



(12) 发明专利

(10) 授权公告号 CN 113283902 B

(45) 授权公告日 2023.05.09

(21) 申请号 202110654086.X

G06N 3/0464 (2023.01)

(22) 申请日 2021.06.11

G06F 18/241 (2023.01)

(65) 同一申请的已公布的文献号

H04L 9/40 (2022.01)

申请公布号 CN 113283902 A

审查员 王琪

(43) 申请公布日 2021.08.20

(73) 专利权人 浙江工业大学

地址 310014 浙江省杭州市下城区潮王路
18号

(72) 发明人 陈晋音 张敦杰 黄国瀚

(74) 专利代理机构 杭州天勤知识产权代理有限

公司 33224

专利代理师 何秋霞 胡红娟

(51) Int. Cl.

G06Q 20/38 (2012.01)

G06Q 40/04 (2012.01)

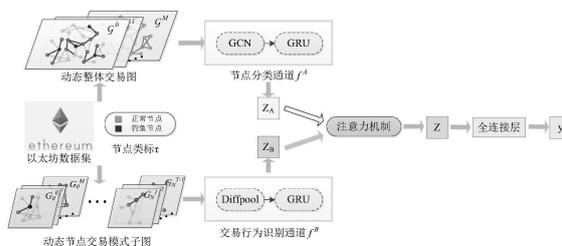
权利要求书3页 说明书8页 附图1页

(54) 发明名称

一种基于图神经网络的多通道区块链钓鱼节点检测方法

(57) 摘要

本发明公开一种基于图神经网络的多通道区块链钓鱼节点检测方法,包括步骤:根据区块链交易信息构建网络序列型动态整体交易网络和动态交易模式子图;获得时间加权交易连边;学习目标节点的用户潜在特征和交易模式特征;获得组合特征输入到检测模块中获得最终检测结果;利用目标节点的真实类标与最终检测结果对检测模型进行训练,获得多通道区块链钓鱼节点检测模型。本发明的钓鱼节点检测方法,采用基于图神经网络的节点分类与节点交易模式识别双通道,可以有效提取用户节点的潜在特征与其交易模式特征,提高区块链钓鱼节点检测模型的有效性,实现对区块链钓鱼节点的检测。



1. 一种基于图神经网络的多通道区块链钓鱼节点检测方法,其特征在于,包括如下步骤:

根据区块链交易信息构建区块链的网络序列型动态整体交易网络图,并针对所述网络序列型动态整体交易网络图中的每个节点,构造目标节点的网络序列型动态交易模式子图;将区块链的所述网络序列型动态整体交易网络图、网络序列型动态交易模式子图和节点真实类标输入到多通道区块链钓鱼节点的检测模型中;

利用可训练的时间重要性系数,将不同网络序列中相同节点对之间的多条交易信息加权求和,作为统一网络序列中唯一的时间加权交易连边;

针对所述网络序列型动态整体交易网络图,利用所述检测模型中动态节点分类通道学习目标节点的用户潜在特征;针对所述网络序列型动态交易模式子图,利用所述检测模型中的动态图分类通道学习目标节点的交易模式特征;

利用注意力机制,为目标节点的所述用户潜在特征和交易模式特征分配不同的注意力系数,得到组合特征;将所述组合特征输入到检测模块中获得目标节点的最终检测结果;

利用目标节点的真实类标与最终检测结果对检测模型进行训练,获得最终的多通道区块链钓鱼节点检测模型;

求取统一网络序列中唯一的时间交易连边的过程具体为:

对于每个交易时间段 $m \in [0, 1, \dots, M]$ 的区块链的网络序列型动态整体交易网络图 G^m 与目标节点的网络序列型动态交易模式子图 G_c^m ,利用可训练的时间重要性系数 $\alpha \in R^M$ 为不同交易时间段间的交易链路分配不同的时间系数,最终合并加权求和得到唯一的时间加权交易连边;

学习目标节点的用户潜在特征和交易模式特征的过程包括:

所述检测模型包括节点分类通道和动态图分类通道,针对不同交易时间段的整体交易网络分片图,采用基于图卷积网络的节点分类通道学习当前时刻目标节点的用户潜在特征;针对不同交易时间段的目标节点交易模式子图,利用基于可微的分层图池化模型的交易模式识别通道提取当前时刻目标节点的交易模式特征;

提取目标节点的用户潜在特征与交易模式特征后,在每个通道中分别采用相同结构的门控循环单元学习整体交易网络与交易模式子图的时间特征,利用网络序列型动态整体交易网络图与交易模式子图对门控循环单元进行训练,采用GRU最后时刻的隐藏层输出作为最终的目标节点的用户潜在特征与目标节点交易模式特征。

2. 根据权利要求1所述的基于图神经网络的多通道区块链钓鱼节点检测方法,其特征在于,构建区块链的网络序列型动态整体交易网络图的过程:

选择一定数量的钓鱼节点与正常节点作为初始节点,根据交易记录提取另一方账户地址作为一阶交易节点,然后分别以各一阶交易节点为中心节点,根据交易记录提取另一方账户地址作为二阶交易节点,得到总节点数为N的时间序列型动态整体交易网络图,表示为 $G_{time} = (V, E, X)$;

其中, $V(|V|=N)$ 表示节点集合, $e \in E$ 表示链路集合, X 表示每条链路对应的交易额 $W(e)$ 与交易时间 $T(e)$ 信息;

