degree of support for the combination of target attribute values=(Occurrence frequency of the combination of target attribute values in abnormal transactions–Occurrence frequency of the combination of target attribute values in normal transactions)/(Quantity of abnormal transactions+Quantity of normal transactions).

[0098] After degrees of support for different combinations of target attribute values are determined, for any combination of target attribute values, whether the combination of target attribute values is one of the N attribute value combinations is determined based on a degree of

support for the combination of target attribute values and a threshold #6. For example, if the degree of support for the combination of target attribute values is less than or equal to the threshold #6, the combination of target attribute values is not one of the N attribute value combinations; or if the degree of support for a combination of target attribute values is greater than the threshold #6, the combination of target attribute values is one of the N attribute value combinations. The threshold #6 can be the same as or different from the threshold #3.

[0099] In conclusion, in Manner 2, the occurrence frequency of the target attribute value in the normal transactions is considered in a process of determining the degree of support for the target attribute value. Therefore, a probability that the N attribute value combinations determined in this manner occur in the normal transaction is relatively low. Subsequently, when the transaction executed by the target group is identified based on the N attribute value combinations, identification accuracy can also be improved to a certain extent.

[0100] S130: Generate a group feature corresponding to the target group based on the N attribute value combinations.

[0101] Based on the two cases in S110, there are two cases in which the group feature corresponding to the target group is generated.

[0102] Case 1

[0103] Case 1 is specific to Case 1 in S110. To be specific, the plurality of abnormal transactions obtained by the mining device 500 in S110 are executed by the target group. In this case, the group feature corresponding to the target group can be directly generated based on the N attribute value combinations. In other words, the group feature of the target group includes the N attribute value combinations.

[0104] Case 2

[0105] Case 2 is specific to Case 2 in S110. To be specific, the plurality of abnormal transactions obtained by the mining device 500 in S110 are not all executed by the target group. In this case, a transaction executed by the target group in the plurality of abnormal transactions needs to be selected from the plurality of abnormal transactions.

[0106] In some embodiments, the transaction information of the abnormal transaction does not include an abnormality label. Therefore, an abnormality label of an abnormal transaction that includes any of the N attribute value combinations in the plurality of abnormal transactions needs to be obtained. If the transaction information of the abnormal transaction includes an abnormality

label, the abnormality label can be directly used, and is not obtained repeatedly.

[0107] For example, for each attribute value combination in the N attribute value combinations, the mining device 500 searches for an abnormal transaction that includes the attribute value combination, and determines, based on an abnormality label corresponding to the abnormal transaction, whether to exclude the attribute value combination from the N attribute value combinations. After some attribute value combinations are excluded, the mining device 500 can generate the group feature corresponding to the target group based on a target attribute value combination that remains after exclusion from the N attribute value combinations.

[0108] Specifically, for an attribute value combination 1 in the N attribute value combinations, at least one target abnormal event is determined from the abnormal event. Event information of the target abnormal event includes the attribute value combination 1. Further, an abnormality probability corresponding to the attribute value combination 1 is determined based on an abnormality label corresponding to the at least one target abnormal event. For example, n1 target abnormal events include the attribute value combination 1, a quantity n2 of target abnormal events executed by the target group in the n1 target abnormal events can be determined based on abnormality labels corresponding to the n1 target abnormal events, n1 is greater than n2, n2 is an integer greater than or equal to 1, and Abnormality probability corresponding to the attribute value combination 1= $n2/n1$. If the abnormality probability corresponding to the attribute value combination 1 is less than a preset threshold #7, the attribute value combination 1 is excluded from the N attribute value combinations.

[0109] By analogy, the mining device 500 determines an abnormality probability corresponding to each attribute value combination in the N attribute value combinations, and excludes an attribute value combination whose abnormality probability is less than the threshold #7 from the N attribute value combinations. A remaining attribute value combination is a target attribute value combination. The mining device 500 can directly generate the group feature corresponding to the target group based on the target attribute value combination. In other words, the group feature of the target group includes the target attribute value combination. It can be understood that the target group can be one or more groups.

