



(12) 发明专利申请

(10) 申请公布号 CN 106302207 A

(43) 申请公布日 2017. 01. 04

(21) 申请号 201510291104. 7

(22) 申请日 2015. 05. 29

(71) 申请人 株式会社日立制作所  
地址 日本东京都

(72) 发明人 姜淼 耿璐

(74) 专利代理机构 北京永新同创知识产权代理  
有限公司 11376

代理人 杨胜军

(51) Int. Cl.

H04L 12/801(2013. 01)

H04L 29/06(2006. 01)

H04L 12/14(2006. 01)

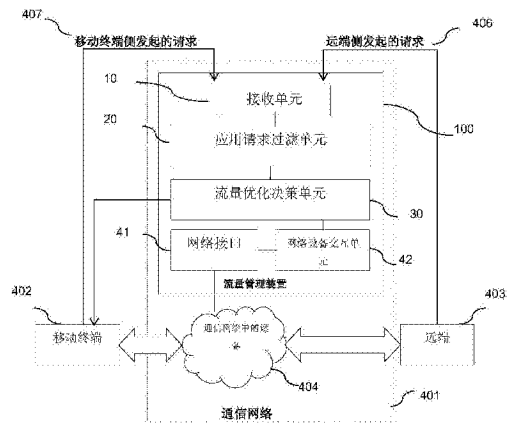
权利要求书2页 说明书15页 附图7页

(54) 发明名称

流量管理装置以及流量管理方法

(57) 摘要

本发明提供一种能够提高网络资源的利用效率的流量管理装置和流量管理方法。流量管理装置包括：应用请求接收单元，接收应用的流量优化请求，上述流量优化请求用于请求对网络中的该应用的流量进行优化；网络信息取得单元，取得与网络有关的网络信息；应用请求过滤单元，按照预定的准则，对上述应用请求接收单元接收到的上述流量优化请求进行过滤，判断是否保留上述流量优化请求；以及流量优化决策单元，针对经过上述应用请求过滤单元的过滤而保留的上述流量优化请求进行响应，利用上述网络信息，生成与经过过滤而保留的上述流量优化请求相对应的流量优化方案。



1. 一种流量管理装置,其特征在于,包括:

应用请求接收单元,接收应用的流量优化请求,上述流量优化请求用于请求对网络中的该应用的流量进行优化;

网络信息取得单元,取得与网络有关的网络信息;

应用请求过滤单元,按照预定的准则,对上述应用请求接收单元接收到的上述流量优化请求进行过滤,判断是否保留上述流量优化请求;以及

流量优化决策单元,针对经过上述应用请求过滤单元的过滤而保留的上述流量优化请求进行响应,利用上述网络信息,生成与经过过滤而保留的上述流量优化请求相对应的流量优化方案。

2. 根据权利要求 1 所述的流量管理装置,其特征在于,

还包括请求队列单元,上述请求队列单元将经过过滤而保留的上述流量优化请求依次放入队列中,

上述流量优化决策单元按照上述请求队列单元中排列的上述流量优化请求的先后顺序,生成相应的流量优化方案。

3. 根据权利要求 2 所述的流量管理装置,其特征在于,

还包括定时器单元,上述定时器单元设定有定时,

上述请求队列单元根据上述定时器设定的定时,丢弃请求队列中超时的流量优化请求。

4. 根据权利要求 1 所述的流量管理装置,其特征在于,

上述预定的准则包括应用提供商与网络运营之间的流量优化签约信息、或者应用侧能够容忍的流量优化请求最长处理时间、或者网络运营商对于应用流量优化的可能性判断结果。

5. 根据权利要求 2 所述的流量管理装置,其特征在于,

上述请求队列单元按照先入先出的方式依次排列上述流量优化请求。

6. 根据权利要求 3 所述的流量管理装置,其特征在于,

上述定时器按照上述流量优化请求中包含的请求最长处理时间来设置定时,

在定时超时的情况下,上述请求队列单元查找队列中是否包括超时的流量优化请求,在包括的情况下,丢弃该超时的流量优化请求。

7. 根据权利要求 1 所述的流量管理装置,其特征在于,

上述网络信息包括网络构成设备的资源占用信息以及资源分配策略,

上述流量优化决策单元根据上述资源占用信息以及上述资源分配策略,生成与经过过滤而保留的上述流量优化请求相对应的流量优化方案。

8. 一种流量管理方法,其特征在于,包括:

应用请求接收步骤,接收应用的流量优化请求,上述流量优化请求用于请求对网络中的该应用的流量进行优化;

网络信息取得步骤,取得与网络有关的网络信息;

应用请求过滤步骤,按照预定的准则,对上述应用请求接收步骤中接收到的上述流量优化请求进行过滤,判断是否保留上述流量优化请求;以及

流量优化决策步骤,针对经过上述应用请求过滤步骤的过滤而保留的上述流量优化请

求进行响应,利用上述网络信息,生成与经过过滤而保留的上述流量优化请求相对应的流量优化方案。

9. 根据权利要求 8 所述的流量管理方法,其特征在于,  
还包括排列步骤,将经过过滤而保留的上述流量优化请求依次放入队列中,  
在上述流量优化决策步骤中,按照上述排列步骤中排列的上述流量优化请求的先后顺序,生成相应的流量优化方案。