随后以该时间序列型动态整体交易网络图中第一笔交易与最后一笔交易的时间间隔

作为总交易时长,将该时长均分为M个交易时间段,并依据划分的时间段对时间序列型动态整体交易网络图进行分片,每个交易时间段的交易信息形成一个网络分片图,每个网络分片图中的节点数与整体交易网络图的节点数相同,最终获得M个网络分片图组成的网络序列型动态整体交易网络图,表示为 $\mathcal{G} = (\mathcal{G}_{time}^0, \mathcal{G}_{time}^1, \dots, \mathcal{G}_{time}^M)$;在同一交易时间段内,节点i与节点j可能存在L条交易链路 $e_{i,j}^l \in E, l = [1, 2, \dots, L]$ 。

3. 根据权利要求1所述的基于图神经网络的多通道区块链钓鱼节点检测方法,其特征在于,

构造目标节点的网络序列型动态交易模式子图的过程包括:

遍历时间序列型动态整体交易网络图中的每个节点,将每个节点作为中心节点,根据交易记录提取其一阶交易节点和二阶交易节点之间的交易信息,得到目标节点的时间序列型交易模式子图 G_{time}^c ,表示为:

$$G_{time} = \{G_{time}^0, G_{time}^1, \dots, G_{time}^N\}, G_{time}^c = (V_c, E_c, X_c);$$

其中, V_c 表示以节点 v_c 作为目标节点构建的交易模式子图中的节点集合, E_c 与 X_c 分别表示 G_{time}^c 中的交易链路集合与每条交易的信息;

随后根据所述时间序列型交易模式子图中第一笔交易与最后一笔交易的时间间隔作为总交易时长,将该时长均分为M个交易时间段,并依据划分的时间段对时间序列型交易模式子图进行分片,每个交易时间段的交易信息形成一个网络分片图,每个网络分片图中的节点数与交易模式子图的节点数相同,最终获得M个网络分片图组成的网络序列型动态交易模式子图集合,表示为 $G = \{G_0, G_1, \dots, G_N\}$,其中第c个网络序列型动态交易模式子图 $G_c = (G_{time}^{c(0)}, G_{time}^{c(1)}, \dots, G_{time}^{c(M)})$ 。

4. 根据权利要求1所述的基于图神经网络的多通道区块链钓鱼节点检测方法,其特征在于,获得最终检测结果的过程具体包括:

所述钓鱼节点检测方法采用以下公式对用户潜在特征与交易模式特征的注意力加权求和,得到组合特征 $Z(\mathcal{G}, G, \Theta^A, \Theta^B)$ 作为钓鱼节点检测模块的输入数据,获得目标节点的最终检测结果,表示为:

$$Z(\mathcal{G}, G, \Theta^A, \Theta^B) = \gamma_A \cdot Z_A + \gamma_B \cdot Z_B$$

其中, $Z_A = f^A(\mathcal{G}, \Theta^A)$ 表示采用模型参数为 Θ^A 的节点分类通道 $f^A(\cdot)$ 对区块链动态整体交易网络图 \mathcal{G} 提取的用户潜在特征, $Z_B = f^B(G, \Theta^B)$ 表示采用模型参数为 Θ^B 的交易模式识别通道 $f^B(\cdot)$ 对动态交易模式子图G提取的交易模式特征, $\gamma = [\gamma_A, \gamma_B]$ 为注意力系数矩阵。

5. 根据权利要求1或4所述的基于图神经网络的多通道区块链钓鱼节点检测方法,其特征在于,在获得组合特征后,钓鱼节点检测模块采用全连接神经网络作为分类器,实现对钓鱼节点的二分类预测。

6. 根据权利要求1所述的基于图神经网络的多通道区块链钓鱼节点检测方法,其特征在于,获得最终的多通道区块链钓鱼节点检测模型的过程包括:利用目标节点的真实类标

与最终检测结果对所述检测模型进行训练,以区块链钓鱼检测模型的预测置信度与节点类标的交叉熵为损失函数,得到最终的以动态整体交易网络图 \mathcal{G} 和动态交易模式子图 G 为输入的多通道区块链钓鱼节点检测模型 $f(\mathcal{G}, G)$ 。

一种基于图神经网络的多通道区块链钓鱼节点检测方法

技术领域

[0001] 本发明涉及区块链网络安全领域,具体涉及一种基于图神经网络的多通道区块链钓鱼节点检测方法。

背景技术

[0002] 区块链技术起源于比特币,其最初提出的目的是为了解决电子支付中过度依赖可信第三方的问题,实现电子支付的去中心化。区块链技术是一种具有匿名性与不可篡改性分布式账本,并成功应用于金融、政治、商业等领域。随着区块链技术在金融领域中的飞速发展,加密货币的盛行也滋生了大量的网络诈骗犯罪。区块链技术虽然能确保加密货币被用户自身掌握,并提供可靠的交易行为。但在另一方面,由于其去中心化的性质,缺乏第三方机构监管的交易环境令用户陷入网络诈骗者的威胁中。一旦向诈骗者的转账行为被区块链确认,用户失去的电子货币几乎不可能再被寻回。

[0003] 在众多的区块链金融诈骗中,有50%以上的可以归类为以加密货币作为钓鱼目标进行的钓鱼诈骗。(参照文献1:M.Conti,E.S.Kumar,C.Lal,and S.Ruj,“A survey on security and privacy issues of bitcoin,”IEEE Commun.Surv.Tut.,vol.20,no.4,pp.3416-3452,2018.)