[0110] In conclusion, in the group feature mining method provided in this specification, an attribute value combination can be determined based on obtained event information of a normal event and obtained event information of an abnormal event, and subsequently, a group feature that

represents a target group is generated based on the attribute value combination. The attribute value combination can highly summarize the abnormal event. In other words, the attribute value combination frequently occurs in the event information of the abnormal event, and does not occur or rarely occurs in the event information of the normal event. It can be learned that in this solution, the event information of the normal event is considered in a process of generating the group feature of the target group, to avoid as far as possible a case in which the determined group feature once occurs in the event information of the normal event, so as to accurately prevent and control the target group, without affecting a normal group. In addition, in this solution, when the group feature of the target group is mined, a plurality of normal events and event information of the plurality of abnormal events can be simultaneously processed, and event information of a plurality of abnormal events can be generated through execution of a plurality of target groups. Therefore, in this solution, group feature mining efficiency can be improved, and prevention and control of the target group are performed in a more timely manner. In addition, a group feature that is of a target group and that is mined on one transaction platform can be uploaded to another transaction platform for use, and can be migrated.

[0111] The foregoing describes a process in which the mining device 500 mines a group feature. With reference to FIG. 9, the following describes in detail a solution in which a server 300 or a terminal device 200 or 700 performs event identification based on a group feature.

[0112] FIG. 9 is a flowchart of an event identification method P200 according to an embodiment of this specification. As described above, a server 300 or a terminal device 200 or 700 can perform the event identification method P200 in this specification. Specifically, a processor 820 in the server 300 or the terminal device 200 or 700 can read an instruction set stored in a local storage medium of the server 300 or the terminal device 200 or 700, and then perform the event identification method P200 in this specification based on a stipulation of the instruction set. The following provides descriptions by using an example in which the server 300 performs the method P200. The method P200 can be applied to event identification in the transaction field, event identification in the network security field, event identification in the financial field, etc. To describe the method P200 clearly and comprehensively, identifying a transaction event of a shopping platform is used as an example for description below, and application of the method P200 to another field can be obtained by analogy.

[0113] As shown in FIG. 9, the event identification method P200 can include the following

steps.

[0114] S210: Obtain event information of a to-be-identified target event in response to receiving an event identification request, where the event information includes a plurality of attribute values for describing the target event.

5 **[0115]** The target event can be an event that has occurred or is to occur in a network. For example, a transaction event that is being executed by a user or a group on the shopping platform is being performed and is not completed. For example, if a user or a group places an order on the shopping platform without payment, a client of the shopping platform can trigger sending of the event identification request. After receiving the event identification request, the server 300 obtains
10 transaction information of the order. For an attribute value that is included in the transaction information and that is used to describe the order, references can be made to an example of the attribute value in the method P100. Details are not described again.

[0116] S220: Obtain a group feature corresponding to a target group, where the group feature is mined in the method P100, and is used to describe an attribute feature satisfied by an event
15 executed by the target group.

[0117] For example, after generating the group feature of the target group in the method P100, the mining device 500 sends the group feature to the server 300, and the server 300 can directly obtain the group feature from the server 300.

[0118] S230: Match the event information of the target event against the group feature, to
20 determine whether the target event is an event executed by the target group.

[0119] In some embodiments, if the event information of the target event includes the group feature, it is determined that the target event is an event executed by the target group. For example, the group feature includes a plurality of attribute value combinations. If the event information of the target event includes one attribute value combination, it is considered that the target event is
25 an event executed by the target group.

[0120] After determining that the target event is an event executed by the target group, the server 300 can send a verification requirement to the client of the shopping platform, to request the target group to perform verification on the client of the shopping platform by swiping a face or inputting information. If verification fails, for example, information entered by the target group
30 is inconsistent with information originally registered with the account on the shopping platform, execution of the transaction is prevented. Alternatively, after determining that the target event is

an event executed by the target group, the server 300 directly prevents or intercepts payment of the order, so that payment of the order fails, and a transaction cannot be successfully completed.

[0121] In conclusion, in the event identification method provided in this specification, a target event can be identified based on the group feature generated in the mining method. After it is identified that the target event is an event executed by the target group, execution of the target event can be impeded by using a means, to prevent and control the target group. According to the event identification solution, prevention and control timeliness and accuracy can be improved.

[0122] It should be noted that in a process of describing the method P100 and the method P200, an application scenario of the shopping platform is used as an example. Therefore, a term used in the shopping scenario is used in specific descriptions, for example, an example of the transaction or an attribute value. If the method P100 and the method P200 are applied to another field mentioned above, the used terms can be adjusted adaptively. The terms are not enumerated one by one herein.

