(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 104994016 A
(43)申请公布日 2015.10.21

(21)申请号 201510019242.X

(22)申请日 2015.01.14

(30)优先权数据
61/927,266 2014.01.14 US

(71)申请人 马维尔国际有限公司
地址 百慕大群岛哈密尔顿

(72)发明人 S·坎皮斯 G·纳丘姆

(74)专利代理机构 北京市金杜律师事务所 11256
代理人 酆迅 庞淑敏

(51)Int.Cl.
H04L 12/70(2013.01)
H04L 29/06(2006.01)

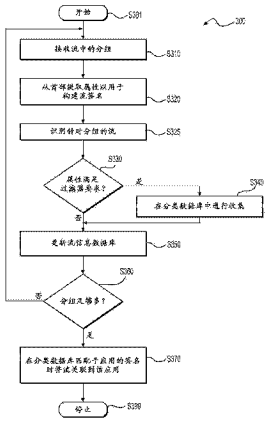
权利要求书2页 说明书7页 附图4页

(54)发明名称

用于分组分类的方法和装置

(57)摘要

本公开的方面提供了一种用于分组分类的方法和装置。该方法包括：在网络设备的计算机存储器中，存储针对网络应用的分组流的特征签名。所述特征签名包括在针对网络应用的分组流中的两个或更多分组的两个或更多分组属性的模式。然后，该方法包括：在网络设备处接收网络分组的流，识别网络分组的流中的一个或多个分组流，在分组处理器处处理分组，以获得在相应的分组流中的分组的分组属性，以及在分组流中的分组的分组属性与给定网络应用的特征签名相匹配时，将分组流识别为与给定网络应用相关联。



CN 104994016 A

1. 一种用于分组分类的方法，包括：

在网络设备的计算机存储器中，存储针对网络应用的分组流的特征签名，所述分组流包括两个或更多分组，并且特征签名包括在针对网络应用的分组流中的所述两个或更多分组的两个或更多分组属性的模式；

在所述网络设备处，接收网络分组的流；

识别网络分组的所述流中的一个或多个分组流；

在分组处理器处，处理所述分组，以获得相应的所述分组流中的分组的分组属性；以及

在分组流中的分组的所述分组属性与给定网络应用的所述特征签名相对应时，将所述分组流识别为与所述给定网络应用相关联。

2. 根据权利要求 1 所述的方法，其中在所述分组处理器处处理所述分组以获得相应的所述分组流中的所述分组的所述分组属性还包括：

从所述分组流中的所述分组的首部提取一个或多个分组属性。

3. 根据权利要求 2 所述的方法，其中从所述分组流中的所述分组的所述首部提取所述一个或多个分组属性还包括：

对分组的首部进行位屏蔽以提取所述分组的有效载荷大小。

4. 根据权利要求 2 所述的方法，其中从所述分组流中的所述分组的所述首部提取所述一个或多个分组属性还包括：

对分组的首部进行位屏蔽以提取所述分组的总长度以及一个或多个首部长度；以及

通过从所述总长度减去所述首部长度来计算有效载荷大小。

5. 根据权利要求 1 所述的方法，其中在所述分组处理器处处理所述分组以获得相应的所述分组流中的所述分组的所述分组属性还包括：

获得所述分组属性而不检查所述分组的有效载荷。

6. 根据权利要求 1 所述的方法，其中在所述网络设备的所述计算机存储器中存储针对所述网络应用的所述分组流的所述特征签名还包括：

在所述网络设备的所述计算机存储器中存储以下各项中的至少一项：在接收到具有特定有效载荷大小的分组时的时间的模式，以及分组流中的索引的模式，在所述索引处的所述分组具有特定有效载荷大小。

7. 根据权利要求 1 所述的方法，其中在所述网络设备的所述计算机存储器中存储针对所述网络应用的所述分组流的所述特征签名还包括：

在所述网络设备的所述计算机存储器中存储指定以下各项的两个或更多分组属性的模式的所述特征签名：具有特定有效载荷大小的分组的字节速率、具有特定大小的分组在时域中的到达速率、具有特定大小的分组在位置域中的到达速率、具有特定大小的分组之间的到达间隔时间、具有特定大小的分组之间的到达间隔间隙、有效载荷大小、分组速率、字节速率、到达间隔时间，分组数目。

8. 根据权利要求 1 所述的方法，其中在所述分组流中的分组的所述分组属性与所述给定网络应用的所述特征签名相匹配时将所述分组流识别为与所述给定网络应用相关联还包括：

根据基于所述分组属性、通过分支决策序列而进行最终决策的决策树，将所述分组流识别为与所述给定网络应用相关联。

9. 根据权利要求 8 所述的方法，还包括 ：

构建基于所述分组属性、通过分支决策序列而将特征签名与所述网络应用相关联的所述决策树。

10. 根据权利要求 1 所述的方法，还包括 ：

基于所识别的与所述分组相关联的所述网络应用，对分组采取动作。

11. 一种网络设备，包括 ：

存储器，被配置为存储针对网络应用的分组流的特征签名，所述分组流包括两个或更多分组，并且特征签名包括在针对网络应用的分组流中的所述两个或更多分组的两个或更多分组属性的模式 ；

端口，被配置为接收分组的流 ；以及

分组处理器，被配置为识别网络分组的所述流中的一个或多个分组流，处理所述分组以获得相应的所述分组流中的分组的分组属性 ；并且在分组流中的分组的所述分组属性与给定网络应用的所述特征签名相对应时，将所述分组流识别为与所述给定网络应用相关联。

12. 根据权利要求 11 所述的网络设备，其中所述分组处理器被配置为从所述分组流中的所述分组的首部提取一个或多个分组属性。

13. 根据权利要求 12 所述的网络设备，其中所述分组处理器被配置为对分组的首部进行位屏蔽以提取所述分组的有效载荷大小。

14. 根据权利要求 12 所述的网络设备，其中所述分组处理器被配置为对分组的首部进行位屏蔽以提取所述分组的总长度以及一个或多个首部长度，并且通过从所述总长度减去所述首部长度来计算有效载荷大小。

