

【DESCRIPTION】**【Title of Invention】**

SYSTEM AND METHOD FOR GENERATING VIRTUAL ASSET WALLET
ADDRESS BLACKLIST DATABASE BASED ON GRAPH ATTENTION NETWORK (GAT)

5 **【Technical Field】**

[1] The present disclosure relates to a system and method for generating a virtual asset wallet address blacklist database, and more specifically, to a system and method for generating a virtual asset wallet address blacklist database based on a graph attention network (GAT).

【Background Art】

10 **[2]** Due to their nature, blockchain-based virtual assets are difficult to trace and are often used for illegal financial transactions. If a virtual asset wallet itself is frequently used for very risky transactions, it is necessary to preemptively identify the virtual asset wallet and prevent any fraudulent use thereof.

[3] However, since it is very easy to create a virtual asset wallet, and the procedures related
15 to security or authentication are lax, anyone is able to create a virtual asset wallet at any time and use the wallet for fraudulent purposes.

[4] Therefore, it is necessary to identify high-risk or harmful virtual asset wallets in advance and prevent such wallets from being used in illegal transactions or fraudulent use.

[5] However, current virtual asset exchanges lack the means to prevent these illegal
20 transactions or misuse.

[6] In particular, there is no properly established standard for determining whether numerous virtual asset wallets pose a high or low risk of illegality. Therefore, an algorithm is needed to accurately estimate the potential illegality of virtual asset wallets.

【Disclosure】25 **【Technical Problem】**

[7] One object of the present disclosure is to provide a system for generating a virtual asset wallet address blacklist database based on a graph attention network (GAT).

[8] Another object of the present disclosure is to provide a method for generating a virtual asset wallet address blacklist database based on a GAT.

5 **【Technical Solution】**

[9] A system for generating a virtual asset wallet address blacklist database based on a graph attention network (GAT) according to an aspect of the present disclosure includes: a GAT AI engine server configured to train a GAT-based AI model using an index of a full node for each virtual asset pre-stored in an index database server, common transaction item information
10 pre-stored in a virtual asset transaction analysis database server, and a main blacklist consisting of high-risk virtual asset wallet addresses pre-stored in a main blacklist server, to calculate GAT scores based on the trained AI model, to estimate high-risk virtual asset wallet addresses using the calculated GAT scores, and to generate a GAT blacklist consisting of the estimated high-risk virtual asset wallet addresses; and a GAT blacklist database server configured to store the
15 GAT blacklist generated by the GAT AI engine server.

[10] Here, the GAT AI engine server may include: a first data preprocessing module configured to preprocess the common transaction item information stored in the virtual asset transaction analysis database server, to label each common transaction item information according to a type of coin, and to query any transaction corresponding to a predetermined
20 virtual asset wallet address; an AI learning module configured to perform GAT learning using transactions queried by the first data preprocessing module; a teacher module configured to perform pseudo-labeling on unlabeled transactions from the first data preprocessing module, and feed the pseudo-labeled transactions to the AI learning module to re-learn the pseudo-labeled transactions; and a first risk calculation module configured to calculate a risk level
25 corresponding to a GAT score of each virtual asset wallet address based on results of the GAT

learning performed by the AI learning module.

[11] In addition, the system may further include: a high-risk wallet address service server configured to receive a virtual asset wallet address from a virtual asset exchange server, to calculate a risk level for the received virtual asset wallet address, and to respond to the virtual asset exchange server with the calculated risk level.

[12] Further, the high-risk wallet address service server may include: a second data preprocessing module configured to receive a virtual asset wallet address from the virtual asset exchange server and perform preprocessing; a transaction inquiry module configured to query transactions of the virtual asset wallet address, which has undergone preprocessing by the second data preprocessing module 331, from the virtual asset transaction analysis database server; a second risk calculation module configured to calculate a risk level corresponding to a GAT score of the virtual asset wallet address using the transactions queried by the transaction query module; and a learning update module configured to request the AI learning module to train an artificial intelligence model based on the risk level calculated by the second risk calculation module.

