[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
H04L 12/26 (2006.01)
H04L 12/56 (2006.01)

[12] 发明专利申请公布说明书

[21] 申请号 200710167935.9

[43] 公开日 2009 年 4 月 29 日

[11] 公开号 CN 101420336A

[22] 申请日 2007.10.26
[21] 申请号 200710167935.9
[71] 申请人 诺基亚西门子通信有限责任两合公司
地址 德国慕尼黑
[72] 发明人 姜 峰

[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 卢 江 王忠忠

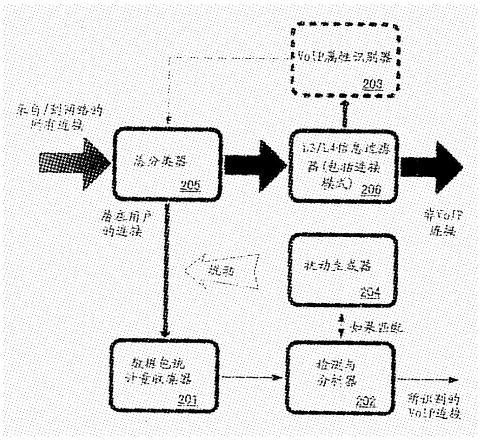
权利要求书 3 页 说明书 9 页 附图 2 页

[54] 发明名称

在网络中识别网络电话流量的方法及其系统

[57] 摘要

本发明涉及一种用于在网络中识别 VoIP 流量的方法及其系统。 根据用户身份将所有数据连接所对应的用户分成潜在的 VoIP 应用用户和普通用户。监控潜在的 VoIP 应用用户的流量并收集感兴趣的统计量。 将收集到的统计量与 VoIP 应用的流量分布的先验知识相关联，并且计算类似度 SI_p。 如果分析揭示一连接展现出所述统计量与 VoIP 应用的流量分布的先验知识之间高度关联，扰动就被施加到所监控的、潜在的 VoIP 应用用户的连接并且观察反馈 SI_a。 类似度 SI_p 和反馈 SI_a 被组合到一起来计算识别结果 D。 如果识别结果 D 大于阈值，则所监控的、潜在的 VoIP 应用用户的连接被识别为 VoIP 应用连接。 本发明不仅提高了识别精度，而且在稳定性、可扩展性和可推广性等方面均体现出其优点。



1. 一种用于在网络中识别VoIP流量的方法，其特征在于，该方法包括下列步骤：

根据用户身份将所有数据连接所对应的用户分成潜在的 VoIP 应用用户和普通用户；

监控潜在的 VoIP 应用用户的流量并收集感兴趣的统计量；

将收集到的统计量与 VoIP 应用的流量分布的先验知识相关联，并且计算类似度 SI_p；

如果分析揭示一连接展现出所述统计量与 VoIP 应用的流量分布的先验知识之间高度关联，扰动就被施加到所监控的、潜在的 VoIP 应用用户的连接并且观察反馈 SI_a；

类似度 SI_p 和反馈 SI_a 被组合到一起来计算识别结果 D；如果识别结果 D 大于阈值，则所监控的、潜在的 VoIP 应用用户的连接被识别为 VoIP 应用连接。

2. 如权利要求 1 所述的方法，将与服务器进行通信的用户标记为潜在的 VoIP 应用用户。

3. 根据权利要求 1 所述的方法，其特征在于，该方法还包括利用启发式经验和逻辑来分析用户登录过程中的交互过程。

4. 根据权利要求 3 所述的方法，其特征在于，所述启发式经验和逻辑包括对用户登录请求包应用深度包检测或者分析用户登录连接模式。

5. 根据权利要求 1－4 中的任一权利要求所述的方法，其特征在于，通过以下方式由类似度 SI_p 和反馈 SI_a 来确定识别结果 D：

$$D = \left\lfloor \frac{\delta \cdot SI_p + (1-\delta) \cdot SI_a}{\eta_0} \right\rfloor ,$$

其中，δ 是根据 VoIP 应用来确定的权重，并且 δ 的值在 1 到 0 之间变化，而 η_0 是根据 VoIP 应用来确定的阈值。

6. 根据权利要求 1－4 中的任一权利要求所述的方法，其特征在于，所述感兴趣的统计量是双层的：在连接层，统计量是持续时间或平均比特率；在数据包层，统计量是数据包大小或时间戳。