[0004] 为了在去中心化的区块链金融系统中为正常用户提供一定的保护,Wu等人(参照文献2:Wu J,Dan Lin,Zibin Zheng and Qi Yuan.Temporal WEighted MultiDiGraph Embedding[J].2019.)研究了区块链生态系统中的反网络钓鱼问题。考虑到金融交易网络中起作用的交易时间和交易额信息,他们将以太坊网络钓鱼地址的识别建模为二分类问题,为识别钓鱼节点,或预测潜在的被诈骗对象提供新的技术手段。

[0005] 图数据由于其强大的表达能力,在现实生活中有着广泛的应用。图神经网络是处理图类型数据的主要方法之一,图神经网络可以巧妙地从业数据中提取特征,而这些被提取出来的特征可以完成许多图数据分析任务,如:节点分类、图分类、链路预测和社区发现等。这些图数据分析任务又被广泛应用于社交网络、推荐系统、电子商务网络等实际场景,并取得了出色的性能。

[0006] 区块链交易系统的公开性为构建区块链交易网络提供了大量的数据基础,这也使得图神经网络成为检测区块链钓鱼诈骗的一种可能的工具。将图神经网络应用于区块链钓鱼诈骗检测的难点在于区块链交易网络往往具有复杂的交易时间与交易额信息,传统的图神经网络难以适用于具有时间信息的多链路交易网络。此外,不同图数据分析任务下的图神经网络关注到的图信息也有所区别,通过单一的节点分类模型难以全面地学习潜在钓鱼诈骗者的特征。因此,针对具有时间信息的多链路区块链交易图数据,如何利用图神经网络更全面地提取用户节点的潜在特征,以便精准识别钓鱼诈骗者有着重要的实践意义。

发明内容

[0007] 本发明目的是设计一种基于图神经网络的多通道区块链钓鱼节点检测方法,利用

可训练的时间重要性系数对不同交易时间段内的多链路交易进行加权求和,采用基于图神经网络的节点分类与节点交易模式识别双通道,分别对区块链动态整体交易网络图与节点交易模式子图进行动态节点分类与动态图分类,最终融合用户节点在不同层次的特征,提高区块链钓鱼节点检测模型的检测有效性。

[0008] 为实现上述目的,本发明采用的技术方案是:

[0009] 一种基于图神经网络的多通道区块链钓鱼节点检测方法,其特征在于,包括如下步骤:

[0010] 根据区块链交易信息构建区块链的网络序列型动态整体交易网络图,并针对所述网络序列型动态整体交易网络图中的每个节点,构造目标节点的网络序列型动态交易模式子图;将区块链的所述网络序列型动态整体交易网络图、网络序列型动态交易模式子图和节点真实类标输入到多通道区块链钓鱼节点的检测模型中;

[0011] 利用可训练的时间重要性系数,将不同网络序列中相同节点对之间的多条交易信息加权求和,作为统一网络序列中唯一的时间加权交易连边;

[0012] 针对所述网络序列型动态整体交易网络图,利用所述检测模型中动态节点分类通道学习目标节点的用户潜在特征;针对所述网络序列型动态交易模式子图,利用所述检测模型中的动态图分类通道学习目标节点的交易模式特征;

[0013] 利用注意力机制,为目标节点的所述用户潜在特征和交易模式特征分配不同的注意力系数,得到组合特征;将所述组合特征输入到检测模块中获得目标节点的最终检测结果;

[0014] 利用目标节点的真实类标与最终检测结果对检测模型进行训练,获得最终的多通道区块链钓鱼节点检测模型。

[0015] 优选地,构建区块链的网络序列型动态整体交易网络图的过程:

[0016] 选择一定数量的钓鱼节点与正常节点作为初始节点,根据交易记录提取另一方账户地址作为一阶交易节点,然后分别以各一阶交易节点为中心节点,根据交易记录提取另一方账户地址作为二阶交易节点,得到总节点数为N的时间序列型动态整体交易网络图,表示为 $G_{time} = (V, E, X)$;

[0017] 其中, $V(|V|=N)$ 表示节点集合, $e \in E$ 表示链路集合, X 表示每条链路对应的交易额 $W(e)$ 与交易时间 $T(e)$ 信息;

[0018] 随后以该时间序列型动态整体交易网络图中第一笔交易与最后一笔交易的时间间隔作为总交易时长,将该时长均分为M个交易时间段,并依据划分的时间段对时间序列型动态整体交易网络图进行分片,每个交易时间段的交易信息形成一个网络分片图,每个网络分片图中的节点数与整体交易网络图的节点数相同,最终获得M个网络分片图组成的网络序列型动态整体交易网络图,表示为 $G = (G_{time}^0, G_{time}^1, \dots, G_{time}^M)$;在同一交易时间段内,节点i与节点j可能存在L条交易链路 $e_{i,j}^l \in E, l = [1, 2, \dots, L]$ 。

[0019] 优选地,构造目标节点的网络序列型动态交易模式子图的过程包括:

[0020] 遍历所述时间序列型动态整体交易网络图中的每个节点,将每个节点作为中心节点,根据交易记录提取其一阶交易节点和二阶交易节点之间的交易信息,得到目标节点的

时间序列型交易模式子图 G_{time}^c ,表示为:

$$[0021] \quad G_{time} = \{G_{time}^0, G_{time}^1, \dots, G_{time}^N\}, G_{time}^c = (V_c, E_c, X_c);$$

