

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2024 年 8 月 15 日 (15.08.2024)



(10) 国际公布号
WO 2024/164828 A1

- (51) 国际专利分类号:
H04L 41/06 (2022.01)
- (21) 国际申请号: PCT/CN2024/073409
- (22) 国际申请日: 2024 年 1 月 22 日 (22.01.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202310129577.1 2023年2月7日 (07.02.2023) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 施大年 (SHI, Danian); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。谢于明 (XIE, Yuming); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。张亮 (ZHANG, Liang); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。王俊 (WANG, Jun); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。侯延祥 (HOU, Yanxiang); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 北京三高永信知识产权代理有限责任公司 (BEIJING SAN GAO YONG XIN INTELLECTUAL

PROPERTY AGENCY CO., LTD.); 中国北京市海淀区上地信息产业基地三街1号楼四层C段457, Beijing 100085 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

(54) Title: DATA STREAM IDENTIFICATION METHOD, APPARATUS AND ELECTRONIC DEVICE

(54) 发明名称: 数据流识别方法及装置、电子设备

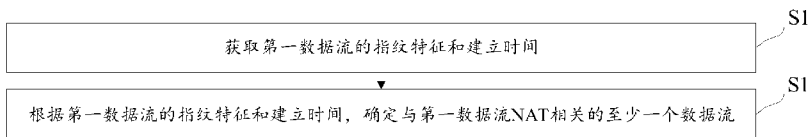


图 4

- S11 Acquire a fingerprint feature and an establishment time of a first data stream
S12 According to the fingerprint feature and the establishment time of the first data stream, determine at least one data stream related to a first data stream NAT

(57) Abstract: Disclosed are a data stream identification method, an apparatus, and an electronic device. The method comprises: acquiring a fingerprint feature and an establishment time of a first data stream; and, according to the fingerprint feature and the establishment time of the first data stream, determining at least one data stream related to a first data stream NAT.

(57) 摘要: 公开了一种数据流识别方法及装置、电子设备。该方法包括: 获取第一数据流的指纹特征和建立时间; 根据所述第一数据流的指纹特征和建立时间, 确定与所述第一数据流NAT相关的至少一个数据流。

WO 2024/164828 A1

数据流识别方法及装置、电子设备

本申请要求于 2023 年 02 月 07 日提交的申请号 202310129577.1、申请名称为“数据流识别方法及装置、电子设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及网络技术领域，特别涉及一种数据流识别方法及装置、电子设备。

背景技术

网络地址转换（network address translation, NAT）的作用是进行内部私有网络地址和公网互联网协议（internet protocol, IP）地址的转换。网络故障分析器通常基于数据流的五元组检测数据流在传输过程中的传输质量，以进行故障诊断。如果一个数据流的传输过程中涉及 NAT 设备，则 NAT 设备转换前后的数据流的五元组将发生变化，这将导致网络故障分析器无法准确精细的定界故障范围。例如，网络故障分析器仅能将 NAT 设备到目的端或源端到 NAT 设备之间的范围确定为故障范围。

相关技术可以读取开启了 NAT 功能的设备（后文简称 NAT 设备）的 NAT 会话表信息，根据 NAT 会话表信息获得 NAT 转换前后两条数据流的对应关系，根据数据流的对应关系确定一条数据流属于同一 NAT 会话的数据流，进而可以在进行故障诊断时同时考虑属于同一 NAT 会话的数据流。

但是，不同厂商和型号的 NAT 设备的 NAT 会话表的导出方式不一，部署上述方案需要逐个 NAT 设备对接，甚至部分 NAT 设备没有现成的管理接口来获取 NAT 会话表，实现难度及代价较大；并且，需要预先获取网络中的 NAT 设备的部署信息，从而能够从 NAT 设备读取 NAT 会话表信息，并且在 NAT 设备的部署变化时需要同步更新部署信息，部署难度较大。

发明内容

本申请提供了一种数据流识别方法及装置、电子设备，能够降低确定与第一数据流 NAT 相关的数据流的实现难度及代价，降低网络部署难度。

第一方面，本申请提供了一种数据流识别方法。该方法包括：获取第一数据流的指纹特征和建立时间；根据第一数据流的指纹特征和建立时间，确定与第一数据流 NAT 相关的至少一个数据流。

在该实现方式中，通过第一数据流的指纹特征确定与第一数据流 NAT 相关的至少一个数据流，不需要从 NAT 设备读取会话表，实现难度及代价小，并且网络部署难度小。

由于实际网络中可以是一级 NAT，也可能存在多级 NAT。一级 NAT 时会产生一个数据流与第一数据流 NAT 相关，多级 NAT 时会产生多个数据流与第一数据流 NAT 相关。因此，可以确定出一个或多个与第一数据流 NAT 相关的数据流。

在本申请的实现方式中，与第一数据流 NAT 相关可以是与第一数据流属于同一个 NAT 会话。或者，与第一数据流 NAT 相关还可以是与第一数据流经过 N 级 NAT 转换后的数据流属于同一个 NAT 会话。

该方法可以由网络中的分析平台、网络设备、采集器、存储平台等执行，网络中的节点上部署有采集器（也即探针），通过采集器采集数据流的信息，数据流的信息包括指纹特征。以分析平台为例，分析平台通过采集器采集到的数据流的指纹特征执行数据流识别方法。

在一些可能的实现方式中，采集器通常采用软件实现。在其他实现方式中，采集器也可以采用硬件实现。

为了保证采集到的指纹特征能够用来识别数据流是否 NAT 相关，可以在启用 NAT 功能的节点的上下游均部署采集器。可选地，也可以采用其他方式部署采集器。作为示例，下方列出了三种采集器的部署方式。

第一种，在每个节点上均部署采集器，从而保证能够采集到每个 NAT 会话的数据流的指纹特征。

第二种，在网络的多个核心节点部署采集器，例如多个 Spine 交换机、多个边界节点、和/或防火墙上部署采集器，从而能够尽量采集到每个 NAT 会话的数据流的指纹特征。

第三种，在具有 NAT 功能的节点上部署采集器，以采集到每个 NAT 会话的数据流的指纹特征。

在第一种和第二种部署方式中，采集器通常部署在节点的接口的入方向，当然也可以部署在接口的出方向。在第三种部署方式中，采集器需要同时在接口的入方向和接口的出方向部署。

在一些可能的实现方式中，部署采集器（软件）的各个节点或者采集器（硬件）自己存储各自采集到的数据流的信息。

在另一些可能的实现方式中，该网络还可以包括存储平台，用于集中存储各个采集器采集到的数据流的信息。

以分析平台为例，在数据流的信息采集完成后，分析平台通过与采集器或者存储平台交互，从而获取数据流的信息并进一步确定与第一数据流 NAT 相关的至少一个数据流。

以存储平台为例，在数据流的信息采集完成后，存储平台通过本地存储的数据流的信息确定与第一数据流 NAT 相关的至少一个数据流。

下面分别对这两种情况进行说明：

第一种，分析平台通过与采集器或者存储平台交互实现该方法的情况：

在这种实施方式中，获取第一数据流的指纹特征和建立时间，包括：

向第一网络节点发送第一查询请求，第一查询请求包括第一数据流的第一流标识；

接收第一网络节点发送的第一查询应答，第一查询应答包括第一数据流的指纹特征和建立时间。

通过从第一网络节点请求第一数据流的指纹特征和建立时间，为后续确定与第一数据流 NAT 相关的至少一个数据流做准备。

这里的第一网络节点可以是采集器、部署有采集器的网络设备或者存储平台。并且，在发送查询请求时，如果是发送给采集器或者部署有采集器的网络设备，可以同时发送给多个采集器或者部署有采集器的网络设备。

示例性地，数据流的第一流标识为数据流的五元组，包括源互联网协议（internet protocol, IP）地址、目的 IP 地址、源端口、目的端口和协议类型，其中协议类型为传输控制协议（transmission control protocol, TCP）。

在上述过程中，第一网络节点接收到查询请求后，采用第一流标识作为索引进行查询，得到第一数据流的指纹特征和建立时间。

由于一条数据流是双向传输的，因此查询时除了按照原本的五元组进行查询外，还可以将源、目的 IP 调换，将源、目的端口调换后进行查询。

可选地，第一查询请求除了包括第一流表示外，还包括查询时间范围，查询时间范围用于限制查询到的数据流的建立时间所处的时间范围。

例如，查询时间范围采用确切的范围表示，如某一天的 0 点至 24 点，或者某一天某个小时的 0 分至 60 分等。再例如，查询时间范围采用时间粒度表示，如一天，则对应的范围是当天，或者一个小时，则对应的范围是当前小时内等。

当然，在第一查询请求中不包括查询时间范围时，查询时间范围可以采用默认值，例如查询当天等。

在一些可能的实现方式中，数据流的指纹特征可以包括一个或多个特征。

示例性地，指纹特征包括如下特征中的至少一项：

首个数据包的互联网协议标识（IP identifier, IPID）、首个数据包的载荷、首个数据包的载荷的哈希值、TCP 三次握手过程中同步序列编号（synchronize sequence numbers, SYN）报文的 IPID、TCP 三次握手过程中 SYN 报文的初始序列码（initial sequence number, ISN）、TCP 三次握手过程中 SYN-ACK 报文的 ISN。

其中，首个数据包（简称首包）是指 TCP 三次握手建立连接后传输的第一个数据包。不同数据流需要传输的内容不同，有的数据流需要多个数据包才能完成内容的传输，而有的数据流仅需要传输一个数据包就可以完成内容的传输。因此，本申请实施例采用首个数据包的信息作为指纹特征，可以为每个需要判断的数据流采集到指纹特征。另一方面，首个数据包即是 TCP 建链后第一个被传输的数据包，而且其序列号等于 SYN 报文携带的 ISN+1，采集器可以很容易地确定首个数据包，因此，采集首个数据包所需的资源和

时间相比后续数据包更少。

采集器在采集时，可以根据数据包的序列号（sequence）来判断是否是首个数据包，通常首个数据包的序列号是 SYN 报文的 ISN+1。而第二个数据包的序列号是 SYN 报文的 ISN+1 再加首个数据包的长度，因此，如果要采集首个数据包之后的数据包的特征，则存在判断数据包的序列号更加困难，占用资源和时间更多的问题。所以优先选用首个数据包进行指纹特征采集。

首个数据包的载荷可以是首个数据包的全部载荷，也可以是首个数据包的部分载荷，例如是从数据包包头开始截取的部分载荷，如 500 或 1000 字节长度的载荷。相应地，首个数据包的哈希值也可以是全部载荷的哈希值，也可以是部分载荷的哈希值，这里不再赘述。

在该实现方式中，各个采集器采集相同的指纹特征。申请人通过理论分析和大量的试验发现，上述所列举的指纹特征在 NAT 转换前后不会变化，并且不同数据流难以发生冲突，即，NAT 转换前后的两个数据流尽管 IP 地址或端口可能不同，但两个数据流中的上述指纹特征相同，因此，上述指纹特征能够标识 NAT 相关的数据流。另外，上述所列举的指纹特征采集过程较为容易实现，占用资源少，有利于实现本申请实施例提供的方法。

在其他可能的实现方式中，数据流的指纹特征还可以包括其他特征，只要能够在 NAT 转换前后不变，并且不同数据流难以发生冲突即可，本申请对此不做限制。

在一些可能的实现方式中，采集器采集的数据流的创建时间可以是接收或发送某个数据包或者报文的时间。

采集器在采集数据流的创建时间时，可以采用如下时间中的一种作为数据流的创建时间：

SYN 报文接收时间、SYN-ACK 报文接收时间、ACK 报文接收时间。此处的 ACK 报文为 TCP 三次握手手中的 ACK 报文。

当然上面针对的是接口的入方向的采集器，如果是接口的出方向的采集器，则采用如下时间中的一种作为数据流的创建时间：

SYN 报文发送时间、SYN-ACK 报文发送时间、ACK 报文发送时间。

在其他可能的实现方式中，采集器在采集数据流的创建时间时，也可以采用其他时间，例如首个数据包的接收或发送时间等。

在这种实施方式中，根据第一数据流的指纹特征和建立时间，确定与第一数据流 NAT 相关的至少一个数据流的方式包括多种：

在一种可能的实现方式中，分析平台利用第一数据流的指纹特征，查询指纹特征相同的第二数据流，然后根据二者的创建时间判断是否 NAT 相关，步骤如下：

向第二网络节点发送第二查询请求，第二查询请求包括第一数据流的指纹特征；

接收第二网络节点发送的第二查询应答，第二查询应答包括与第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；

在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。

在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值不小于阈值的情况下，确定第一数据流和第二数据流不 NAT 相关。

其中，第二网络节点和第一网络节点可以是相同的网络节点，也可以是不同的网络节点。

第一网络节点和第二网络节点可以是分布式的多个采集器或部署有采集器的网络设备。

在该实现方式中，通过第二次查询指纹特征相同的第二数据流的创建时间。然后以指纹特征相同、创建时间相近作为判断标准，判断 NAT 相关的数据流。由于 NAT 转换以及数据流在不同采集器之间传输存在时间延迟，因此，NAT 转换前采集到的数据流的建立时间和 NAT 转换后采集到的数据流的建立时间有时间延迟，通过限定两个数据流的建立时间的时间差的绝对值小于阈值，能够确定两个相同指纹特征的数据流是经过 NAT 转换的两个数据流。

不同网络节点上的采集器采集时存在同步精度问题，不同的采集器之间还会存在传输时延，因此，可以基于时间同步精度和/或传输时延确定上述阈值。

在另一种可能的实现方式中，第一查询应答还包括与第一数据流的指纹特征相同的第二数据流的第一

流标识以及建立时间；此时，分析平台直接根据两个数据流的创建时间判断是否 NAT 相关，步骤如下：

在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。

在这种情况下，分析平台发送第一查询请求时，第一网络节点除了能够将第一数据流的信息反馈给分析平台外，还可以同时查询与第一数据流指纹相同的第二数据流的信息，一起反馈给分析平台。这样，分析平台可以直接根据第一查询应答中的内容，确定两个数据流是否 NAT 相关。

在上述几种确定与第一数据流 NAT 相关的至少一个数据流的方式中，第二数据流可以是一个，也可以是多个，如果是多个，则分别确定每一个第二数据流是否与第一数据流 NAT 相关。