10. 根据权利要求 9 所述的流量管理方法,其特征在于,  
上述排列步骤中,根据定时器中设定的定时,丢弃请求队列中超时的流量优化请求。

11. 根据权利要求 8 所述的流量管理方法,其特征在于,  
上述预定的准则包括应用提供商与网络运营商之间的流量优化签约信息、或者应用侧能够容忍的流量优化请求最长处理时间、或者网络运营商对于应用流量优化的可能性判断结果。

12. 根据权利要求 9 所述的流量管理方法,其特征在于,  
上述排列步骤中,按照先入先出的方式依次排列上述流量优化请求。

13. 根据权利要求 10 所述的流量管理方法,其特征在于,  
按照上述流量优化请求中包含的请求最长处理时间来设置上述定时器的定时,  
上述排列步骤中,在定时超时的情况下,查找队列中是否包括超时的流量优化请求,在包括的情况下,丢弃该超时的流量优化请求。

14. 根据权利要求 8 所述的流量管理方法,其特征在于,  
上述网络信息包括网络构成设备的资源占用信息以及资源分配策略,  
上述流量优化决策步骤中,根据上述资源占用信息以及上述资源分配策略,生成与经过过滤而保留的上述流量优化请求相对应的流量优化方案。

## 流量管理装置以及流量管理方法

### 技术领域

[0001] 本发明涉及一种在通信网络中用于流量优化的流量管理装置以及流量管理方法。特别是涉及一种承载第三方服务提供商流量的通信网络中的流量管理装置以及流量管理方法。

### 背景技术

[0002] 随着信息技术的不断发展,通信网络的利用越来越广泛。通信网络是指包含 LTE(Long Term Evolution, 长期演进)网络、3G(3rd Generation, 第三代数字通信)(WCDMA、cdma2000 等)网络等由网络运营商提供网络传输并维护的现有网络。

[0003] 在以数据流量为最大比重业务的通信网络(例如目前的 3G 及 LTE 网络)中,绝大多数的流量来自于互联网业务提供商,即非网络运营商自营的业务(也称为来自于第三方服务提供商的业务)。

[0004] 以 LTE 网络为例,当用户需要通过 LTE 网络访问因特网或者某个应用的远端服务器时,网络会为用户建立一条承载(Bearer)。承载从用户处连接到数据网关,中间涉及若干个网络设备,如基站、服务网关、数据网关等。在这条承载上,会产生相应的流量。

[0005] 所谓“承载”,表示网络支持业务的程度,表现为承载 QoS(Quality of Service, 服务质量)参数,并且,根据对业务支持的质量,能够分为多个类型。

[0006] 承载的 QoS 参数包含 GBR(Guaranteed Bit Rate, 保证比特速率)、QCI(QoS Class Identifier, QoS 等级标识符)、ARP(Address Resolution Protocol, 地址解析协议)、MBR(Maximum Bit Rate, 最大比特速率)、或 AMBR(Aggregated Maximum Bit Rate, 聚合最大比特率)(包括 UE-AMBR 和 APN-AMBR)。

[0007] 其中,GBR 承载和 Non-GBR 承载是两种不同的承载资源类型。GBR 承载保证为网络分配一定的带宽用于流量传输,而 Non-GBR 承载意味着网络“尽力而为”地传输流量,并不能保证有足够的带宽。QCI 是表示 QoS 等级的标识,每个 QCI 等级对应着相应的资源类型(GBR/Non-GBR)、优先级、时延、误包率以及这类参数对应的典型业务。MBR 用于为 GBR 类型的承载提供最大带宽的限制。由于 GBR 类型的承载有一定的网络带宽保证,但是为了限制这种类型的承载对网络资源的无限占用,MBR 用于为 GBR 承载定义一个带宽上限。AMBR 是 LTE 系统为了限制一组 Non-GBR 对网络资源的无限占用而新设计的参数。AMBR 的实现形式又有 APN(Access Point Name, 接入点名称)-AMBR 和 UE(User Equipment, 用户设备)-AMBR。

[0008] 从 3G 技术开始,网络中主要的业务类型从传统的语音、短信发展到被数据业务替代。特别是社交网络应用逐渐成为 3G 网络最主要的流量产生源。

[0009] 另外,当前正在中国逐步布设起来的 4G 网络,主要采用 LTE 技术,其网络速率和 3G 技术相比又有大幅度的提升,所以使得一些上传视频类的应用(例如智能安防)可以被网络传输所支持,进一步丰富了网络中的流量类型。

[0010] 另一方面,与社交网络类应用产生的相对低数据量多频次的数据量相比,智能安

防类等应用产生的流量的数据量更大,或展示出多样性,即,可能在平常情况下以低清晰度(对应产生低数据量)进行将摄像头拍摄的视频实时地通过移动网络传输到另一端的监控中心的传输,在需要重点观看某区域监控视频等特殊情况下以高清晰度(对应产生高数据量)进行传输。