15. 根据权利要求 11 所述的网络设备，其中所述分组处理器被配置为获得所述分组属性而不检查所述分组的有效载荷。

16. 根据权利要求 11 所述的网络设备，其中所述存储器被配置为存储以下各项中的至少一项 ：在接收到具有特定有效载荷大小的分组时的时间的模式，以及分组流中的索引的模式，在所述索引处的所述分组具有特定有效载荷大小。

17. 根据权利要求 11 所述的网络设备，其中所述存储器被配置为存储指定以下各项的两个或更多分组属性的模式的所述特征签名 ：具有特定有效载荷大小的分组的字节速率、具有特定大小的分组在时域中的到达速率、具有特定大小的分组在位置域中的到达速率、具有特定大小的分组之间的到达间隔时间、具有特定大小的分组之间的到达间隔间隙、有效载荷大小、分组速率、字节速率、到达间隔时间，以及分组数目。

18. 根据权利要求 11 所述的网络设备，其中所述分组处理器被配置为根据基于所述分组属性、通过分支决策序列而进行最终决策的决策树，将所述分组流识别为与所述给定网络应用相关联。

19. 根据权利要求 18 所述的网络设备，其中所述分组处理器被配置为构建基于所述分组属性、通过分支决策序列而将特征签名与所述网络应用相关联的所述决策树。

20. 根据权利要求 11 所述的网络设备，其中所述分组处理器被配置为基于所识别的与所述分组相关联的所述网络应用，对分组采取动作。

用于分组分类的方法和装置

[0001] 引用并入

[0002] 本公开要求在2014年1月14日提交的、题为“Packet Capture by Size in a Packet Processor”的第61/927,266号美国临时申请的权益，其整体以引用的方式并入于此。

背景技术

[0003] 这里提供的背景技术描述是为了一般性地呈现本公开的上下文的目的。当前指定的发明人的工作（到在此背景技术部分描述的工作的程度）以及在提交时可能以其他方式不适合作为现有技术的本描述的诸多方面，既并非明确地也并非隐含地被承认为是本公开的现有技术。

[0004] 在各种场景中，网络设备基于层7分类来对分组采取动作。在一个示例中，网络设备接收分组，并且执行深度分组有效载荷检查，以搜索分组中的特定字符串，从而确定所述分组所关联的网络应用，并且然后根据网络应用的策略对分组采取动作。

发明内容

[0005] 本公开的方面提供了一种用于分组分类的方法。该方法包括：在网络设备的计算机存储器中，存储针对网络应用的分组流的特征签名。分组流包括两个或更多分组，并且特征签名包括在针对网络应用的分组流中的两个或更多分组属性的模式。然后，该方法包括：在网络设备处，接收网络分组的流；识别网络分组的流中的一个或多个分组流；在分组处理器处，处理分组，以获得相应的分组流中的分组的分组属性；以及，在分组流中的分组的分组属性与给定网络应用的特征签名相对应时，将分组流识别为与给定网络应用相关联。

[0006] 根据本公开的一个方面，该方法包括：从分组流中的分组的首部提取一个或多个分组属性。在一个示例中，该方法包括：对分组的首部进行位屏蔽(bit masking)以提取分组的有效载荷大小。在另一示例中，该方法包括：对分组的首部进行位屏蔽以提取分组的总长度以及一个或多个首部长度，以及通过从总长度减去首部长度来计算有效载荷大小。

[0007] 在一个实施例中，该方法包括：在网络设备的计算机存储器中存储以下各项中的至少一项：在接收到具有特定有效载荷大小的分组时的时间的模式，以及分组流中的索引的模式，在索引处的分组具有特定有效载荷大小。在另一实施例中，该方法包括：在网络设备的计算机存储器中存储指定以下各项的两个或更多分组属性的模式的特征签名：具有特定有效载荷大小的分组的字节速率、具有特定大小的分组在时域中的到达速率、具有特定大小的分组在位置域中的到达速率、具有特定大小的分组之间的到达间隔时间、具有特定大小的分组之间的到达间隔间隙、有效载荷大小、分组速率、字节速率、到达间隔时间、分组数目。

[0008] 根据本公开的一个方面，该方法包括：根据基于分组属性、通过分支决策序列而进行最终决策的决策树，将分组流识别为与给定网络应用相关联。在一个示例中，该方法包括：构建基于分组属性、通过分支决策序列而将特征签名与网络应用相关联的决策树。

[0009] 本公开的方面提供了一种网络设备，所述网络设备包括存储器、端口以及分组处理器。所述存储器被配置为存储针对网络应用的分组流的特征签名，分组流包括两个或更多分组，并且特征签名包括在针对网络应用的分组流中的两个或更多分组的两个或更多分组属性的模式。所述端口被配置为接收分组的流。所述分组处理器被配置为识别网络分组的流中的一个或多个分组流，处理分组以获得相应的分组流中的分组的分组属性，并且在分组流中的分组的分组属性与给定网络应用的特征签名相对应时，将分组流识别为与给定网络应用相关联。

附图说明

[0010] 将参照以下附图来详细描述本公开中作为示例而提出的各个实施例，其中相同的附图标记指代相同的元件，并且其中：

[0011] 图 1 示出了根据本公开的实施例的网络系统 100 的框图；

[0012] 图 2A-2B 示出了根据本公开的实施例的与两个网络应用相关联的签名；

[0013] 图 3 示出了概述根据本公开的实施例的用于业务 (traffic) 分类的过程 300 的流程图；并且

[0014] 图 4 示出了概述根据本公开的实施例的用于签名学习的过程 400 的流程图。

具体实施方式

[0015] 图 1 示出了根据本公开的实施例的网络系统 100 的框图。网络系统 100 包括网络设备 130，该网络设备 130 将第一网络 110 和第二网络 120 通信耦合在一起，如图 1 所示。网络设备 130 包括基于流签名的应用识别单元 140，该应用识别单元 140 被配置为基于识别属性签名，根据分组的首部，执行对诸如分组的流之类的网络业务流的层 7 分类，而不检查分组的有效载荷。