这种情况下，第一网络设备可以为存储平台。

第二种，存储平台基于本地存储的采集信息实现该方法的情况：

在这种实施方式中，获取第一数据流的指纹特征和建立时间，包括：

接收第一查询请求，第一查询请求包括第一数据流的第一流标识；根据第一数据流的第一流标识确定第一数据流的指纹特征和建立时间。

在这种情况下，方法由存储平台执行，存储平台可以根据查询请求中的流标识查询第一数据流的指纹特征和建立时间。

在这种实施方式中，根据第一数据流的指纹特征和建立时间，确定与第一数据流 NAT 相关的至少一个数据流的步骤包括：

根据第一数据流的指纹特征确定第二数据流，第二数据流的指纹特征与第一数据流的指纹特征相同；

在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。

其中，第二数据流的建立时间可以是查询第一数据流的指纹特征找到相同指纹特征的第二数据流，然后得到的。

在本申请的实现方式中，由于指纹特征采用了冲突可能性小的特征来实现，但为了保证确定 NAT 相关的数据流的准确性，在上述过程中，还可以将指纹冲突的情况排除掉，指纹冲突是指非 NAT 相关的数据流的指纹特征相同。

在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关，包括：

在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值，第一数据流和第二数据流不存在指纹冲突的情况下，确定第一数据流和第二数据流 NAT 相关。

在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值，第一数据流和第二数据流存在指纹冲突的情况下，确定第一数据流和第二数据流非 NAT 相关，也即存在指纹冲突。

示例性地，确定第一数据流和第二数据流指纹冲突，可以通过如下步骤：

确定与第二数据流具有相同第二流标识的至少一条第三数据流；

确定与每条第三数据流具有相同指纹特征的至少一条第四数据流；

在至少一条第四数据流中存在一条第四数据流的建立时间和对应第三数据流的建立时间的绝对值小于阈值，且这一条第四数据流的第二流标识和第一数据流的第二流标识相同的情况下，确定第一数据流和第二数据流不存在指纹冲突。

其中，数据流的第二流标识可以是四元组，包括源 IP 地址、目的 IP 地址、目的端口和协议类型。

在上述确定指纹冲突的过程中，使用的原理是查询与该数据流具有相同四元组的数据流在历史记录中是否进行过 NAT 转换，如果存在具有相同四元组的数据流经过 NAT 转换，则说明本次也是 NAT 转换，而非指纹冲突。

实施时，先确定与第二数据流的第二流标识相同的第三数据流，然后查询与第三数据流指纹特征相同且建立时间差小于阈值的第四数据流，如果第四数据流与第一数据流的第二流标识相同，则说明是 NAT 转换，而非指纹冲突。其中，第二流标识相同说明，两条数据流是从相同的源地址出发去往相同的目的地，并且在 NAT 转换的同一侧（转换前或后）。如果第三数据流和第四数据流存在 NAT 转换，那么分别与之对应的第二数据流和第一数据流也存在 NAT 转换。

在本申请的实现方式中，第一数据流可以是用户选定的数据流，用户在观测到第一数据流出现故障时，确定与第一数据流 NAT 相关的数据流。当然，在其他实现方式中，第一数据流也可以是诊断设备自动选取的，诊断设备检测到第一数据流的故障，然后确定与第一数据流 NAT 相关的数据流。

在确定存在与第一数据流 NAT 相关的数据流之后，该方法还包括：

在第一数据流发生故障的情况下，对第一数据流以及与第一数据流 NAT 相关的至少一个数据流进行故障诊断。

通过对 NAT 相关数据流同时进行故障诊断，能够保证故障诊断的准确性。

在确定不存在与第一数据流 NAT 相关的数据流之后，该方法还包括：

在第一数据流发生故障的情况下，对第一数据流进行故障诊断。

其中，数据流发生故障是指基于该数据流的应用发生故障，包括但不限于数据流丢包率高、延迟大等情况。

第二方面，本申请提供了一种数据流识别方法。该方法包括：接收查询请求，查询请求包括第一数据流的第一流标识；发送查询应答，查询应答包括第一数据流的指纹特征和建立时间，第一数据流的指纹特征和建立时间用于确定与第一数据流 NAT 相关的至少一个数据流。

第二方面提供的方法与第一方面提供的方法的区别在于：第二方面提供的方法仅进行数据流的指纹特征的查询和反馈，而不进行与第一数据流 NAT 相关的至少一个数据流的判断。

执行该方法的可以是采集器、部署有采集器的网络设备或者存储平台。

在一些可能的实现方式中，查询应答还包括与第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间。

可选地，该方法还包括：

采集各个数据流的建立时间、第一流标识和指纹特征；

存储各个数据流的建立时间、第一流标识和指纹特征。

在一些可能的实现方式中，采集器以旁路的方式接入。此时，采集各个数据流的建立时间、第一流标识和指纹特征，包括：

通过镜像功能抓取数据流的报文；

分析抓取到的数据流的报文，以获取数据流的建立时间、第一流标识和指纹特征。

在本申请的实现方式中，当指纹特征采用 TCP 三次握手中的 SYN 报文的特征时，通过镜像功能抓取数据流的报文，包括：

通过配置访问控制列表（access control lists, ACL）抓取 SYN 报文和/或 SYN-ACK 报文。

通过配置 ACL 实现 SYN 报文和/或 SYN-ACK 报文的特征的采集，可以减少需要抓取的报文数量，降低采集负担。

在另一些可能的实现方式中，采集器以随路的方式接入。此时，直接分析数据流的报文，获取数据流的建立时间、第一流标识和指纹特征即可。

第三方面，提供了一种数据流识别装置，该装置包括：

获取单元，用于获取第一数据流的指纹特征和建立时间；

确定单元，用于根据第一数据流的指纹特征和建立时间，确定与第一数据流 NAT 相关的至少一个数据流。

可选地，获取单元，用于向第一网络节点发送第一查询请求，第一查询请求包括第一数据流的第一流标识；接收第一网络节点发送的第一查询应答，第一查询应答包括第一数据流的指纹特征和建立时间。

可选地，确定单元，用于向第二网络节点发送第二查询请求，第二查询请求包括第一数据流的指纹特征；接收第二网络节点发送的第二查询应答，第二查询应答包括与第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。

可选地，第一查询应答还包括与第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；

确定单元，用于在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。

可选地，获取单元，用于接收第一查询请求，第一查询请求包括第一数据流的第一流标识；根据第一数据流的第一流标识确定第一数据流的指纹特征和建立时间。

可选地，确定单元，用于根据第一数据流的指纹特征确定第二数据流，第二数据流的指纹特征与第一数据流的指纹特征相同；在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。

可选地，指纹特征包括如下特征中的至少一项：

首个数据包的互联网协议标识 IPID、首个数据包的载荷、首个数据包的载荷的哈希值、传输控制协议 TCP 三次握手过程中 SYN 报文的 IPID、TCP 三次握手过程中 SYN 报文的初始序列码 ISN、TCP 三次握手过程中 SYN-ACK 报文的 ISN。

可选地，第一查询请求还包括查询时间范围，查询时间范围用于限制查询到的数据流的建立时间所处的时间范围。

可选地，确定单元，用于在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值，第一数据流和第二数据流不存在指纹冲突的情况下，确定第一数据流和第二数据流 NAT 相关，指纹冲突是指非 NAT 相关的数据流的指纹特征相同。

可选地，确定单元，还用于确定与第二数据流具有相同第二流标识的至少一条第三数据流；确定与每条第三数据流具有相同指纹特征的至少一条第四数据流；在至少一条第四数据流中存在一条第四数据流的建立时间和对应第三数据流的建立时间的时间差的绝对值小于阈值，且这一条第四数据流的第二流标识和第一数据流的第二流标识相同的情况下，确定第一数据流和第二数据流不存在指纹冲突。

可选地，该装置还包括：

诊断单元，用于在第一数据流发生故障的情况下，对第一数据流以及与第一数据流 NAT 相关的至少一个数据流进行故障诊断。

第四方面，提供了一种数据流识别装置，装置包括：

接收单元，用于接收查询请求，查询请求包括第一数据流的第一流标识；

发送单元，用于发送查询应答，查询应答包括第一数据流的指纹特征和建立时间，第一数据流的指纹特征和建立时间用于确定与第一数据流 NAT 相关的至少一个数据流。

可选地，查询应答还包括与第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间。

可选地，该装置还包括：

采集单元，用于采集各个数据流的建立时间、第一流标识和指纹特征；

存储单元，用于存储各个数据流的建立时间、第一流标识和指纹特征。

可选地，采集单元，用于通过镜像功能抓取数据流的报文；分析抓取到的数据流的报文，以获取数据流的建立时间、第一流标识和指纹特征。

第五方面，提供了一种电子设备。所述电子设备包括处理器和存储器。所述存储器用于存储软件程序以及模块。所述处理器通过运行或执行存储在所述存储器内的软件程序和/或模块实现上述第一方面或第一方面的任一种可能的实施方式中的方法，或者实现上述第二方面或第二方面的任一种可能的实施方式中的方法。

可选地，所述处理器为一个或多个，所述存储器为一个或多个。

可选地，所述存储器可以与所述处理器集成在一起，或者所述存储器与处理器分离设置。

在具体实现过程中，存储器可以为非瞬时性(non-transitory)存储器，例如只读存储器(read only memory, ROM)，其可以与处理器集成在同一块芯片上，也可以分别设置在不同的芯片上，本申请对存储器的类型以及存储器与处理器的设置方式不做限定。

第六方面，提供了一种数据流识别系统，所述数据流识别系统包括采集器和分析平台，所述分析平台用于执行上述第一方面或第一方面的任一种可能的实施方式中的方法，所述采集器用于执行上述第二方面或第二方面的任一种可能的实施方式中的方法。

可选地，所述数据流识别系统还包括存储平台，用于存储所述采集器采集的数据流的信息。

第七方面，提供了一种计算机程序产品。所述计算机程序产品包括计算机程序代码，当所述计算机程序代码被计算机运行时，使得所述计算机执行上述第一方面或第一方面的任一种可能的实施方式中的方法，

或者执行上述第二方面或第二方面的任一种可能的实施方式中的方法。

第八方面，本申请提供了一种计算机可读存储介质，所述计算机可读存储介质用于存储处理器所执行的程序代码，所述程序代码包括用于实现上述第一方面任一种可能的实施方式中的方法，或者实现上述第二方面或第二方面的任一种可能的实施方式中的方法。

第九方面，提供了一种芯片，包括处理器，处理器用于从存储器中调用并运行所述存储器中存储的指令，使得安装有该芯片的通信设备执行上述第一方面任一种可能的实施方式中的方法，或者执行上述第二方面或第二方面的任一种可能的实施方式中的方法。

第十方面，提供另一种芯片。所述另一种芯片包括输入接口、输出接口、处理器和存储器。所述输入接口、输出接口、所述处理器以及所述存储器之间通过内部连接通路相连。所述处理器用于执行所述存储器中的代码，当所述代码被执行时，所述处理器用于执行上述第一方面任一种可能的实施方式中的方法，或者执行第二方面或第二方面的任一种可能的实施方式中的方法。

附图说明

- 图 1 是本申请实施例提供的一种应用场景的结构示意图；
- 图 2 是本申请实施例提供的另一种应用场景的结构示意图；
- 图 3 是本申请实施例提供的一种数据中心的网络拓扑示意图；
- 图 4 是本申请实施例提供的一种数据流识别方法的流程图；
- 图 5 是本申请实施例提供的一种数据流识别方法的流程图；
- 图 6 是本申请实施例提供的一种数据流信息的采集方法的流程图；
- 图 7 是本申请实施例提供的一种数据流识别方法的流程图；
- 图 8 是本申请实施例提供的一种数据流识别方法的流程图；
- 图 9 是本申请实施例提供的一种数据流识别方法的流程图；
- 图 10 是本申请实施例提供的一种数据流识别装置的框图；
- 图 11 是本申请实施例提供的一种数据流识别装置的框图；
- 图 12 示出了本申请实施例提供的一种电子设备的结构示意图。

具体实施方式

为使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请实施方式作进一步地详细描述。

图 1 是本申请实施例提供的一种应用场景的结构示意图。参见图 1，该应用场景包括多个网络设备 11，多个网络设备 11 构成网络，例如数据中心网络。网络设备 11 可以是交换机、路由设备、防火墙设备、服务器等。

多个网络设备 11 中的至少部分网络设备 11 上部署并启用 NAT 功能。如图 1 所示，多个网络设备 11 中的至少部分网络设备 11 上部署有采集器 12，用于采集数据流的信息。

图 2 是本申请实施例提供的另一种应用场景的结构示意图。参见图 2，该应用场景与图 1 的区别在于，还包括存储平台 13，存储平台 13 同时与各个采集器 12 连接，用于集中存储各个采集器采集到的数据流的信息。

在本申请实施例中，采集器和存储平台可以均对外提供数据查询接口。

可选地，在图 1 和图 2 的场景中，还可以包括分析平台 14，分析平台 14 同时与各个采集器 12 连接，或者与存储平台 13 连接，通过分析数据流的信息以确定 NAT 相关的数据流。可选地，存储平台可以和分析平台集成到同一个设备，例如，分析平台包括存储模块，或，存储平台包括分析模块等。

在本申请实施例中，网络设备、采集器、存储平台或分析平台可以通过数据流的信息确定 NAT 相关的数据流。

图 3 是本申请实施例提供的一种数据中心的网络拓扑示意图。参见图 3，数据中心包括 Spine 交换机 14 和虚拟化服务器集群 15，Spine 交换机 14 分别连接虚拟化服务器集群 15。

可选地，数据中心还可以包括分析器网络 16，分析器网络 16 包括采集器集群 161、分析器集群 162 和

Leaf 交换机 163, Spine 交换机 14 通过 Leaf 交换机 163 分别连接采集器集群 161 和分析器集群 162, 采集器集群 161 包括前述存储平台, 分析器集群 162 包括前述分析平台。在其他实现方式中, 分析器网络 16 也可以独立于数据中心之外, 相应地, Leaf 交换机 163 可以是其他网络设备。

其中, Spine 交换机 14 和虚拟化服务器集群 15 都属于前述网络设备 11, 在 Spine 交换机 14 上、以及 Spine 交换机 14 和虚拟化服务器集群 15 之间的网络设备 11 上部署采集器 12。在数据流 17 传输过程中, 三个采集器 12 采集数据流的信息, 然后通过虚线所示的路径上传到采集器集群 161 的存储平台中。

