

CLAIMS

What is claimed is:

1. A method, comprising:

5 obtaining an account medium network including a plurality of nodes and at least one edge, a node of the plurality of nodes representing one of an account or a medium, an edge of the at least one edge indicating a login behavior relationship between nodes of the plurality of nodes connected by the edge;

obtaining a feature data and a risk labeling data of a node of the plurality of nodes, the feature data reflecting a login behavior of the node in a period
10 of time, the risk labeling data indicating a first possibility of the node being a target node; and

training a graph structure model based on the account medium network, the feature data, and the risk labeling data.

15 2. The method according to claim 1, wherein the medium comprises a device.

3. The method according to claim 1, comprising calculating an embedded vector of the node in a hidden feature space using the graph structure
20 model and based on the feature data of the node and a topology corresponding to the node in the account medium network.

4. The method according to claim 3, wherein the calculating the embedded vector of the node in the hidden feature space includes a plurality of
25 iterations, the embedded vector of the node in the hidden feature space after a t^{th} iteration being calculated based on the feature data of the node, the topology corresponding to the node in the account medium network, and the embedded vector of the node in the hidden feature space after a $t - 1^{\text{th}}$ iteration of the plurality of iterations.

30 5. The method according to claim 4, wherein the calculating the embedded vector of the node in the hidden feature space comprises:

calculating the embedded vector of the node in the hidden feature space with the plurality of iterations based on an equation:

$$\Phi^{(t+1)} = \sigma \left(X W_1 + G \Phi^{(t)} W_2 \right);$$

wherein:

5 $\Phi^{(t+1)}$ indicates the embedded vector of the node in the hidden feature space after a $t+1$ th iteration of the plurality of iterations;

σ indicates a non-linear transformation function;

W_1 and W_2 each indicate weight matrices;

X indicates the feature data of the node; and

10 G indicates the topology corresponding to the node in the account medium network.

6. The method according to claim 5, comprising calculating a prediction data indicating a second possibility of the node being a target node based on an equation:

$$pred_i = w^T \phi_i$$

wherein:

$pred_i$ indicates the prediction data of the node;

20 ϕ_i indicates the embedded vector of the node in the hidden feature space;

w^T indicates a parameter vector that is obtained by quantizing ϕ_i ;

and

T indicates a transposition operation.

25

7. The method according to claim 6, wherein the training the graph structure model comprises:

training the graph structure model to improve a consistency between the prediction data and the risk labeling data of the node.

30

8. The method according to claim 6, wherein the training the graph

structure model comprises:

optimizing $\arg \min_{w_1, w_2, w} \sum_i L(y_i, pred_i)$ based on a

backpropagation algorithm and the risk labeling data of the node to obtain w_1 , w_2 ,
and w ,

5 wherein:

y_i indicates the risk labeling data of the node; and

L indicates a loss function for measuring an inconsistency between
the prediction data and the risk labeling data.

10 9. The method according to claim 3, comprising calculating, based on
the embedded vector, a prediction data of the node indicating a second possibility of
the node being a target node.

15 10. The method according to claim 1, wherein the obtaining the feature
data of the node comprises:

obtaining a login behavior data of the node within the period of time;

dividing the period of time to obtain a time sequence; and

generating, based on a distribution of the login behavior data in the
time sequence, a feature vector as the feature data of the node.

20

11. The method according to claim 1, comprising:

obtaining a feature data of a to-be-identified account;

obtaining the account medium network that contains the
to-be-identified account as a first node;

25 inputting the feature data of the to-be-identified account and a topology
corresponding to the first node in the account medium network into the graph
structure model for calculation of a first prediction data of the first node; and

determining whether the to-be-identified account is a target account
based on the first prediction data.

30

12. An apparatus, comprising:

a first acquisition module, configured to obtain an account medium

network including a plurality of nodes and at least one edge, a node of the plurality of nodes representing one of an account or a medium, an edge of the at least one edge indicating a login behavior relationship between nodes of the plurality of nodes connected by the edge;

5 a second acquisition module, configured to obtain a feature data and a risk labeling data of a node of the plurality of nodes, the feature data reflecting a login behavior of the node in a period of time, the risk labeling data indicating a first possibility of the node being a target node; and

10 a training and identification module, configured to train a graph structure model based on the account medium network, the feature data, and the risk labeling data.