[0016] 第一网络 110 和第二网络 120 中的每个网络可以是单个网络或者是相同或不同类型的多个网络，比如数据网络、电信网络、视频分发（例如，线缆、地面广播，或者卫星）网络、电信、视频 / 音频分发和数据网络的组合、全球网络、国家网络、区域网络、广域网、局域网、家用网络等等。网络设备 130 可以是任何适合的网络设备，比如路由器、交换机、调制解调器、因特网协议 (IP) 机顶盒等等。

[0017] 出于讨论目的，在一个示例中，第一网络 110 是因特网，并且第二网络 120 是局域网 (LAN)，并且网络设备 130 是被配置为提供至第二网络 120 的进入点的边缘设备，比如边缘路由器。

[0018] 在一个实施例中，第一网络 110 包括被配置为提供资源和 / 或服务的多个服务器主机，并且第二网络 120 包括被配置为请求资源或服务的多个客户端设备，比如桌上型计算机，膝上型计算机等。在一个示例中，每个服务器主机执行服务器程序以将服务器主机的资源向客户端分享。另外，在该示例中，每个客户端设备执行一个或多个网络应用软件以发起通信会话，从而请求和接收来自服务器主机的资源和服务。

[0019] 根据本公开的一个方面，网络设备 130 被配置为根据发起分组的网络应用对分组执行动作。在一个实施例中，网络设备 130 被配置为向与不同的网络应用相关联的分组应用不同的策略。

[0020] 在一个示例中，网络设备 130 被配置为阻挡从特定网络应用发起的业务（例如分组）。例如，网络设备 130 被配置为阻挡社交媒体应用或已知包含色情或恶意内容的应用的业务（例如，丢弃分组）。在另一示例中，网络设备 130 被配置为基于发起分组的网络应用来对分组定义服务质量。例如，网络设备 130 被配置为对因特网协议语音 (VOIP) 电话应用的分组定义相对高的服务质量，并且对在线视频游戏应用的分组定义相对低的服务质量。

[0021] 另外，根据本公开的一个方面，网络设备 130 被配置为使用接收的分组的首部中的信息来进行分组分类，该分组分类将分组与发起该分组的网络应用相关联，而无需深度分组检查 (packet inspection)。在一个实施例中，网络设备 130 基于从首部或者从诸如接收的时间之类的分组特性获得的信息，存储针对一个或多个网络应用的分组流的签名。在一个示例中，针对网络应用的签名指定针对网络应用的一个或多个属性的特定模式。在一个示例中，从分组首部或者从分组的外在特性来提取属性，而无需由网络设备 130 中的中央处理单元 (CPU)（未示出）进行的密集计算。例如，在网络设备 130 接收到分组的流时，网络设备 130 从分组的首部提取一个或多个属性，将分组分类成分组流，确定诸如分组的到达时间之类的一个或多个外在属性，并且基于从在给定分组流中的分组集合中选择的属性，形成针对新接收的流的模式。在该模式匹配所存储的针对网络应用的特定模式时，分组流与网络应用相关联。然后，基于相关联的网络应用，网络设备 130 对分组流中的分组执行一个或多个动作。

[0022] 在一个实施例中，根据协议发送针对网络应用的网络业务，并且不同的网络应用通常使用不同的可分开识别的协议。因而，在一个示例中，在会话发起期间，与应用相对应的分组基于应用所使用的协议具有形成特定的可区分的模式（例如，签名）的属性，比如有效载荷大小、分组速率、字节速率，到达间隔时间、分组数目等等。在一个实施例中，这样的属性因而指示网络应用。在一个示例中，流内具有特定有效载荷大小的分组以特定的定义时间或以流中的特定索引来发送，并且因而指示应用所使用的协议。在另一示例中，具有特定有效载荷大小的分组在流内的到达速率（例如，每三秒发送大小为零的分组）指示协议和网络应用。在另一示例中，在序列内的具有特定有效载荷大小的分组在流内的到达速率（例如，每三个分组发送大小为零的分组）指示协议和网络应用。在另一示例中，在流内的具有特定有效载荷大小的分组之间的间隔时间指示被用于发送分组的协议和网络应用。在另一示例中，在流内具有特定有效载荷大小的分组之间的到达间隔间隙指示协议和网络应用。

[0023] 具体地，在图 1 的示例中，网络设备 130 包括：用于存储一个或多个网络应用的签名的存储器 160；被配置为接收属性的接收单元 135；用于提取分组的首部中的属性的首部属性提取单元 140；用于识别分组分别所属的流的流识别单元 145，用于基于属性模式进行层 7 分类的应用识别单元 150，以及用于根据发起分组的网络应用来对分组进行操作的基于应用的策略执行单元 170。这些元件如图 1 所示被耦合在一起。

[0024] 存储器 160 可以任何适合的存储设备，比如静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、闪速存储器、固态驱动器、硬盘驱动器、光盘驱动器等。可以按照任何适合的数据结构、诸如表格、列表等来存储签名。在一个示例中，按照促进决策树的形式存储签名，该决策树基于分组属性、通过分支决策序列而进行最终决策。

[0025] 接收单元 135 被配置为接收分组并将接收信息与分组相关联。在一个示例中，接

收单元 135 包括用于接收分组的进入端口，并且将进入端口号与经由进入端口进入网络设备 130 的分组相关联。另外，在一个示例中，接收单元 135 包括适合的定时电路，该定时电路被配置为捕捉分组的到达时间，并且适合地将到达时间与分组相关联。