数据中心在边界会部署防火墙 (firewall, FW) (图中未示出), 启用 NAT 功能, 保护数据中心的网络。

在本申请实施例中, 为了保证采集到的指纹特征能够用来识别数据流是否 NAT 相关, 可以在启用 NAT 功能的节点的上下游均部署采集器。可选地, 也可以采用其他方式部署采集器。作为示例, 下方列出了三种采集器的部署方式。

第一种, 在每个节点上均部署采集器, 从而保证能够采集到每个 NAT 会话的数据流的指纹特征。

第二种, 在网络的多个核心节点部署采集器, 例如多个 Spine 交换机、多个边界节点、和/或防火墙上部署采集器, 从而能够尽量采集到每个 NAT 会话的数据流的指纹特征。通常, 开启 NAT 功能的设备两侧都部署有上述核心节点, 因此在这些核心节点上部署采集器, 能够采集到每个 NAT 会话的数据流的指纹特征。以图 3 所示的数据中心为例, 开启 NAT 功能的防火墙通常连接两个 Spine 交换机, 数据流先经过一个 Spine 交换机, 然后在防火墙中进行 NAT 转换, 然后经过另一个 Spine 交换机, 因此, 在 2 个 Spine 交换机上部署采集器, 能够采集到该数据流 NAT 转换前后的指纹特征。

第三种, 在具有 NAT 功能的节点上部署采集器, 以采集到每个 NAT 会话的数据流的指纹特征。

在第一种和第二种部署方式中, 采集器通常部署在节点的接口的入方向, 当然也可以部署在接口的出方向。在第三种部署方式中, 采集器需要同时在接口的入方向和接口的出方向部署。

在一些可能的实现方式中, 部署采集器 (软件) 的各个节点或者采集器 (硬件) 自己存储各自采集到的数据流的信息。

在一些可能的实现方式中, 采集器将采集到的数据流信息发送给存储平台, 由存储平台集中存储。

图 4 是本申请实施例提供的一种数据流识别方法的流程图。在一些可能的实现方式中, 该方法可以由图 1~图 3 所示的应用场景中的分析平台执行, 分析平台与各个采集器连接, 或者与存储平台连接, 以获取数据流信息以实现该方法, 或者, 分析平台获取自身存储的数据流信息以实现该方法。在另一些可能的实现方式中, 该方法还可以由图 1~图 3 所示的应用场景中的网络设备执行。下文以该方法由分析平台执行为例进行说明。如图 4 所示, 该方法包括如下步骤。

S11: 获取第一数据流的指纹特征和建立时间。

在本申请实施例中, 指纹特征包括如下特征中的至少一项:

首个数据包的 IPID、首个数据包的载荷、首个数据包的载荷的哈希值、TCP 三次握手过程中 SYN 报文的 IPID、TCP 三次握手过程中 SYN 报文的 ISN、TCP 三次握手过程中 SYN-ACK 报文的 ISN。

示例性地, 指纹特征为上述特征中的任一个, 例如, 指纹特征为首个数据包的 IPID, 或者, 首个数据包的载荷的哈希值, 或者, SYN 报文的 ISN。采用单独字段作为指纹特征, 减少了计算和存储负担。

示例性地, 指纹特征包括上述特征中的多个, 例如, 指纹特征包括 SYN 报文的 ISN+首个数据包的 IPID, 或者, 首个数据包的载荷+SYN 报文的 ISN, 或者, 首个数据包的载荷+SYN 报文的 ISN+首个数据包的 IPID 等。采用多个字段的组合作为指纹特征, 减小了指纹冲突的可能性。

其中, IPID 是 IP 层中用来标识 IP 数据包的字段, 用来标识主机发出的每一个 IP 报文, 长度 16bit。ISN 是在 TCP 建链的 3 次握手过程的 SYN 报文携带的用于通知给对端的初始的序列码, 长度 32bit。数据包的载荷是数据包内携带的应用层内容 (也即数据包中 TCP 头之后的内容)。

如果采集过程中 SYN 报文或者首个数据包 (简称首包) 发生重传, 采用重传过程中最后一次收到的 SYN 报文或数据包作为采集对象, 采集指纹特征, 保证各个采集器采集的指纹特征的一致。也就是说, 针对一个数据流, 若采集器在一段时间内接收多个 SYN 报文或多个首包, 则采集器从最后一个接收到的 SYN 报文或最后一个接收到的首包获取指纹特征。

其中, 首个数据包是指 TCP 三次握手建立连接后传输的第一个数据包。不同数据流需要传输的内容不

同,有的数据流需要多个数据包才能完成内容的传输,而有的数据流仅需要传输一个数据包就可以完成内容的传输。因此,本申请实施例采用首个数据包的信息作为指纹特征,能够保证可以为每个需要判断的数据流采集到指纹特征。另一方面,首个数据包即是TCP建链后第一个被传输的数据包,而且其序列号等于SYN报文携带的ISN+1,采集器可以很容易地确定首个数据包,因此,采集首个数据包所需的资源和时间相比后续数据包更少。

采集器在采集时,可以根据数据包的序列号来判断是否是首个数据包,通常首个数据包的序列号是SYN报文的ISN+1。而第二个数据包的序列号是SYN报文的ISN+1再加首个数据包的长度,因此,如果要采集首个数据包之后的数据包的特征,则存在判断数据包的序列号更加困难,占用资源和时间更多的问题。所以优先选用首个数据包进行指纹特征采集。

首个数据包的载荷可以是首个数据包的全部载荷,也可以是首个数据包的部分载荷,例如是从数据包包头开始截取的部分载荷,如500或1000字节长度的载荷。相应地,首个数据包的哈希值也可以是全部载荷的哈希值,也可以是部分载荷的哈希值,这里不再赘述。

由于首个数据包的载荷或首个数据包的载荷的哈希值可能仅对应部分长度载荷,因此,首个数据包的载荷除了可以包括载荷内容,还可以包括载荷长度,首个数据包的载荷的哈希值除了可以包括哈希值,还可以包括载荷长度。其中,载荷长度可以作为一个标识,标识载荷内容的长度或者哈希值对应的载荷内容的长度,属于可选项。该载荷长度可以用于筛选指纹特征,例如,在分析平台获取到各个采集器采集的载荷内容或者载荷的哈希值时,可以根据载荷长度判断各个采集器采集配置的采集载荷的长度是否相同,如果相同则可以使用采集的载荷内容或者载荷的哈希值作为指纹特征,否则不使用采集的载荷内容或者载荷的哈希值作为指纹特征。如果指纹特征仅针对部分载荷,在采集过程中,通过端口镜像直接截取所需长度的载荷进行特征提取即可。

在该实现方式中,各个采集器采集相同的指纹特征。申请人通过理论分析和大量的试验发现,上述所列举的指纹特征在NAT转换前后不会变化,并且不同数据流难以发生冲突,即,NAT转换前后的两个数据流尽管IP地址或端口可能不同,但两个数据流中的上述指纹特征相同,因此,上述指纹特征能够标识NAT相关的数据流。另外,上述所列举的指纹特征采集过程较为容易实现,占用资源少,有利于实现本申请实施例提供的方法。

建立时间指示第一数据流的建立时刻,例如,开始创建第一数据流的连接的时刻,或者,第一数据流的连接建立完成的时刻,或者开始发送第一数据流的数据报文的时刻。例如,建立时间可以是SYN报文接收时间、SYN-ACK报文接收时间、ACK报文接收时间、SYN报文发送时间、SYN-ACK报文发送时间、ACK报文发送时间或首个数据包的接收或发送时间等。此处的SYN、SYN-ACK和ACK报文为第一数据流对应的TCP三次握手中的报文。首个数据包是第一数据流对应的TCP连接创建完成后发送或接收的第一个数据包。

S12:根据第一数据流的指纹特征和建立时间,确定与第一数据流NAT相关的至少一个数据流。

与第一数据流NAT相关可以是与第一数据流属于同一个NAT会话。

一个NAT会话包括NAT转换前的数据流和NAT转换后的数据流。例如,第一数据流为NAT设备A转换前的数据流,第二数据流为NAT设备A对第一数据流进行NAT转换后获得的数据流,则第一数据流和第二数据流属于同一个NAT会话。NAT转换前的数据流和NAT转换后的数据流具有相同的指纹特征。例如,第一数据流的SYN报文的IPID和第二数据流的SYN报文的IPID相同,第一数据流的SYN报文的ISN和第二数据流的SYN报文的ISN相同,第一数据流的首包的IPID和第二数据流的首包的IPID相同,第一数据流的首包的载荷和第二数据流的首包的载荷相同,第一数据流的首包载荷的哈希值和第二数据流的首包载荷的哈希值相同,第一数据流的SYN-ACK报文的ISN和第二数据流的SYN-ACK报文的ISN。

与第一数据流NAT相关还可以是与第一数据流经过N次NAT转换后的数据流属于同一个NAT会话。

其中,N为大于等于1的自然数。例如,业务数据流从业务源端到业务目的端可能会经过多次NAT(例如,传输路径上部署了M个NAT设备)转换,假设第一数据流为从业务源端发出的数据流,则该数据流经过第一个NAT设备的NAT转换后变为第二数据流,第二数据流与第一数据流属于同一个NAT会话,第二数据流经过第二个NAT设备的NAT转换后变成第三数据流,第三数据流与第二数据流属于同一

个 NAT 会话，第三数据流与第一数据流 NAT 相关，以此类推，第 M-1 数据流经过第 M-1 个 NAT 设备转换后变成第 M 数据流，第 M 数据流与第 M-1 数据流属于同一个 NAT 会话，第 M 数据流与第一数据流 NAT 相关，第 M 数据流经过第 M 个 NAT 设备转换后变成第 M+1 数据流，第 M+1 数据流与第 M 数据流属于同一个 NAT 会话，第 M+1 数据流也与第一数据流 NAT 相关。其中，第二数据流是第一数据流经过一次 NAT 转换后的数据流，第三数据流与第二数据流属于同一个 NAT 会话，即，第三数据流与第一数据流经过一次 NAT 转换后的数据流属于同一个 NAT 会话。其中，第 M-1 数据流是第一数据流经过 M-1 次 NAT 转换后的数据流，第 M 数据流与第 M-1 数据流属于同一个 NAT 会话，即，第 M 数据流与第一数据流经过 M-1 次 NAT 转换后的数据流属于同一个 NAT 会话。其中，第 M 数据流是第一数据流经过 M 次 NAT 转换后的数据流，第 M+1 数据流与第 M 数据流属于同一个 NAT 会话，即，第 M+1 数据流与第一数据流经过 M 次 NAT 转换后的数据流属于同一个 NAT 会话。NAT 转换前后的数据流的指纹特征相同，即，属于同一个 NAT 会话的数据流的指纹特征相同，而与第一数据流以及与第一数据流 NAT 相关的至少一个数据流中的每两个数据流均属于同一个 NAT 会话，因此，第一数据流和与第一数据流 NAT 相关的数据流的指纹特征均相同。

分析平台可以将指纹特征与第一数据流的指纹特征相同，且建立时间与第一数据流的建立时间接近的数据流确定为与第一数据流 NAT 相关的数据流。其中，建立时间与第一数据流的建立时间接近例如为建立时间与第一数据流的建立时间的差值的绝对值小于阈值。

在本申请实施例中，通过第一数据流的指纹特征确定与第一数据流 NAT 相关的至少一个数据流，不需要从 NAT 设备读取会话表，实现难度及代价小，并且网络部署难度小。

图 5 是本申请实施例提供的一种数据流识别方法的流程图。该方法可以由图 1~图 3 所示的应用场景中的采集器、部署有采集器的网络设备或存储平台执行。如图 5 所示，该方法包括如下步骤。

S21：接收查询请求，查询请求包括第一数据流的第一流标识。

S22：发送查询应答，查询应答包括第一数据流的指纹特征和建立时间，第一数据流的指纹特征和建立时间用于确定与第一数据流 NAT 相关的至少一个数据流。

在接收到第一数据流的第一流标识后，以第一流标识为索引进行查询，以获取第一数据流的指纹特征和建立时间，然后反馈查询应答。

可选地，查询请求还可以包括查询时间范围，当查询请求包括查询时间范围时，查询在该查询时间范围内建立的数据流。

查询时间范围是用来辅助查找数据流的记录条件，如果不指定，系统可以采用默认值，如当天。

如果查询时间范围较大，查询到多条数据流，可以返回多条数据流的指纹特征作为应答，也可以通过从多条数据流中选择一条数据流的指纹特征作为应答，选择的依据可以是流的创建时间。例如，由采集器或存储平台自动选择创建时间最近的数据流的指纹特征作为应答。创建时间最近的数据流例如为采集器或存储平台最新记录的数据流。当查询请求还包括查询时间范围时，创建时间最近的数据流例如为查询时间范围内最新记录的数据流。或者，当查找到多条数据流时，将多条数据流的创建时间反馈给分析平台，分析平台提供选择界面由用户根据创建时间从多条数据流中选择，然后将选择结果反馈给采集器或存储平台，采集器或存储平台再根据选择反馈对应数据流的指纹特征。

图 6 是本申请实施例提供的一种数据流信息的采集方法的流程图。该方法由采集器执行。如图 6 所示，该方法包括如下步骤。

S31：采集各个数据流的建立时间、第一流标识和指纹特征。

在一些可能的实现方式中，采集器以旁路的方式接入。此时，采集各个数据流的建立时间、第一流标识和指纹特征，包括：

通过镜像功能抓取数据流的报文；

分析抓取到的数据流的报文，以获取数据流的建立时间、第一流标识和指纹特征。

其中，镜像功能可以是报文端口镜像或者远程镜像（例如封装式远程交换端口分析仪技术（encapsulated remote switch port analyzer, ERSPAN）远程镜像）。其中，报文端口镜像是对端口报文原封不动进行复制，

而远程镜像则是将端口报文封装后进行远程复制。

在本申请的实现方式中，当指纹特征采用 TCP 三次握手中的 SYN 报文和/或 SYN-ACK 报文的特征时，通过镜像功能抓取数据流的报文，包括：

通过配置 ACL 抓取 SYN 报文和/或 SYN-ACK 报文。

在配置镜像功能时，可以配置镜像功能的多个可选参数，其中可以包括 ACL，通过 ACL 能够对指定报文或者指定报文长度进行过滤，实现 SYN 报文和/或 SYN-ACK 报文采集。

