

CLAIMS

1. A method for abnormal account prevention and control based on a graph structure model, comprising:
 - obtaining (S302) a graph structure model trained by using labeled samples, wherein the graph structure model is trained based on an account relationship network, and the account relationship network comprises a plurality of nodes, and edges corresponding to relationships between the nodes, wherein the plurality of nodes include an account, device information, network information, location information, and identity information, wherein each labeled sample comprises a label indicating whether a corresponding node of the graph structure model is an abnormal node, wherein training the graph structure model based on the account relationship network using labeled samples comprises:
 - for each node included in the plurality of nodes of the account relationship network,
 - the graph structure model is used to calculate an embedding vector of the node after multiple iterations in a hidden feature space based on an original feature of the node and/or a feature of an edge of the node that is associated with another node and/or an embedding vector of the node in a previous iteration in the hidden feature space,
 - wherein the trained graph structure model maximizes a consistency between
 - (i) a predicted result of a node based on the embedding vector of the node
 - and (ii) the label of the labeled sample corresponding to the node;
 - calculating (S304), by using the trained graph structure model, an embedding vector that corresponds to a to-be-tested sample, and a prediction probability that corresponds to the to-be-tested sample based on the embedding vector that corresponds to the to-be-tested sample; and
 - performing (S306) abnormal account prevention and control on the to-be-tested sample based on the embedding vector that corresponds to the to-be-tested sample, and based on the prediction probability that corresponds to the to-be-tested sample,

wherein the to-be-tested sample is the account and if the calculated prediction probability that corresponds to the account is greater than a specified threshold, the account is determined as an abnormal account, and then the abnormal account is deregistered;

wherein the labeled samples are labeled with information indicating whether a node that corresponds to the labeled sample is an abnormal node.

2. The method according to claim 1, wherein the account relationship network comprises one or a combination of more of the following networks: an association network of an account and an international mobile equipment identity (IMEI), an association network of an account and a media access control (MAC) address, an association network of an account and an Internet protocol (IP) address, an association network of an account and a mobile phone number, an association network of an account and an international mobile subscriber identity (IMSI), an association network of an account and an email address, an association network of an account and an identity card number, an association network of an account and a bank card number, and an association network of an account and location information.
3. The method according to claim 1, wherein the original feature comprises at least one of the following types of data: node type, the number of registered accounts, the number of authenticated accounts, the number of abnormal accounts, and time-related statistical data; and the feature of the edge comprises at least one of the following types of data involved in a node that is connected to the edge: behavior occurrence time, behavior consumption time, and behavior failure times, wherein the behavior comprises registration behavior and/or authentication behavior.
4. The method according to claim 1, wherein an embedding vector of the node after the t^{th} iteration in the hidden feature space is calculated based on the original feature of the node, a feature of an edge of the node that is associated with another node, and an embedding vector of the node after the $(t-1)^{\text{th}}$ iteration in

the hidden feature space.

5. The method according to claim 1, wherein the calculating an embedding vector of the node after multiple iterations in a hidden feature space based on the original feature of the node and the feature of the edge of the node that is associated with another node comprises:
calculating the embedding vector of the node after multiple iterations in the hidden feature space based on the following equation:

$$H_i^t = \sigma \left(W_1 F_i + W_2 \sum_{j \in N_i} E_{ij} + W_3 \sum_{j \in N_i} \alpha_{ij} H_i^{t-1} \right);$$

wherein H_i^t represents an embedding vector of an i^{th} node after the t^{th} iteration in the hidden feature space, σ represents a nonlinear transform function, W_1 , W_2 , and W_3 represent weight matrices, F_i represents an original feature of the i^{th} node, E_{ij} represents a feature of an edge of the i^{th} node that is associated with a j^{th} node, N_i represents a set comprising neighboring nodes of the i^{th} node, and α_{ij} represents a weight matrix for the i^{th} node and the j^{th} node.

6. The method according to claim 5, wherein the calculating a prediction probability of the node based on the embedding vector comprises:
calculating the prediction probability of the node based on the following equation:

$$pred_i = softmax(W_4 * H_i^T);$$

wherein $pred_i$ represents a prediction probability of the i^{th} node after T iterations, and W_4 represents a weight matrix.

7. The method according to claim 1, wherein the training the graph structure model by using the labeled samples comprises:
training the graph structure model by using the labeled samples and targeting

the maximization of consistency between the prediction probability and a corresponding sample labeling result.

8. The method according to claim 6, wherein the training the graph structure model by using the labeled samples comprises:

using a back-propagation algorithm and the labeled samples to optimize

$\arg \min_{w_1, w_2, w_3, w_4} \sum_i corss_entrep(pred_i, y_i)$, to obtain optimal W_1 , W_2 ,

W_3 , and W_4 ; wherein

$corss_entrep$ represents a function used to calculate cross entropy.

9. A device for abnormal account prevention and control based on a graph structure model, comprising:

at least one processor; and

a memory communicatively connected to the at least one processor; wherein

the memory stores an instruction that can be executed by the at least one

processor, and the instruction is executed by the at least one processor, to

enable the at least one processor to perform a method according to any one of the preceding claims.