7. 根据权利要求 6 所述的方法，其特征在于，根据 VoIP 应用程序

内的编解码器来选择感兴趣的统计量。

8. 根据权利要求 1 - 4 中的任一权利要求所述的方法，其特征在于，所述扰动是数据包丢弃率。

9. 根据权利要求 1 - 4 中的任一权利要求所述的方法，其特征在于，所述用户身份是网络用于区分不同用户的标识。

10. 根据权利要求 9 所述的方法，其特征在于，所述用户身份是 WCDMA 移动通信网络的 PDP 上下文、固定通信网络的 IP 地址或者 PPP 会话号中的至少一种。

11. 一种用于在网络中识别 VoIP 流量的系统，该系统包括总分类器和第三层/第四层信息过滤器，该第三层/第四层信息过滤器连接在总分类器之后，用于过滤关于 VoIP 的信息并将过滤结果通过内部控制消息机制返回给总分类器，使得总分类器根据用户身份将所有数据连接所对应的用户分成潜在的 VoIP 应用用户和普通用户，

其特征在于，该系统还包括:

数据包统计量收集器，用于监控经总分类器所分出的潜在的 VoIP 应用用户的流量并收集感兴趣的统计量，该数据包统计量收集器输出所收集到的统计量;

检测与分析器，用于将由数据包统计量收集器所输出的统计量与 VoIP 应用的流量分布的先验知识相关联来决定，一个连接是否包含 VoIP 流量;

扰动生成器，用于和检测与分析器交互作用，并且生成施加到所监控的、潜在的 VoIP 应用用户的连接的统计扰动。

12. 根据权利要求 11 所述的系统，其特征在于，该系统还包括 VoIP 属性识别器，用于利用启发式经验和逻辑来分析用户登录过程中的交互过程，该 VoIP 属性识别器接受被第三层/第四层信息过滤器过滤后的连接，并通过内部控制消息机制将结果返回给总分类器。

13. 根据权利要求 12 所述的系统，其特征在于，所述启发式经验和逻辑包括对用户登录请求包应用深度包检测或者分析用户登录连接模式。

14. 根据权利要求 11 - 13 中的任一权利要求所述的系统，其特征在于，在检测与分析器中，所收集到的统计量被关联到 VoIP 应用的流量分布的先验知识，并且计算类似度 SI_p; 如果分析揭示一连接展现出

所述统计量与 VoIP 应用的流量分布的先验知识之间高度关联，扰动就被施加到所监控的、潜在的 VoIP 应用用户的连接并且观察反馈 SI_a；类似度 SI_p 和反馈 SI_a 被组合到一起来计算识别结果 D；如果识别结果 D 大于阈值，则所监控的、潜在的 VoIP 应用用户的连接被识别为 VoIP 应用连接。

15. 根据权利要求 14 所述的系统，其特征在于，通过以下方式由类似度 SI_p 和反馈 SI_a 来确定识别结果 D：

$$D = \left\lfloor \frac{\delta \cdot SI_p + (1-\delta) \cdot SI_a}{\eta_0} \right\rfloor ,$$

其中，δ 是根据 VoIP 应用来确定的权重，并且 δ 的值在 1 到 0 之间变化，而 η_0 是根据 VoIP 应用来确定的阈值。

16. 根据权利要求 11－13 中的任一权利要求所述的系统，其特征在于，所述感兴趣的统计量是双层的：在连接层，统计量是持续时间或平均比特率；在数据包层，统计量是数据包大小或时间戳。

17. 根据权利要求 16 所述的系统，其特征在于，根据 VoIP 应用程序内的编解码器来选择感兴趣的统计量。

18. 根据权利要求 11－13 中的任一权利要求所述的系统，其特征在于，所述扰动是数据包丢弃率。

19. 根据权利要求 11－13 中的任一权利要求所述的系统，其特征在于，所述用户身份是网络用于区分不同用户的标识。

20. 根据权利要求 19 所述的系统，其特征在于，所述用户身份是 WCDMA 移动通信网络的 PDP 上下文、固定通信网络的 IP 地址或者 PPP 会话号中的至少一种。