[0022] 其中, V_c 表示以节点 v_c 作为目标节点构建的交易模式子图中的节点集合, E_c 与 X_c 分别表示 G_{time}^c 中的交易链路集合与每条交易的信息;

[0023] 随后根据所述时间序列型交易模式子图中第一笔交易与最后一笔交易的时间间隔作为总交易时长,将该时长均分为M个交易时间段,并依据划分的时间段对时间序列型交易模式子图进行分片,每个交易时间段的交易信息形成一个网络分片图,每个网络分片图中的节点数与交易模式子图的节点数相同,最终获得M个网络分片图组成网络序列型动态交易模式子图集合,表示为 $G = \{G_0, G_1, \dots, G_N\}$,其中第c个网络序列型动态交易模式子图

$$G_c = (G_{time}^{c(0)}, G_{time}^{c(1)}, \dots, G_{time}^{c(M)}).$$

[0024] 优选地,求取统一网络序列中唯一的时间交易连边的过程具体为:

[0025] 对于每个交易时间段 $m \in [0, 1, \dots, M]$ 的区块链的网络序列型动态整体交易网络图 G^m 与目标节点的网络序列型动态交易模式子图 G_c^m ,利用可训练的时间重要性系数 $\alpha \in R^M$ 为不同交易时间段间的交易链路分配不同的时间系数,最终合并加权求和得到唯一的时间加权交易连边。

[0026] 优选地,学习目标节点的用户潜在特征和交易模式特征的过程包括:

[0027] 所述检测模型包括节点分类通道和动态图分类通道,针对不同交易时间段的整体交易网络分片图,采用基于图卷积网络(GCN)的节点分类通道学习当前时刻目标节点的用户潜在特征;针对不同交易时间段的的目标节点交易模式子图,利用基于可微的分层图池化模型(Diffpool)的交易模式识别通道提取当前时刻目标节点的交易模式特征;

[0028] 优选地,提取目标节点的用户潜在特征与交易模式特征后,在每个通道中分别采用相同结构的门控循环单元(GRU)学习整体交易网络与交易模式子图的时间特征,利用网络序列型动态整体交易网络图与交易模式子图对门控循环单元进行训练,采用GRU最后时刻的隐藏层输出作为最终的目标节点的用户潜在特征与目标节点交易模式特征。

[0029] 优选地,获得最终检测结果的过程具体包括:

[0030] 所述钓鱼节点检测方法采用以下公式对用户潜在特征与交易模式特征的注意力加权求和,得到组合特征 $Z(G, G, \Theta^A, \Theta^B)$ 作为钓鱼节点检测模块的输入数据,获得目标节点的最终检测结果,表示为:

$$[0031] \quad Z(G, G, \Theta^A, \Theta^B) = \gamma_A \cdot Z_A + \gamma_B \cdot Z_B$$

[0032] 其中, $Z_A = f^A(G, \Theta^A)$ 表示采用模型参数为 Θ^A 的节点分类通道 $f^A(\cdot)$ 对区块链动态整体交易网络图 G 提取的用户潜在特征, $Z_B = f^B(G, \Theta^B)$ 表示采用模型参数为 Θ^B 的交易模式识别通道 $f^B(\cdot)$ 对动态交易模式子图 G 提取的交易模式特征, $\gamma = [\gamma_A, \gamma_B]$ 为注意力系数矩阵。

[0033] 优选地,在获得组合特征后,钓鱼节点检测模块采用全连接神经网络作为分类器,实现对钓鱼节点的二分类预测。

[0034] 获得最终的多通道区块链钓鱼节点检测模型的过程包括：利用目标节点的真实类标与最终检测结果对所述检测模型进行训练，以区块链钓鱼检测模型的预测置信度与节点类标的交叉熵为损失函数，得到最终的以动态整体交易网络图 \mathcal{G} 和动态交易模式子图 G 为输入的多通道区块链钓鱼节点检测模型 $f(\mathcal{G}, G)$ 。

[0035] 与现有技术相比，本发明具有以下有益效果：

[0036] 本发明提供的基于图神经网络的多通道区块链钓鱼节点检测方法，利用可训练的时间重要性系数融合区块链交易网络中多条交易链路的时间与交易额信息。采用基于图神经网络的节点分类与节点交易模式识别双通道，可以有效提取用户节点的潜在特征与其交易模式特征，提高区块链钓鱼节点检测模型的有效性，实现对区块链钓鱼节点的检测。

附图说明

[0037] 图1为本发明基于图神经网络的多通道区块链钓鱼节点检测方法的系统框图。

[0038] 图2为实施例中以太坊整体交易网络图构建过程示意图。

[0039] 图3为实施例中以太坊节点交易模式子图构建过程示意图。

具体实施方式

[0040] 为了使本发明的目的、技术方案及优点更加清楚明白，以下结合实施例，对本发明进行进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本发明，并不用于限定本发明。本领域技术人员在理解本发明的技术方案基础上进行修改或等同替换，而未脱离本发明技术方案的精神和范围，均应涵盖在本发明的保护范围内。