[0026] 首部属性提取单元 140 被配置为从分组的首部提取属性。在一个实施例中，首部属性提取单元 140 被配置为从分组的首部提取每个分组的有效载荷大小。在一个示例中，在分组根据 IPv4 协议编排格式时，分组的首部包括指示分组的总长度的第一字段、指示 IP 首部长度的第二字段以及指示 TCP 首部长度的第三字段。在一个示例中，首部属性提取单元 140 使用位屏蔽，从首部的第一字段、第二字段和第三字段提取分组的总长度、IP 首部长度和 TCP 首都长度。然后首部属性提取单元 140 从总长度减去 IP 首部长度和 TCP 首部长度以计算分组的有效载荷大小。在另一示例中，在分组根据 IPv6 协议编排格式时，首部具有用于有效载荷长度的专用字段。继而，首部属性提取单元 140 使用位屏蔽来提取该有效载荷长度。

[0027] 在一个实施例中，首部属性提取单元 140 提取其他适合的信息，比如首部中的序列号、端口号等。另外，首部属性提取单元 140 基于提取信息计算适合的属性值。此外，在一个示例中，记录相应的分组到达时间。首部属性提取单元 140 计算例如具有特定有效载荷大小的分组的字节速率、具有特定有效载荷大小的分组在时域中的到达速率、具有特定有效载荷大小的分组在位置域中的到达速率、具有特定大小的分组之间的到达间隔时间、具有特定大小的分组之间的到达间隔间隙、分组速率、字节速率、到达间隔时间、分组数目等。

[0028] 流分类单元 145 被配置为基于首部属性和其他适合的信息，将分组分类成流，所述其他适合的信息诸如为源地址、目的地址、源端口、目的端口、协议类型等。

[0029] 在一个实施例中，应用识别单元 150 被配置为基于针对网络应用的签名来执行分组分类。在一个示例中，应用识别单元 150 构建来自分组流中的分组的属性的模式，并且将该模式与存储的签名进行比较。在模式匹配于所存储的针对网络应用的签名时，应用识别单元 150 将分组流与匹配的网络应用相关联。在另一示例中，按照决策树的形式来存储签名以区分网络应用。应用识别单元 150 使用决策树来执行分组分类。

[0030] 策略执行单元 170 然后根据分类，基于对所识别的应用适用的策略，对分组采取动作。在一个示例中，策略执行单元 170 丢弃特定网络应用的分组。在另一示例中，策略执行单元 170 向特定网络应用的分组分配服务质量。在另一示例中，策略执行单元 170 制作从特定网络应用发起的分组的副本，并且将该副本发送给在第二网络 120 中的监测设备（未示出），以例如用于进一步的监测和分类。

[0031] 根据本公开的一个方面，网络设备 130 包括签名学习单元 180，该签名学习单元 180 被配置为对先前未学习的网络应用的签名进行学习。在一个示例中，另外在网络设备 130 并不繁忙时，网络设备 130 接收由网络应用发起的分组流，并且经由机器学习来学习网络应用的签名。例如，首部属性提取单元 140 从分组的首部提取分组属性并且提供给签名学习单元 180。签名学习单元 180 接收分组属性，检查分组流中的分组的有效载荷以识别与分组流相关联的网络应用，并且构建基于分组属性、将所述网络应用与其他先前已学习的网络应用进行区分的决策树。

[0032] 注意，首部属性提取单元 140、流分类单元 145、应用识别单元 150、策略执行单元 170 和签名学习单元 180 可以分别以各种技术来实现。在一个实施例中，使用分组处理器

中的电路来实现首部属性提取单元 140、流分类单元 145、应用识别单元 150、策略执行单元 170 和签名学习单元 180。在另一实施例中，首部属性提取单元 140、流分类单元 145、应用识别单元 150、策略执行单元 170 和签名学习单元 180 被实现为由诸如中央处理单元之类的处理器执行的软件指令。

[0033] 注意，在一个实施例中，网络设备 130 使用其他适合的技术来确定应用特定信息，而不执行深度分组有效载荷检查，并且使用应用特定信息和分组属性模式信息来共同地识别层 7 应用。在一个示例中，网络设备 130 被配置为对特定字节值的出现进行计数，比如在申请人于 2014 年 5 月 12 日提交、并转让给 Marvell 的共同待决申请 14/275,332 中所公开的那样，其通过引用的方式整体并入于此。

[0034] 图 2A-2B 示出了与两个不同的网络应用相关联的签名的示例。在图 2A 和图 2B 中的每幅图中，X 轴表示以秒为单位的时间，并且 Y 轴表示具有有效载荷大小为 X(X 是自然数）的分组的数目，并且在一时间处的竖直线的高度指示在对应时间处接收的具有有效载荷大小为 X 的分组的数目。

[0035] 在一个实施例中，具有在 X 周围的范围内的有效载荷的分组被认为是有效载荷大小为 X 的分组。在一个示例中，X 为零，并且在分组具有在从零到 8 字节的范围内的大小的有效载荷时，该分组被认为有效载荷大小为零。

[0036] 在图 2A 的示例中，第一网络应用具有较小数目的有效载荷大小为 X 的分组，并且在接收有效载荷大小为 X 的分组时，具有较小数目的持续时间，比如从 0 秒到 120 秒少于 5 个持续时间。在图 2B 的示例中，第二网络应用具有较大数目的有效载荷大小为 X 的分组，并且在从时间 170 秒到 220 秒接收有效载荷大小为 X 的分组时，具有较大数目的持续时间。

[0037] 图 3 示出了概述根据本公开的实施例的用于分组分类的过程 300 的流程图。在一个示例中，在诸如网络设备 130 之类的网络设备中执行过程 300，以将分组的流与网络应用相关联。在一个实施例中，执行过程 300 以在接收并识别分组的新流时构建首部属性模式。该过程在 S301 处开始并进行至 S310。

