RISK CONTROL MODEL TRAINING METHOD AND SYSTEM, AND RISK CONTROL METHOD AND SYSTEM

TECHNICAL FIELD

[0001] This specification relates to the field of artificial intelligence technologies, and in particular, to a risk control model training method and system, and a risk control method and system.

BACKGROUND

[0002] Some service systems can provide payment services, for example, provide payment services between users and merchants. After accumulating a large quantity of users and merchants, the service system further considers developing a credit service, for example, issuing credit products to the users or merchants.

[0003] The credit service usually has a high risk. For example, some credit users cannot repay credit products of the credit users in a timely manner after applying for the credit products, causing a fund loss to the service system. Therefore, the service system needs to perform risk prevention and control on a risk in a credit scenario. However, at an initial stage of developing the credit service, the service system accumulates a small amount of credit data, which is insufficient to train a risk control model. Therefore, the service system can only manually make some risk control rules based on expert experience, and perform risk prevention and control based on these risk control rules. However, a prevention and control effect of the risk control rule is not good.

[0004] Content of the background part is merely information learned of by the inventor, and neither means that the information has entered a public domain before an application date of this disclosure, nor means that the information can become a conventional technology of this disclosure.

SUMMARY

[0005] This specification provides a risk control model training method and system, and a risk control method and system. When a quantity of samples in a credit scenario is relatively small, a

risk control model applicable to the credit scenario can be obtained through training based on historical transaction data in a payment scenario. Therefore, a risk prevention and control effect can be improved by using the risk control model at an initial stage of developing a credit service.

[0006] According to a first aspect, this specification provides a risk control model training method, including: obtaining historical transaction data in a payment scenario; generating a plurality of training samples based on the historical transaction data, where one training sample includes transaction behavior data of a sample transaction participant in the payment scenario; and obtaining a target risk control model through training based on the plurality of training samples, where the target risk control model is used to identify, in a credit scenario, whether a target transaction participant has a transaction risk.

[0007] According to a second aspect, this specification further provides a risk control method, including: obtaining transaction behavior data of a credit user in a payment scenario in response to a credit request initiated by the credit user; performing risk prediction on the credit user based on the transaction behavior data by using a trained target risk control model, to obtain a risk prediction result, where the target risk control model is obtained through training in the risk control model training method provided in the first aspect; and processing the credit request based on the risk prediction result.

[0008] According to a third aspect, this specification further provides a risk control model training system, including: at least one storage medium, storing at least one instruction set, to train a risk control model; and at least one processor, communicatively connected to the at least one storage medium, where when the training system runs, 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 historical transaction data in a payment scenario; generating a plurality of training samples based on the historical transaction data, where one training sample includes transaction behavior data of a sample transaction participant in the payment scenario; and obtaining a target risk control model through training based on the plurality of training samples, where the target risk control model is used to identify, in a credit scenario, whether a target transaction participant has a transaction risk.

[0009] According to a fourth aspect, this specification further provides a risk control system, including: at least one storage medium, storing at least one instruction set, to perform risk control; and at least one processor, communicatively connected to the at least one storage medium, where

when the risk control system runs, 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 transaction behavior data of a credit user in a payment scenario in response to a credit request initiated by the credit user; performing risk prediction on the credit user based on the transaction behavior data by using a trained target risk control model, to obtain a risk prediction result, where the target risk control model is obtained through training in the risk control model training method provided in the first aspect; and processing the credit request based on the risk prediction result.

[0010] According to a fifth aspect, this specification further provides a computer-readable nonvolatile storage medium. The computer-readable nonvolatile storage medium stores at least one instruction set, and when the at least one instruction set is executed by at least one processor, the risk control model training method provided in the first aspect is implemented, or the risk control method provided in the second aspect is implemented.

[0011] Other functions of the risk control model training method and system, and the risk control method and system provided in this specification are partially listed in the following descriptions. The creative aspects of the risk control model training method and system, and the risk control method and system provided in this specification can be fully explained by practice or by using the method, apparatus, and combination described in the following detailed example.

BRIEF DESCRIPTION OF DRAWINGS

[0012] To describe the technical solutions in the embodiments of this specification more clearly, the following briefly describes the accompanying drawings needed for describing the embodiments. Clearly, the accompanying drawings in the following descriptions show merely some embodiments of this specification, and a person of ordinary skill in the art can still derive other drawings from these accompanying drawings without creative efforts.

[0013] FIG. 1 is a schematic diagram of an application scenario of performing risk control in a credit scenario according to an embodiment of this specification;

[0014] FIG. 2 is a diagram of a hardware structure of a computing device according to an embodiment of this specification;

[0015] FIG. 3 is a flowchart of a risk control model training method according to an

embodiment of this specification;

[0016] FIG. 4 is a schematic diagram of generating a training sample according to an embodiment of this specification;

[0017] FIG. 5 is a schematic diagram of a first subgraph according to an embodiment of this specification;

[0018] FIG. 6 is a schematic diagram of another first subgraph according to an embodiment of this specification;

[0019] FIG. 7 is a schematic diagram of a second subgraph according to an embodiment of this specification;

[0020] FIG. 8 is a schematic diagram of a process of training a target risk control model according to an embodiment of this specification; and

[0021] FIG. 9 is a flowchart of a risk control method according to an embodiment of this specification.

DESCRIPTION OF EMBODIMENTS

[0022] 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 clear, and the general principles defined herein can be applied to other 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.

[0023] 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 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.

[0024] 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.

[0025] 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.

[0026] For ease of description, the terms that appear below in this specification are first explained.

[0027] Term 1: Centered graph: The centered graph is a graph in which a node serves as a central node and other nodes in the graph are directly or indirectly connected to the central node. The centered graph can reflect a clustering feature of the central node.

[0028] Term 2: Path graph: The path graph is a graph used in this specification to represent all possible paths from one node (start) and another node (end). The path graph can represent an interaction feature between the start and the end.

[0029] Term 3: Hop: "One hop" is usually a step of directly arriving at a neighboring node of a node from the node. In this specification, if an edge exists between two nodes, a distance between the two nodes is one hop.

[0030] Term 4: Edge: The edge is a line segment connecting two nodes, and a plurality of edges can exist between the two nodes. Each edge represents a relationship or connection the between two nodes. In this specification, the edge is used to represent a transaction relationship between the two nodes.

[0031] Term 5: Large-scale sampling: Large-scale sampling is selecting a relatively small subset from a dataset. The subset is used to represent a key feature of the dataset. In a risk control scenario, massive data exists in the dataset. Due to a limitation of a performance bottleneck and computing resources, it is impossible to perform real-time analysis on all data. Therefore, a data

amount needs to be reduced through large-scale sampling, while as much information as possible useful for risk control is retained. A sampling process can be performed based on a plurality of standards such as random sampling or sampling after sorting by a transaction amount.

[0032] Term 6: Ego-model: The ego-model is a subgraph model centered on a single node (central node, Ego). The ego-model is mainly configured to extract a clustering feature between the central node and a neighboring node (one-hop neighbor) directly connected to the central node.

[0033] Term 7: Path-model: The path-model focus on a path between two nodes in a graph, and a relationship between all nodes and edges in the path. The path-model is mainly used to reveal causality between variables (nodes), including direct and indirect impact. This model is applicable to analysis of fund flowing or a transaction sequence, because the path-model can capture a fund interaction pattern between a user role and a merchant role.