21. 根据权利要求 11－13 中的任一权利要求所述的系统，其特征在于，所述系统被安装在被识别的流量在网络中的汇聚处。

22. 根据权利要求21所述的系统，其特征在于，所述被识别的流量在网络中的汇聚处是运营商的核心网络的因特网边界或者网关。

在网络中识别网络电话流量的方法及其系统

技术领域

本发明通常涉及网络流量识别，并且更特别地，本发明涉及一种用于在网络中识别网络电话（Voice over IP，VoIP）流量的方法及其系统。

背景技术

近十年来，通信技术急剧发展。特别是，因特网的快速发展不仅降低了保持连接的成本，而且其广泛的可用性促进传统地基于专用网络（例如电话、传真、电视等等）的通信方式的迁移。因特网应用是实现这些通信方式的计算机软件，这通常遵循其副本在专用网络中的设计哲学：一个复杂（并且通常是强壮的）服务器由多个简单的客户机包围。这样存在的主要问题是，服务器成为关键点，并且一旦该服务器故障或者被阻塞，整个服务就会立即崩溃。

为了解决这种危机，对等（P2P）通信和覆盖网络自 1999 年起不断发展（以 Napster 为标志），并且现在，包括文件共享、内容递送、电话和视频广播的各种各样的因特网应用已经大大地利用这个突破的好处。技术上，P2P 网络实质上通过使通信参与者的角色均衡来分散，这些通信参与者形成覆盖网络。在这个网络中，任何一方都能从网络接收服务（扮演客户机的角色），同时为网络中的其它方提供服务（扮演服务器的角色），使得工作负荷平均分布并且有效利用资源，而且单个参与者的故障仅仅具有局部影响并且将自动通过网络自适应的机制来恢复。

最近，基于 P2P 的因特网电话应用（网络电话，VoIP）已吸引了来自各方的显著注意。对于网络运营商，第三方 P2P 电话应用特别令人感兴趣，因为这些应用直接与传统电话服务竞争但是通常难于控制、进行计费，甚至难于测量，从而导致严重损害运营商的利益。例如，拥有数百万注册用户的、最流行的 P2P 电话应用 Skype(一种免费的语音通信程序，参见 http://www.skype.com)提供了优质的语音质量和软件易用性。并且，在固定网络中成功之后，Skype 向移动用户提供服务仅仅是

一个时间问题。

传统上，运营商对服务控制存在极大兴趣。为实施服务控制，运营商需要在网络中安装服务控制设备（比如防火墙）。对于这些设备，流量识别是随后的处理的先决条件。为此，本发明提出了一种新颖的方法来识别网络中的 VoIP 流量，此方法不依赖于具体的应用实现，从而能稳定地识别 VoIP 流量及其高性能地实现。

传统的 VoIP 应用或者使用公知的端口、集中式服务器或者包含可被服务控制设备辨别的签名（signature，也可被称为应用程序特征字符串，等）。相反，具有 P2P 和加密技术的新一代 VoIP 应用对这类设备提出了极大挑战，给予那些应用逃避被识别的能力。

在 P2P 技术方面，基于 P2P 的网络架构允许 VoIP 呼叫在任何呼叫方和被呼叫方之间进行，而不涉及特定的服务器，并且允许 VoIP 呼叫经由任何端口发送数据包，甚至通过其它协议（例如 HTTP）来隧穿。这阻止服务控制设备通过检查其联系对象来知道 VoIP 呼叫的存在。例如，两个 Skype 用户只能在其间建立呼叫连接，并且如果直接连接不是最佳选择，则 Skype SN(也就是超级节点)将动态地被选择来传递呼叫的流量。一种值得注意的现象是，在其启动阶段，许多应用联系已知服务器，例如联系已知服务器来注册或登录。然而，这对于识别 VoIP 流量是不够的，因为这些应用通常建立另一对等连接来递送实际的数据内容（例如递送文本消息、语音等等）。

在加密方面，许多 P2P 应用已集成这种技术，其中数据包有效载荷以端到端的方式被加密，从而阻止其间的检查者观看他们承载的实际内容，并且实质上使基于签名的装置失效。再次参考 Skype，它的软件实现中包含了一个加密层，以对其发送的（几乎）所有数据包施以高强度的 TLS/AES 加密。

此外，另一实际问题来自于这些 VoIP 应用对传统的“一种类型的通信一个连接”规则的违反。如图 1 中所示，新的 VoIP 应用意图承载一个连接中的任何事物，这不同于将其流量（也就是信令、文本消息、多媒体等等）分到不同的套接字连接的传统应用。这个事实给试图进行服务区分的运营商带来了一个新问题：与连接相比，目前的服务具有更精细的粒度，一个连接可包含多种不同类型的服务。对于这些服务，现有的方法和设备不能处理，因为这些方法和设备通常假设套接字连接是