[0038] 在 S310 处，接收分组。在图 1 的示例中，接收单元 135 接收分组并且将接收信息与分组相关联。在一个示例中，接收单元 135 将进入端口号与经由进入端口进入网络设备 130 的分组相关联。另外，在一个示例中，接收单元 135 将到达时间与分组相关联。

[0039] 在 S320 处，从分组的首部提取属性。在图 1 的示例中，首部属性提取单元 140 从分组的首部提取属性。在一个示例中，首部属性提取单元 140 从分组的首部提取有效载荷大小。在一个示例中，在分组根据 IPv4 协议编排格式时，分组的首部包括用于分组的总长度的第一字段、用于 IP 首部长度的第二字段以及用于 TCP 首部长度的第三字段。在一个示例中，首部属性提取单元 140 使用位屏蔽从首部的第一字段、第二字段和第三字段提取分组的总长度、IP 首部长度和 TCP 首都长度。然后首部属性提取单元 140 从总长度减去 IP 首部长度和 TCP 首部长度，以计算分组的有效载荷大小。在另一示例中，在分组根据 IPv6 协议编排格式时，首部具有用于有效载荷长度的字段。继而，首部属性提取单元 140 使用位屏蔽来提取有效载荷长度。

[0040] 另外，首部属性提取单元 140 计算其他适合的信息，用于生成签名，所述其他适合的信息比如具有特定有效载荷大小的分组的字节速率、具有特定大小的分组在时域中的到达速率、具有某大小的分组在位置域中的到达速率、具有特定大小的分组之间的到达间隔

时间、具有特定大小的分组之间的到达间隔间隙、有效载荷大小、分组速率、字节速率、到达间隔时间、分组数目等。

[0041] 在一个实施例中，首部属性提取单元 140 解析分组的首部，并且从分组的首部提取其他适合的信息，比如进入端口、出口端口、源地址、目的地址、分组类型（例如，IP、TCP、UDP 等）、服务水平等。

[0042] 在 S325 处，识别针对分组的分组流。在图 1 的示例中，流分类单元 145 能够基于分组的首部信息，诸如源地址、目的地址、源端口、目的端口、协议类型等，唯一地确定分组所属的分组流。

[0043] 在 S330 处，网络设备确定所提取的属性是否满足过滤器要求。在一个示例中，网络设备 130 确定所提取的有效载荷长度是否在一范围内，比如在一个示例中等于或低于 8 字节。在有效载荷长度处于该指定范围内时，过程进行至 S340；否则，过程进行至 S350。

[0044] 在 S340 处，在分类数据库中收集分组的信息。在一个示例中，网络设备 130 在分类数据库的新记录中，存储用于分组流的分组计数器的当前值（该当前值指示分组在分组流中的位置）、分配给分组流的定时器的当前时间，以及当前分组的分组长度。

[0045] 在 S350 处，更新分组流的信息。在一个示例中，网络设备 130 更新用于分组流的分组计数器（例如，对于在分组流中每个接收的分组增加一）、定时器以及分组流中的总字节数。

[0046] 在 S360 处，网络设备确定出于对正与应用相关联的流的基于签名的分类的目的，是否已经接收分组流中足够多的分组，并且网络设备相应地继续进行。例如，在网络设备 130 具有足够多的分组以用于分类时，过程进行至 S370；否则，过程返回 S310 以等待接收分组流的更多分组。

[0047] 在 S370 中，在分组分类中使用分类数据库以将分组流与网络应用相关联。在一个示例中，存储器 160 以决策树的形式存储网络应用的签名。然后，应用识别单元 150 使用决策树，基于分类数据库，将分组流关联到网络应用。在一个实施例中，在分组流与网络应用相关联时，策略执行单元 170 对分组并且也对分组流中的后续分组，应用针对该网络应用的策略。该过程进行至 S399 并终止。

[0048] 注意，在一个实施例中，可以预先确定基于协议的应用签名，并且提供给网络设备 130 从而在存储器 160 中进行存储。在另一实施例中，在第一次接收到由特定应用发起的分组的流时，由网络设备 130 学习基于协议的应用签名。

[0049] 图 4 示出了概述根据本公开的实施例的用于学习签名的过程 400 的流程图。在一个示例中，在网络设备 130 并不繁忙时，在网络设备 130 中执行过程 400。过程在 S401 处开始并进行至 S410。

[0050] 在 S410 处，接收分组流。在一个示例中，网络设备 130 接收从网络应用发起的分组的流。

[0051] 在 S420 处，提取来自分组的首部中的属性。在一个示例中，由首部属性提取单元 140 在接收到分组之时提取属性，并且所提取的属性被存储在分类数据库中，如在步骤 S320 中所描述的。

[0052] 在 S425 处，识别分组流。在一个示例中，流分类单元 145 能够基于分组的首部信息，诸如源地址、目的地址、源端口、目的端口、协议类型等，唯一地确定分组流。

[0053] 在 S430 处，识别与分组流相关联的网络应用。在一个示例中，因为网络设备 130 并不繁忙，因此网络设备 130 执行对分组流中的一个或多个分组的深度有效载荷检查，以识别与分组流相关联的网络应用。网络设备 130 可以使用任何适合的技术，诸如搜索特定字符串等，来识别与分组流相关联的网络应用。

[0054] 在 S440 处，学习基于两个或更多属性的签名并将其映射到网络应用。在一个实施例中，两个或更多属性为相互独立的。在一个示例实施例中，与应用相对应的流中大小为 X 的分组的数目并不取决于大小为 X 的这种分组在应用的发起期间在交换设备处被接收的时间。在一个示例中，网络设备 130 执行用于机器学习的软件指令，以基于分类数据库和所识别的网络应用来构建决策树。该决策树能够基于与分组相关的两个或更多类型的不相关的属性信息，将该网络应用与其他网络应用进行区分。属性信息包括包含在首部中的属性信息，以及诸如进入端口和 / 或接收到分组的时间之类的其他属性信息。