[0034] Term 8: Continuous-time and context-aware graph attention network (C2GAT): The C2GAT is a graph neural network model, and can process graph data of continuous time and capture context information of nodes in a graph. The C2GAT has characteristics of continuous time sensing and context sensing. With these two characteristics, the C2GAT can better process dynamic graph data, and can capture the context information of the node, to improve prediction performance of the model.

[0035] The following describes an application scenario of this specification.

[0036] The technical solutions provided in this specification can be applied to a service system. The service system can provide a payment service and a credit service. For example, the service system can first develop the payment service, for example, provide the payment service between a user and a merchant. After accumulating a large quantity of users and merchants, the service system can develop the credit service, for example, provide a credit product to the user or the merchant.

[0037] This specification provides a risk control model training method and a risk control method. The risk control model training method and the risk control method can be applied to a credit scenario. In this scenario, a service system can first obtain a target risk control model through training based on historical transaction data in a payment scenario. Then, when a credit user initiates a credit request, the service system can obtain the credit user and transaction behavior data in the payment scenario, and perform risk prediction on the credit user based on the transaction behavior data by using the target risk control model obtained through training, to obtain a risk

prediction result. Further, the service system can process the credit request based on the risk prediction result, for example, grant the request, reject the request, increase a credit limit, or decrease the credit limit. The above-mentioned training process of the target risk control model can be performed in the risk control model training method provided in this specification, and the above-mentioned process of performing risk prediction on the credit user based on the transaction behavior data, to obtain the risk prediction result can be performed in the risk control method provided in this specification.

[0038] It should be noted that the credit scenario is merely one of a plurality of use scenarios provided in this specification. The risk control model training method and the risk control method provided in this specification not only can be applied to the credit scenario, but also can be applied to all scenarios in which risk control needs to be performed on a user, for example, a credit analysis scenario or a lease scenario. A person skilled in the art should understand that, when the risk control model training method and the risk control method provided in this specification are applied to another use scenario, an implementation and a technical effect thereof are similar.

[0039] FIG. 1 is a schematic diagram of an application scenario of performing risk control in a credit scenario according to an embodiment of this specification.

[0040] As shown in FIG. 1, an application scenario 100 can include a training device 11 and an execution device 12. The application scenario 100 involves a model training phase and a model application phase. In the model training phase, the training device 11 can obtain a target risk control model through training based on historical transaction data in a payment scenario. The trained target risk control model can be deployed in the execution device 12. The execution device 12 can perform risk control in the credit scenario. For example, when a credit user initiates a credit request, the execution device 12 can perform risk prediction on the credit user based on transaction behavior data of the credit user in the payment scenario by using the target risk control model, to obtain a risk prediction result of the credit user.

[0041] The training device 11 can be an electronic device that has a specific computing capability. The training device 11 can perform a risk control model training method provided in this specification. The training device 11 can store data or instructions for performing the risk control model training method described in this specification, and can execute or be configured to execute the data or the instructions. The training device 11 can include a hardware device with a data information processing function and a necessary program required for driving the hardware

device to work.

[0042] The target risk control model is a machine learning model that is obtained through training by the training device 11 and that can be used to identify, in the credit scenario, whether a target transaction participant (for example, the credit user) has a transaction risk. This specification sets no limitation on a specific structure of the target risk control model.

[0043] In an example, the training device 11 can be a server, for example, a local server, a cloud server, a rack server, or a blade server; or can be a desktop computer, a notebook computer, etc.

[0044] The execution device 12 can be an electronic device that has a specific computing capability. The execution device 12 can perform the risk control method provided in this specification. The execution device 12 can store data or instructions for performing the risk control method described in this specification, and can execute or be configured to execute the data or the instructions. The execution device 12 can include a hardware device with a data information processing function and a necessary program required for driving the hardware device to work.

[0045] A well-trained target risk control model is deployed on the execution device 12. After receiving the credit request initiated by the credit user, the execution device 12 can input the transaction behavior data of the credit user in the transaction scenario into the target risk control model, to obtain the risk prediction result of the credit user.

[0046] In a system architecture of a specific scenario, the execution device 12 can be a specific device in the system architecture. For example, the credit scenario is used as an example. A system architecture of the scenario usually includes a terminal device (corresponding to a transaction client) and a server (corresponding to a transaction server).

[0047] In some embodiments, the execution device 12 can be a server. In this case, the target risk control model is deployed in the server. The credit user (for example, an individual user or a merchant) initiates the credit request by using a terminal device, and uploads the credit request of the credit user to the server by using the terminal device. The server obtains the transaction behavior data of the credit user based on the credit request, and obtains the risk prediction result of the credit user based on the transaction behavior data by using the target risk control model.

[0048] In some embodiments, the execution device 12 can alternative be a terminal device. In this case, the target risk control model is deployed in the terminal device. The credit user initiates the credit request by using the terminal device, and the terminal device obtains the transaction

behavior data of the credit user based on the credit request, and further obtains the risk prediction result of the credit user based on the transaction behavior data by using the target risk control model.

[0049] In some embodiments, the terminal device can include a mobile device, a tablet computer, 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.

[0050] In some embodiments, one or more application programs (app) can be installed on the terminal device. 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, an instant messaging tool, a mailbox client, social platform software, etc.

[0051] In some embodiments, a target app can be installed on the terminal device. The target app can provide a human-machine interaction interface, and generates a credit request based on data that is entered by a user and that is received in the human-machine interaction interface. The target app can further send the credit request to the server, and receive a risk prediction result fed back by the server and a processing result of processing the credit request.

[0052] In some embodiments, the target app can further obtain the transaction behavior data of the credit user based on the credit request. The target app can obtain the risk prediction result of the credit user based on the transaction behavior data by using the target risk control model, and then send the credit request and the risk prediction result to a server that provides a credit service, and receive a processing result that is of processing the credit request and that is fed back by the

server.

[0053] It should be noted that the training device 11 and the execution device 12 can correspond to the same entity device, or can correspond to different entity devices. This specification sets no limitation thereto.

[0054] FIG. 2 is a diagram of a hardware structure of a computing device according to an embodiment of this specification.

[0055] In some embodiments, the computing device 200 can serve as the training device 11 in FIG. 1 to perform a risk control model training method provided in this specification. The computing device 200 can further serve as a risk control model training system provided in this specification. In some embodiments, the computing device 200 can serve as the execution device 12 in FIG. 1 to perform a risk control method provided in this specification. The computing device 200 can further serve as a risk control system provided in this specification.

[0056] As shown in FIG. 2, the computing device 200 can include at least one storage medium 230 and at least one processor 220. In some embodiments, the computing device 200 can further include a communication port 250 and an internal communication bus 210. The computing device 200 can further include an I/O component 260.

[0057] The internal communication bus 210 can be connected to different system components. For example, the internal communication bus 210 can be connected to the storage medium 230, the processor 220, the communication port 250, and the I/O component 260.

[0058] The I/O component 260 supports input/output between the computing device 200 and other components.