要处理的最小单元。

这类问题属于一个被称为 P2P 流量识别的更广泛的研究领域中。现有技术通常被分成两大类：基于签名的方法和基于行为的方法。这些方法的主要缺点如下：

基于签名的方法检查某些凭经验推知的签名的连接的数据包(有时仅检查前若干个数据包)，通常是检查控制关键字。如果这些控制关键字匹配，则断定该连接通过具有那些签名的某种应用来生成，[1] [2]。例如正则表达式的某些技术能被用来帮助展示复杂签名。然而，使得这些方法不那么有效的主要缺点是这些方法对签名的完全依赖。这些方法必须频繁地更新其签名数据库，以反映软件升级，从而导致增加的维护成本。此外，这些方法无法识别端到端加密后的数据包，这种情况将变得越来越普遍，因为软件经销商意识到被阻塞的威胁。最后，逐字节的检查要求高性能的 CPU 和大容量的存储器，从而提高了这种产品的成本和售价。

基于行为的方法是以统计学方式分析形成P2P覆盖网络的对等体之间的数据包交互作用。K. Suh 在[3]中提出通过观察给定网络的进入点和出口点处的每个连接的比特率的上升沿和下降沿来识别那个网络内的 Skype 中继节点的存在。X. Wang[4]建议通过增加时间扰动来标记潜在的连接，以便追踪那个连接的最终目的地。关于性能，由于这些方法需要比较和关联每两个连接的统计信息，所以这些方法的复杂度为 o(nlogn)，并且如果监控下的连接数量极大（这在运营商级别网络中是典型情况），则这些方法具有严重的可扩展性问题。另外，这些方法假设，两条接入链路（连接的第一跳和最后一跳）的流量统计数据可以被同步到服务识别装置。这种假设在现实情形下不太可能，因为 VoIP 呼叫能跨越不同的网络、运营商或者甚至国家进行。因此，即便存在这种可能，单个网络的运营商也需要花费大量成本来获得其它运营商的统计数据。

发明内容

以高稳定性和低复杂度为目标，提出了本发明来通过利用 VoIP 流量的本质行为特征而有效地识别 VoIP 流量。由于宽带自适应 VoIP 编解码器被认为对于未来的语音应用是有前途的，所以在本发明中通过在现

有的被动识别中结合主动识别来提高识别精度。

根据本发明的一个方面，提供了一种用于在网络中识别VoIP流量的方法。该方法包括下列步骤：根据用户身份将所有数据连接所对应的用户分成潜在的VoIP应用用户和普通用户；监控潜在的VoIP应用用户的流量并收集感兴趣的统计量；将收集到的统计量与VoIP应用的流量分布的先验知识相关联，并且计算类似度SI_p；如果分析揭示某些连接展现出所述统计量与VoIP应用的流量分布的先验知识之间高度关联，就将扰动施加到所监控的、潜在的VoIP应用用户的连接并且观察反馈SI_a；类似度SI_p和反馈SI_a被组合到一起来计算识别结果D；如果识别结果D大于阈值，则所监控的、潜在的VoIP应用用户的连接被识别为VoIP应用连接。

如果用户与已知服务器进行通信，则可将该用户视为潜在的VoIP应用用户。

在有些情况下，准确定义“已知服务器”是困难的。此时，可将所有用户定义为潜在的VoIP应用用户，其连接将受到流量监控。

此外，本发明方法还可以利用启发式经验和逻辑，通过分析用户登录过程中的交互过程，获得更高的性能。这些启发式经验和逻辑包括但不限于：对用户登录请求包应用深度包检测（DPI）、分析用户登录连接模式等等。

所述感兴趣的统计量包括两类：在连接层，这些统计量包括持续时间和平均比特率；在数据包层，这些统计量包括数据包大小和时间戳。根据具体的 VoIP 应用程序内的编解码器来选择感兴趣的统计量。

在本发明方法中，扰动例如是数据包丢弃率。用户身份是网络用于区分不同用户的标识。例如，用户身份是 WCDMA 移动通信网络的 PDP 上下文、固定通信网络的 IP 地址或者 PPP 会话号中的至少一种。