通过配置 ACL 仅抓取 TCP 建立过程中的 3 次握手报文以采集 SYN 报文和/或 SYN-ACK 报文的特征，可以减少需要抓取的镜像报文数量，减轻了网络设备的镜像负担和采集器的报文处理负担，有利于大规模部署。相比于首包的特征，使用 SYN 报文或 SYN-ACK 报文的特征作为指纹特征，处理负担更小。

在另一些可能的实现方式中，采集器以随路的方式接入，例如，转发设备也包含采集功能。此时，采集器直接分析数据流的报文，以获取数据流的建立时间、第一流标识和指纹特征即可。

示例性地，采集器在采集数据流的创建时间时，可以采用如下时间中的一种作为数据流的创建时间：

SYN 报文接收时间、SYN-ACK 报文接收时间、ACK 报文接收时间。此处的 ACK 报文为 TCP 三次握手手中的 ACK 报文。

其中，采集器在采集 TCP 三次握手手中的 ACK 报文的接收时间时，可以以 SYN 报文或 SYN-ACK 报文为基准，采集 SYN 报文或 SYN-ACK 报文后的第一个 ACK 报文的接收时间，即为 TCP 三次握手手中的 ACK 报文的接收时间。

当然上面针对的是接口的入方向的采集器，如果是接口的出方向的采集器，则采用如下时间中的一种作为数据流的创建时间：

SYN 报文发送时间、SYN-ACK 报文发送时间、ACK 报文发送时间。

S32：存储各个数据流的建立时间、第一流标识和指纹特征。

在一些可能的实现方式中，采集器本地存储采集到数据，或者存储到部署所在的网络设备中。

在另一些可能的实现方式中，采集器将采集到的数据发送到存储平台进行集中式存储。

示例性地，数据流的第一流标识为数据流的五元组，包括源 IP 地址、目的 IP 地址、源端口、目的端口和协议类型，其中协议类型为 TCP。

在本申请的实施例中，数据流的信息可以按照下表 1 的格式进行存储：

表 1

记录时间	目的 IP 地址	目的端口	源 IP 地址	源端口	协议类型	建立时间	指纹特征 1	指纹特征 2
11:24:32	100.1.1.1	80	200.1.1.1	65432	TCP	11:24:31.312	A	B

其中，记录时间可以是采集时间，也可以是存储时间。由于记录时间和建立时间通常较为接近，因此，查询时间范围除了可以限定建立时间，也可以限定的是记录时间，对此不做限定。

数据流的指纹特征可以由 1 个特征或多个特征组成，表 1 中包括两个特征，仅为举例，不作为本申请的限制。

当然，采集器除了采集和存储上述流的信息外，还可以采集和记录包数、速率、丢包率等信息。

图 7 是本申请实施例提供的一种数据流识别方法的流程图。该方法以分析平台、第一网络节点和第二网络节点共同执行为例进行说明，示例性地，第一网络节点和第二网络节点为采集器或部署有采集器的网络设备。如图 7 所示，该方法包括如下步骤。

S41：分析平台向第一网络节点发送第一查询请求，第一查询请求包括第一数据流的第一流标识；第一网络节点接收第一查询请求。

这里的第一网络节点可以是一个也可以是多个。

第一查询请求除了包括流标识外，还可以包括指示请求的参数字段，例如第一查询请求可以通过携带不同标识指示要请求的不同参数，比如标识 a 指示请求数据流的建立时间，标识 b 指示请求数据流的指纹

特征。当然，前述指示请求的方式仅为一种示例，不作为本申请的限制。

S42：第一网络节点以第一流标识为索引进行查询。当查询到结果时，获取第一数据流的指纹特征和建立时间，然后执行 S43，否则不执行后续步骤。

这里的网络节点是采集器或部署有采集器的网络设备，存储有如表 1 所示的数据流的信息，因此可以采用第一流标识（五元组）为索引进行查询。

若分析平台向多个第一网络节点发送第一查询请求，则每个第一网络节点均执行步骤 S42。

可选地，在没有查询到结果时，该方法还可以包括：第一网络节点向分析平台反馈查询失败消息，以通知分析平台未查询到结果。

S43：第一网络节点向分析平台发送第一查询应答，第一查询应答包括第一数据流的指纹特征和建立时间；分析平台接收第一查询应答。

S44：分析平台向第二网络节点发送第二查询请求，第二查询请求包括第一数据流的指纹特征；第二网络节点接收第二查询请求。

这里的第二网络节点可以是一个也可以是多个。第二网络节点和第一网络节点可以至少部分相同，也可以完全不同。

S45：第二网络节点以第一数据流的指纹特征为索引进行查询。当查询到结果时，获取与第一数据流的指纹特征相同的第二数据流的第一流标识和建立时间，然后执行 S46，否则不执行后续步骤。

若分析平台向多个第二网络节点发送第二查询请求，则每个第二网络节点均执行步骤 S45。

可选地，在没有查询到结果时，该方法还可以包括：第二网络节点向分析平台反馈查询失败消息，以通知分析平台未查询到结果。

S46：第二网络节点向分析平台发送第二查询应答，第二查询应答包括与第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；分析平台接收第二查询应答。

若分析平台向多个第二网络节点发送第二查询请求，则分析平台有可能接收到不同的第二网络节点发送的第二查询应答。每个第二查询应答可能包含相同的第二数据流的第一流标识，这对应于第一数据流的源端和目的端的传输路径上仅存在一个 NAT 设备的场景。不同的第二查询应答可能包含不同的第二数据流的第一流标识，这对应于第一数据流的源端和目的端的传输路径上存在多个 NAT 设备的场景，相应地，分析平台可能获取到多个第二数据流的信息。

S47：分析平台根据第一数据流的建立时间和至少一个第二数据流的建立时间，确定与第一数据流 NAT 相关的数据流。

当分析平台仅获取到一个第二数据流时，在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值的情况下，分析平台确定第一数据流和第二数据流属于同一个 NAT 会话，也即，第二数据流是第一数据流 NAT 相关的数据流。否则，分析平台确定第一数据流和第二数据流不属于同一个 NAT 会话。

由于 NAT 会话可能有多级，因此可能存在多个数据流和第一数据流 NAT 相关，因此在查询到多个指纹相同的数据流时，可以分别针对每个数据流进行判断，确定是否与第一数据流 NAT 相关。

其中，对于多级 NAT，多级 NAT 的转换和传输延迟较小，第一数据流和经过多级 NAT 转换后的数据流的建立时间的时间差的绝对值小于设定的阈值，因此，采用上述方案能够准确判断两个数据流是否 NAT 相关。

不同网络节点上的采集器采集时存在同步精度问题，不同的采集器之间还会存在传输时延，因此，可以基于时间同步精度和/或传输时延确定上述阈值。

在一些可能的实现方式中，NAT 相关的一个或多个数据流经过的路程短，传输延迟低，这种情况下，可以仅基于采集器的时间同步精度确定阈值。例如，采集器均不属于高速数据中心，采集器间的传输时延为微妙级，采集器之间采用网络时间协议（network time protocol, NTP）进行同步，同步精度在毫秒级，相对于同步精度，传输时延可以忽略不计，因此，可以直接基于同步精度确定阈值，例如，阈值可以为 10 毫秒等。再例如，采集器之间的同步精度更高时，阈值可以设计得更小，例如 1 毫秒，2 毫秒等。其中，一个时间同步协议的同步精度指示网络中的设备基于该时间同步协议进行时间同步时可能存在的时间误差，例如，设备 A 和设备 B 基于 NTP 协议进行时间同步，若同步精度为毫秒级，其表示设备 A 和设备 B

的时间可能会相差几毫秒。

在另一些可能的实现方式中，NAT 相关的数据流在采集点间的传输路径长，传输延迟高，相对于传输延迟，时间同步精度问题可以忽略不计，这种情况下，可以仅基于传输时延确定上述阈值。例如，两个采集点跨广域网，采集点间的传输延迟约为 50~60ms，远超同步精度的几毫秒，此时可以直接基于传输延迟确定阈值，例如，阈值为 100ms。

在本申请另一些可能的实现方式中，采集点间的传输延迟与时间同步精度接近，这种情况下，除了考虑采集器的时间同步精度，还需要考虑传输延迟，可以基于时间同步精度和传输时延确定阈值。例如，同步精度在毫秒级，传输延迟也是毫秒级，则阈值可以为 20ms。

随着网络规模的增大和网络中部署业务的增多，有可能出现指纹冲突的情况，指纹冲突是指非 NAT 相关的数据流的指纹特征相同。为进一步提高准确性，可选地，在确定 NAT 相关的数据流时，还可以排除指纹冲突的数据流。可选地，在本申请实施例中，在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关，包括：

在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值，第一数据流和第二数据流不存在指纹冲突的情况下，确定第一数据流和第二数据流 NAT 相关，指纹冲突是指非 NAT 相关的数据流的指纹特征相同。

例如，两个数据流不是同一 NAT 会话中的数据流，但两个数据流的指纹特征相同，此时这两个数据流存在指纹冲突。

再例如，与第一数据流指纹特征相同的多个数据流，不是第一数据流经过 N 级 NAT 转换后的数据流，此时这多个数据流与第一数据流指纹冲突。

分析平台可以分别对每个第二数据流进行指纹冲突判断。

示例性地，确定指纹冲突的方式如下：

确定与第二数据流具有相同第二流标识的至少一条第三数据流；

确定与每条第三数据流具有相同指纹特征的至少一条第四数据流；

在至少一条第四数据流中的存在一条第四数据流的建立时间和对应第三数据流的建立时间的时间差的绝对值小于阈值，且这一条第四数据流的第二流标识和第一数据流的第二流标识相同的情况下，确定第一数据流和第二数据流不存在指纹冲突。

其中，数据流的第二流标识可以是四元组，包括源 IP 地址、目的 IP 地址、目的端口和协议类型。

该方案不仅可以判断仅存在单级 NAT 的场景下的指纹冲突，还可以判断存在多级 NAT 的场景下的指纹冲突。因为，上述确定指纹冲突的过程中，使用的原理是：

开启 NAT 功能的设备会对从相同主机发出到相同目的地的相同端口的数据流（也即四元组相同的数据流）都进行 NAT 转换。故，查询历史记录中四元组相同的数据流是否经过 NAT 转换，如果查询到存在四元组相同的数据流经过 NAT 转换，则认为本次也是 NAT 转换，而非指纹冲突。对应到存在多级 NAT 的场景下，如果存在四元组相同的数据流在历史记录中存在过多级 NAT 转换，则认为本次也是多级 NAT 转换。故上述判断过程能够确定每条第二数据流与第一数据流是指纹冲突，还是 NAT 相关。

其中，确定第三数据流时，不限定数据流的源端口是因为第二数据流和第三数据流可以是同一主机的不同应用产生的数据流，但由于目的地址和端口相同，传输时会进行同样的 NAT 转换。

例如，数据流 a 经过 NAT 转换得到数据流 b，数据流 b 经过 NAT 转换得到数据流 c。在历史记录中，与数据流 a 具有相同第二流标识的数据流 A 经过 NAT 转换得到数据流 B，数据流 B 经过 NAT 转换得到数据流 C。

在数据流识别时，确定出数据流 b 和 c 的指纹特征和 a 相同，并且建立时间的时间差的绝对值也小于阈值。此时进行指纹冲突判断：

第一数据流为数据流 a，第二数据流为数据流 c，确定与数据流 c 具有相同第二流标识的第三数据流为数据流 C，确定与数据流 C 具有相同指纹特征的第四数据流为数据流 B 和数据流 A。

在第四数据流中，数据流 A 的建立时间和数据流 C 的建立时间的时间差的绝对值小于阈值，并且数据流 A 的第二流标识和数据流 a 相同，此时确定不存在指纹冲突。

下面结合表 2 和表 3 提供的的数据流信息来对指纹冲突进行说明。

输入的第一数据流为 flow11, 根据指纹匹配能查询到第二数据流 flow21 和 flow31 可能是 NAT 相关的数据流, 如下表 2 所示:

表 2

	目的 IP 地址	目的端口	源 IP 地址	源端口	建立时间	指纹特征
flow11	100.1.1.1	80	200.1.1.1	65432	11:24:31.312	A
flow21	10.10.1.2	8080	10.20.1.1	52347	11:24:31.351	A
flow31	10.10.1.3	8081	10.20.1.2	56475	11:24:31.987	A

分别查询具有相同第二流标识的其他流, 可以得到 3 组流信息, 如下表 3 所示:

表 3

	目的 IP 地址	目的端口	源 IP 地址	源端口	建立时间	指纹特征
flow11	100.1.1.1	80	200.1.1.1	65432	11:24:31.312	A
flow12	100.1.1.1	80	200.1.1.1	76543	11:24:16.678	B
flow13	100.1.1.1	80	200.1.1.1	63542	11:24:23.547	C
flow21	10.10.1.2	8080	10.20.1.1	52347	11:24:31.351	A
flow22	10.10.1.2	8080	10.20.1.1	34214	11:24:17.121	B
flow31	10.10.1.3	8081	10.20.1.2	56475	11:24:31.987	A
flow32	10.10.1.3	8081	10.20.1.2	43265	11:20:21.334	D

根据第二数据流 flow21 和 flow31 分别查询到第三数据流 flow22 和 flow32。

其中, 第三数据流 flow22 对应的至少一条第四数据流为 flow12, 第三数据流 flow32 没有对应的第四数据流。

由于第三数据流 flow22 对应的至少一条第四数据流为 flow12 符合建立时间的时间差的绝对值小于阈值, 且与第一数据流的第二流标识相同的条件, 因此可以判断第二数据流 flow21 和第一数据流 flow11 不存在指纹冲突。而第二数据流 flow31 和第一数据流 flow11 存在指纹冲突。

可选地, 该分析平台可以是网络诊断设备, 相应地, 该方法还包括:

S48: 分析平台在第一数据流发生故障的情况下, 对第一数据流以及与第一数据流 NAT 相关的至少一个数据流进行故障诊断。

例如, 故障诊断设备在确定第一数据流存在丢包率或延迟较高的故障后, 先确定与第一数据流 NAT 相关的至少一个第二数据流, 然后针对第一数据流和至少一个第二数据流, 分析丢包率或延迟高的原因, 确定链路上造成丢包或延迟的故障点, 完成故障诊断。

在一些场景中, 网络诊断设备在第一数据流根据数据包的 ACK 报文, 确定存在丢包, 但丢包的位置并非在第一数据流经过的链路中, 此时针对第一数据流和第二数据流进行故障诊断, 确定出丢包的位置在第二数据流经过的链路中, 并且确定出造成丢包的故障点位置, 即实现了更精确地定位或定界网络故障。