[0059] The communication port 250 is used for data communication between the computing device 200 and the outside. For example, the communication port 250 can be used for data communication between the computing device 200 and a network 13. The communication port 250 can be a wired communication port, or can be a wireless communication port.

[0060] The storage medium 230 can include a data storage apparatus. The data storage 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 232, a read-only storage medium (ROM) 234, or a random access storage medium (RAM) 236. The storage medium 230 further includes at least one instruction set stored in the data storage apparatus. The instruction set can include computer program code. The computer program code can include a program, a routine,

an object, a component, a data structure, a process, a module, etc.

[0061] The at least one processor 220 can be communicatively connected to the at least one storage medium 230. When the computing device 200 runs, the at least one processor 220 reads the at least one instruction set, and performs, based on an indication of the at least one instruction set, the risk control model training method and the risk control method provided in this specification. The processor 220 can perform steps included in the risk control model training method or the risk control method. The processor 220 can be in a form of one or more processors. In some embodiments, the processor 220 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.

[0062] To merely describe a problem, only one processor 220 is shown in the computing device 200 in the accompanying drawings. However, it should be noted that, the computing device 200 in this specification can further include a plurality of processors. Therefore, the operations and/or method steps disclosed in this specification can be performed by one processor, or can be jointly performed by a plurality of processors. For example, if the processor 220 in the computing device 200 described 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 220 (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).

[0063] FIG. 3 is a flowchart of a risk control model training method according to an embodiment of this specification. As described above, a training device 11 can perform the risk control model training method in this specification.

[0064] As shown in FIG. 3, the risk control model training method can include the following steps.

[0065] S310: Obtain historical transaction data in a payment scenario.

[0066] The payment scenario is a scenario in which a service system provides a payment service. For example, when a transaction is generated between a user and a merchant, the service

system can provide the payment service for the transaction.

[0067] The historical transaction data can include transaction data corresponding to a plurality of transactions generated by the service system in a historical time period. Transaction data corresponding to one transaction can include the following parameters: a transaction time of a transaction event, a transaction amount, a transaction participant, a transaction location, logistics information, etc. A transaction participant involved in one transaction can include a transaction participant (for example, a user role) that initiates the transaction and a transaction participant (for example, a merchant role) that executes a transaction event.

[0068] The historical transaction data includes a positive sample and a negative sample. The positive sample is a sample in which a risk event occurs, and a negative sample is a sample in which no risk event occurs. The risk event includes a cash-out event, a fraud event, a false transaction, etc. The historical transaction data in the payment scenario can be stored in a local database, or can be stored in a database at a location such as the cloud or a server.

[0069] In some embodiments, to prevent a data distribution from tilting due to various influence factors, transaction data in a relatively long time period can be obtained, to smooth a possible short-term data distribution fluctuation. For example, the historical transaction data can be real transaction data generated by the service system in the past one year.

[0070] In some embodiments, due to a low probability that a risk event occurs in transaction data within one year, a proportion of positive samples to negative samples is imbalanced. In this case, a quantity of negative samples can be reduced, to increase a relative proportion of the positive samples, to ensure a training effect of the model. For example, the negative samples can be partially downsampled, until the proportion of the positive samples to the negative samples is any proportion from 1:10 to 1:100.

[0071] In some embodiments, the historical transaction data can be represented in a form of graph data. The graph data is of a data structure including a node (vertices) and an edge, and is used to represent an entity (for example, a user or a merchant) and a relationship (for example, a transaction relationship) between the entities. The graph data is used to represent the historical transaction data, and can effectively show a complex relationship in a transaction, for example, a transaction relationship between the user and the merchant or fund flowing.

[0072] In this specification, the training device can generate a corresponding transaction graph based on the historical transaction data. One node corresponds to one transaction participant (the

user role or the merchant role). One edge connects two nodes, and represents a transaction (transaction relationship) between the two nodes.

[0073] In some embodiments, in the historical transaction data, data of each transaction further includes a transaction occurrence time. The transaction graph generated based on the historical transaction data is a dynamic graph. The transaction graph can show a transaction relationship between a plurality of transaction participants at a specific moment.

[0074] S320: Generate a plurality of training samples based on the historical transaction data, where one training sample includes transaction behavior data of a sample transaction participant in the payment scenario.

[0075] In some embodiments, the transaction behavior data includes transaction relationship data, used to represent a transaction relationship between the sample transaction participant and at least one another transaction participant. The transaction relationship can be used to indicate that the sample transaction participant and another transaction participant have a transaction at a specific moment. The transaction relationship can include a direct transaction relationship and an indirect transaction relationship.

[0076] The direct transaction relationship can indicate that the sample transaction participant and the another transaction participant directly have a transaction. The indirect transaction relationship includes a plurality of direct transaction relationships. The plurality of direct transaction relationships include a transaction relationship between the sample transaction participant and one another transaction participant, and a transaction relationship between the another transaction participant and another transaction participant. The transaction relationship data can reflect a behavior clustering feature of the sample transaction participant.

[0077] In some embodiments, the transaction behavior data further includes fund path data, used to represent a fund interaction path related to the sample transaction participant. The fund interaction path can represent a fund flowing path including a transaction relationship between the sample transaction participant and another transaction participant. The fund interaction path can be used to reflect a fund interaction feature between the sample transaction participant and another transaction participant.

[0078] In some embodiments, the transaction relationship data and the fund path data can also be represented in a form of graph data. In this case, the transaction relationship data and the fund path data are subgraphs of the transaction graph corresponding to the historical transaction data in

S310. The subgraph is a part of a graph snapshot of the transaction graph at a specific moment.

[0079] For example, the training device can obtain the graph snapshot of the transaction graph at a specific moment. Then, the training device obtains the subgraphs corresponding to the transaction relationship data and the fund path data from the graph snapshot.

[0080] In some embodiments, FIG. 4 is a schematic diagram of generating a training sample according to an embodiment of this specification.

[0081] As shown in FIG. 4, the training device can generate the transaction graph based on the historical transaction data in the payment scenario. Further, the training device can generate the plurality of training samples based on the transaction graph. In each training sample, transaction relationship data corresponds to a first subgraph in the transaction graph, and fund path data corresponds to a second subgraph in the transaction graph.

[0082] In some embodiments, the training device can obtain some preset selection conditions based on a risk event that needs to be identified, and perform selection in the transaction graph based on these selection conditions, to obtain the first subgraph used to indicate the transaction relationship data and the second subgraph used to indicate the fund path data.

[0083] In some embodiments, the preset condition can include a hop count, a quantity of edges, and a threshold of the parameter in the transaction data.

[0084] In some embodiments, the first subgraph is a centered graph. When the sample transaction participant corresponds to the user role, the first subgraph can include a user centered graph centered on the user role. When the sample transaction participant corresponds to the merchant role, the first subgraph can include a merchant centered graph centered on the merchant role.

[0085] The following describes a generation manner of the first subgraph by using an example in which the sample transaction participant corresponds to the user role.

[0086] In some embodiments, the first subgraph can be generated in the following manner: The training device can first determine a plurality of target nodes in the transaction graph. The plurality of target nodes can include a first node (a central node) corresponding to the sample transaction participant and a second node connected to the first node. For the centered graph, there is one first node, a quantity of second nodes is not limited, and each second node is directly or indirectly connected to the first node. When the second node is determined, selection can be further performed based on the preset condition, to ensure that the determined second node can be used

to indicate a clustering feature of the first node. For example, the preset condition can be that a hop count between the first node and the second node is less than or equal to a preset hop count.