[13] Further, the system may further include a high-risk wallet address management server configured to receive a virtual asset wallet address from an administrator terminal, to calculate a risk level for the received virtual asset wallet address, and to respond to the administrator terminal with the calculated risk level.

[14] A method for generating a virtual asset wallet address blacklist database based on a graph attention network (GAT) according to another aspect of the present disclosure includes: training, by a GAT AI engine server, a GAT-based AI model using an index of a full node for each virtual asset pre-stored in an index database server, common transaction item information pre-stored in a virtual asset transaction analysis database server, and a main blacklist consisting of high-risk virtual asset wallet addresses pre-stored in a main blacklist server; calculating GAT

scores based on the trained AI model by the GAT AI engine server; estimating high-risk virtual asset wallet addresses using the calculated GAT scores and generating a GAT blacklist consisting of the estimated high-risk virtual asset wallet addresses by the GAT AI engine server; and storing, by a GAT blacklist database server, the GAT blacklists generated by the GAT AI engine server.

[15] Here, the method may further include receiving a virtual asset wallet address from a virtual asset exchange server, calculating a risk level for the received virtual asset wallet address, and responding to the virtual asset exchange server with the calculated risk level by a high-risk wallet address service server.

[16] Further, the method may further include receiving a virtual asset wallet address from an administrator terminal, calculating a risk level for the received virtual asset wallet address, and responding to the administrator terminal with the calculated risk level by a high-risk wallet address management server.

【Advantageous Effects】

[17] According to the above-described system and method for generating a virtual asset wallet address blacklist database based on a graph attention network (GAT), it is configured to collect a full node and corresponding transaction data for each virtual asset network, and to perform GAT learning based on a database generated by automatically extracting training data to proactively detect a full node with a high potential for fraudulent use, etc., from the collected full node and transaction data, and a database built by generating a blacklist and whitelist for each virtual asset wallet address, and this allows for more accurate estimation and monitoring of a probability of illegal use or fraudulent transactions regarding all virtual asset wallet addresses.

【Description of Drawings】

[18] FIG. 1 is a block diagram of a system for generating a virtual asset wallet address

blacklist database based on a graph attention network (GAT) according to an embodiment of the present disclosure.

[19] FIG. 2 is a detailed configuration diagram of a system for generating a virtual asset wallet address blacklist database based on a GAT according to an embodiment of the present disclosure.

[20] FIG. 3 is a flowchart of a method for generating a virtual asset wallet address blacklist database based on a GAT according to an embodiment of the present disclosure.

[21] FIGS. 4 and 6 are detailed flowcharts of a method for generating a virtual asset wallet address blacklist database based on a GAT according to an embodiment of the present disclosure.

[22] FIG. 5 is an example of risk calculation data according to an embodiment of the present disclosure.

[23]

【Best Mode for Carrying Out the Invention】

[24] Various modifications may be made to exemplary embodiments of the present disclosure, and specific exemplary embodiments will be described below in detail with reference to attached drawings. However, it should be understood that the present disclosure is not to be limited to the specific embodiments, but includes all modifications, equivalents, and alternatives falling within the spirit and scope of the disclosure. The like reference numerals are used for similar components in describing each of the drawings.

[25] It will be understood that, although the terms “first”, “second”, “A”, “B”, and the like may be used herein in explaining various components of the present disclosure, such components should not be limited by these terms. The above terms are only used to distinguish one component from another. For example, without departing from the scope of the present disclosure, a first component may be referred to as a second component, and similarly, the second component may also be referred to as the first component. The

expression “and/or” encompasses any one of a combination of a plurality of associated items as described, and a plurality of associate items as described.