[0041] 图1为本发明提供的基于图神经网络的多通道区块链钓鱼节点检测方法的系统框图。实施例提供的区块链钓鱼节点检测方法包括以下步骤：

[0042] 步骤1，根据区块链交易信息构建区块链的网络序列型动态整体交易网络图，并针对所述网络序列型动态整体交易网络图中的每个节点，构造目标节点的网络序列型动态交易模式子图；将区块链的所述网络序列型动态整体交易网络图、网络序列型动态交易模式子图和节点真实类标输入到多通道区块链钓鱼节点的检测模型中；

[0043] 本实例从公开的以太坊数据查询网站etherscan获取以太坊交易数据。该网站中的每笔交易数据包含数十种属性，仅保留其中的交易节点(from和to)，交易时间(timeStamp)和交易额(value)信息作为构建以太坊交易网络的基础。

[0044] 针对构建网络序列型区块链整体交易网络图，原始以太坊的交易数据过于庞大，且钓鱼节点仅占有所有用户节点的极小部分，本发明选择一定数量的钓鱼节点与正常节点作为初始节点，通过查询原始交易记录中与初始节点的交易记录，构建小规模以太坊交易子网络进行训练。以给定的初始节点作为中心节点，根据交易记录提取另一方账户地址作为一阶交易节点。接着分别以各个一阶交易节点为中心节点，根据交易记录提取另一方账户地址作为二阶交易节点，得到总节点数为 N 的时间序列型的区块链动态整体交易网络图，表示为 $\mathcal{G}_{time} = (V, E, X)$ 。其中， $V(|V|=N)$ 表示节点集合， $e \in E$ 表示链路集合， X 表示每条链路对应的交易额 $W(e)$ 与交易时间 $T(e)$ 信息。

[0045] 随后以该交易网络图中第一笔交易与最后一笔交易的时间间隔作为总交易时长，

将该时长均分为M个交易时间段,并依据划分的时间段对时间序列型区块链交易网络图进行分片。每个交易时间段的交易信息形成一个网络分片图,每个网络分片图中的节点数与整体交易网络图的节点数相同。最终获得M个网络分片图组成的网络序列型区块链动态整体交易网络图,表示为 $\mathcal{G} = (\mathcal{G}_{time}^0, \mathcal{G}_{time}^1, \dots, \mathcal{G}_{time}^M)$ 。在同一交易时间段内,节点i与节点j可能存在L条交易链路 $e_{i,j}^l \in E, l = [1, 2, \dots, L]$ 。在本发明中,以数量比为1:1的钓鱼节点和等量随机选取的正常节点作为中心节点,获取一个具有时间与交易额信息的二阶网络。图2为示例性的给出的一个二阶网络的构建过程。

[0046] 针对构建网络序列型区块链动态交易模式子图,本发明历遍上述时间序列型的区块链动态整体交易网络图中的每个节点,将每个节点作为中心节点,根据交易记录提取其一阶交易节点,以及各个一阶交易节点与其对应的二阶交易节点之间的交易信息,得到时间序列型的动态交易模式子图 $G_{time} =$

$\{G_{time}^0, G_{time}^1, \dots, G_{time}^N\}, G_{time}^c = (V_c, E_c, X_c)$ 。其中, V_c 表示以节点 v_c 作为目标节点构建的交易模式子图中的节点集合, E_c 与 X_c 分别表示 G_{time}^c 中的交易链路集合与每条交易的信息。随后根据构造网络序列型动态整体交易网络图的方法将时间序列型的动态交易模式子图转化为具有M网络分片图组成的网络序列型区的动态交易模式子图 $G = \{G_0, G_1, \dots, G_N\}, G_c = (G_{time}^{c(0)}, G_{time}^{c(1)}, \dots, G_{time}^{c(M)})$ 。图3为示例性的给出的一个二阶交易模式子图的构建过程。

[0047] 步骤2,利用可训练的时间重要性系数,将不同网络序列中相同节点对之间的多条交易信息加权求和,作为统一网络序列中唯一的时间加权交易连边;

[0048] 对于每个交易时间段 $m \in [0, 1, \dots, M]$ 的区块链动态整体交易网络图 \mathcal{G}^m 与目标节点动态交易模式子图 G_c^m ,利用可训练的时间重要性系数 $\alpha \in R^M$ 为不同交易时间段间的交易链路分配不同的时间系数,最终合并得到唯一的时间加权交易连边:

$$[0049] \quad e_{i,j}^{\mathcal{G}^m} = \sum_{e_{i,j}^l \in E_{i,j}^{\mathcal{G}^m}} \alpha_{node}^m \cdot W(e_{i,j}^l) \cdot T(e_{i,j}^l) \quad (1)$$

$$[0050] \quad e_{i,j}^{G_c^m} = \sum_{e_{i,j}^l \in E_{i,j}^{G_c^m}} \alpha_{graph}^m \cdot W(e_{i,j}^l) \cdot T(e_{i,j}^l) \quad (2)$$

[0051] 其中 α_{node} 与 α_{graph} 分别表示用于动态整体交易网络图与目标节点动态交易模式子图的时间重要性系数。 $e_{i,j}^{\mathcal{G}^m}$ 表示节点i与节点j在 \mathcal{G}^m 中唯一的时间加权交易连边, $e_{i,j}^{G_c^m}$ 表示节点i与节点j在 G_c^m 中唯一的时间加权交易连边。 $E_{i,j}^{\mathcal{G}^m}$ 与 $E_{i,j}^{G_c^m}$ 分别表示节点i与节点j在 \mathcal{G}^m 与 G_c^m 中的交易链路集合, $W(e)$ 与 $T(e)$ 表示交易链路e的交易额与交易时间。