[0087] Then, the training device determines at least one target edge from edges between the plurality of target nodes. The target edge can be an edge directly connecting the first node and the second node, or can be an edge that is passed through when the second node is indirectly connected to the first node. For example, the training device can sample an edge between any two of the plurality of target nodes based on a sequence of transactions corresponding to different edges, to determine at least one target edge from edges between the plurality of target nodes.

[0088] In some embodiments, for the user centered graph, when the at least one target edge is determined, if a quantity of edges between two target nodes is greater than a first preset quantity, latest edges whose quantity is the first preset quantity are selected as the target edge from the edges between the two target nodes in a sequence of transactions corresponding to all edges. Alternatively, if a quantity of edges between two target nodes is less than or equal to a first preset quantity, an edge between the two target nodes is used as the target edge. When a quantity of target edges exceeds the first preset quantity, that the latest edge is selected as the target edge can clearly indicate a clustering feature of the first node in the latest time scale, and further, the latest data is used to train the risk control model, so that the trained risk control model can have a better risk control effect.

[0089] In an example, if the first subgraph is used to represent a feature of a cash-out risk, the preset condition is that the hop count between the first node and the second node is less than or equal to two. The first preset quantity can be 30. In other words, a quantity of edges between two nodes is not greater than 30.

[0090] Finally, the training device obtains the first subgraph from the graph snapshot of the transaction graph based on the plurality of target nodes (the first node and a plurality of second nodes) and the at least one target edge. In other words, the training device uses, as the first subgraph, a subgraph that is in the graph snapshot of the transaction graph and that is determined by the plurality of target nodes and the at least one target edge. FIG. 5 is a schematic diagram of a first subgraph according to an embodiment of this specification. In an example, the first subgraph is obtained from the graph snapshot of the transaction graph in the above-mentioned method, to obtain the first subgraph shown in FIG. 5.

[0091] As shown in FIG. 5, the first node is a user role 1, and the second node includes a user

role 2, a user role 3, a user role 4, a user role 5, a user role 6, a user role 7, a user role 8, a merchant role 1, a merchant role 2, a merchant role 3, a merchant role 4, and a merchant role 5. A hop count between the second nodes and the first node is not greater than 2.

[0092] In FIG. 5, one target edge exists between the user role 1 and each of the merchant role 1, the merchant role 3, and the user role 8, and three target edges exist between the user role 1 and the merchant role 2. One target edge exists between the merchant role 1 and the user role 2, and two target edges exist between the merchant role 1 and the user role 3. One target edge exists between the merchant role 2 and each of the user role 4 and the user role 5. One target edge exists between the merchant role 3 and each of the user role 6 and the user role 7. One target edge exists between the user role 8 and each of the merchant role 4 and the merchant role 5.

[0093] As shown in FIG. 5, the target edge between the user role 1 and each of the merchant role 1, the merchant role 2, the merchant role 3, and the user role 8 indicates that a direct transaction relationship exists between the user role 1 and the user role 1, and three target edges exist between the user role 1 and the merchant role 2. In other words, three direct transaction relationships exist between the user role 1 and the merchant role 2.

[0094] As shown in FIG. 5, a direct transaction relationship exists between the user role 1 and the merchant role 1, and a direct transaction relationship exists between the merchant role 1 and each of the user role 2 and the user role 3. An indirect transaction relationship exists between the user role 1 and the user role 2, and an indirect transaction relationship exists between the user role 1 and the user role 3.

[0095] As shown in FIG. 5, a direct transaction relationship exists between the user role 1 and the merchant role 2, and a direct transaction relationship exists between the merchant role 2 and each of the user role 4 and the user role 5. An indirect transaction relationship exists between the user role 1 and the user role 4, and an indirect transaction relationship exists between the user role 1 and the user role 5.

[0096] As shown in FIG. 5, a direct transaction relationship exists between the user role 1 and the merchant role 3, and a direct transaction relationship exists between the merchant role 3 and each of the user role 6 and the user role 7. An indirect transaction relationship exists between the user role 1 and the user role 6, and an indirect transaction relationship exists between the user role 1 and the user role 7.

[0097] As shown in FIG. 5, a direct transaction relationship exists between the user role 1 and

the user role 8, and a direct transaction relationship exists between the user role 8 and each of the merchant role 4 and the merchant role 5. An indirect transaction relationship exists between the user role 1 and the merchant role 4, and an indirect transaction relationship exists between the user role 1 and the merchant role 5.

[0098] The following describes a generation manner of the first subgraph by using an example in which the sample transaction participant corresponds to the merchant role.

[0099] In some embodiments, the training device can first find the first node corresponding to the sample transaction participant in the transaction graph. The first node is a merchant role. When the first node is a merchant role, the first node is usually connected to a relatively large quantity of edges. For example, the merchant role may have a transaction amount of 10000 or even a larger transaction amount in a short period of time. In this case, there may be 10000 edges. Therefore, the training device can sample the edges connected to the first node, to determine a plurality of target edges from the edges connected to the first node. For example, if a quantity of edges connected to the first node is greater than a second preset quantity, the edges need to be selected through large-scale sampling, to retain an edge that is more helpful for the clustering feature, and the quantity of edges is less than or equal to the second preset quantity.

[0100] In some embodiments, the training device can perform large-scale sampling, based on amounts of transactions corresponding to different edges, on the edges connected to the first node, to determine a plurality of target edges from the edges connected to the first node. For example, sampling can be performed after sorting by a transaction amount. For example, an edge whose amount satisfies a preset condition can be selected as the target edge from the edges connected to the first node based on an amount of a transaction corresponding to each edge.

[0101] In some embodiments, large-scale sampling can alternatively be random sampling. For example, random sampling can be randomly determining, as the target edge, edges whose quantity is the second preset quantity in the edges.

[0102] In some embodiments, large-scale sampling can alternatively be sampling performed based on an abnormal value. An abnormal value condition can be set as the preset condition when sampling is performed based on the abnormal value, and a transaction whose transaction amount deviates from a normal range is selected. Then, an edge that satisfies the abnormal value condition is used as the target edge. The abnormal value condition can be a threshold of the abnormal value, for example, can be a percentage at which a transaction amount is greater than or less than a value

or a transaction amount is greater than an average transaction amount, etc.

[0103] In some other embodiments, large-scale sampling can alternatively be performed based on a time window. For example, if a time window (for example, one hour) is set, all edges within one hour before sampling are used as the target edge by using a sampling time as a reference during sampling.

[0104] When the quantity of edges connected to the first node exceeds the second preset quantity, the edges are selected through large-scale sampling, to ensure, to the maximum extent, that an edge left after large-scale sampling is an edge that contributes to identifying the risk event. The clustering feature of the first node can be represented more clearly. These edges are used as the target edge to generate the first subgraph, and the risk control model is trained based on the first subgraph, so that the risk control model obtained through training can quickly respond to the risk event, and has a better risk control effect.

