

CLAIMS

What is claimed is:

1. A method, comprising:

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

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

calculating 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
15 corresponding to the node in the account medium network, wherein the calculating the
embedded vector of the node in the hidden feature space includes 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
20 the node in the hidden feature space after a $t-1^{\text{th}}$ iteration of the plurality of iterations.

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

25 3. The method according to claim 1, 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:

$$\Phi^{(t+1)} = \sigma(XW_1 + G\Phi^{(t)}W_2);$$

30 wherein:

$\Phi^{(t+1)}$ indicates the embedded vector of the node in the hidden feature

space after a $t+1^{\text{th}}$ 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

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

4. The method according to claim 3, comprising calculating a prediction data indicating a second possibility of the node being the target node based on an equation:

$$pred_i = w^T \phi_i$$

wherein:

$pred_i$ indicates the prediction data of the node;

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

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

T indicates a transposition operation.

5. The method according to claim 4, wherein the training the graph structure model comprises:

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

6. The method according to claim 4, wherein the training the graph structure model comprises:

25 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 ,

wherein:

y_i indicates the risk labeling data of the node; and

30 L indicates a loss function for measuring an inconsistency between the

prediction data and the risk labeling data.

7. The method according to claim 1, comprising calculating, based on the embedded vector, a prediction data of the node indicating a second possibility of the node being the target node.

8. 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.

9. 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;
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.

10. 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;
a second acquisition module, configured to obtain a feature data and a risk labeling data of the 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;
a training and identification module, configured to: (i) train a graph

structure model based on the account medium network, the feature data, and the risk labeling data, and (ii) calculate 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, 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.

11. The apparatus according to claim 10, wherein the medium comprises a device.

12. The apparatus according to claim 10, wherein the training and identification module calculates the embedded vector of the node in the hidden feature space based on an equation:

$$\Phi^{(t+1)} = \sigma(XW_1 + G\Phi^{(t)}W_2);$$

wherein:

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

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

13. The apparatus according to claim 12, wherein the training and identification module calculates a prediction data indicating a second possibility of the node being the target node based on an equation:

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

wherein:

$pred_i$ indicates the prediction data of the node;

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.

5

14. The apparatus according to claim 13, 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.

10

15. The apparatus according to claim 13, 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 ;

15

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.

20

16. The apparatus according to claim 10, 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 the target node.

25

17. The apparatus according to claim 10, 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

generating, based on a distribution of the login behavior data in the time

30

sequence, a feature vector as the feature data of the node.

18. The apparatus according to claim 10, comprising:
a third 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;
5 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
a determining module, configured to determine whether the to-be-identified account is a target account based on the first prediction data.
10

19. A device, comprising:
at least one processor; and
a memory communicatively connected to the at least one processor and
15 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
20 between nodes of the plurality of nodes connected by the edge;
obtaining a feature data and a risk labeling data of the 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;
training a graph structure model based on the account medium network,
25 the feature data, and the risk labeling data; and
calculating 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, wherein the calculating the embedded vector of the node in the hidden feature space includes a plurality of
30 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.

20. The device according to claim 19, 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
5 account as a first node;
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
10 based on the first prediction data.