根据本发明的另一方面，提供了一种用于在网络中识别 VoIP 流量的系统。该系统的三个关键模块为：数据包统计量收集器、检测与分析器和扰动生成器。数据包统计量收集器被用来监控经总分类器所分出的潜在的 VoIP 应用用户的流量并收集感兴趣的统计量。检测与分析器将由数据包统计量收集器所输出的统计量与 VoIP 应用的流量分布的先验知识相关联来决定，一个连接是否包含 VoIP 流量。扰动生成器和检测与分析器交互作用并且生成到所监控的、潜在的 VoIP 应用用户的连接的统计量扰动。

在上述三个模块的基础上，该系统还包括可以提高性能的可选模块：VoIP 属性识别器。该 VoIP 属性识别器利用启发式经验和逻辑来分析用户登录过程中的交互过程。并且该 VoIP 属性识别器接受被第三层/第四层信息过滤器过滤后的连接，并通过内部控制消息机制将结果返回给总分类器。这些启发式经验和逻辑包括但不限于：对用户登录请求包应用深度包检测（DPI）、分析用户登录连接模式等等。

该系统还需要利用以下两个公用模块的功能：总分类器和第三层/第四层信息过滤器。该第三层/第四层信息过滤器连接在总分类器之后，用于过滤关于 VoIP 的信息并将过滤结果通过内部控制消息机制返回给总分类器，使得总分类器根据用户身份将所有数据连接所对应的用户分成潜在的 VoIP 应用用户和普通用户。这两个模块作为整个服务区分系统的基础功能而供其上所有的具体协议识别模块使用。

在检测与分析器中，所收集到的统计量被关联到 VoIP 应用的流量分布的先验知识，并且计算类似度 SI_p。如果分析揭示某些连接展现出所述统计量与 VoIP 应用的流量分布的先验知识之间高度关联，扰动就被施加到所监控的、潜在的 VoIP 应用用户的连接并且观察反馈 SI_a。类似度 SI_p 和反馈 SI_a 被组合到一起来计算识别结果 D。如果识别结果 D 大于阈值，则所监控的、潜在的 VoIP 应用用户的连接被识别为 VoIP 应用连接。

本发明的系统能被安装在被识别的流量在网络中的汇聚处。

本发明的优点包括以下五个方面：

· 精度

通过将主动扰动结合到被动监控，本发明能识别出以 P2P 和加密为特征的 VoIP 流量，而其对于任何现有技术是不可能的。公知的 VoIP 流量分布提供基本精度，此精度通过可变比特率（VBR）VoIP 编解码器的自适应特性来加强。

· 稳定性

在本发明所提出的方法中，对 VoIP 编解码器特性的依赖性而不是对 VoIP 应用（及其实现方案）的依赖性极大地贡献于本方法的稳定性。由于 VoIP 编解码器比 VoIP 应用更不易于变化，所以这种方法能识别出大量的 VoIP 应用，而不必要求逐应用地、甚至逐应用版本地加以识别。

· 可扩展性

收集和分析数据包层的统计量比传统的逐个字节的有效载荷检查更有效，并且更适于大规模的部署。如果有必要进行离线分析，则本发明还可以节约计算机存储器或者磁盘存储空间。

·可推广性

如果相对应的应用的流量分布是已知的，则本发明能被推广来识别由许多类型的网络（例如固定电话网络）中的其它应用所产生的通用VoIP流量。

·法律、隐私、物流和财政优点

因为本方法不要求有效载荷检查，所以本方法没有在许多类似产品中所使用的其它方法面对的法律问题和隐私问题。而且，由于其稳定性，所以本方法能通过消除频繁更新签名/服务器（超级节点）数据库来降低运营商的维护成本。另外，由于本发明中所提出的方法并不要求任何专用硬件，所以该方法能通过一般的硬件平台来实现，从而使得运营商在为硬件故障时的替换所作的备份仓库中仅需维护数量和品种较少的存货。