[0105] In some other embodiments, for some merchant roles with a small transaction amount, the quantity of edges connected to the first node in the entire transaction graph is relatively small. In this case, all edges connected to the first node can be directly used as the target edge without selection. For example, if the quantity of edges connected to the first node is less than or equal to the second preset quantity, all the edges connected to the first node are used as the target edge.

[0106] In an example, if the first subgraph is used to represent a feature of the cash-out risk, the second preset quantity can be 100. In other words, a total quantity of edges connected to the first node is not greater than 100.

[0107] After the first node and the plurality of target edges are determined, the training device can use, as the first subgraph, a subgraph that is in the transaction graph and that is determined by the first node, the plurality of target edges, and second nodes respectively connected to the plurality of target edges.

[0108] FIG. 6 is a schematic diagram of another first subgraph according to an embodiment of this specification. In an example, the first subgraph is obtained from the graph snapshot of the transaction graph in the above-mentioned method, to obtain the first subgraph shown in FIG. 6.

[0109] As shown in FIG. 6, the first node is a merchant role 1, and the second node includes a user role 1, a user role 2, a user role 3, a user role 4, and a user role 5, …, and a user role 100. Because the hop count is set to one, no indirect transaction exists in FIG. 6, and one target edge exists between each user role and the merchant role 1.

[0110] In an example, it is assumed that the merchant role 1 has 10000 transactions in the transaction graph. Sampling can be performed after sorting by a transaction amount. The 10000 transactions are sorted based on transaction amounts, and 100 transactions with the largest transaction amount are selected. Transaction participants corresponding to the 100 transactions are the user role 1, the user role 2, the user role 3, the user role 4, the user role 5, …, and the user role 100. An edge between the merchant role 1 and each of the user role 1, the user role 2, the user role 3, the user role 4, the user role 5, …, and the user role 100 can be used as the target edge.

[0111] In some embodiments, the second subgraph is a path graph, and displays the fund interaction path between the user role and the merchant role. The second subgraph can be used to represent a fund interaction feature between the user role and the merchant role.

[0112] In some embodiments, the second subgraph can be obtained in the following manner: The training device finds, in the transaction graph, the first node corresponding to the sample transaction participant, and determines, in the transaction graph, at least one third node connected to the first node. A hop count between the third node and the first node is less than or equal to a preset hop count, and a fund return path exists between the third node and the first node. Further, the training device can determine the second subgraph in the transaction graph based on the first node and the at least one third node. For example, the training device can use, as the second subgraph, a subgraph that is in the transaction graph and that is determined by the first node, the at least one third node, and an edge between the above-mentioned nodes.

[0113] The cash-out risk is used as an example. In the second subgraph, the first node can be a user role corresponding to the sample transaction participant. The third node can be another transaction participant that has a fund return path with the sample transaction participant. In this case, the another transaction participant can be a merchant role.

[0114] In some embodiments, there can be a plurality of fund return paths between the first node and the third node. The fund return path can include a direct return path and an indirect return path. The direct return path can correspond to one direct transaction relationship (edge), and the indirect return path can correspond to one indirect transaction relationship. When the indirect transaction relationship exists, a hop count of a path through which the indirect transaction relationship passes is less than or equal to the preset hop count. For example, when the preset hop count is 3, a maximum of two nodes between the first node and the third node form the fund return path.

[0115] FIG. 7 is a schematic diagram of a second subgraph according to an embodiment of this specification.

[0116] In an example, if the preset hop count is 3, and the second subgraph can be obtained from the graph snapshot of the transaction graph in the above-mentioned method, the second subgraph shown in FIG. 7 may be obtained.

[0117] As shown in FIG. 7, the first node is a user role 1, and the third node includes a merchant role 1. Seven fund return paths exist between the first node and the third node. There are five direct return paths, represented by five edges between the first node and the third node. There are two indirect return paths. One is a fund return path including the user role 2 and the user role 3, and is a three-hop path, and the other one is a fund return path including the user role 4, and is a two-hop path

[0118] A corresponding path graph is obtained based on a type of the risk event, to more accurately obtain the fund interaction feature between the user role and the merchant role in the risk event, so that a risk control model obtained through training based on the path graph can more accurately identify the risk event, and has a better risk control effect.

[0119] S330: Obtain a target risk control model through training based on the plurality of training samples, where the target risk control model is used to identify, in a credit scenario, whether a target transaction participant has a transaction risk.

[0120] FIG. 8 is a schematic diagram of a process of training a target risk control model according to an embodiment of this specification.

[0121] In some embodiments, as shown in FIG. 8, the target risk control model can be trained in a supervised learning manner or an unsupervised learning manner. The following uses the supervised learning manner as an example to describe training of the target risk control model.

[0122] In some embodiments, when the target risk control model is trained in the supervised learning manner, the training sample further includes label information of the sample transaction participant. The obtaining a target risk control model through training based on the plurality of training samples includes: performing a plurality of times of iterative training based on the plurality of training samples until the target risk control model is obtained, where a process of one time of iterative training includes:

inputting the transaction relationship data (first subgraph) and the fund path data (second subgraph) that correspond to the sample transaction participant into the target risk control

model, to obtain risk prediction information obtained by performing risk prediction on the sample transaction participant by the target risk control model; and updating a parameter of the target risk control model with a training objective of minimizing a difference between the risk prediction information and the label information.

[0123] In some embodiments, the target risk control model includes a first feature extraction module, a second feature extraction module, and a prediction module; and

the inputting the transaction relationship data and the fund path data that correspond to the sample transaction participant into the target risk control model, to obtain risk prediction information obtained by performing risk prediction on the sample transaction participant by the target risk control model includes: inputting the transaction relationship data into the first feature extraction module, to obtain a clustering feature between the sample transaction participant and the another transaction participant; inputting the fund path data into the second feature extraction module, to obtain an interaction feature of the fund exchange path related to the sample transaction participant; and inputting the clustering feature and the interaction feature into the prediction module, to obtain the risk prediction information.

[0124] In some embodiments, the first feature extraction module can be a model configured to extract the clustering feature based on the centered graph, and the second feature extraction module can be a model configured to extract the interaction feature based on the path graph. For example, the first feature extraction module can be an ego-model, and the second feature extraction module can be a path-model.

[0125] In some embodiments, the first feature extraction module can be configured to: extract a timing feature and a topology feature of the transaction relationship based on the transaction relationship data; and generate the clustering feature corresponding to the transaction relationship data based on the timing feature and the topology feature of the transaction relationship data. The ego-model is used as an example. The Ego-model can include a time encoding module and a C2GAT module. The time encoding module can perform, based on a preset formula, encoding processing on the first subgraph (transaction relationship data) input into the first feature extraction module, and encode the transaction data in the first subgraph in a time sequence, to extract a timing feature of the transaction relationship data. Then, the timing feature and the topology feature are input to the CAGAT module, to obtain the clustering feature of the transaction relationship data. The topology feature is used to represent network structure features such as degree centrality and proximity centrality of a user or a merchant. The topology feature can be obtained based on a graph

analysis library. For example, the graph analysis library can be a preset software package (for example, NetworkX) or an open-source graph library (for example, igraph).