在另一些场景中, 网络诊断设备发现第一数据流的数据包传输延迟较大, 此时针对第一数据流和第二数据流进行故障诊断, 确定出造成传输延迟的故障点位置在第二数据流所在链路上, 即实现了更精确地定位或定界网络故障。

本申请实施例以数据流为粒度进行故障诊断, 原因在于一些场景下, 造成数据包丢包或者传输的延迟的原因是第二数据流所在链路上的设备对该数据流的突发流量进行了限速, 这种情况不涉及链路上的其他数据流, 以数据流为粒度诊断更准确。

可见, 通过本申请提供的数据流识别方案能够确定出 NAT 相关的数据流, 这使得分析平台不仅可以针对原数据流进行故障诊断, 还可以继续针对 NAT 相关的数据流进行故障诊断, 可以更精确地定位或定界网络故障。

可选地, 在确定不存在与第一数据流 NAT 相关的数据流时, 可以确定不存在与第一数据流属于同一个 NAT 会话的数据流。可选地, 在确定不存在与第一数据流 NAT 相关的数据流时, 该方法还包括:

分析平台在第一数据流发生故障的情况下, 对第一数据流进行故障诊断。

图 8 是本申请实施例提供的一种数据流识别方法的流程图。该方法以分析平台、第一网络节点和第二网络节点共同执行为例进行说明，示例性地，第一网络节点为存储平台。如图 8 所示，该方法包括如下步骤。

S51: 分析平台向第一网络节点发送第一查询请求，第一查询请求包括第一数据流的第一流标识；第一网络节点接收第一查询请求。

S52: 第一网络节点以第一流标识为索引进行查询。当查询到结果时，获取第一数据流流的指纹特征和建立时间，然后执行 S53，否则不执行后续步骤。

可选地，在没有查询到结果时，该方法还可以包括：第一网络节点向分析平台反馈查询失败消息，以通知分析平台未查询到结果。

S53: 第一网络节点以第一数据流的指纹特征为索引进行查询。当查询到结果时，获取与第一数据流的指纹特征相同的第二数据流的第一流标识和建立时间，然后执行 S54，否则不执行后续步骤。

可选地，在没有查询到结果时，该方法还可以包括：第一网络节点向分析平台反馈查询失败消息，以通知分析平台未查询到结果。

S54: 第一网络节点根据第一数据流和第二数据流的建立时间，确定第一数据流和第二数据流是否 NAT 相关。当确定第一数据流和第二数据流 NAT 相关时，执行 S55，否则不执行后续步骤。

在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。否则，确定第一数据流和第二数据流非 NAT 相关。

可选地，在本申请实施例中，在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关，包括：

在第二数据流的建立时间与第一数据流的建立时间的时间差的绝对值小于阈值，第一数据流和第二数据流不存在指纹冲突的情况下，确定第一数据流和第二数据流 NAT 相关，指纹冲突是指非 NAT 相关的数据流的指纹特征相同。

确定指纹冲突的详细过程参见步骤 S47，在此不做赘述。

可选地，在确定第一数据流和第二数据流非 NAT 相关时，该方法还可以包括：第一网络节点向分析平台反馈通知消息，以通知分析平台没有与第一数据流 NAT 相关的数据流。

S55: 第一网络节点向分析平台发送第三查询应答，第三查询应答包括与第一数据流 NAT 相关的第二数据流的第一流标识；分析平台接收第三查询应答。

S56: 分析平台根据第三查询应答，确定第一数据流和第二数据流 NAT 相关。

S57: 分析平台在第一数据流发生故障的情况下，对第一数据流以及与第一数据流 NAT 相关的至少一个数据流进行故障诊断。

步骤 S57 的详细过程参考步骤 S48。

图 9 是本申请实施例提供的一种数据流识别方法的流程图。该方法以分析平台和第一网络节点共同执行为例进行说明，示例性地，第一网络节点为存储平台。如图 9 所示，该方法包括如下步骤。

S61: 分析平台向第一网络节点发送第一查询请求，第一查询请求包括第一数据流的第一流标识；第一网络节点接收第一查询请求。

S62: 第一网络节点以第一流标识为索引进行查询。当查询到结果时，获取第一数据流流的指纹特征和建立时间，然后执行 S63，否则不执行后续步骤。

S63: 第一网络节点以第一数据流的指纹特征为索引进行查询。当查询到结果时，获取与第一数据流的指纹特征相同的第二数据流的第一流标识和建立时间，然后执行 S64，否则不执行后续步骤。

可选地，在没有查询到结果时，该方法还可以包括：第一网络节点向分析平台反馈查询失败消息，以通知分析平台未查询到结果。

S64: 第一网络节点向分析平台发送第一查询应答，第一查询应答包括第一数据流的指纹特征和建立时间、与第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；分析平台接收第一查询应答。

S65: 分析平台根据第一数据流和第二数据流的建立时间，确定第一数据流和第二数据流是否 NAT 相

关。

在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。否则，确定第一数据流和第二数据流非 NAT 相关。

可选地，在本申请实施例中，在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关，包括：

在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值，第一数据流和第二数据流不存在指纹冲突的情况下，确定第一数据流和第二数据流 NAT 相关，指纹冲突是指非 NAT 相关的数据流的指纹特征相同。

确定指纹冲突的详细过程参见步骤 S47，在此不做赘述。

S66：分析平台在第一数据流发生故障的情况下，对第一数据流以及与第一数据流 NAT 相关的至少一个数据流进行故障诊断。

步骤 S66 的详细过程参考步骤 S48。

图 10 是本申请实施例提供的一种数据流识别装置的框图。该数据流识别装置可以通过软件、硬件或者两者的结合实现成为分析平台、网络设备、采集器或存储平台的全部或者一部分。该数据流识别装置可以包括：获取单元 701 和确定单元 702。

其中，获取单元 701，用于获取第一数据流的指纹特征和建立时间；

确定单元 702，用于根据第一数据流的指纹特征和建立时间，确定与第一数据流 NAT 相关的至少一个数据流。

可选地，获取单元 701，用于向第一网络节点发送第一查询请求，第一查询请求包括第一数据流的第一流标识；接收第一网络节点发送的第一查询应答，第一查询应答包括第一数据流的指纹特征和建立时间。

可选地，确定单元 702，用于向第二网络节点发送第二查询请求，第二查询请求包括第一数据流的指纹特征；接收第二网络节点发送的第二查询应答，第二查询应答包括与第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。

可选地，第一查询应答还包括与第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；

确定单元 702，用于在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。

可选地，获取单元 701，用于接收第一查询请求，第一查询请求包括第一数据流的第一流标识；根据第一数据流的第一流标识确定第一数据流的指纹特征和建立时间。

可选地，确定单元 702，用于根据第一数据流的指纹特征确定第二数据流，第二数据流的指纹特征与第一数据流的指纹特征相同；在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值的情况下，确定第一数据流和第二数据流 NAT 相关。

可选地，指纹特征包括如下特征中的至少一项：

首个数据包的互联网协议标识 IPID、首个数据包的载荷、首个数据包的载荷的哈希值、传输控制协议 TCP 三次握手过程中 SYN 报文的 IPID、TCP 三次握手过程中 SYN 报文的初始序列码 ISN、TCP 三次握手过程中 SYN-ACK 报文的 ISN。

可选地，第一查询请求还包括查询时间范围，查询时间范围用于限制查询到的数据流的建立时间所处的时间范围。

可选地，确定单元 702，用于在第二数据流的建立时间与第一数据流的建立时间的绝对值小于阈值，第一数据流和第二数据流不存在指纹冲突的情况下，确定第一数据流和第二数据流 NAT 相关，指纹冲突是指非 NAT 相关的数据流的指纹特征相同。

可选地，确定单元 702，还用于确定与第二数据流具有相同第二流标识的至少一条第三数据流；确定与每条第三数据流具有相同指纹特征的至少一条第四数据流；在至少一条第四数据流中存在一条第四数据流的建立时间和对应第三数据流的建立时间的绝对值小于阈值，且这一条第四数据流的第二流标识和第一数据流的第二流标识相同的情况下，确定第一数据流和第二数据流不存在指纹冲突。

可选地，该装置还包括：

诊断单元 703，用于在第一数据流发生故障的情况下，对第一数据流以及与第一数据流 NAT 相关的至少一个数据流进行故障诊断。

需要说明的是，上述实施例提供的数据流识别装置在进行数据流识别时，仅以上述各功能单元的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能单元完成，即将设备的内部结构划分成不同的功能单元，以完成以上描述的全部或者部分功能。另外，上述实施例提供的数据流识别装置与数据流识别方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

图 11 是本申请实施例提供的一种数据流识别装置的框图。该显示装置可以通过软件、硬件或者两者的结合实现成为采集器、部署有采集器的网络设备或存储平台的全部或者一部分。该显示装置可以包括：接收单元 801 和发送单元 802。

其中，接收单元 801，用于接收查询请求，查询请求包括第一数据流的第一流标识；

发送单元 802，用于发送查询应答，查询应答包括第一数据流的指纹特征和建立时间，第一数据流的指纹特征和建立时间用于确定与第一数据流 NAT 相关的至少一个数据流。

可选地，查询应答还包括与第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间。

可选地，该装置还包括：

采集单元 803，用于采集各个数据流的建立时间、第一流标识和指纹特征；

存储单元 804，用于存储各个数据流的建立时间、第一流标识和指纹特征。

可选地，采集单元 803，用于通过镜像功能抓取数据流的报文；分析抓取到的数据流的报文，以获取数据流的建立时间、第一流标识和指纹特征。

需要说明的是，上述实施例提供的数据流识别装置在进行数据流识别时，仅以上述各功能单元的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能单元完成，即将设备的内部结构划分成不同的功能单元，以完成以上描述的全部或者部分功能。另外，上述实施例提供的数据流识别装置与数据流识别方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

上述各个附图对应的流程的描述各有侧重，某个流程中没有详述的部分，可以参见其他流程的相关描述。

图 12 示出了本申请实施例提供的电子设备 150 的结构示意图。该电子设备可以为分析平台、网络设备、采集器或存储平台。图 12 所示的电子设备 150 用于执行上述图 4 至图 9 任一幅所示的数据流识别方法所涉及的操作。该电子设备 150 可以由一般性的总线体系结构来实现。

如图 12 所示，电子设备 150 包括至少一个处理器 151、存储器 153 以及至少一个通信接口 154。

处理器 151 例如是通用中央处理器（central processing unit, CPU）、数字信号处理器（digital signal processor, DSP）、网络处理器（network processor, NP）、数据处理单元（Data Processing Unit, DPU）、微处理器或者一个或多个用于实现本申请方案的集成电路。例如，处理器 151 包括专用集成电路（application-specific integrated circuit, ASIC）、可编程逻辑器件（programmable logic device, PLD）或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件或者其任意组合。PLD 例如是复杂可编程逻辑器件（complex programmable logic device, CPLD）、现场可编程逻辑门阵列（field-programmable gate array, FPGA）、通用阵列逻辑（generic array logic, GAL）或其任意组合。其可以实现或执行结合本申请实施例公开内容所描述的各种逻辑方框、模块和电路。所述处理器也可以是实现计算功能的组合，例如包括一个或多个微处理器组合，DSP 和微处理器的组合等等。

可选的，电子设备 150 还包括总线。总线用于在电子设备 150 的各组件之间传送信息。总线可以是外设部件互连标准（peripheral component interconnect, 简称 PCI）总线或扩展工业标准结构（extended industry standard architecture, 简称 EISA）总线等。总线可以分为地址总线、数据总线、控制总线等。为便于表示，图 12 中仅用一条粗线表示，但并不表示仅有一根总线或一种类型的总线。

存储器 153 例如是只读存储器（read-only memory, ROM）或可存储静态信息和指令的其它类型的静态存储设备，又如是随机存取存储器（random access memory, RAM）或者可存储信息和指令的其它类型

的动态存储设备，又如是电可擦可编程只读存储器（electrically erasable programmable read-only Memory, EEPROM）、只读光盘（compact disc read-only memory, CD-ROM）或其它光盘存储、光碟存储（包括压缩光碟、激光碟、光碟、数字通用光碟、蓝光光碟等）、磁盘存储介质或者其它磁存储设备，或者是能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其它介质，但不限于此。存储器 153 例如是独立存在，并通过总线与处理器 151 相连接。存储器 153 也可以和处理器 151 集成在一起。

通信接口 154 使用任何收发器一类的装置，用于与其它设备或通信网络通信，通信网络可以为以太网、无线接入网（RAN）或无线局域网（wireless local area networks, WLAN）等。通信接口 154 可以包括有线通信接口，还可以包括无线通信接口。具体的，通信接口 154 可以为以太（Ethernet）接口、快速以太（Fast Ethernet, FE）接口、千兆以太（Gigabit Ethernet, GE）接口，异步传输模式（Asynchronous Transfer Mode, ATM）接口，无线局域网（wireless local area networks, WLAN）接口，蜂窝网络通信接口或其组合。以太网接口可以是光接口，电接口或其组合。在本申请实施例中，通信接口 154 可以用于电子设备 150 与其他设备进行通信。

在具体实现中，作为一种实施例，处理器 151 可以包括一个或多个 CPU，如图 12 中所示的 CPU0 和 CPU1。这些处理器中的每一个可以是一个单核（single-CPU）处理器，也可以是一个多核（multi-CPU）处理器。这里的处理器可以指一个或多个设备、电路、和/或用于处理数据（例如计算机程序指令）的处理核。

在具体实现中，作为一种实施例，电子设备 150 可以包括多个处理器，如图 12 中所示的处理器 151 和处理器 155。这些处理器中的每一个可以是一个单核处理器（single-CPU），也可以是一个多核处理器（multi-CPU）。这里的处理器可以指一个或多个设备、电路、和/或用于处理数据（如计算机程序指令）的处理核。

在具体实现中，作为一种实施例，电子设备 150 还可以包括输出设备和输入设备。输出设备和处理器 151 通信，可以以多种方式来显示信息。例如，输出设备可以是液晶显示器（liquid crystal display, LCD）、发光二极管（light emitting diode, LED）显示设备、阴极射线管（cathode ray tube, CRT）显示设备或投影仪（projector）等。输入设备和处理器 151 通信，可以以多种方式接收用户的输入。例如，输入设备可以是鼠标、键盘、触摸屏设备或传感设备等。