[26] When a component is referred to as being “connected” or “accessed” to other component, it should be understood that not only is the component directly connected or accessed to the other component, but also, another component may exist therebetween. On the other hand, when a component is referred to as being “directly connected” or “directly accessed” to other component, it should be understood that there is no component therebetween.

[27] The terms used in this application are only used to describe specific embodiments and are not intended to limit the present disclosure. The singular forms are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms “comprise”, “include”, “have”, etc. when used in this specification, specify the presence of stated features, integers, steps, operations, elements, components, and/or combinations of them but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or combinations thereof.

[28] Unless otherwise defined, all terms including technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the present disclosure belongs. It will be further understood that terms, such as those defined in commonly used dictionaries, should be interpreted as having a meaning that is consistent with their meaning in the context of the relevant art and the present disclosure, and will not be interpreted in an idealized or overly formal sense unless expressly so defined herein.

[29] Hereinafter, preferred embodiments according to the present disclosure will be described in detail with reference to the attached drawings.

[30] FIG. 1 is a block diagram of a system for generating a virtual asset wallet address blacklist database based on a graph attention network (GAT) according to an embodiment of the present disclosure, and FIG. 2 is a diagram illustrating the detailed configuration of a system

for generating a virtual asset wallet address blacklist database based on a GAT according to an embodiment of the present disclosure.

[31] Referring to FIGS. 1 and 2, a system 300 for generating a virtual asset wallet address blacklist database based on a GA may include a GAT AI engine server 310, a GAT blacklist database server 320, a high-risk wallet address service server 330, and a high-risk wallet address management server 340.

[32] Hereinafter, the detailed configuration will be described.

[33] The GAT AI engine server 310 may be configured to train a GAT-based artificial intelligence model using an index of a full node for each virtual asset pre-stored in an index database server 101, common transaction item information pre-stored in a virtual asset transaction analysis database server 102, and a main blacklist comprising high-risk virtual asset wallet addresses pre-stored in a main blacklist server 201.

[34] Here, the index database server 101 is configured to store the index of the full node for each virtual asset, and the index may be generated in advance by sequentially indexing each virtual asset wallet address. In addition, the common transaction item information may be a sender's wallet address, a recipient's wallet address, a virtual asset transfer amount, a transfer fee, a transaction time, a transaction hash, a previous transaction output, and signatures for all transactions.

[35] In addition, the GAT AI engine server 310 may be configured to calculate GAT scores based on the artificial intelligence model trained as above, estimate high-risk virtual asset wallet addresses using the calculated GAT scores, and generate a GAT blacklist consisting of the estimated high-risk virtual asset wallet addresses.

[36] The GAT AI engine server 310 may be configured to include a first data preprocessing module 311, an AI learning module 312, a teacher module 313, and a first risk calculation module 314.

[37] Hereinafter, the detailed configuration will be described.

[38] The first data preprocessing module 311 may be configured to preprocess the common transaction item information stored in the virtual asset transaction analysis database server 102, label each common transaction item information according to a type of coin, and query any transaction corresponding to a predetermined virtual asset wallet address.

[39] The AI learning module 312 may be configured to perform GAT learning using the transaction queried by the first data preprocessing module 311.

[40] At this point, the AI learning module 312 may execute a GAT learning model through the following processes to generate a GAT blacklist.

[41] First, the AI learning module 312 may extract a GAT graph structure for all transactions based on the GAT using the index of the full node for each virtual asset pre-stored in the index database server 101 and the common transaction item information pre-stored in the virtual asset transaction analysis database server 102. In addition, an adjacency matrix corresponding to connections within the GAT graph structure as edges may be generated.

[42] Then, the AI learning module 312 may generate a feature matrix by assigning a preset weight to each online source site during crawl collection for data matrixed by risk category, using a main blacklist consisting of high-risk virtual asset wallet addresses pre-stored in the main blacklist server 201.