附图说明

下面结合附图更详细地说明本发明，其中：

图1示出VoIP应用可能包含的多种不同服务。

图2示出用于识别类似Skype的VoIP流量的系统结构。

具体实施方式

为了更好地理解本发明，现将本发明中所使用到的术语解释如下：

VoIP：网络电话，通过IP网络发送数字化语音的技术。

P2P：对等的、分散式网络架构，常常是构建于现有的IP网络之上的。

SN：超级节点，P2P网络中的专用对等体，该对等体提供共享服务，例如索引、中继等等。P2P网络能包含多个这种超级节点。

HTTP：超文本传输协议，通常用来递送网页内容（web content）的协议。

编解码器：编码器和解码器，用于数字媒体与专用格式之间的转换的计算机程序。

VBR：可变比特率技术，其能在工作时改变媒体编码的输出比特率。

DPI：深度包检测，一种根据数据包中所包含的签名来对流量进行分类的技术。

注意，VoIP 应用的实施细节因类型的不同而不同。以下针对目前最流行的应用 Skype，来展示一种有代表性的实现方案。图 2 提供用于识别类似 Skype 的 VoIP 流量的框图，该框图由三个被动识别模块（包括一个可选模块）、一个主动识别模块和两个系统公用模块组成：

· 数据包统计量收集器 201（被动模块）

这个模块监控潜在的 VoIP 用户的流量并收集稍后分析所需要的统计量。感兴趣的统计量是双层的：在连接层，这些统计量包括持续时间和平均比特率；在数据包层，这些统计量包括数据包大小和时间戳。根据具体的 VoIP 应用程序内的编解码器来选择这些感兴趣的统计量。

· 检测与分析器 202（被动模块）

将这些统计量作为输入，这个模块通过将输入与某些应用（例如 Skype）的流量分布的先验知识相关联来决定，一个连接是否可能包含 VoIP 流量。网络排队对于数据包在网络装置上经历的延迟有影响，特别是当网络严重负荷时。为了解决这个问题，将数据包到达之间的时间波动考虑为噪声，并且利用统计学信号处理技术来识别期望的 VoIP 流量。

· VoIP 属性识别器 203（被动模块，可选）

考虑到在文献中已报告与某些应用（例如 Skype）相关的一些线索，如果可得到这些线索，则利用这些线索作为启发式经验和逻辑有助于区分 VoIP 应用。本方法可以利用这些启发式经验和逻辑，通过分析用户登录过程中的交互过程，从而获得更高的精确度。可能的启发式经验和逻辑包括对用户登录请求包应用深度包检测（DPI）、分析用户登录连接模式等等。然而，这些线索高度遭受软件实现方案变化的影响，并且不能单独识别出 VoIP 流量，因为 VoIP 应用可以在登录完成之后打开一个新的连接以传输实际的语音数据，而这个新的连接往往没有前述线索可供识别之用。这个模块是可选的，其对第三层/第四层信息过滤器 206 的过滤结果进行进一步处理，其处理结果通过内部控制消息机制递交给总分类器 205。如果这个模块不存在，则来自第三层/第四层信息过滤器的结果被直接供给总分类器用来定义潜在的用户。

·扰动生成器 204（主动模块）

通过和“检测与分析器”交互作用，这个模块生成到目标连接的统计量扰动，比如数据包丢弃率。目标连接所对应的语音编解码器将这种扰动视为网络状况的改变，并触发相应的自适应机制。关于目标连接如何对扰动作出反应的结果随后被本方法中的“检测与分析器”记录和分析。

·总分类器 205（被动模块，公用的系统功能）

许多系统/平台为了进行流分类而具有这个模块。本发明利用这个总分类器，根据用户身份将所有数据连接分成两类：潜在的 VoIP 用户和普通用户。所述用户身份是网络用于区分不同用户的标识，例如 WCDMA 移动通信网络的 PDP 上下文、固定通信网络的 IP 地址或 PPP 会话号等。属于潜在的 VoIP 用户的连接被递送给数据包统计量收集器，用于后续处理。

·第三层/第四层信息过滤器 206（被动模块，公用的系统功能）

如前所述，当客户端登录到相应的服务网络时，Skype 以及大多数基于 SIP 的应用需要与该服务网络中的已知服务器之一交换信息。因此，如果发现这种情况，则这个模块将该用户标记为潜在的 VoIP 用户。然后，其过滤结果通过图 2 中虚线示出的内部控制消息机制被返回给总分类器。考虑到潜在的 VoIP 用户比较少，这种行为有助于将监控任务限定为仅仅小部分流量。