[0126] In an example, the preset formula of the time encoding module can include the following formula:

$$v, t \mapsto \Phi_{\omega}^{\mathcal{M}}(v,t) \coloneqq \left[\sqrt{c_1(v)}, \sqrt{c_2(v)}, \ldots, \sqrt{c_{2j}(v)}\cos\left(\frac{j\pi t}{\omega}\right), \sqrt{c_{2j+1}(v)}\sin\left(\frac{j\pi t}{\omega}\right), \ldots\right]$$

[0127] The formula is used to define a time encoding vector, and encode transaction data corresponding to a node in the first subgraph based on a time step, to generate a corresponding time encoding vector. The formula maps the time step t to a two-dimensional vector including a cosine function and a sine function, to encode the time step t based on different frequency values $c_n(v)$.

[0128] Here, v represents a node (the first node or the second node) in the first subgraph, and t represents the time step. $\Phi_{\omega}^{\mathcal{M}}(v,t)$represents a function, an input into the function is the node v and the time step t, and an output is a time encoding vector. Here, $\mathcal{M}$ represents a dimension of time encoding, ω is an adjustable hyperparameter, and is used to control a periodicity of a time signal, and j represents a dimension index, and is used to mark each element in the feature vector. $c_n(v)$is a frequency value corresponding to the nth node (for example, can be a transaction occurrence frequency). $\sqrt{c_n(v)}$ is a square root of the frequency value corresponding to the nth node, and is used to control an amplitude of time encoding.

[0129] In an example, the preset formula of the time encoding module can further include the following formula:

$$v, t \mapsto \Phi_{d}^{\mathcal{M}}(v,t) \coloneqq \left[\Phi_{\omega_1,d}^{\mathcal{M}}(v,t) \middle\| \Phi_{\omega_2,d}^{\mathcal{M}}(v,t) \middle\| \ldots \middle\| \Phi_{\omega_k,d}^{\mathcal{M}}(v,t)\right]$$

[0130] The formula is used to define a time encoding vector of a node v, and the time encoding vector is represented in detail by using a Fourier series. Φ represents a specific operation or function. In this specification, Φ used to map an original feature state of the node to another space, including encoding of time information, and $\mathcal{M}$ represents a dimension of time encoding. Here, d is a learnable location encoding vector, and ω is also an adjustable hyperparameter, and is used to control a periodicity of a time signal. $\Phi_{\omega_k,d}^{\mathcal{M}}(v,t)$ represents a correspondence between a node state obtained through conversion by performing an operation Φ and an original node state, and is used to calculate a weight of attention or other related metrics at a subsequent network layer. Double vertical lines are used to indicate that all results obtained by

calculating $\Phi_{\omega_k,d}^{\mathcal{M}}(v,t)$ are connected to obtain the time encoding vector (timing feature). Based on the formula, the time signal can be directly encoded based on the time step t and a learnable location code d, to obtain the time encoding vector of the node v.

[0131] In some embodiments, the timing feature extracted based on the above-mentioned formula includes a time sequence corresponding to each node in the first subgraph. A time sequence of a node can be used to indicate a transaction frequency, a transaction amount increase/decrease trend, etc. of the user role or the merchant role corresponding to the node in a time window. The time sequence of each node is used as an input, and the CAGAT module learns impact of each transaction time point on a current node state.

[0132] In some embodiments, the timing feature and the topology feature of the first subgraph can be input into the trained CAGAT module in the ego-model, to obtain the clustering feature of the first subgraph. When the first subgraph is a user centered graph, a clustering feature of the user role is used to describe a behavior pattern of the user role in the service system, for example, a transaction frequency, a transaction amount, or a transaction time. When the first subgraph is a merchant centered graph, a clustering feature of the merchant role is used to describe a behavior pattern of the merchant role in the service system, for example, a transaction amount of a merchant role or a transaction relationship with the user role.

[0133] In some embodiments, the second feature extraction module is configured to: extract a timing feature and a topology feature of a fund path based on the fund path data; and generate the interaction feature corresponding to the fund path data based on the timing feature and the topology feature of the fund path.

[0134] In some embodiments, the path-model can also include a time encoding module and a C2GAT module. The time encoding module and the C2GAT module are similar to those in the ego-model. Details are omitted here for simplicity.

[0135] In some embodiments, the timing feature extracted based on the above-mentioned formula includes a time sequence corresponding to each node in the second subgraph. The timing feature and the topology feature of the second subgraph can be input into the trained CAGAT module in the path-model, to obtain the interaction feature of the second subgraph. The interaction feature of the second subgraph can include an interaction feature of a user role corresponding to the first node and an interaction feature of a merchant role corresponding to the third node. The interaction feature of the user role is used to describe interaction between the user role and another

transaction participant (including another user role and the merchant role), for example, a fund flowing direction or a transaction path. The interaction feature of the merchant role is used to describe interaction between the merchant role and another transaction participant, and may involve fund flowing, a transaction partner, etc. of the merchant role.

[0136] In this specification, the clustering feature and the interaction feature are separately extracted from different subgraphs by using the first feature extraction module and the second feature extraction module, to indicate behavior and interaction patterns of the user role and the merchant role in different aspects in the service system. Based on these features, whether the transaction event has a risk can be predicted more accurately.

[0137] In some embodiments, the prediction module can include a concatenation (concat) submodule, a perception submodule (MLP), and a classification submodule.

[0138] In some embodiments, the concatenation submodule is configured to perform concatenation processing on the clustering feature and the interaction feature, to obtain a risk feature of the sample transaction participant in the training sample. The risk feature includes information about the behavior and interaction patterns of the user role and the merchant role in different aspects in the service system, and can provide richer data to support risk assessment.

[0139] In some embodiments, the perception submodule can be a feedforward neural network, and is configured to: learn of a complex relationship between risk features, and extract risk key information that facilitates risk determining.

[0140] In some embodiments, the classification submodule is configured to determine risk prediction information based on the risk key information. For example, whether a current sample transaction participant has a fraud risk, a cash-out risk, etc. is output based on the risk key information.

[0141] In some embodiments, the first feature extraction module, the second feature extraction module, and the prediction module can be jointly trained, or can be separately trained. During joint training, a parameter of the target risk control model can be updated with a training objective of minimizing a difference between the risk prediction information and the label information. After a specific quantity of iterations, the trained target risk control model is obtained.

[0142] Alternatively, training can be separately performed. For example, the first feature extraction module and the second feature extraction module can be jointly trained. During joint training, supervised learning or unsupervised learning can be performed for the two modules. This

is not limited here. Then, the clustering feature and the interaction feature are input into the prediction module based on the trained first feature extraction module and the trained second feature extraction module, to obtain the risk prediction information output by the prediction module. Finally, a parameter of the prediction module is updated with a training objective of minimizing a difference between the risk prediction information and the label information, and the trained target risk control model is obtained after a specific quantity of iterations.

[0143] In conclusion, in the risk control model training method and system provided in this specification, the training device can generate a plurality of training samples based on the historical transaction data in the transaction scenario. The training sample includes the transaction behavior data of the sample transaction participant in the payment scenario. Then, the training device obtains the target risk control model through training based on the plurality of training samples. The target risk control model has a capability of identifying, based on the historical transaction data of the target transaction participant in the payment scenario, whether the target transaction participant has a transaction risk in the credit scenario. In the above-mentioned training process, the target risk control model applicable to the credit scenario can be obtained through training without using sample data in the credit scenario, to resolve a problem that no enough sample data in the credit scenario is used for model training at an initial stage of developing the credit service. The target risk control model is obtained through training in the risk control model training method provided in this specification, without a need to use credit data in the credit scenario. In other words, risk control in the credit scenario can be implemented based on the historical transaction data of the target transaction participant in the payment scenario. The target risk control model satisfies a requirement for risk control of a new user in the credit scenario.