[43] The AI learning module 312 may pass the feature matrix and adjacency matrix, which constitute the nodes and edges of the GAT graph structure, through an attention layer to generate a pre-trained model for the labeled training data of high-risk virtual asset wallet addresses. Using this pre-trained model, the AI learning module 312 may predict unlabeled data for high-risk virtual asset wallet addresses and perform a main learning model which is based on semi-supervised learning, and in which a prediction result is used as a GAT score to estimate a high-risk virtual asset wallet address.

[44] At this point, the AI learning module 312 may be configured to automatically generate and update a GAT blacklist database, which contains GAT scores for estimating high-risk virtual asset wallet addresses, by repeatedly executing the main learning model.

[45] Meanwhile, the AI learning module 312 may be configured to generate a feature matrix with category classification values of high-risk virtual asset wallet addresses or transactions, by mapping all transactions to nodes in the graph structure for a BTC network and by mapping all wallet addresses to nodes in a graph structure for an ETH network.

[46] Furthermore, the AI learning module 312 may be configured to generate an adjacency matrix by mapping the connections between transactions as edges for the BTC network and by mapping the connections between wallet addresses as edges for the ETH network.

[47] On the other hand, a process of generating a GAT blacklist proceeds as follows.

[48] First, an arbitrary virtual asset wallet address is entered. Then, all transactions adjacent to the arbitrary virtual asset wallet address are queried. Next, a risk category classified for each transaction and an estimated probability value for each transaction to be a high-risk virtual asset wallet address are queried. The risk category classified for each transaction and the estimated probability value for each transaction to be a high-risk virtual asset wallet address are then combined to calculate an overall risk level for the arbitrary virtual asset wallet address.

[49] Here, the process of querying the probability value includes the steps of: querying all transactions related to the arbitrary virtual asset wallet address from the GAT blacklist database; predicting unlabeled data for high-risk virtual asset wallet addresses by utilizing transaction data labeled with GAT scores, which are used to estimate high-risk virtual asset wallet addresses among all the transactions, and using the prediction results as risk probability values to estimate the high-risk virtual asset wallet addresses; and storing a risk category classified for each transaction and an estimated probability value for each transaction to be a high-risk virtual asset

wallet address.

[50] The teacher module 313 may be configured to perform pseudo-labeling on the unlabeled transactions from the first data preprocessing module 311, and feed the pseudo-labeled transactions to the AI learning module to re-learn the transactions.

5 **[51]** The first risk calculation module 314 may be configured to calculate a risk level corresponding to a GAT score of each virtual asset wallet address based on a result of the GAT learning performed by the AI learning module 312.

[52] The GAT blacklist database server 320 may be configured to store the GAT blacklist generated by the GAT AI engine server 310.

10 **[53]** The high-risk wallet address service server 330 may be configured to receive a virtual asset wallet address from a virtual asset exchange server, calculate a risk level for the received virtual asset wallet address, and respond to the virtual asset exchange server with the calculated risk level.

[54] The high-risk wallet address service server 330 may be configured to include a second
15 data preprocessing module 331, a transaction query module 332, a second risk calculation module 333, and a learning update module 334.

[55] Hereinafter, the detailed configuration will be described.

[56] The second data preprocessing module 331 may be configured to receive a virtual asset wallet address from the virtual asset exchange server and perform preprocessing.

20 **[57]** The transaction query module 332 may be configured to query transactions of the virtual asset wallet address, which has undergone preprocessing by the second data preprocessing module 331, from the virtual asset transaction analysis database server 102.

[58] The second risk calculation module 333 may be configured to calculate a risk level
25 by the transaction query module 332.

[59] The training update module 334 may be configured to request the AI learning module 312 to train an artificial intelligence model based on the risk level calculated by the second risk calculation module 333.

[60] The high-risk wallet address management server 340 may be configured to receive a virtual asset wallet address from an administrator terminal 400, calculate a risk level for the received virtual asset wallet address, and respond to the administrator terminal 400 with the calculated risk level.