根据第三层/第四层过滤的结果，所有网络用户（图 2 的左输入）被分成潜在的 VoIP 用户和普通用户。潜在的 VoIP 用户的每个连接被监控并且收集其统计量，而不考虑数据包有效载荷中承载的实际内容。在分析统计量之后，将这些分析结果与对应应用的已知流量分布进行关联，并且计算类似度 SI_p。一旦分析揭示这两者之间高度关联，扰动就被增加到该连接并且观察被定义为 SI_a 的其反馈。最后，这两个量被组合到一起来计算识别结果 D（参见等式（1））。例如，SI_a 和 SI_p 加权 δ 求和，并且如果该识别结果 D 大于阈值 η_0，则这个连接被识别为 VoIP 连接（参见等式（2））。

$$D = D(SI_p, SI_a) \quad (1)$$

$$D = \left\lfloor \frac{\delta \cdot SI_p + (1-\delta) \cdot SI_a}{\eta_0} \right\rfloor \quad (2)$$

在等式（2）中，根据具体的 VoIP 应用来确定权重δ和阈值η_0。并且权重δ的值在 1 到 0 之间变化。

本发明所对应的系统以软件模块的形式实现，能被安装在被识别的流量在网络中的汇聚处。例如，可能的安装点是运营商的核心网络的因特网边界（特别是完成服务区分/控制功能的设备）或者一般的网关。现有的设备平台，比如 Cisco CSG/SSG、Nokia ISN、Siemens IPS，能为本发明系统的实现方案提供充分的软硬件平台支持。

尽管结合附图所示的实例对本发明进行了以上的描述，但是显然本发明不是局限于此，而可在随附的权利要求所公开的范围之内以多种方式进行修改。

参考文献

[1] S. Ehlert, S. Petgang. Analysis and signature of Skype VoIP session traffic. Technical Report. 2006 年 7 月.

[2] Ipp2p, http://www.ipp2p.org/. 17filter, http://17-filter.sourceforge.net/.

[3] K. Suh 等人的 Characterizing and detecting Skype-relayed traffic. IEEE Infocom'06, 2006 年 4 月.

[4] X. Wang 等人的 Tracking anonymous peer-to-peer VoIP calls on the Internet. ACM CCS'05, 2005 年 11 月.