[0055] 在 S450 处，存储被映射到网络应用的属性的签名。在一个示例中，网络设备以促进决策树的形式存储签名，该决策树基于分组属性、通过分支决策序列而进行最终决策。然后，在网络设备 130 随后接收由该网络应用发起的新的分组流时，网络设备 130 基于分组属性而不是有效载荷大小检验来识别网络应用，属性中的一些属性是从分组首部中提取的，例如参照图 3 所描述的。然后，过程进行至 S499 并终止。

[0056] 当以硬件实现时，硬件可以包括一个或多个分立的部件、集成电路、专用集成电路 (ASIC) 等。

[0057] 虽然已经结合作为示例提出的本公开内容的具体实施例描述了本公开内容的各方面，但是可以对示例做出替换、修改和变化。因此，在此所叙述的实施例意在是说明性而非限制性的。在不偏离以下记载的权利要求的范围情况下，存在可以做出的变化。

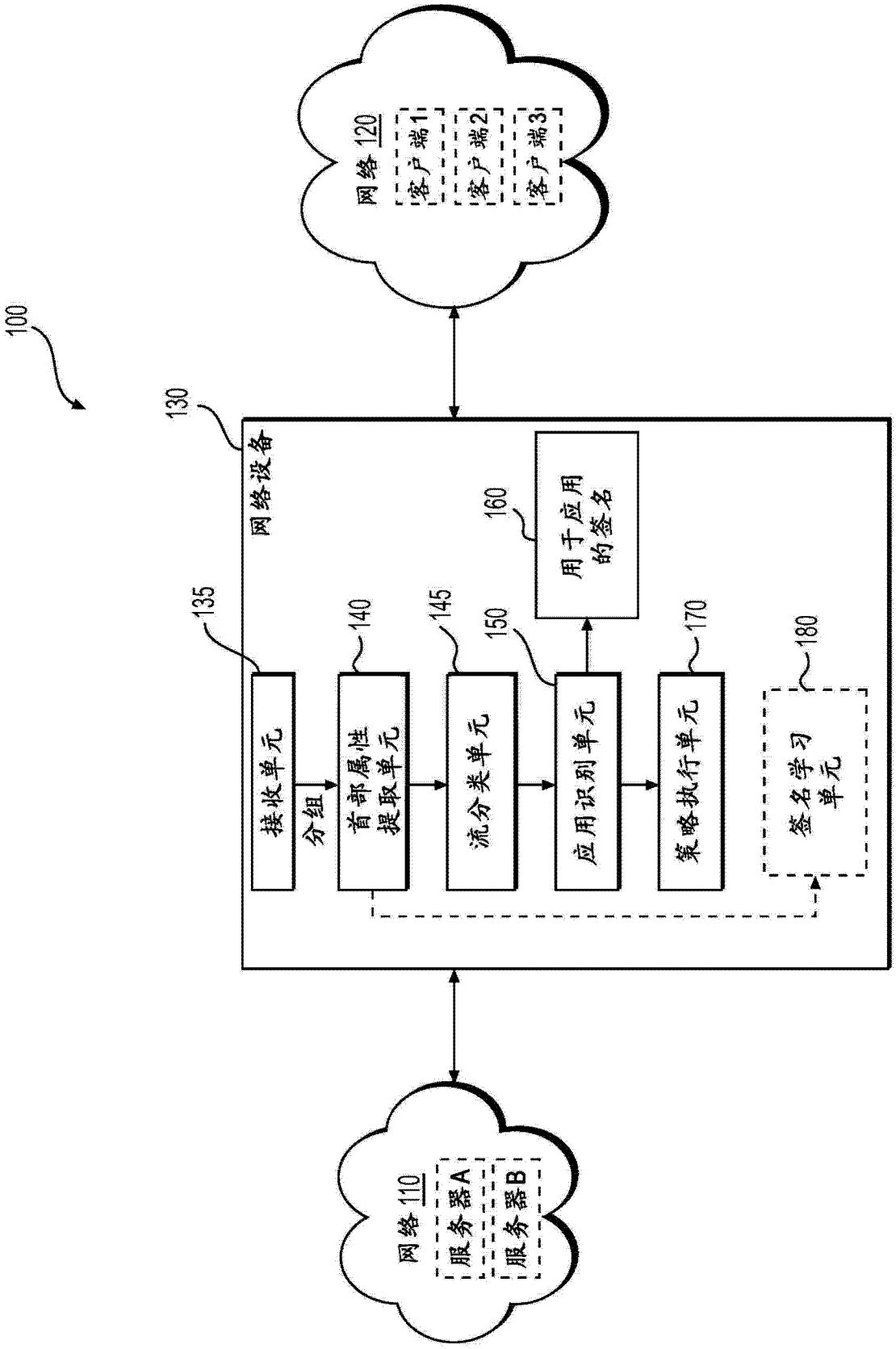

网络 120
客户端1
客户端2
客户端3
100
130
网络设备
160
用于应用的签名
135
140
145
150
170
180
接收单元
分组
首部属性提取单元
流分类单元
应用识别单元
策略执行单元
签名学习单元
网络 110
服务器A
服务器B