[61] FIG. 3 is a flowchart of a method for generating a virtual asset wallet address blacklist database based on a GAT according to an embodiment of the present disclosure.

[62] Referring to FIG. 3, a GAT AI engine server 310 trains a GAT-based artificial intelligence model using an index of a full node for each virtual asset pre-stored in an index database server 101, common transaction item information pre-stored in a virtual asset transaction analysis database server 102, and a main blacklist consisting of high-risk virtual asset wallet addresses pre-stored in a main blacklist server 201 (S10).

[63] Next, the GAT AI engine server 310 calculates GAT scores based on the trained artificial intelligence model (S20).

[64] Next, the GAT AI engine server 310 estimates high-risk virtual asset wallet addresses using the calculated GAT scores and generates a GAT blacklist consisting of the estimated high-risk virtual asset wallet addresses (S30).

[65] Next, the GAT blacklist generated by the GAT AI engine server 310 is stored in a GAT blacklist database server 320 (S40).

[66] Next, a high-risk wallet address service server 330 receives a virtual asset wallet address from a virtual asset exchange server, calculates a risk level for the received virtual asset wallet address, and responds to the virtual asset exchange server with the calculated risk level (S50).

[67] Next, a high-risk wallet address management server 340 receives a virtual asset wallet address from an administrator terminal 400, calculates a risk level for the received virtual asset wallet address, and responds to the administrator terminal 400 with the calculated risk level (S60).

5 **[68]** FIGS. 4 and 6 are detailed flowcharts of a method for generating a virtual asset wallet address blacklist database based on a GAT according to an embodiment of the present disclosure, and FIG. 5 illustrates an example of risk level calculation data according to an embodiment of the present disclosure.

[69] Here, FIG. 4 relates to a detailed sequence of generating a GAT-based learning model, where all transactions for a predetermined wallet address are queried and learning is performed based on a GAT algorithm. At this point, even for data that is not labeled during a data preprocessing process, the accuracy of the data is improved through re-learning by the teacher module 313. Also, by calculating a risk level corresponding to a GAT score, a GAT-based blacklist may be generated. Furthermore, a process of immediately accepting unlabeled data through a temporary teacher and performing immediate relearning is exemplified.

[70] FIG. 5 illustrates the result of searching for transactions of a specific BTC wallet address A, calculating a sum of A's transactions by category, and monitoring a total amount sent and received by A.

[71] For example, the total amount sent and received in the "Threat" category for A is 12.4, and the total amount sent and received in the "Tumble" category is 20.1531. The total amount sent and received in all categories for A is 100.

[72] Here, the risk level of A is calculated by dividing a maximum value of the amount sent and received by category by the total amount sent and received by A. Here, A's risk is calculated by dividing the maximum transmission and reception volume for each category by A's total transmission and reception volume. That is, a risk level of A's Tumble category is

20.1531/100, which is 20.1531%.

[73] FIG. 6 illustrates a process of calculating and responding with a risk level of a virtual asset wallet address requested by a virtual asset exchange server. At this point, a result of risk level calculation is updated in a GAT blacklist database server 320 to build a database.

5 **[74]** As described above, the present disclosure is configured to collect a full node and corresponding transaction data for each virtual asset network, and to perform GAT learning based on a database generated by automatically extracting training data to proactively detect a full node with a high potential for fraudulent use, etc., from the collected full node and transaction data, and a database built by generating a blacklist and whitelist for each virtual
10 asset wallet address, and this allows for more accurate estimation and monitoring of a probability of illegal use or fraudulent transactions regarding all virtual asset wallet addresses.

[75] Although the description has been made with reference to the above examples, those skilled in the art can understand that various modifications and changes can be made to the present disclosure without departing from the spirit and scope of the present disclosure as set
15 forth in the claims below.