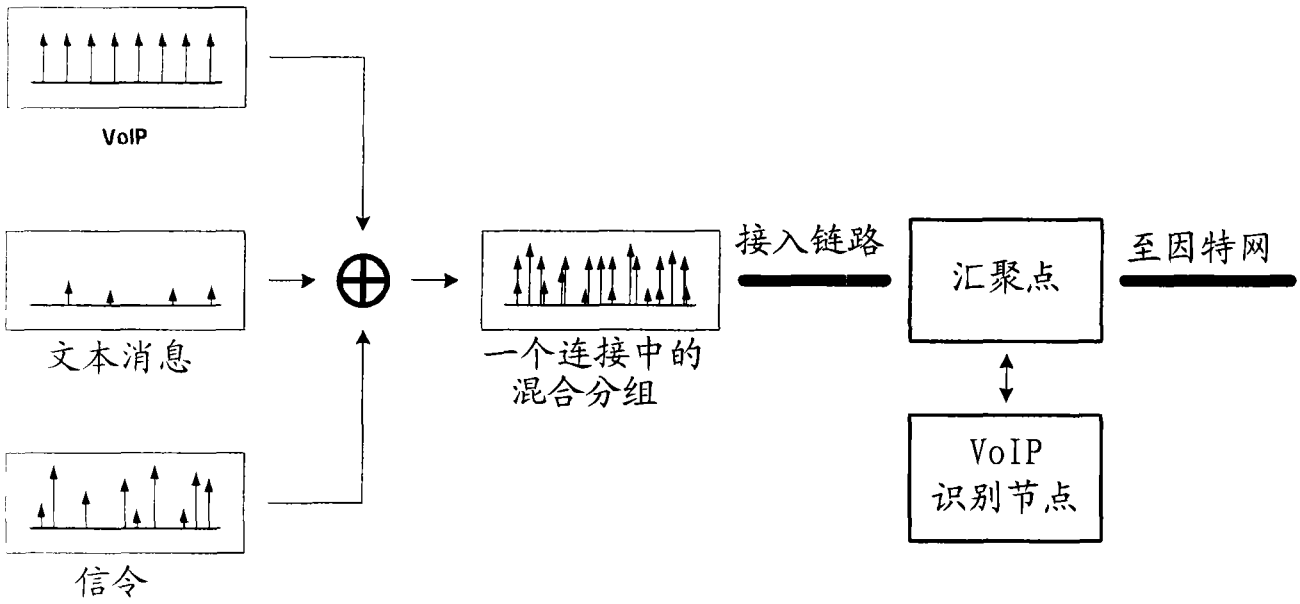


图 1

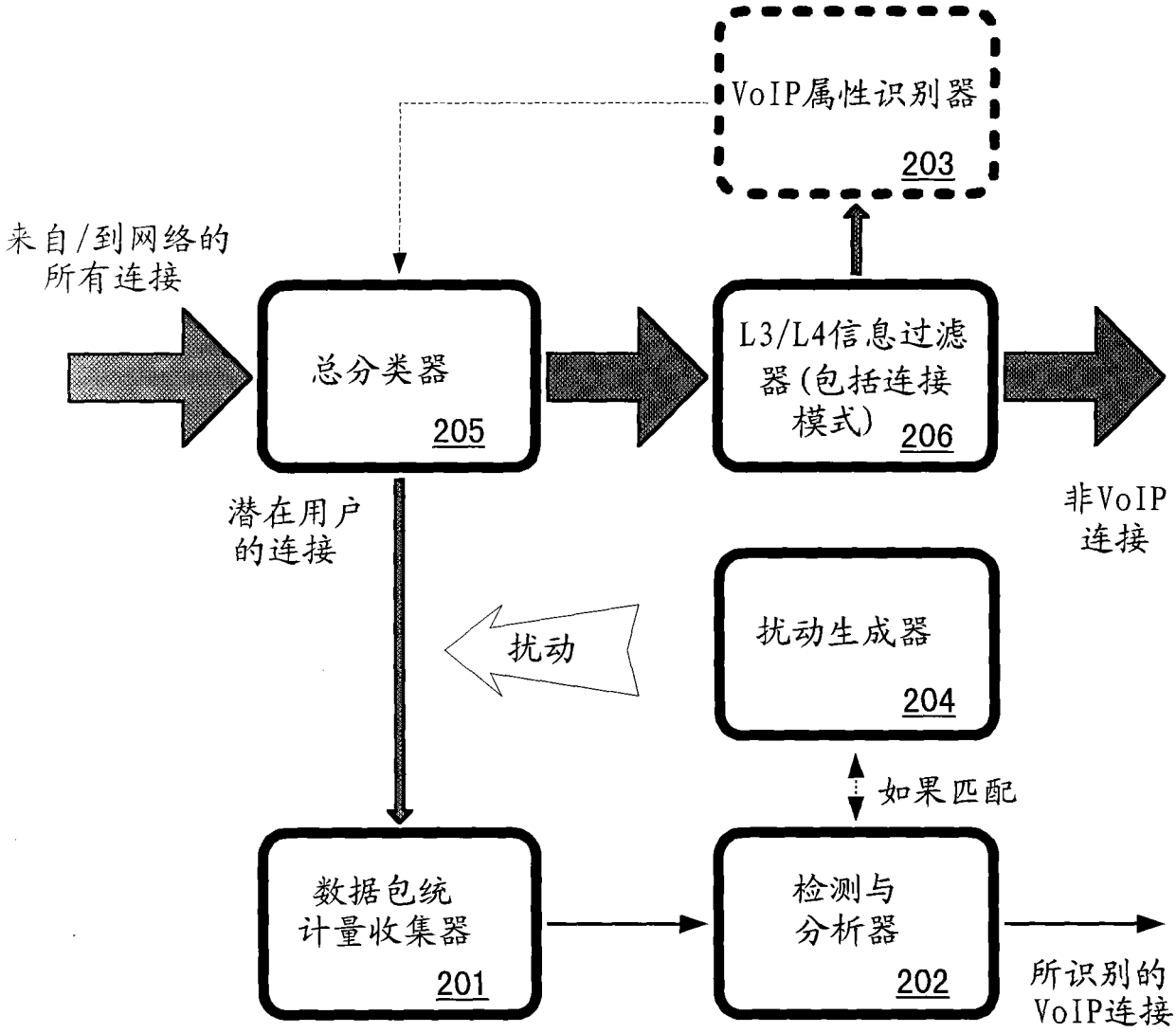

VoIP属性识别器
203
来自/到网络的
所有连接
总分类器
205
L3/L4信息过滤
器(包括连接
模式) 206
非VoIP
连接
潜在用户
的连接
扰动
扰动生成器
204
如果匹配
数据包统
计量收集器
201
检测与
分析器
202
所识别的
VoIP连接

图 2