[0144] Another aspect of this specification provides a risk control method. FIG. 9 is a flowchart of a risk control method according to an embodiment of this specification. In an example, the execution device 12 in FIG. 1 can perform the risk control method in this specification.

[0145] S410: Obtain transaction behavior data of a credit user in a payment scenario in response to a credit request initiated by the credit user.

[0146] In some embodiments, the transaction behavior data of the credit user is similar to transaction behavior data of a sample transaction participant, and also includes transaction relationship data and fund path data. The transaction relationship data is used to represent a transaction relationship between the credit user and at least one another transaction participant,

and is used to represent a fund exchange path related to the credit user.

[0147] In some embodiments, the obtaining transaction behavior data of a credit user in a payment scenario includes: obtaining a transaction graph of the credit user in the payment scenario, where the transaction graph includes a plurality of nodes and an edge between different nodes, one node corresponds to one transaction participant, and one edge corresponds to one transaction; and obtaining the transaction behavior data corresponding to the credit user based on the transaction graph, where in the transaction behavior data, the transaction relationship data corresponds to a first subgraph in the transaction graph, and the fund path data corresponds to a second subgraph in the transaction graph.

[0148] In some embodiments, a specific manner of obtaining the first subgraph and the second subgraph that correspond to the credit user is similar to a specific manner of obtaining a first subgraph and a second subgraph that correspond to the sample transaction participant. Details are omitted here for simplicity.

[0149] S420: Perform risk prediction on the credit user based on the transaction behavior data by using a trained target risk control model, to obtain a risk prediction result.

[0150] The target risk control model is obtained through training in the above-mentioned training method for the target risk control model. A manner and a procedure of obtaining the risk prediction result are similar to those in a training process. Details are omitted here for simplicity.

[0151] S430: Process the credit request based on the risk prediction result.

[0152] In some embodiments, the risk prediction result is used to indicate whether the transaction behavior data of the credit user has a risk. The credit request can be granted when there is no risk. The credit request can be rejected when there is a risk.

[0153] In conclusion, in the risk control method and system provided in this specification, when the credit user initiates the credit request, the target risk control model is used to analyze the transaction behavior data of the credit user in the payment scenario, to obtain the risk prediction result of the credit user, and further, process the credit request based on the risk prediction result. The target risk control model can predict a risk in the credit scenario based on the transaction behavior data of the credit user in the payment scenario. Therefore, the credit request can be processed more accurately based on the risk prediction result, to avoid occurrence of a credit risk, and satisfy a requirement for more accurate risk control of the new user in the credit scenario. In addition, the target risk control model is obtained through training based on the historical

transaction data in the payment scenario, and depends on sample data in the credit scenario. Therefore, even at an initial stage of developing a credit service, the target risk control model can accurately perform risk prediction on the credit user.

[0154] Another aspect of this specification provides a computer-readable non-transitory storage medium. The computer-readable non-transitory storage medium stores at least one instruction set, to perform risk control model training or risk control. When the at least one instruction set is executed by a processor, the at least one instruction set instructs the processor to implement steps of the risk control model training method or the risk control method 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 a computing device 200, the program code is used to enable the computing device 200 to perform steps of the risk control model training method or the risk control method described in this specification. The program product configured to implement the foregoing method can use a portable compact disc read-only memory (CD-ROM) that includes program code, and can run on the computing device 200. 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. The program code can be completely executed on the computing device 200, partially executed on the computing device 200, executed as an independent software package, partially executed on the computing device 200 and partially executed on a remote computing device, or completely executed on a remote computing device.

[0155] 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.

[0156] 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.

[0157] In addition, some terms in this specification are used to describe the embodiments of this specification. For example, "an embodiment", "embodiments", and/or "some embodiments" mean that a particular feature, structure, or feature described with reference to the embodiments can be included in at least one embodiment of this specification. Therefore, it can be emphasized and understood that two or more of "embodiments", "one embodiment", or "alternative

embodiments" in each part of this specification do not necessarily indicate the same embodiment. In addition, specific features, structures, or features can be properly combined in one or more embodiments of this specification.

[0158] It should be understood that in the above-mentioned descriptions of the embodiments of this specification, to help understand a feature, for the purpose of simplifying this specification, various features are combined in a single embodiment, accompanying drawing, or description thereof in this specification. However, this does not mean that the combination of these features is necessary. It is entirely possible for a person skilled in the art to mark some devices and understand the devices as separate embodiments when reading this specification. That is, the embodiments in this specification can also be understood as an integration of a plurality of secondary embodiments. Content of each secondary embodiment is also true when features are less than all features in one of the above-mentioned disclosed embodiments.

[0159] Each patent, patent application, publication of a patent application, and other materials, for example, articles, books, instructions, publications, documents, and products, that are cited in this specification and that are other than any document that is inconsistent or conflict with this document or that imposes limiting impact on the widest scope of the claims can be incorporated here by reference and used for all purposes now or later associated with this document. In addition, the term in this document shall prevail when there is any inconsistency or conflict between the description, definition and/or use of a related term in any material and the description, definition and/or use of a related terms in this document.

[0160] Finally, it should be understood that the principles of the implementation solutions of this specification are described in the implementation solutions of the application disclosed in this specification. Other modified embodiments also fall within the scope of this specification. Therefore, the embodiments disclosed in this specification are merely examples and not limitations. A person skilled in the art can use an alternative configuration according to the embodiments of this specification to implement the application in this specification. Therefore, the embodiments of this specification are not limited to the embodiments accurately described in the application.

CLAIMS

What is claimed is:

1. A risk control model training method, comprising:

obtaining historical transaction data in a payment scenario;

generating a plurality of training samples based on the historical transaction data, wherein one training sample comprises transaction behavior data of a sample transaction participant in the payment scenario; and

obtaining a target risk control model through training based on the plurality of training samples, wherein the target risk control model is used to identify, in a credit scenario, whether a target transaction participant has a transaction risk.

2. The method according to claim 1, wherein the transaction behavior data comprises:

transaction relationship data, used to represent a transaction relationship between the sample transaction participant and at least one another transaction participant; and

fund path data, used to represent a fund exchange path related to the sample transaction participant.

3. The method according to claim 2, wherein the generating a plurality of training samples based on the historical transaction data comprises:

obtaining a transaction graph corresponding to the historical transaction data, wherein the transaction graph comprises a plurality of nodes and an edge between different nodes, one node corresponds to one transaction participant, and one edge corresponds to one transaction; and

generating the plurality of training samples based on the transaction graph, wherein in one training sample, the transaction relationship data corresponds to a first subgraph in the transaction graph, and the fund path data corresponds to a second subgraph in the transaction graph.