图 1

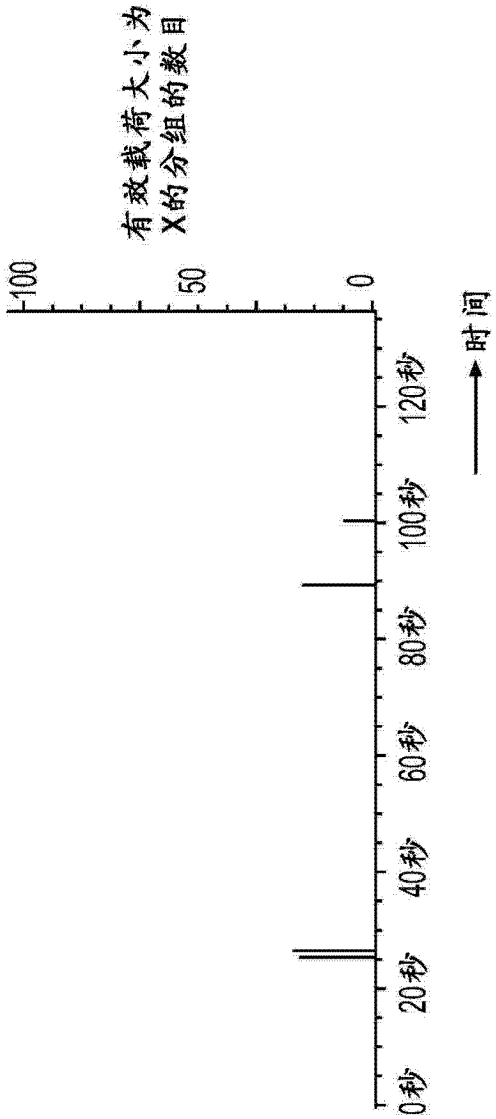

有效载荷大小为X的分组的数目
100
50
0
时间
0秒
20秒
40秒
60秒
80秒
100秒
120秒

图 2A

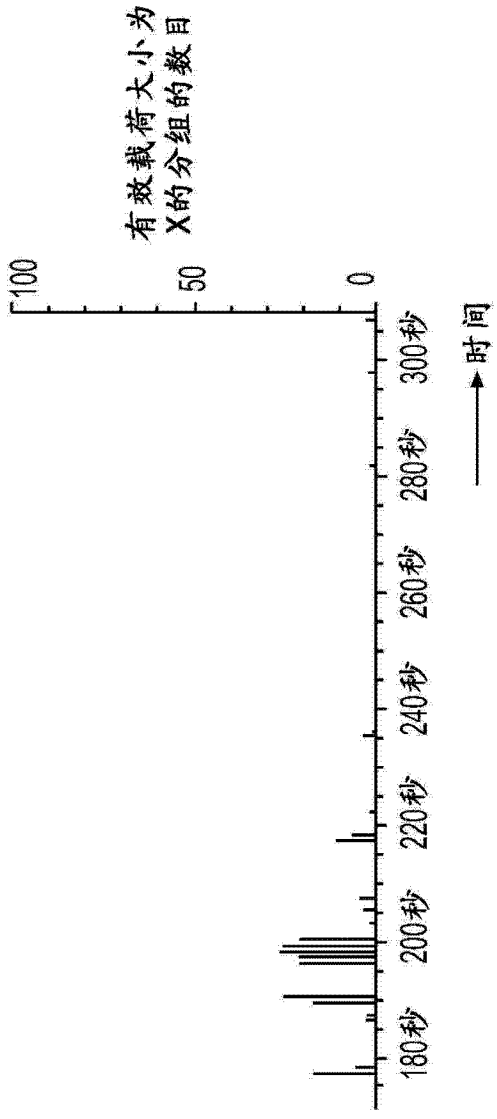

有效载荷大小为X的分组的数目
100
50
0
时间
180秒
200秒
220秒
240秒
260秒
280秒
300秒

图 2B

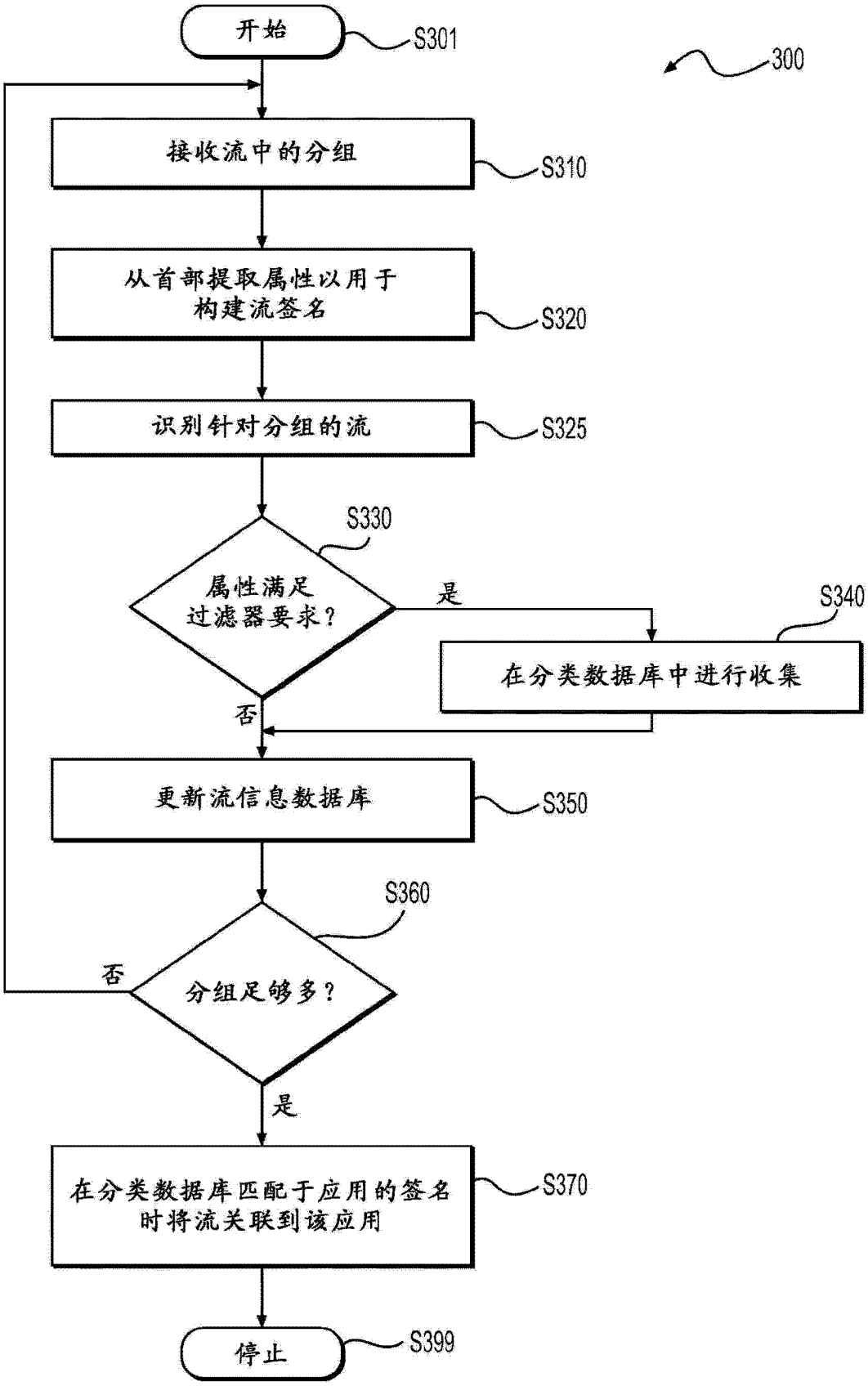

开始
S301
300
接收流中的分组
S310
从首部提取属性以用于构建流签名