[0123] Another aspect of this specification provides a non-transitory storage medium, storing at least one group of executable instructions used to mine the group feature or at least one group of executable instructions used to perform event identification. When the executable instructions are executed by a processor, the executable instructions instruct the processor to implement steps of the group feature mining method P100 or the event identification method P200 described in this specification. In some possible implementations, aspects of this specification can further be implemented in a form of a program product, including program code. When the program product runs on the mining device 500 or on the server 300 or the terminal device 200 or 700, the program code is used to enable these devices to perform the steps of the group feature mining method P100 or the event identification method P200 described in this specification. The program product used to implement the above-mentioned method can be a portable compact disk read-only memory (CD-ROM), including program code, and can run on the mining device 500 or on the server 300 or the terminal device 200 or 700. However, the program product in this specification is not limited thereto. In this specification, a readable storage medium can be any tangible medium that includes or stores a program, and the program can be used by or in combination with an instruction execution system. The program product can be any combination of one or more readable media. The readable medium can be a readable signal medium or a readable storage medium. The readable storage medium can be, for example, an electrical system, apparatus, or device, a magnetic system,

apparatus, or device, an optical system, apparatus, or device, an electromagnetic system, apparatus, or device, an infrared system, apparatus, or device, or a semiconductor system, apparatus, or device, or any combination thereof. More specific examples of the readable storage medium include: an electrical connection including one or more conducting wires, a portable disk, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM or flash memory), an optical fiber, a portable compact disc read-only memory (CD-ROM), an optical storage component, a magnetic storage component, or any proper combination thereof. The computer-readable storage medium can include a data signal propagated in a baseband or as part of a carrier, and carries readable program code. The propagated data signal can be in a plurality of forms, and includes but is not limited to an electromagnetic signal, an optical signal, or any proper combination thereof. The readable storage medium can further be any readable medium other than the readable storage medium, and the readable medium can be used to send, propagate, or transmit a program used by or in combination with an instruction execution system, apparatus, or component. The program code included in the readable storage medium can be transmitted in any proper medium, including but not limited to a wireless medium, a wired medium, optical cable, an RF medium, or any proper combination thereof. Program code for performing an operation of this specification can be written in any combination of one or more program design languages. The program design language includes an object-oriented program design language, for example, Java or C++, and further includes a conventional procedural program design language, for example, "C" language, or a similar program design language.

[0124] Specific embodiments of this specification are described above. Other embodiments fall within the scope of the appended claims. In some cases, the actions or steps described in the claims can be performed in an order different from that in the embodiments, and the desired results can still be achieved. In addition, the process depicted in the accompanying drawings does not necessarily need a particular order or consecutive order to achieve the desired results. In some implementations, multitasking and parallel processing are possible or may be advantageous.

[0125] In conclusion, after reading the detailed disclosure, a person skilled in the art can understand that the detailed disclosure can be presented only in an example manner, and can impose no limitation. Although it is not explicitly stated here, a person skilled in the art can understand that the requirements of this specification include various proper changes, improvements, and modifications to the embodiments. These changes, improvements, and

modifications are intended to be presented in this specification, and fall within the spirit and scope of the example embodiments of this specification.

CLAIMS

What is claimed is:

1. A group feature mining method, comprising:

obtaining event information of a plurality of normal events and event information of a
5 plurality of abnormal events, wherein event information of each event comprises a plurality of
attribute values for describing the event, and an operator of the plurality of abnormal events
comprises a target group;

determining N attribute value combinations based on the event information of the plurality of
normal events and the event information of the plurality of abnormal events, wherein N is an
10 integer greater than or equal to 1, each attribute value combination comprises at least one attribute
value, an occurrence frequency of each attribute value combination in the plurality of abnormal
events is greater than a first preset threshold, and an occurrence frequency of each attribute value
combination in the plurality of normal events is 0; and

generating a group feature corresponding to the target group based on the N attribute value
15 combinations.