13. The apparatus according to claim 12, wherein the medium comprises a device.

15 14. The apparatus according to claim 12, wherein the training and identification module calculates an embedded vector of the node in a hidden feature space using the graph structure model and based on the feature data of the node and a topology corresponding to the node in the account medium network.

20 15. The apparatus according to claim 14, wherein the training and identification module calculates the embedded vector of the node in the hidden feature space with a plurality of iterations, the embedded vector of the node in the hidden feature space after a t^{th} iteration being calculated based on the feature data of the node, the topology corresponding to the node in the account medium network, and the embedded vector of the node in the hidden feature space after a $t - 1^{\text{th}}$ iteration of the plurality of iterations.

30 16. The apparatus according to claim 15, wherein the training and identification module calculates the embedded vector of the node in the hidden feature space based on an equation:

$$\Phi^{(t+1)} = \sigma \left(X W_1 + G \Phi^{(t)} W_2 \right);$$

wherein:

$\Phi^{(t+1)}$ indicates the embedded vector of the node in the hidden feature space after a $t+1$ th iteration of the plurality of iterations;

σ indicates a non-linear transformation function;

W_1 and W_2 each indicate weight matrices;

5 X indicates the feature data of the node; and

G indicates the topology corresponding to the node in the account medium network.

17. The apparatus according to claim 16, wherein the training and
10 identification module calculates a prediction data indicating a second possibility of the node being a target node based on an equation:

$$pred_i = w^T \phi_i;$$

wherein:

$pred_i$ indicates the prediction data of the node;

15 ϕ_i indicates the embedded vector of the node in the hidden feature space;

w^T indicates a parameter vector that is obtained by quantitizing ϕ_i ;
and

T indicates a transposition operation.

20

18. The apparatus according to claim 17, wherein the training and identification module trains the graph structure model to improve a consistency between the prediction data and the risk labeling data of the node.

25 19. The apparatus according to claim 17, wherein that the training and identification module trains the graph structure model by:

optimizing $\arg \min_{W_1, W_2, w} \sum_i L(y_i, pred_i)$ based on a

backpropagation algorithm and the risk labeling data of the node to obtain w_1 , w_2 ,
and w ;

30 wherein:

y_i indicates the risk labeling data of the node; and

L indicates a loss function for measuring an inconsistency between the prediction data and the risk labeling data.

5 20. The apparatus according to claim 14, wherein the training and identification module calculates a prediction data of the node based on the embedded vector, the prediction data indicating a second possibility of the node being a target node.

10 21. The apparatus according to claim 12, wherein the second acquisition module obtains the feature data of the node by:

obtaining a login behavior data of the node within the period of time;

dividing the period of time to obtain a time sequence; and

15 generating, based on a distribution of the login behavior data in the time sequence, a feature vector as the feature data of the node.

22. The apparatus according to claim 12, comprising:

20 an acquisition module, configured to obtain a feature data of a to-be-identified account, and obtain the account medium network that contains the to-be-identified account as a first node;

an input module, configured to input the feature data of the to-be-identified account and a topology corresponding to the first node in the account medium network into the graph structure model for calculation of a first prediction data of the first node; and

25 a determining module, configured to determine whether the to-be-identified account is a target account based on the first prediction data.

23. A device, comprising:

at least one processor; and

30 a memory communicatively connected to the at least one processor and storing instructions that when executed by the at least one processor, configure the at least one processor to implement acts including:

obtaining an account medium network including a plurality of nodes

and at least one edge, a node of the plurality of nodes representing one of an account or a medium, an edge of the at least one edge indicating a login behavior relationship between nodes of the plurality of nodes connected by the edge;

5 obtaining a feature data and a risk labeling data of a node of the plurality of nodes, the feature data reflecting a login behavior of the node in a period of time, the risk labeling data indicating a first possibility of the node being a target node; and

training a graph structure model based on the account medium network, the feature data, and the risk labeling data.

10

24. The device according to claim 23, wherein the acts include:

obtaining a feature data of a to-be-identified account;

obtaining the account medium network that contains the to-be-identified account as a first node;

15

inputting the feature data of the to-be-identified account and a topology corresponding to the first node in the account medium network into the graph structure model for calculation of a first prediction data of the first node; and

determining whether the to-be-identified account is a target account based on the first prediction data.