S320
识别针对分组的流
S325
S330
属性满足过滤器要求？
是
S340
在分类数据库中进行收集
否
更新流信息数据库
S350
S360
否
分组足够多？
是
在分类数据库匹配于应用的签名时将流关联到该应用
S370
停止
S399

图 3

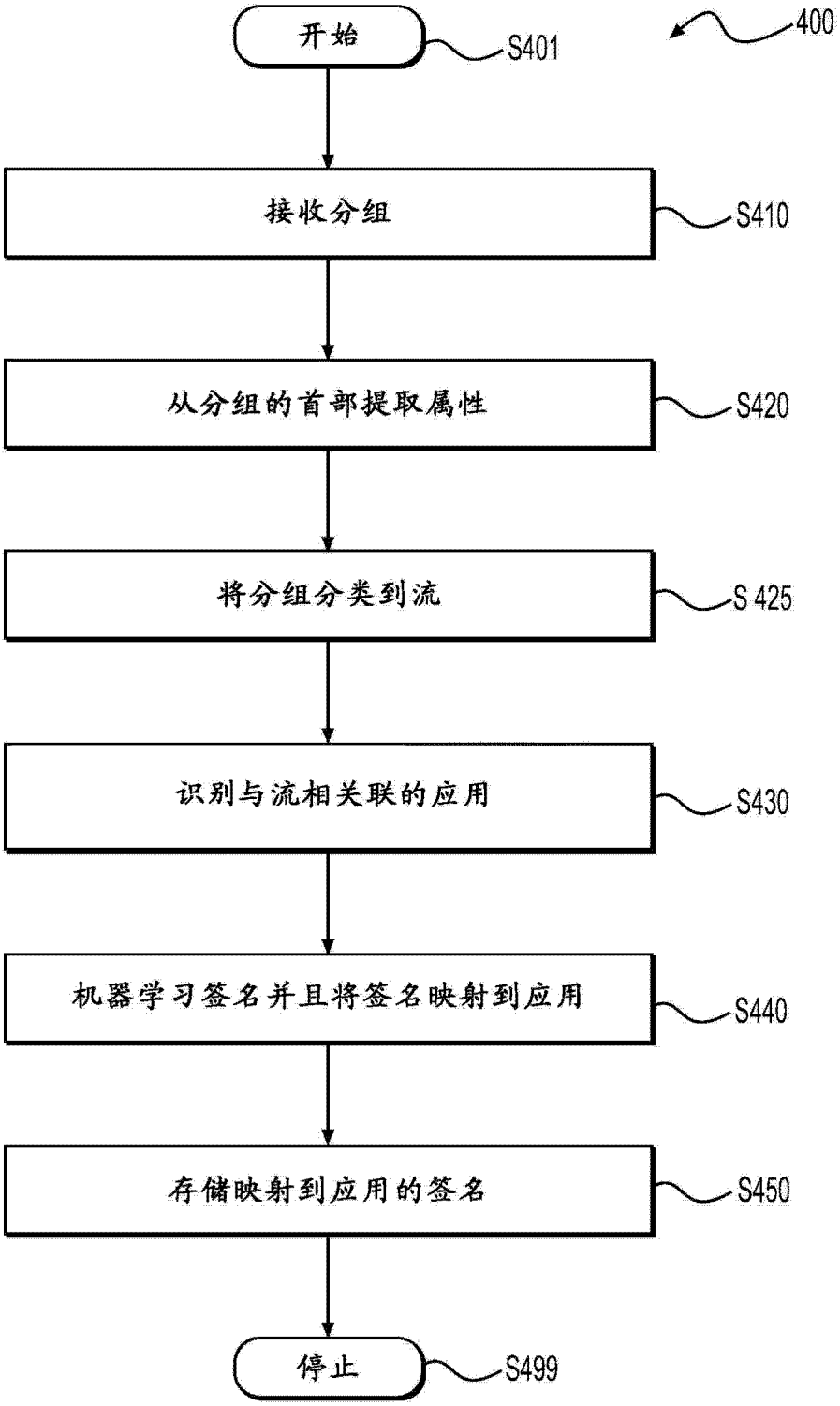

开始
S401
400
接收分组
S410
从分组的首部提取属性
S420
将分组分类到流
S 425
识别与流相关联的应用
S430
机器学习签名并且将签名映射到应用
S440
存储映射到应用的签名
S450
停止
S499

图 4