2. The method according to claim 1, wherein the determining N attribute value combinations
based on the event information of the plurality of normal events and the event information of the
plurality of abnormal events comprises:

determining M candidate attribute value combinations based on the event information of the
20 plurality of abnormal events, wherein an occurrence frequency of the candidate attribute value
combinations in the plurality of abnormal events is greater than the first preset threshold, and M
is an integer greater than or equal to N; and

determining the N attribute value combinations from the M candidate attribute value
combinations based on the event information of the plurality of normal events.

25 3. The method according to claim 2, wherein the determining M candidate attribute value
combinations based on the event information of the plurality of abnormal events comprises:

performing de-duplication on attribute values in the event information of the plurality of
abnormal events, to obtain a plurality of target attribute values;

for each target attribute value in the plurality of target attribute values, determining a degree
30 of support for the target attribute value based on an occurrence frequency of the target attribute

value in the plurality of abnormal events; and

determining the M candidate attribute value combinations based on the event information of the plurality of abnormal events and degrees of support for the plurality of target attribute values.

4. The method according to claim 2, wherein the determining the N attribute value combinations from the M candidate attribute value combinations based on the event information of the plurality of normal events comprises:

obtaining the N attribute value combinations by excluding a first attribute value combination from the M candidate attribute value combinations based on the event information of the plurality of normal events, wherein

the first attribute value combination has the same attribute value as event information of a normal event in the plurality of normal events, and a quantity of same attribute values is greater than or equal to a second preset threshold.

5. The method according to claim 1, wherein the determining N attribute value combinations based on the event information of the plurality of normal events and the event information of the plurality of abnormal events comprises:

performing de-duplication on attribute values in the event information of the plurality of abnormal events, to obtain a plurality of target attribute values;

for each target attribute value in the plurality of target attribute values, determining a degree of support for the target attribute value based on the event information of the plurality of normal events and the event information of the plurality of abnormal events; and

determining the N attribute value combinations based on the event information of the plurality of abnormal events and degrees of support for the plurality of target attribute values.

6. The method of claim 5, wherein the determining a degree of support for the target attribute value based on the event information of the plurality of normal events and the event information of the plurality of abnormal events comprises:

obtaining a first occurrence frequency of the target attribute value in the plurality of abnormal events, and obtaining a second occurrence frequency of the target attribute value in the plurality of normal events; and

determining the degree of support for the target attribute value based on the first occurrence frequency and the second occurrence frequency.

7. The method according to claim 6, wherein the determining the degree of support for the

target attribute value based on the first occurrence frequency and the second occurrence frequency comprises:

determining a difference between the first occurrence frequency and the second occurrence frequency; and

5 determining the degree of support for the target attribute value based on the difference.

8. The method according to claim 1, wherein the generating a group feature corresponding to the target group based on the N attribute value combinations comprises:

obtaining an abnormality label corresponding to an abnormal event that comprises the N attribute value combinations in the plurality of abnormal events, wherein the abnormality label
10 represents a degree of abnormality of the abnormal event;

obtaining at least one target attribute value combination by excluding some attribute value combinations from the N attribute value combinations based on the abnormality label corresponding to the abnormal event; and

generating the group feature corresponding to the target group based on the at least one target
15 attribute value combination.

9. The method according to claim 8, wherein the obtaining at least one target attribute value combination by excluding some attribute value combinations from the N attribute value combinations based on the abnormality label corresponding to the abnormal event comprises:

for each of the N attribute value combinations:

20 determining at least one target abnormal event from the abnormal event, wherein event information of the target abnormal event comprises the attribute value combination;

determining an abnormality probability corresponding to the attribute value combination based on an abnormality label corresponding to the at least one target abnormal event; and

excluding the attribute value combination from the N attribute value combinations if the
25 abnormality probability is less than a third preset threshold.

10. An event identification method, comprising:

obtaining event information of a to-be-identified target event in response to receiving an event identification request, wherein the event information comprises a plurality of attribute values for describing the target event;

30 obtaining a group feature corresponding to a target group, wherein the group feature is mined in the method according to any one of claims 1 to 9, and is used to describe an attribute feature

satisfied by an event executed by the target group; and

matching the event information of the target event against the group feature, to determine whether the target event is an event executed by the target group.

11. The method according to claim 10, wherein the matching the event information of the target event against the group feature, to determine whether the target event is an event executed by the target group comprises:

if the event information of the target event comprises the group feature, determining that the target event is an event executed by the target group.