在一些实施例中，存储器 153 用于存储执行本申请方案的程序代码 1510，处理器 151 可以执行存储器 153 中存储的程序代码 1510。也即是，电子设备 150 可以通过处理器 151 执行存储器 153 中的程序代码 1510，来实现方法实施例提供的的数据流识别方法。程序代码 1510 中可以包括一个或多个软件模块。可选地，处理器 151 自身也可以存储执行本申请方案的程序代码或指令。

在具体实施例中，本申请实施例的电子设备 150 可对应于上述各个方法实施例中的控制器，电子设备 150 中的处理器 151 读取存储器 153 中的指令，使图 12 所示的电子设备 150 能够执行控制器所执行的全部或部分操作。

具体的，处理器 151 用于获取第一数据流的指纹特征和建立时间；根据所述第一数据流的指纹特征和建立时间，确定与所述第一数据流 NAT 相关的至少一个数据流。

其他可选的实施方式，为了简洁，在此不再赘述。

其中，图 4 至图 9 任一幅所示的数据流识别方法的各步骤通过电子设备 150 的处理器中的硬件的集成逻辑电路或者软件形式的指令完成。结合本申请实施例所公开的方法的步骤可以直接体现为硬件处理器执行完成，或者用处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器，处理器读取存储器中的信息，结合其硬件完成上述方法的步骤，为避免重复，这里不再详细描述。

本申请实施例还提供了一种芯片，包括：输入接口、输出接口、处理器和存储器。输入接口、输出接口、处理器以及存储器之间通过内部连接通路相连。处理器用于执行存储器中的代码，当代码被执行时，处理器用于执行上述任一种的数据流识别方法。

应理解的是，上述处理器可以是 CPU，还可以是其他通用处理器、DSP、ASIC、FPGA 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者是任何常规

的处理器等。值得说明的是，处理器可以是支持 ARM 架构的处理器。

进一步地，在一种可选的实施例中，上述处理器为一个或多个，存储器为一个或多个。可选地，存储器可以与处理器集成在一起，或者存储器与处理器分离设置。上述存储器可以包括只读存储器和随机存取存储器，并向处理器提供指令和数据。存储器还可以包括非易失性随机存取存储器。例如，存储器还可以存储参考块和目标块。

该存储器可以是易失性存储器或非易失性存储器，或可包括易失性和非易失性存储器两者。其中，非易失性存储器可以是 ROM、PROM、EPROM、EEPROM 或闪存。易失性存储器可以是 RAM，其用作外部高速缓存。通过示例性但不是限制性说明，许多形式的 RAM 可用。例如，SRAM、DRAM、SDRAM、DDR SDRAM、ESDRAM、SLDRAM 和 DR RAM。

本申请实施例中，还提供了一种计算机可读存储介质，计算机可读存储介质存储有计算机指令，当计算机可读存储介质中存储的计算机指令被电子设备执行时，使得电子设备执行上述所提供的数据流识别方法。

本申请实施例中，还提供了一种包含指令的计算机程序产品，当其在电子设备上运行时，使得电子设备执行上述所提供的数据流识别方法。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线）或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，DVD）、或者半导体介质（例如固态硬盘 Solid State Disk）等。

本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成，也可以通过程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，上述提到的存储介质可以是只读存储器，磁盘或光盘等。

以上所述仅为本申请的可选实施例，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到的变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应该以权利要求的保护范围为准。

除非另作定义，此处使用的技术术语或者科学术语应当为本申请所属领域内具有一般技能的人士所理解的通常意义。本申请专利申请说明书以及权利要求书中使用的“第一”、“第二”、“第三”以及类似的词语并不表示任何顺序、数量或者重要性，而只是用来区分不同的组成部分。同样，“一个”或者“一”等类似词语也不表示数量限制，而是表示存在至少一个。“包括”或者“包含”等类似的词语意指出现在“包括”或者“包含”前面的元件或者物件涵盖出现在“包括”或者“包含”后面列举的元件或者物件及其等同，并不排除其他元件或者物件。

以上仅为本申请一个实施例，并不用以限制本申请，凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

权 利 要 求 书

1.一种数据流识别方法，其特征在于，所述方法包括：

获取第一数据流的指纹特征和建立时间；

根据所述第一数据流的指纹特征和建立时间，确定与所述第一数据流网络地址转换 NAT 相关的至少一个数据流。

2.根据权利要求1所述的方法，其特征在于，所述获取第一数据流的指纹特征和建立时间，包括：

向第一网络节点发送第一查询请求，所述第一查询请求包括所述第一数据流的第一流标识；

接收所述第一网络节点发送的第一查询应答，所述第一查询应答包括所述第一数据流的指纹特征和建立时间。

3.根据权利要求1或2所述的方法，其特征在于，所述根据所述第一数据流的指纹特征和建立时间，确定与所述第一数据流 NAT 相关的至少一个数据流，包括：

向第二网络节点发送第二查询请求，所述第二查询请求包括所述第一数据流的指纹特征；

接收所述第二网络节点发送的第二查询应答，所述第二查询应答包括与所述第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；

在所述第二数据流的建立时间与所述第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定所述第一数据流和所述第二数据流 NAT 相关。

4.根据权利要求2所述的方法，其特征在于，所述第一查询应答还包括与所述第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；

所述根据所述第一数据流的指纹特征和建立时间，确定与所述第一数据流 NAT 相关的至少一个数据流，包括：

在所述第二数据流的建立时间与所述第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定所述第一数据流和所述第二数据流 NAT 相关。

5.根据权利要求1所述的方法，其特征在于，所述获取第一数据流的指纹特征和建立时间，包括：

接收第一查询请求，所述第一查询请求包括所述第一数据流的第一流标识；

根据所述第一数据流的第一流标识确定所述第一数据流的指纹特征和建立时间。

6.根据权利要求1或5所述的方法，其特征在于，所述根据所述第一数据流的指纹特征和建立时间，确定与所述第一数据流 NAT 相关的至少一个数据流，包括：

根据所述第一数据流的指纹特征确定第二数据流，所述第二数据流的指纹特征与所述第一数据流的指纹特征相同；

在所述第二数据流的建立时间与所述第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定所述第一数据流和所述第二数据流 NAT 相关。

7.根据权利要求1至6任一项所述的方法，其特征在于，所述指纹特征包括如下特征中的至少一项：

首个数据包的互联网协议标识 IPID、首个数据包的载荷、首个数据包的载荷的哈希值、传输控制协议 TCP 三次握手过程中同步序列号 SYN 报文的 IPID、TCP 三次握手过程中 SYN 报文的初始序列码 ISN、TCP 三次握手过程中 SYN-ACK 报文的 ISN。

8.根据权利要求2、4或5所述的方法，其特征在于，所述第一查询请求还包括查询时间范围，所述查询时间范围用于限制查询到的数据流的建立时间所处的时间范围。

9.根据权利要求3、4或6所述的方法，其特征在于，所述在所述第二数据流的建立时间与所述第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定所述第一数据流和所述第二数据流 NAT 相关，包括：

在所述第二数据流的建立时间与所述第一数据流的建立时间的时间差的绝对值小于阈值，所述第一数据流和所述第二数据流不存在指纹冲突的情况下，确定所述第一数据流和所述第二数据流 NAT 相关，所述指纹冲突是指非 NAT 相关的数据流的指纹特征相同。

10.根据权利要求9所述的方法，其特征在于，所述方法还包括：

确定与所述第二数据流具有相同第二流标识的至少一条第三数据流；

确定与每条所述第三数据流具有相同指纹特征的至少一条第四数据流；

在所述至少一条第四数据流中存在一条第四数据流的建立时间和对应所述第三数据流的建立时间的时间差的绝对值小于阈值，且所述一条第四数据流的第二流标识和所述第一数据流的第二流标识相同的情况下，确定所述第一数据流和所述第二数据流不存在指纹冲突。

11.根据权利要求1至10任一项所述的方法，其特征在于，所述方法还包括：

在所述第一数据流发生故障的情况下，对所述第一数据流以及与所述第一数据流 NAT 相关的至少一个数据流进行故障诊断。

12.一种数据流识别方法，其特征在于，所述方法包括：

接收查询请求，所述查询请求包括第一数据流的第一流标识；

发送查询应答，所述查询应答包括所述第一数据流的指纹特征和建立时间，所述第一数据流的指纹特征和建立时间用于确定与所述第一数据流 NAT 相关的至少一个数据流。

13.根据权利要求12所述的方法，其特征在于，所述查询应答还包括与所述第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间。

14.根据权利要求12或13所述的方法，其特征在于，所述方法还包括：

采集各个数据流的建立时间、第一流标识和指纹特征；

存储所述各个数据流的建立时间、第一流标识和指纹特征。

15.根据权利要求14所述的方法，其特征在于，所述采集各个数据流的建立时间、第一流标识和指纹特征，包括：

通过镜像功能抓取数据流的报文；

分析抓取到的所述数据流的报文，以获取所述数据流的建立时间、第一流标识和指纹特征。

16.一种数据流识别装置，其特征在于，所述装置包括：

获取单元，用于获取第一数据流的指纹特征和建立时间；

确定单元，用于根据所述第一数据流的指纹特征和建立时间，确定与所述第一数据流 NAT 相关的至少一个数据流。

17.根据权利要求16所述的装置，其特征在于，所述获取单元，用于向第一网络节点发送第一查询请求，所述第一查询请求包括所述第一数据流的第一流标识；接收所述第一网络节点发送的第一查询应答，所述第一查询应答包括所述第一数据流的指纹特征和建立时间。

18.根据权利要求16或17所述的装置，其特征在于，所述确定单元，用于向第二网络节点发送第二查询请求，所述第二查询请求包括所述第一数据流的指纹特征；接收所述第二网络节点发送的第二查询应答，所述第二查询应答包括与所述第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；在所述第二数据流的建立时间与所述第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定所述第一数据流和所述第二数据流 NAT 相关。

19.根据权利要求17所述的装置，其特征在于，所述第一查询应答还包括与所述第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间；

所述确定单元，用于在所述第二数据流的建立时间与所述第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定所述第一数据流和所述第二数据流 NAT 相关。

20.根据权利要求16所述的装置，其特征在于，所述获取单元，用于接收第一查询请求，所述第一查询请求包括所述第一数据流的第一流标识；根据所述第一数据流的第一流标识确定所述第一数据流的指纹特征和建立时间。

21.根据权利要求16或20所述的装置，其特征在于，所述确定单元，用于根据所述第一数据流的指纹特征确定第二数据流，所述第二数据流的指纹特征与所述第一数据流的指纹特征相同；在所述第二数据流的建立时间与所述第一数据流的建立时间的时间差的绝对值小于阈值的情况下，确定所述第一数据流和所述第二数据流 NAT 相关。

22.根据权利要求16至21任一项所述的装置，其特征在于，所述指纹特征包括如下特征中的至少一项：

首个数据包的互联网协议标识 IPID、首个数据包的载荷、首个数据包的载荷的哈希值、传输控制协议 TCP 三次握手过程中 SYN 报文的 IPID、TCP 三次握手过程中 SYN 报文的初始序列码 ISN、TCP 三次握手过程中 SYN-ACK 报文的 ISN。

23.根据权利要求 17、19 或 20 所述的装置，其特征在于，所述第一查询请求还包括查询时间范围，所述查询时间范围用于限制查询到的数据流的建立时间所处的时间范围。

24.根据权利要求 18、19 或 21 所述的装置，其特征在于，所述确定单元，用于在所述第二数据流的建立时间与所述第一数据流的建立时间的时间差的绝对值小于阈值，所述第一数据流和所述第二数据流不存在指纹冲突的情况下，确定所述第一数据流和所述第二数据流 NAT 相关，所述指纹冲突是指非 NAT 相关的数据流的指纹特征相同。

25.根据权利要求 24 所述的装置，其特征在于，所述确定单元，还用于确定与所述第二数据流具有相同第二流标识的至少一条第三数据流；确定与每条所述第三数据流具有相同指纹特征的至少一条第四数据流；在所述至少一条第四数据流中存在一条第四数据流的建立时间和对应所述第三数据流的建立时间的时间差的绝对值小于阈值，且所述一条第四数据流的第二流标识和所述第一数据流的第二流标识相同的情况下，确定所述第一数据流和所述第二数据流不存在指纹冲突。

26.根据权利要求 16 至 25 任一项所述的装置，其特征在于，所述装置还包括：

诊断单元，用于在所述第一数据流发生故障的情况下，对所述第一数据流以及与所述第一数据流 NAT 相关的至少一个数据流进行故障诊断。

27.一种数据流识别装置，其特征在于，所述装置包括：

接收单元，用于接收查询请求，所述查询请求包括第一数据流的第一流标识；

发送单元，用于发送查询应答，所述查询应答包括所述第一数据流的指纹特征和建立时间，所述第一数据流的指纹特征和建立时间用于确定与所述第一数据流 NAT 相关的至少一个数据流。

28.根据权利要求 27 所述的装置，其特征在于，所述查询应答还包括与所述第一数据流的指纹特征相同的第二数据流的第一流标识以及建立时间。

29.根据权利要求 27 或 28 所述的装置，其特征在于，所述装置还包括：

采集单元，用于采集各个数据流的建立时间、第一流标识和指纹特征；

存储单元，用于存储所述各个数据流的建立时间、第一流标识和指纹特征。

30.根据权利要求 29 所述的装置，其特征在于，所述采集单元，用于通过镜像功能抓取数据流的报文；分析抓取到的所述数据流的报文，以获取所述数据流的建立时间、第一流标识和指纹特征。

31.一种电子设备，其特征在于，所述电子设备包括处理器和存储器，所述存储器用于存储软件程序，所述处理器通过运行或执行存储在所述存储器内的软件程序，以使所述电子设备实现如权利要求 1 至 15 任一项所述的方法。

32.一种数据流识别系统，其特征在于，所述数据流识别系统包括采集器和分析平台，所述分析平台用于执行如权利要求 1 至 11 任一项所述的方法，所述采集器用于执行如权利要求 12 至 15 任一项所述的方法。

33.根据权利要求 32 所述的系统，其特征在于，所述数据流识别系统还包括存储平台，用于存储所述采集器采集的数据流的信息。

34.一种计算机可读存储介质，其特征在于，所述计算机可读存储介质用于存储处理器所执行的程序代码，所述程序代码包括用于实现如权利要求 1 至 15 任一项所述的方法的指令。