[0052] 在对多交易链路进行加权求和后,最终获得M个网络分片图组成的网络序列型区块链动态整体交易网络图 $\mathcal{G} = \{A^0, A^1, \dots, A^M\}$ 与目标节点动态交易模式子图 $G = \{G_0,$

$G_1, \dots, G_N\}$, $G_c = (A_c^0, A_c^1, \dots, A_c^M)$, $c \in [0, 1 \dots, N]$ 。其中, $A^m \in R^{N \times N}$ 表示第 m 个整体交易网络分片图的邻接矩阵, $A_c^m \in R^{|V_c| \times |V_c|}$ 表示目标节点 v_c 的第 m 个交易模式分片子图的邻接矩阵, $A_{i,j}^m$ 表示 A^m 中第 i 个节点与第 j 个节点间的时间加权交易额。

[0053] 步骤3, 针对所述网络序列型动态整体交易网络图, 利用所述检测模型中动态节点分类通道学习目标节点的用户潜在特征; 针对所述网络序列型动态交易模式子图, 利用所述检测模型中的动态图分类通道学习目标节点的交易模式特征;

[0054] 针对用户节点的潜在特征, 如图1所示, 以不同交易时间段内的整体交易网络分片图为输入, 采用基于图卷积网络 (GCN) 的节点分类通道, 通过公式 (3) 学习当前时刻的节点潜在特征:

$$[0055] \quad Z_A^m = f^A(X, A^m) = f(\hat{A}^m \sigma(\hat{A}^m X W_0^A) W_1^A) \quad (3)$$

[0056] 其中, $Z_A^m \in R^{N \times d}$ 为第 m 个网络分片中的节点潜在特征, N 为网络节点个数, d 为网络节点潜在特征的特征维度; f 和 σ 分别为 softmax 函数和 Relu 函数;

$\hat{A}^m = \tilde{D}^m^{-\frac{1}{2}} \tilde{A}^m \tilde{D}^m^{-\frac{1}{2}}$, $\tilde{A}^m = A^m + I_N$ 是添加自连接的第 m 个整体交易网络分片的邻接矩阵, I_N 是网络的自连边矩阵, A^m 是网络的邻接矩阵; X 表示节点属性, 在本发明中, $X = I_N$; $\tilde{D}^m_{ii} = \sum_j \tilde{A}^m_{ij}$ 是 \tilde{A}^m 的度值矩阵 \tilde{D}^m 对角线上的值; $W_0^A \in R^{N \times H_1}$ 、 $W_1^A \in R^{H_1 \times d}$ 分别为节点分类通道的输入层到隐层、隐层到输出层的权重矩阵。

[0057] 针对用户节点的交易模式特征, 如图1所示, 以不同交易时间段内的交易模式子图分片图为输入, 采用基于可微的分层图池化模型 (Diffpool) 的交易模式识别通道, 通过公式 (4) - 公式 (6) 提取当前时刻的目标节点交易模式特征。

$$[0058] \quad \begin{cases} H_c^m = f(\hat{A}_c^m \sigma(\hat{A}_c^m X W_0^B) W_1^B) \\ C_c^m = f(\hat{A}_c^m \sigma(\hat{A}_c^m X W_2^B) W_3^B) \end{cases} \quad (4)$$

$$[0059] \quad \begin{cases} H_{c(pool)}^m = C_c^{mT} H_c^m \in R^{n_0 \times d} \\ A_{c(pool)}^m = C_c^{mT} H_c^m C_c^m \in R^{n_0 \times n_0} \end{cases} \quad (5)$$

$$[0060] \quad Z_{B(c)}^m = f(A_{c(pool)}^m \sigma(A_{c(pool)}^m X W_4^B) W_5^B) \quad (6)$$

[0061] 其中, $Z_{B(c)}^m \in R^{1 \times d}$ 为目标节点 v_c 的第 m 个网络分片中的交易模式特征;

$\hat{A}_c^m = \tilde{D}_c^m^{-\frac{1}{2}} \tilde{A}_c^m \tilde{D}_c^m^{-\frac{1}{2}}$, $\tilde{A}_c^m = A_c^m + I_N$ 是目标节点 v_c 的添加自连接的第 m 个交易模式网络分片的邻接矩阵, A_c^m 是第 m 个交易模式子图的邻接矩阵; $\tilde{D}_c^m_{ii} = \sum_j \tilde{A}_c^m_{ij}$ 是 \tilde{A}_c^m 的度值矩阵 \tilde{D}_c^m 对角线上的值; $W_0^B \in R^{N \times H_1}$ 、 $W_1^B \in R^{H_1 \times d}$ 为 Diffpool 中第一层图卷积层的权重矩阵; $W_2^B \in R^{N \times H_1}$ 、 $W_3^B \in R^{H_1 \times n_0}$ 为 Diffpool 中图池化层的权重矩阵, 在本发明中, $n_0 = 0.1 * |V_c|$, 表明池化图的尺寸; $W_4^B \in R^{n_0 \times H_1}$ 、 $W_5^B \in R^{H_1 \times d}$ 为 Diffpool 中输出图卷

积层的权重矩阵。

[0062] 利用上述步骤计算得到目标节点 v_c 的第 m 个网络分片中的交易模式特征 $Z_{B(c)}^m$ 后,将所有节点的交易模式特征拼接成最终的节点交易模式特征 $Z_B^m \in R^{N \times d}$ 。