4. The method according to claim 3, wherein the sample transaction participant corresponds to a user role in the payment scenario, and the first subgraph is obtained in the following manner:

finding a plurality of target nodes in the transaction graph, wherein the plurality of target nodes comprise a first node corresponding to the sample transaction participant and a second node connected to the first node, and a hop count between the first node and the second node is less than or equal to a preset hop count;

sampling an edge between any two of the plurality of target nodes based on a sequence of

transactions corresponding to different edges, to determine at least one target edge from edges between the plurality of target nodes; and

using, as the first subgraph, a subgraph that is in the transaction graph and that is determined by the plurality of target nodes and the at least one target edge.

5. The method according to claim 3, wherein the sample transaction participant corresponds to a merchant role in the payment scenario, and the first subgraph is obtained in the following manner:

finding a first node corresponding to the sample transaction participant in the transaction graph;

sampling, based on amounts of transactions corresponding to different edges, edges connected to the first node, and determining a plurality of target edges from the edges connected to the first node; and

using, as the first subgraph, a subgraph that is in the transaction graph and that is determined by the first node, the plurality of target edges, and second nodes respectively connected to the plurality of target edges.

6. The method according to claim 3, wherein the second subgraph is obtained in the following manner:

finding a first node corresponding to the sample transaction participant in the transaction graph;

determining, in the transaction graph, at least one third node connected to the first node, wherein a hop count between the third node and the first node is less than or equal to a preset hop count, and a fund return path exists between the third node and the first node; and

determining the second subgraph in the transaction graph based on the first node and the at least one third node.

7. The method according to claim 2, wherein the training sample further comprises label information of the sample transaction participant, and the label information represents whether the sample transaction participant has a transaction risk; and

the obtaining a target risk control model through training based on the plurality of training samples comprises:

performing a plurality of times of iterative training based on the plurality of training samples until the target risk control model is obtained, wherein a process of one time of iterative training

comprises:

inputting the transaction relationship data and the fund path data that correspond to the sample transaction participant into the target risk control model, to obtain risk prediction information obtained by performing risk prediction on the sample transaction participant by the target risk control model; and

updating a parameter of the target risk control model with a training objective of minimizing a difference between the risk prediction information and the label information.

8. The method according to claim 7, wherein the target risk control model comprises a first feature extraction module, a second feature extraction module, and a prediction module; and

the inputting the transaction relationship data and the fund path data that correspond to the sample transaction participant into the target risk control model, to obtain risk prediction information obtained by performing risk prediction on the sample transaction participant by the target risk control model comprises:

inputting the transaction relationship data into the first feature extraction module, to obtain a clustering feature between the sample transaction participant and the another transaction participant;

inputting the fund path data into the second feature extraction module, to obtain an interaction feature of the fund exchange path related to the sample transaction participant; and

inputting the clustering feature and the interaction feature into the prediction module, to obtain the risk prediction information.

9. The method according to claim 8, wherein the first feature extraction module is configured to:

extract a timing feature and a topology feature of the transaction relationship based on the transaction relationship data; and

generate the clustering feature corresponding to the transaction relationship data based on the timing feature and the topology feature of the transaction relationship.

10. The method according to claim 8, wherein the second feature extraction module is configured to:

extract a timing feature and a topology feature of a fund path based on the fund path data; and

generate the interaction feature corresponding to the fund path data based on the timing feature and the topology feature of the fund path.

11. The method according to claim 8, wherein the prediction module is configured to:

perform concatenation processing on the clustering feature and the interaction feature, to obtain a risk feature of the sample transaction participant; and

extract risk key information from the risk feature, and determine the risk prediction information based on the risk key information.

12. A risk control method, comprising:

obtaining transaction behavior data of a credit user in a payment scenario in response to a credit request initiated by the credit user;

performing risk prediction on the credit user based on the transaction behavior data by using a trained target risk control model, to obtain a risk prediction result, wherein the target risk control model is obtained through training in the method according to any one of claims 1 to 11; and

processing the credit request based on the risk prediction result.

13. The method according to claim 12, wherein the transaction behavior data comprises:

transaction relationship data, used to represent a transaction relationship between the credit user and at least one another transaction participant; and

fund path data, used to represent a fund exchange path related to the credit user.

14. The method according to claim 13, wherein the obtaining transaction behavior data of a credit user in a payment scenario comprises:

obtaining a transaction graph of the credit user in the payment scenario, wherein the transaction graph comprises a plurality of nodes and an edge between different nodes, one node corresponds to one transaction participant, and one edge corresponds to one transaction; and

obtaining the transaction behavior data corresponding to the credit user based on the transaction graph, wherein in the transaction behavior data, the transaction relationship data corresponds to a first subgraph in the transaction graph, and the fund path data corresponds to a second subgraph in the transaction graph.

15. The method according to claim 14, wherein the credit user corresponds to a user role in the payment scenario, and the first subgraph is obtained in the following manner:

finding a plurality of target nodes in the transaction graph, wherein the plurality of target nodes comprise a first node corresponding to the credit user and a second node connected to the first node, and a hop count between the first node and the second node is less than or equal to a preset hop count;

sampling an edge between any two of the plurality of target nodes based on a sequence of transactions corresponding to different edges, to determine at least one target edge from edges between the plurality of target nodes; and

using, as the first subgraph, a subgraph that is in the transaction graph and that is determined by the plurality of target nodes and the at least one target edge.

16. The method according to claim 14, wherein the credit user corresponds to a merchant role in the payment scenario, and the first subgraph is obtained in the following manner:

finding a first node corresponding to the credit user in the transaction graph;

sampling, based on amounts of transactions corresponding to different edges, edges connected to the first node, and determining a plurality of target edges from the edges connected to the first node; and

using, as the first subgraph, a subgraph that is in the transaction graph and that is determined by the first node, the plurality of target edges, and second nodes respectively connected to the plurality of target edges.

17. The method according to claim 14, wherein the second subgraph is obtained in the following manner:

finding a first node corresponding to the credit user in the transaction graph;

determining, in the transaction graph, at least one third node connected to the first node, wherein a hop count between the third node and the first node is less than or equal to a preset hop count, and a fund return path exists between the third node and the first node; and

determining the second subgraph in the transaction graph based on the first node and the at least one third node.

18. A risk control model training system, comprising:

at least one storage medium, storing at least one instruction set, to train a risk control model; and

at least one processor, communicatively connected to the at least one storage medium, wherein when the training system runs, 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 historical transaction data in a payment scenario;

generating a plurality of training samples based on the historical transaction data, wherein one training sample comprises transaction behavior data of a sample transaction participant in the

payment scenario; and

obtaining a target risk control model through training based on the plurality of training samples, wherein the target risk control model is used to identify, in a credit scenario, whether a target transaction participant has a transaction risk.

19. A risk control system, comprising:

at least one storage medium, storing at least one instruction set, to perform risk control; and

at least one processor, communicatively connected to the at least one storage medium, wherein when the risk control system runs,

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 transaction behavior data of a credit user in a payment scenario in response to a credit request initiated by the credit user;

performing risk prediction on the credit user based on the transaction behavior data by using a trained target risk control model, to obtain a risk prediction result, wherein the target risk control model is obtained through training in the method according to any one of claims 1 to 11; and

processing the credit request based on the risk prediction result.

20. A computer-readable nonvolatile storage medium, wherein the computer-readable nonvolatile 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 11 or the method according to any one of claims 12 to 17 is implemented.