35.一种计算机程序产品，其特征在于，包括程序代码，当计算机运行所述计算机程序产品时，使得所述计算机执行如权利要求 1 至 15 任一项所述的方法。

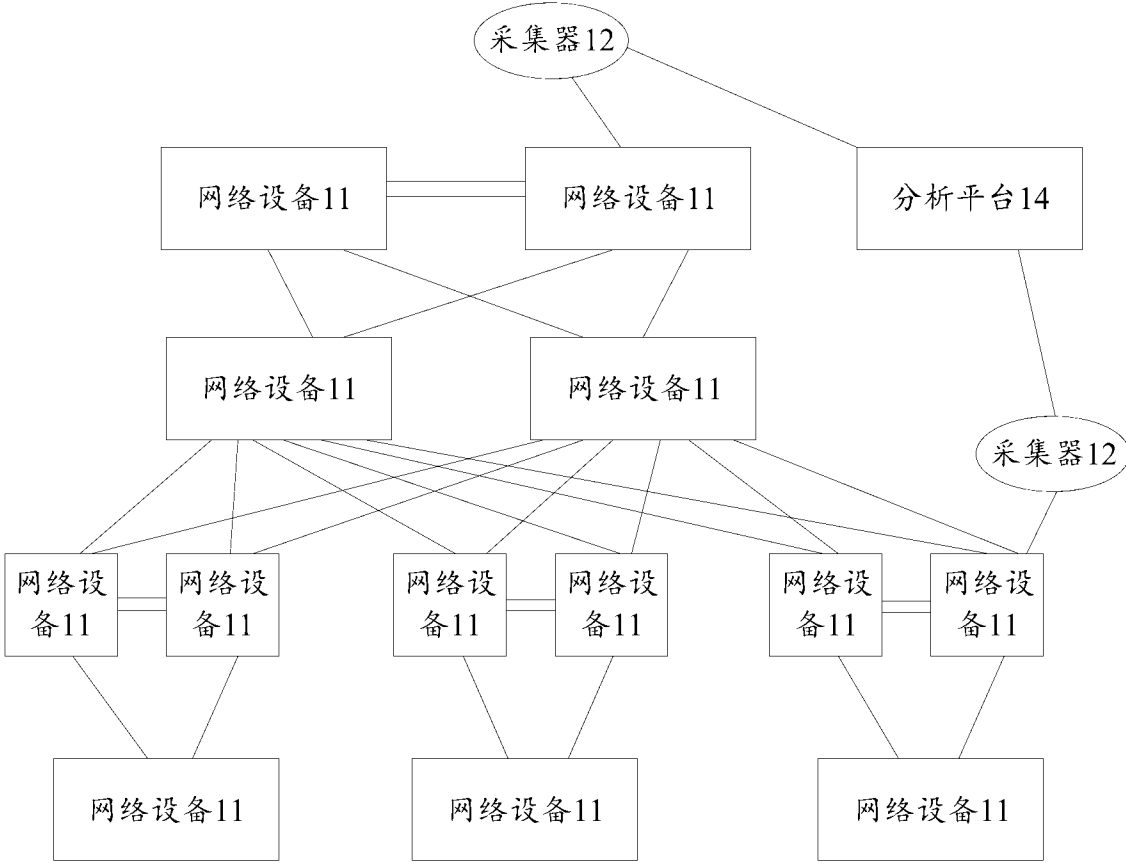


图 1

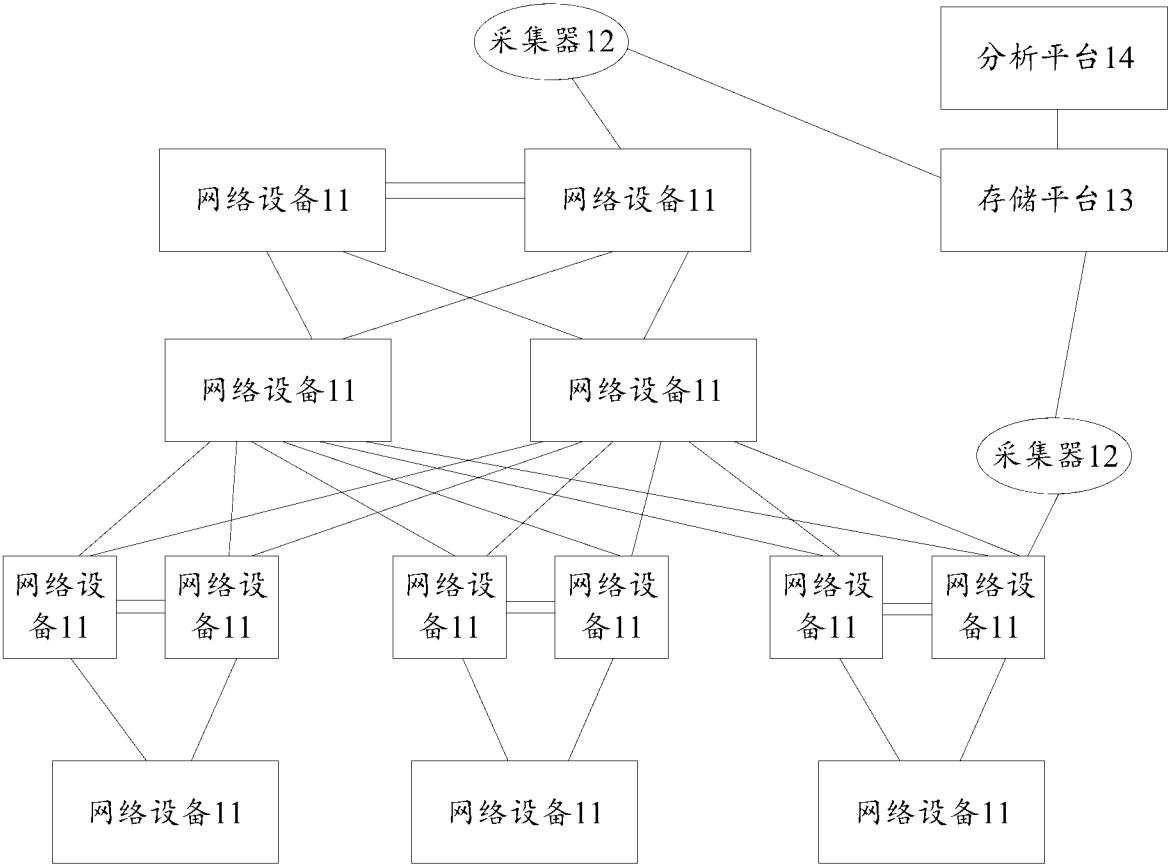


图 2

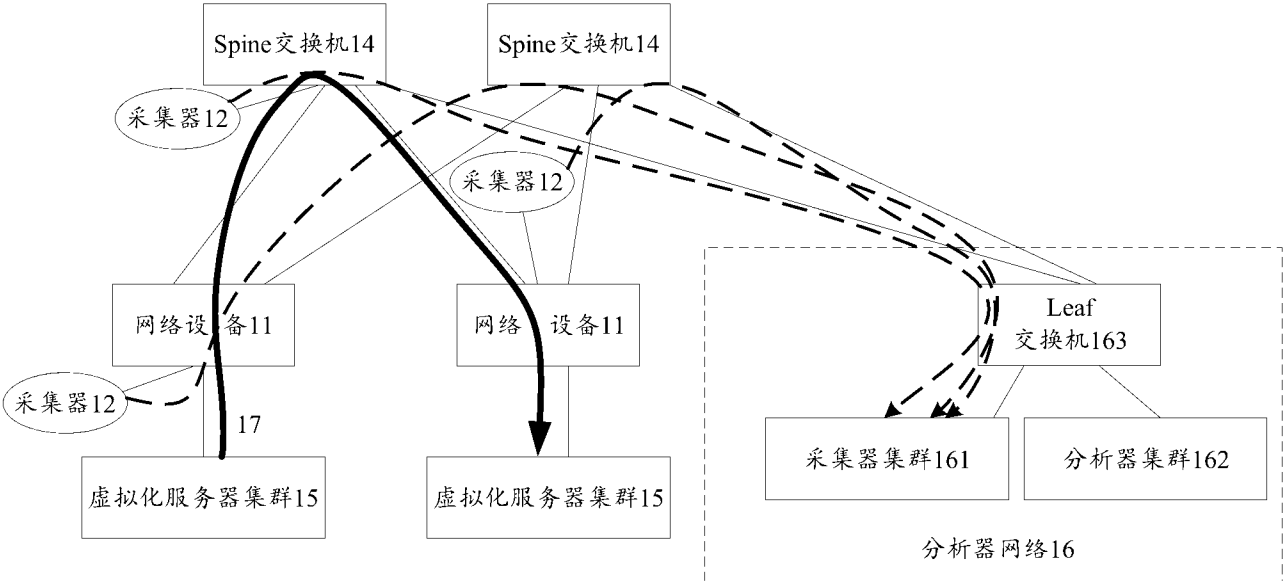


图 3

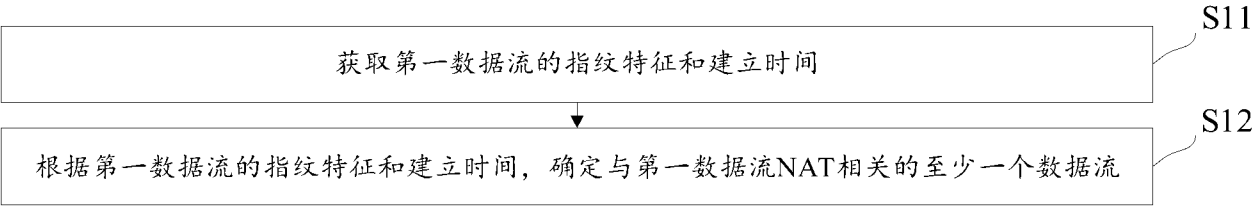


图 4

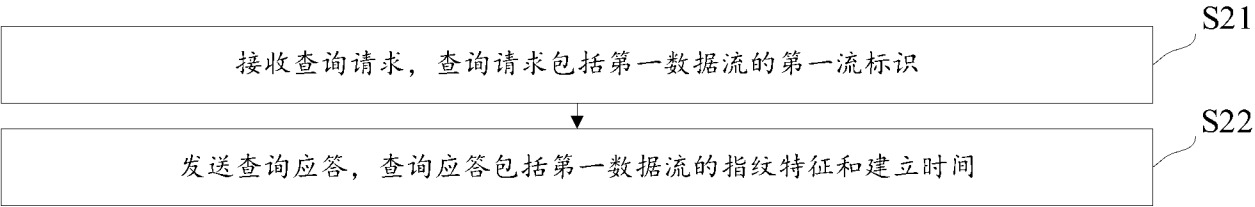


图 5

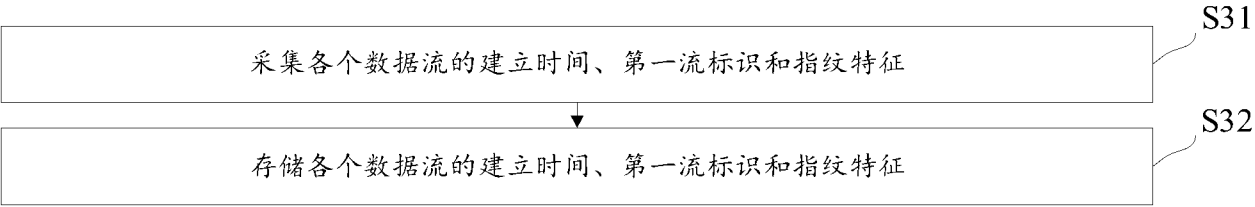


图 6

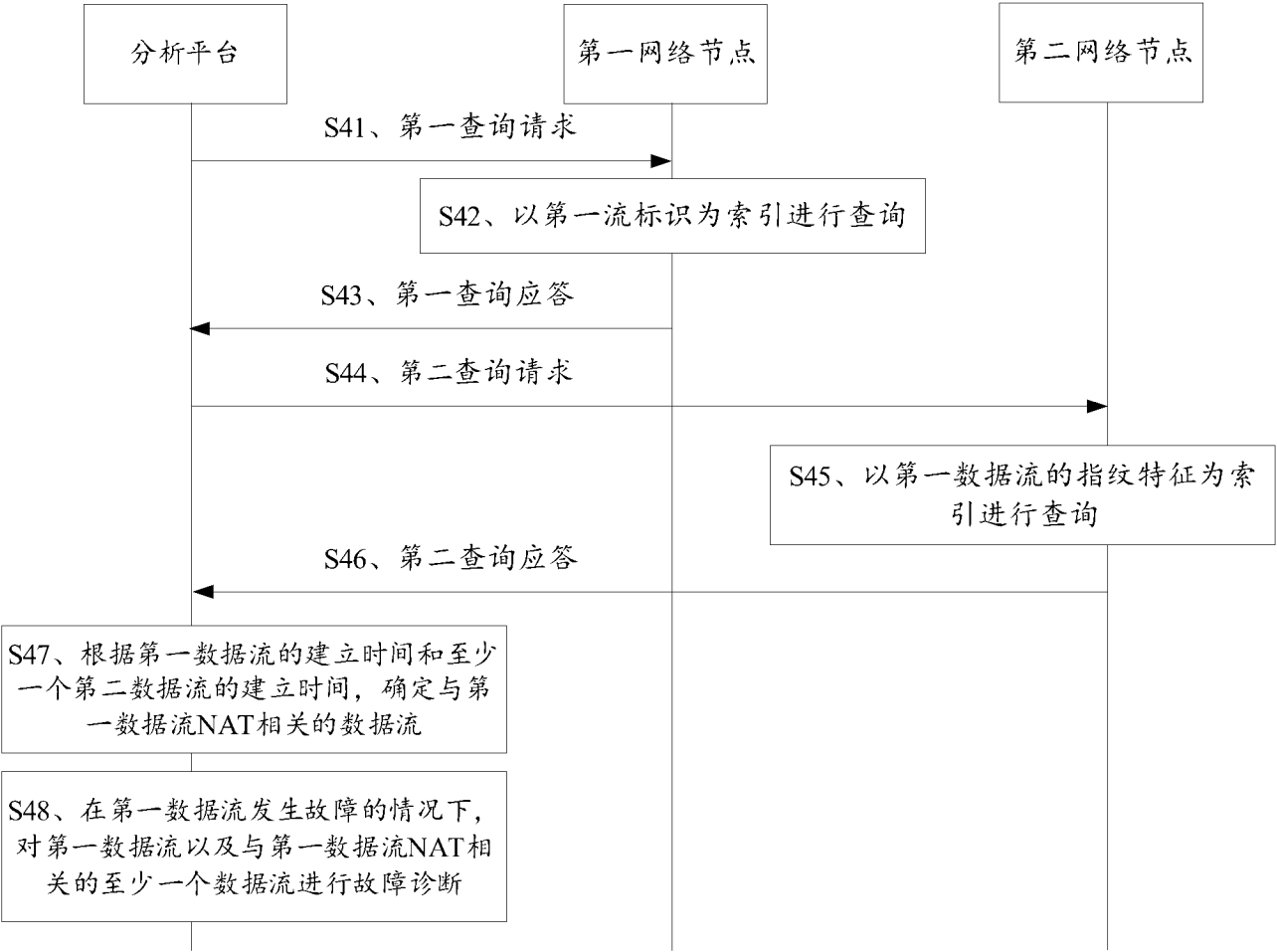


图 7

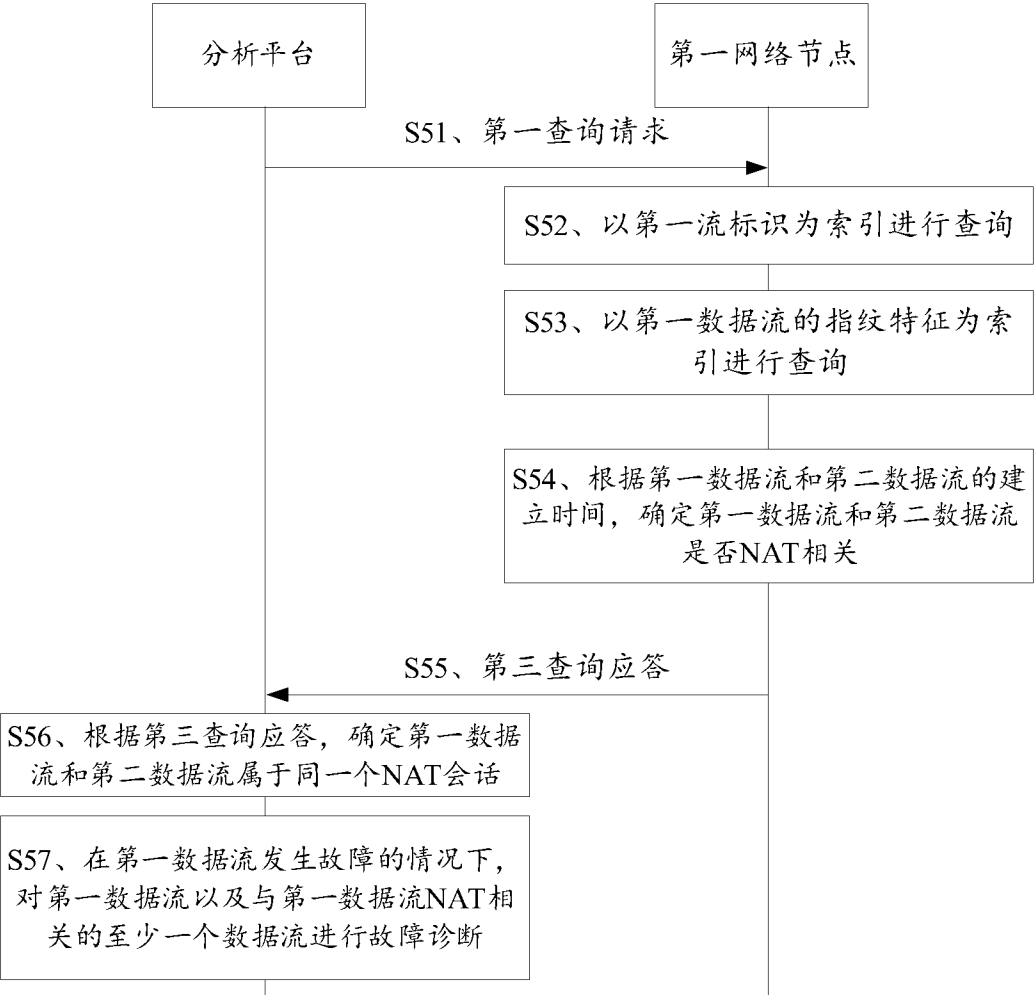


图 8

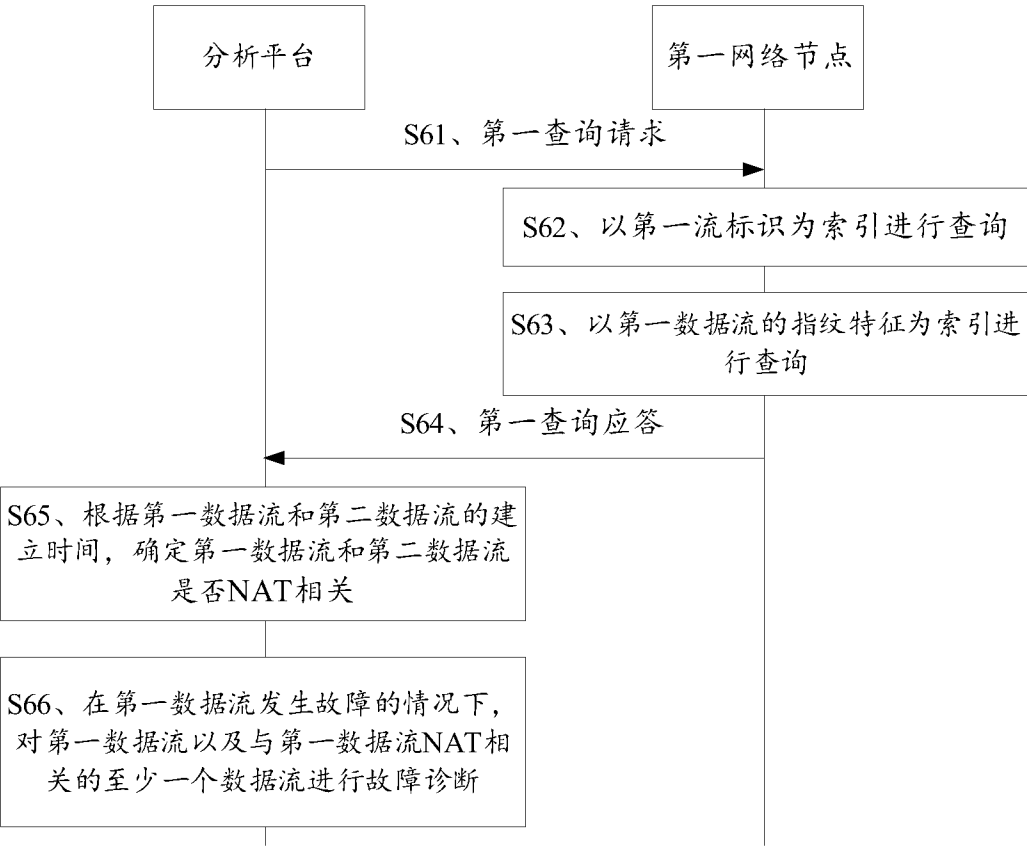


图 9

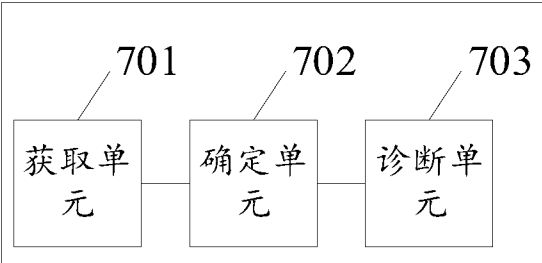


图 10

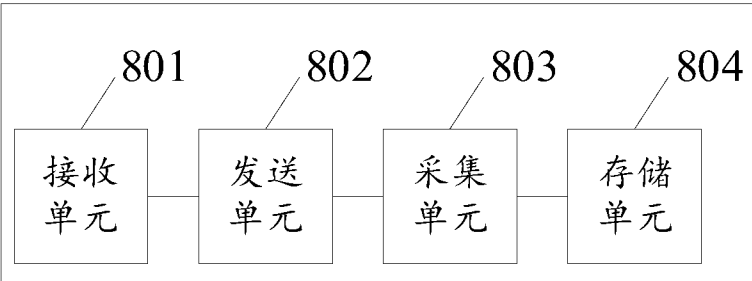


图 11

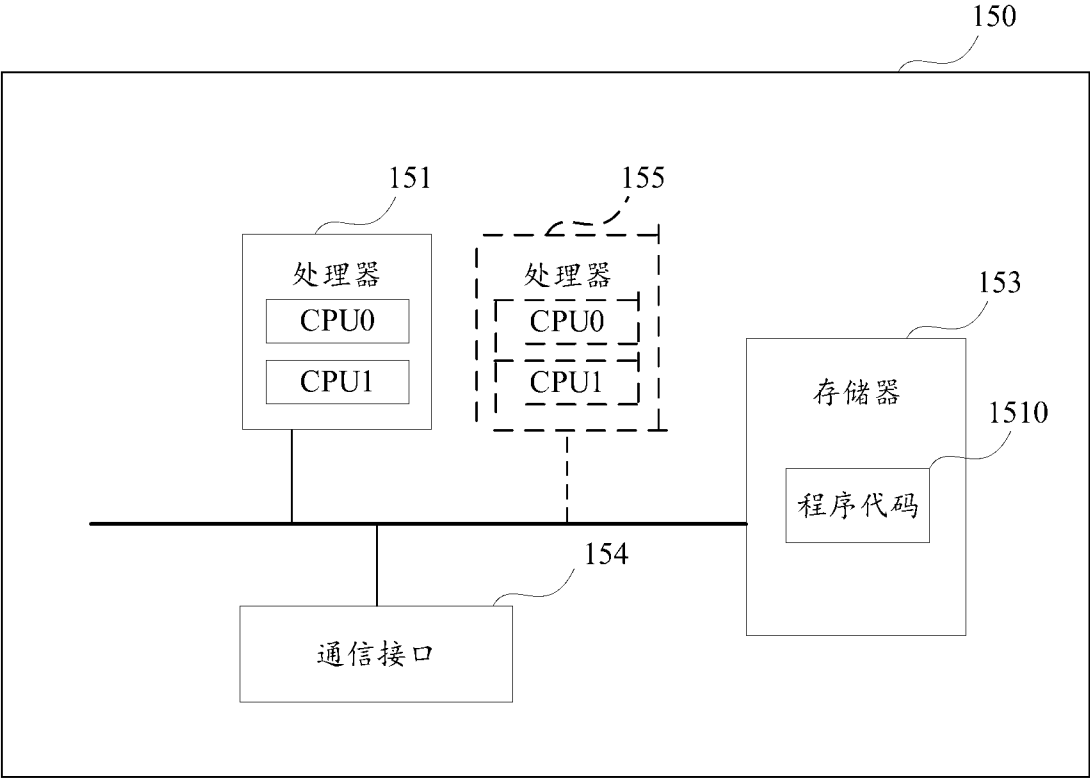


图 12

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/073409

A. CLASSIFICATION OF SUBJECT MATTER

H04L41/06(2022.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04L, H04W

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI, ENTXT, DWPI: 数据流, 识别, 指纹, 特征, 建立, 时间, 转换, 局域网, 内网, 载荷, 哈希, 标识, 协议, 阈值; NAT, stream, recogni+, fingerprint, establish, feature, characteristic, network, address, translat+, time, payload, hash, threshold

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 112468373 A (WUHAN ZHIYI TECHNOLOGY CO., LTD.) 09 March 2021 (2021-03-09) description, paragraphs 23-28	1-2, 5, 7-8, 11-17, 20, 22-23, 26-35
Y	宫博; 王汝传 (GONG, Bo; WANG, Ruchuan). "基于会话的局域网P2P流量识别方法的研究 (Research on Method of Session-Based P2P Network Traffic Identification)" <i>计算机技术与发展 (Computer Technology and Development)</i> , No. 03, 10 March 2010 (2010-03-10), section 0-2, and figure 4	1-2, 5, 7-8, 11-17, 20, 22-23, 26-35
A	CN 101150505 A (HANGZHOU H3C TECHNOLOGIES CO., LTD.) 26 March 2008 (2008-03-26) entire document	1-35
A	CN 114915566 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 16 August 2022 (2022-08-16) entire document	1-35



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

23 April 2024

Date of mailing of the international search report