【CLAIMS】**【Claim 1】**

A system for generating a virtual asset wallet address blacklist database based on a graph attention network (GAT), the system comprising:

5 a GAT AI engine server configured to train a GAT-based AI model using an index of a full node for each virtual asset pre-stored in an index database server, common transaction item information pre-stored in a virtual asset transaction analysis database server, and a main blacklist consisting of high-risk virtual asset wallet addresses pre-stored in a main blacklist server, to calculate GAT scores based on the trained AI model, to estimate high-risk virtual asset
10 wallet addresses using the calculated GAT scores, and to generate a GAT blacklist consisting of the estimated high-risk virtual asset wallet addresses; and

 a GAT blacklist database server configured to store the GAT blacklist generated by the GAT AI engine server.

【Claim 2】

15 The system of claim 1, wherein the GAT AI engine server comprises:

 a first data preprocessing module configured to preprocess the common transaction item information stored in the virtual asset transaction analysis database server, to label each common transaction item information according to a type of coin, and to query any transaction corresponding to a predetermined virtual asset wallet address;

20 an AI learning module configured to perform GAT learning using transactions queried by the first data preprocessing module;

 a teacher module configured to perform pseudo-labeling on unlabeled transactions from the first data preprocessing module, and feed the pseudo-labeled transactions to the AI learning module to re-learn the pseudo-labeled transactions; and

25 a first risk calculation module configured to calculate a risk level corresponding to a

GAT score of each virtual asset wallet address based on results of the GAT learning performed by the AI learning module.

【Claim 3】

The system of claim 1, further comprising a high-risk wallet address service server
5 configured to receive a virtual asset wallet address from a virtual asset exchange server, to calculate a risk level for the received virtual asset wallet address, and to respond to the virtual asset exchange server with the calculated risk level.

【Claim 4】

The system of claim 3, wherein the high-risk wallet address service server comprises:
10 a second data preprocessing module configured to receive a virtual asset wallet address from the virtual asset exchange server and perform preprocessing;

a transaction inquiry module configured to query transactions of the virtual asset wallet address, which has undergone preprocessing by the second data preprocessing module 331, from the virtual asset transaction analysis database server;

15 a second risk calculation module configured to calculate a risk level corresponding to a GAT score of the virtual asset wallet address using the transactions queried by the transaction query module; and

a learning update module configured to request the AI learning module to train an artificial intelligence model based on the risk level calculated by the second risk calculation
20 module.

【Claim 5】

The system of claim 4, further comprising a high-risk wallet address management server configured to receive a virtual asset wallet address from an administrator terminal, to calculate a risk level for the received virtual asset wallet address, and to respond to the
25 administrator terminal with the calculated risk level.

【Claim 6】

A method for generating a virtual asset wallet address blacklist database based on a graph attention network (GAT), the method comprising:

training, by a GAT AI engine server, a GAT-based AI model using an index of a full node for each virtual asset pre-stored in an index database server, common transaction item information pre-stored in a virtual asset transaction analysis database server, and a main blacklist consisting of high-risk virtual asset wallet addresses pre-stored in a main blacklist server;

calculating GAT scores based on the trained AI model by the GAT AI engine server;

estimating high-risk virtual asset wallet addresses using the calculated GAT scores and generating a GAT blacklist consisting of the estimated high-risk virtual asset wallet addresses by the GAT AI engine server; and

storing, by a GAT blacklist database server, the GAT blacklists generated by the GAT AI engine server.

【Claim 7】

The method of claim 6, further comprising receiving a virtual asset wallet address from a virtual asset exchange server, calculating a risk level for the received virtual asset wallet address, and responding to the virtual asset exchange server with the calculated risk level by a high-risk wallet address service server.

【Claim 8】

The method of claim 7, further comprising receiving a virtual asset wallet address from an administrator terminal, calculating a risk level for the received virtual asset wallet address, and responding to the administrator terminal with the calculated risk level by a high-risk wallet address management server.