[0063] 步骤4,利用注意力机制,为目标节点的所述用户潜在特征和交易模式特征分配不同的注意力系数,得到组合特征;将所述组合特征输入到检测模块中获得目标节点的最终检测结果;

[0064] 所述检测模型包括节点分类通道和动态图分类通道,针对不同交易时间段的整体交易网络分片图,采用基于图卷积网络(GCN)的节点分类通道学习当前时刻目标节点的用户潜在特征;针对不同交易时间段的目标节点交易模式子图,利用基于可微的分层图池化模型(Diffpool)的交易模式识别通道提取当前时刻目标节点的交易模式特征;

[0065] 提取目标节点的用户潜在特征与交易模式特征后,在每个通道中分别采用相同结构的门控循环单元(GRU)学习整体交易网络与交易模式子图的时间特征,利用网络序列型动态整体交易网络图与交易模式子图对门控循环单元进行训练,采用GRU最后时刻的隐藏层输出作为最终的目标节点的用户潜在特征与目标节点交易模式特征。

[0066] GRU结构如公式(7)所示;

$$[0067] \quad \begin{cases} z^t = \sigma(W_z x_t + U_z h^{t-1}) \\ r^t = \sigma(W_r x_t + U_r h^{t-1}) \\ \tilde{h}^t = \tanh(W x_t + U(r^t \odot h^{t-1})) \\ h^t = (1 - z^t) \odot h^{t-1} + z^t \odot \tilde{h}^t \end{cases} \quad (7)$$

[0068] 其中, W_z 、 U_z 、 W_r 、 U_r 、 W 和 U 是可训练的权重, \tilde{h}^t 表示计算中的中间量, σ 是sigmoid激活函数。GRU以前一时刻的细胞状态 h^{t-1} 和当前时刻的输入 x_t 作为输入,输出当前的细胞状态 h^t ,并作为当前时刻的输出。将GRU的每次细胞状态计算过程简化表示为公式(8):

$$[0069] \quad h^t = \text{GRU}(h^{t-1}, x^t) \quad (8)$$

[0070] 利用GRU学习节点潜在特征或交易模式特征的过程统一表示为:

$$[0071] \quad \begin{cases} Z_i^m = \text{GRU}(Z_i^{m-1}, A^m(i, :)), m = \{0, \dots, M\} \\ Z_i = [Z_i^0, \dots, Z_i^M] \end{cases} \quad (9)$$

[0072] 其中, Z_i^m 是节点 v_i 在第 m 时刻的GRU细胞状态;时, $Z_i^0 = A^0(i, :)$ 或 $Z_i^0 = A_i^0$,表示第0个整体交易网络分片图的邻接矩阵的第 i 行,或节点 v_i 的交易模式子图的邻接矩阵。节点分类通道 $f_i^A(\cdot)$ 中,GCN与GRU的参数记为 Θ^A ;交易模式识别通道 $f_i^B(\cdot)$ 中,Diffpool与GRU的参数记为 Θ^B 。以 $Z_i \in R^{1 \times d}$ 统一表示 Z_A 或 Z_B 的第 i 行,表明所有时刻的细胞状态的拼接向量,表示节点 v_i 的节点潜在特征或节点交易模式特征。

[0073] 步骤5,采用公式(10)对节点用户潜在特征与节点交易模式特征的注意力加权求和,得到组合特征 $Z(G, G, \Theta^A, \Theta^B)$,将得到组合特征作为钓鱼节点检测模块的输入数据;

$$[0074] \quad Z(G, G, \Theta^A, \Theta^B) = \gamma_A \cdot Z_A + \gamma_B \cdot Z_B \quad (10)$$

[0075] 其中, $\gamma = [\gamma_A, \gamma_B]$ 为注意力系数矩阵, $Z(G, G, \Theta^A, \Theta^B)$ 表示节点融合特征。

[0076] 步骤6, 实施例, 在获得组合特征后, 钓鱼节点检测模块采用全连接神经网络作为二分类器, 实现对钓鱼节点的二分类预测; 利用目标节点的真实类标与最终检测结果对所述检测模型进行训练, 以区块链钓鱼检测模型的预测置信度与节点类标的交叉熵为损失函数, 给定输入的融合特征 $z_i \in Z(G, G, \Theta^A, \Theta^B)$, 神经网络分类器中的隐藏层与输出层分别为:

$$[0077] \quad y_i^{(1)} = \sigma(W^{(1)}z_i + b^{(1)}) \quad (11)$$

$$[0078] \quad y_i^{(m)} = \sigma(W^{(m)}y_i^{(m-1)} + b^{(m)}), \quad m = 2, \dots, M-1 \quad (12)$$

$$[0079] \quad y_i^{(M)} = \text{sigmoid}(W^{(M)}y_i^{(M-1)} + b^{(M)}) \quad (13)$$

[0080] 其中, M 是神经网络隐藏层数量, $W^{(m)}$ 和 $b^{(m)}$ 分别是第 m 层的权重和偏置, $y_i^{(m)}$ 是节点 v_i 的第 m 层的隐藏层输出。神经网络的输出层只有一个神经元, $y_i^{(M)} \in R^{1 \times 2}$ 为节点 v_i 的最终分类结果。