12. A group feature mining device, comprising:

at least one storage medium, storing at least one instruction set, to perform group feature mining; and

at least one processor, communicatively connected to the at least one storage medium, wherein when running, the at least one processor reads the at least one instruction set, and performs the following operations based on an indication of the at least one instruction set:

obtaining event information of a plurality of normal events and event information of a plurality of abnormal events, wherein event information of each event comprises a plurality of attribute values for describing the event, and an operator of the plurality of abnormal events comprises a target group;

determining N attribute value combinations based on the event information of the plurality of normal events and the event information of the plurality of abnormal events, wherein N is an integer greater than or equal to 1, each attribute value combination comprises at least one attribute value, an occurrence frequency of each attribute value combination in the plurality of abnormal events is greater than a first preset threshold, and an occurrence frequency of each attribute value combination in the plurality of normal events is 0; and

generating a group feature corresponding to the target group based on the N attribute value combinations.

13. The device according to claim 12, wherein the at least one processor performs the following operations, to determine the N attribute value combinations:

determining M candidate attribute value combinations based on the event information of the plurality of abnormal events, wherein an occurrence frequency of the candidate attribute value combinations in the plurality of abnormal events is greater than the first preset threshold, and M

is an integer greater than or equal to N; and

determining the N attribute value combinations from the M candidate attribute value combinations based on the event information of the plurality of normal events.

14. The device according to claim 13, wherein the at least one processor performs the following operations, to determine the M candidate attribute value combinations:

performing de-duplication on attribute values in the event information of the plurality of abnormal events, to obtain a plurality of target attribute values;

for each target attribute value in the plurality of target attribute values, determining a degree of support for the target attribute value based on an occurrence frequency of the target attribute value in the plurality of abnormal events; and

determining the M candidate attribute value combinations based on the event information of the plurality of abnormal events and degrees of support for the plurality of target attribute values.

15. The device according to claim 13, wherein the at least one processor performs the following operations, to determine the N attribute value combinations:

obtaining the N attribute value combinations by excluding a first attribute value combination from the M candidate attribute value combinations based on the event information of the plurality of normal events, wherein

the first attribute value combination has the same attribute value as event information of a normal event in the plurality of normal events, and a quantity of same attribute values is greater than or equal to a second preset threshold.

16. The device according to claim 12, wherein the at least one processor performs the following operations, to determine the N attribute value combinations:

performing de-duplication on attribute values in the event information of the plurality of abnormal events, to obtain a plurality of target attribute values;

for each target attribute value in the plurality of target attribute values, determining a degree of support for the target attribute value based on the event information of the plurality of normal events and the event information of the plurality of abnormal events; and

determining the N attribute value combinations based on the event information of the plurality of abnormal events and degrees of support for the plurality of target attribute values.

17. The device according to claim 16, wherein the at least one processor performs the following operations, to determine the degree of support for the target attribute value:

obtaining a first occurrence frequency of the target attribute value in the plurality of abnormal events, and obtaining a second occurrence frequency of the target attribute value in the plurality of normal events; and

determining the degree of support for the target attribute value based on the first occurrence
5 frequency and the second occurrence frequency.

18. The device according to claim 17, wherein the at least one processor performs the following operations, to determine the degree of support for the target attribute value:

determining a difference between the first occurrence frequency and the second occurrence
frequency; and

10 determining the degree of support for the target attribute value based on the difference.

19. An event identification device, comprising:

at least one storage medium, storing at least one instruction set, to perform event identification;
and

at least one processor, communicatively connected to the at least one storage medium,
15 wherein when running, the at least one processor reads the at least one instruction set, and performs the following operations based on an indication of the at least one instruction set:

obtaining event information of a to-be-identified target event in response to receiving an event
identification request, wherein the event information comprises a plurality of attribute values for
describing the target event;

20 obtaining a group feature corresponding to a target group, wherein the group feature is mined in the method according to any one of claims 1 to 9, and is used to describe an attribute feature satisfied by an event executed by the target group; and

matching the event information of the target event against the group feature, to determine
whether the target event is an event executed by the target group.

25 20. A computer-readable non-transitory storage medium, wherein the computer-readable non-transitory storage medium stores at least one instruction set, and when the at least one instruction set is executed by at least one processor, the method according to any one of claims 1 to 9 is implemented, or the method according to claim 10 or 11 is implemented.