24 April 2024

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/073409

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2019114350 A1 (SALESFORCE.COM, INC.) 18 April 2019 (2019-04-18) entire document	1-35

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/073409

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112468373	A	09 March 2021	None			
CN	101150505	A	26 March 2008	None			
CN	114915566	A	16 August 2022	None			
US	2019114350	A1	18 April 2019	US	10685034	B2	16 June 2020

A. 主题的分类

H04L41/06(2022.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: H04L, H04W

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, CNTXT, CNKI, ENTXT, DWPI: 数据流, 识别, 指纹, 特征, 建立, 时间, 转换, 局域网, 内网, 载荷, 哈希, 标识, 协议, 阈值; NAT, stream, recogni+, fingerprint, establish, feature, characteristic, network, address, translat+, time, payload, hash, threshold

C. 相关文件

类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	CN 112468373 A (武汉蚬易科技有限公司) 2021年3月9日 (2021 - 03 - 09) 说明书第23-28段	1-2,5,7-8,11-17 ,20,22-23,26-35
Y	官博;王汝传; "基于会话的局域网P2P流量识别方法的研究" 计算机技术与发展, 第03期, 2010年3月10日 (2010 - 03 - 10), 第0-2部分, 图4	1-2,5,7-8,11-17 ,20,22-23,26-35
A	CN 101150505 A (杭州华三通信技术有限公司) 2008年3月26日 (2008 - 03 - 26) 全文	1-35
A	CN 114915566 A (腾讯科技(深圳)有限公司) 2022年8月16日 (2022 - 08 - 16) 全文	1-35
A	US 2019114350 A1 (SALESFORCE COM INC) 2019年4月18日 (2019 - 04 - 18) 全文	1-35

☐ 其余文件在C栏的续页中列出。

☒ 见同族专利附件。

★ 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“p” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2024年4月23日

国际检索报告邮寄日期

2024年4月24日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

曹元娜

电话号码 (+86) 010-62089367

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/073409

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	112468373	A	2021年3月9日	无			
CN	101150505	A	2008年3月26日	无			
CN	114915566	A	2022年8月16日	无			
US	2019114350	A1	2019年4月18日	US	10685034	B2	2020年6月16日