[0081] 步骤7, 分类器的参数记为 Θ^{NN} , 在每轮训练中, 以公式 (14) 作为目标函数, 通过随机梯度下降法优化 Θ^{NN} , 训练所述的多通道区块链钓鱼检测模型;

$$[0082] \quad \min_{\Theta^{NN}} L = - \sum_{l=1}^{|T_s|} \sum_{k=1}^{|F|} Y_{lk} \ln(y_{lk}^{(M)}) \quad (14)$$

[0083] 其中, T_s 为带类标的区块链交易网络节点集合; $F = [\tau_1, \dots, \tau_N]$ 表示网络节点类标集合, 节点 v_i 的类标 $\tau_i \in [0, 1]$, $\tau_i = 1$ 表示节点 v_i 为钓鱼节点; 若网络节点 v_i 属于类 τ_k , 则 $Y_{ik} = 1$, 否则 $Y_{ik} = 0$; $y_{lk}^{(M)}$ 是由公式 (13) 计算出的分类置信度输出。

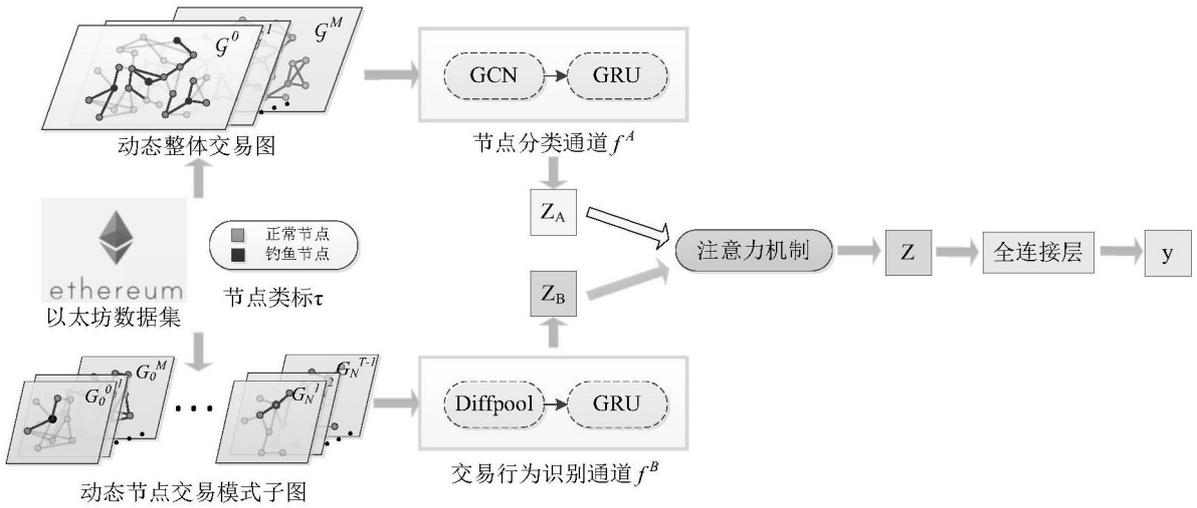


图1

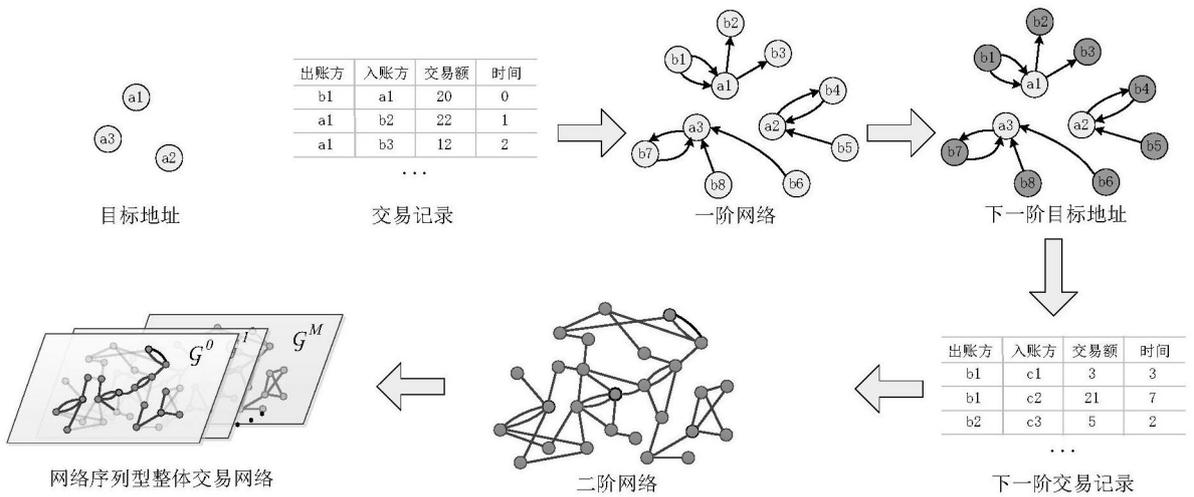


图2

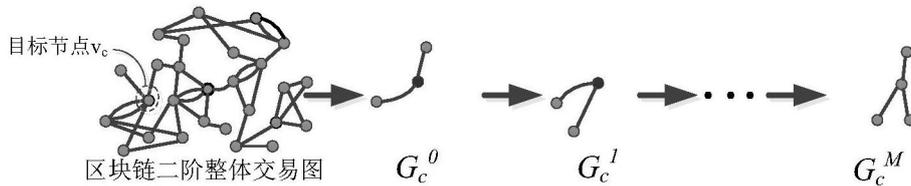


图3