[0011] 相应地,在 4G 网络支持此类应用时,也可能展示出承载的多样性,即,平常情况下用 Non-GBR 类承载进行传输,特别情况下用 GBR 类承载进行传输。通过根据应用的状况对承载的 QoS 参数进行变更,能够对网络中的流量进行优化。

[0012] 但是,由于应用何时处于平常情况或何时处于特殊情况,完全由应用侧(用户终端或应用的远端服务器)决定,网络并不“知情”。所以,网络通过接口接收应用侧提供的此类信息来据此对网络中的相应流量进行优化更具有实用性,也能在应用性能下降时更加及时地做出反馈。

[0013] 现有 3G 网络与 LTE 网络相比较,流量优化机制并不完全相同。但类似地,两种技术均由网络根据用户签约信息以及流量的具体需求(实时性、数据率、时延等)去建立相应的承载以实现流量交付。

[0014] 当网络中出现拥塞时,Non-GBR 承载这种“尽力而为”的交付可能由于网络资源此时非常有限而出现明显的性能下降,如时延大幅增加,甚至出现连接中断,或者数据速率大幅下降,可表现为上传/下载图片或视频变得非常缓慢。在现有网络中,这种性能下降的情况只能在网络拥塞自行缓解后(例如由于用户数减少或传输需求减少,亦可能由于网络内部实施了某些流量优化的措施)出现改善。

[0015] 这意味着,只有网络侧能够通过网络自身的判断去优化应用的流量,常常不能准确反馈到应用本身的需求。例如以上述智能安防应用为例,何时只需要通过 Non-GBR 承载交付其上传的视频流量,何时需要通过更有保障的 GBR 承载交付流量,需要应用侧自行判断后告知给网络,网络依据此信息进行可能的流量优化,使得应用达到更好的效果:普通情况时网络“尽力而为”地承载流量,特殊情况时网络“有保障”地承载流量。

[0016] 根据标准化组织 3GPP 的标准化文档 TS 23.203,在现有网络中,已经存在了网络联合应用反馈的性能指标信息进行流量优化的相关技术。

[0017] 图 11 是现有的非漫游状态下的策略和计费控制(PCC:Policy and Charging Control)的架构图。如图 11 所示,该框架包括:应用服务器 302、策略与计费规则决策单元(PCRF)301、用户签约数据库 303、承载绑定和事件上报单元 304、包括策略和计费执行单元(PCEF)的数据网关 305、在线计费系统 306 以及离线计费系统 307。

[0018] 策略与计费规则功能 301 需要综合应用服务器 302、用户签约数据库 303 以及承载绑定和事件上报单元 304 的信息去设计实现。其中应用服务器 302(AF:Application Function)为 PCRF 301 提供用户的业务信息,提供基于动态的策略与计费功能。AF 302 功能实体根据与用户之间的业务协商,向 PCRF 301 动态提供应用层的会话信息,PCRF 301 根据 AF 302 提供的业务信息以及用户签约信息、本地运营商策略,制定出策略和计费控制策略,并将该策略下发给分组网络中的网关 305, PCEF 305 执行该策略,由于策略中包含流标识信息(IP 五元组)以及针对该流的 QoS 要求以及计费要求,PCEF 305 可以根据这些信息实现基于流的 QoS 和计费,在全 IP 网络实现端到端 QoS。并且,在线计费系统 306 和离线计费系统 307 分别实现流计费问题。

[0019] 但这种基于应用本身的需求进行流量优化的解决方案的实现范围,目前只可能是运营商自营业务的 IP 多媒体子系统 (IMS:IP Multi-media Sub-system)。即,只有运营商自营的业务才能通过已定义的接口与网络进行流量优化的沟通,包括流量优化的请求,动态的流量策略与计费。而在实际的网络中,运营商自营的业务所占比重非常小。以上提到的社交网络类应用、智能安防等新兴应用几乎全部来自于第三方服务提供商。

[0020] 并且,目前还存在第三方服务提供商替代用户向网络运营商支付费用的趋势,从而在与其他同类型应用竞争时更有优势。因此,网络运营商与第三方服务提供商签约变得十分可能。因此,考虑到通过动态的策略实现第三方应用流量交付的必要性。

[0021] 但是这些技术侧重于网络如何在 PCRF(策略与计费规则功能)处增加动态调整策略这一功能,并没有涉及网络如何接收应用反馈的流量优化请求信息(或关于应用性能下降的告知)。

[0022] 此外,现有技术提到的网络可对第三方应用的流量动态调整,意味着无论何种第三方应用,何种性能下降的告知消息,都会被网络受理,并且通过网络内部的策略调整尝试改善这一问题。

[0023] 但实际网络中,由于存在大量的第三方应用并不断地产生流量,而网络资源的有限性特别是在中国尤为突出,因此相当于很频繁地出现应用性能下降这样的问题。在这种情况下,随时受理应用反馈的性能提升请求会耗费大量的网络资源,也势必会浪费相应的时间。并且,这种资源与时间的消耗,会导致需要“尽快”改善性能的实时应用的请求被延误。

## 发明内容

[0024] 本发明是鉴于以上问题而完成的。本发明的目的在于提供一种能够提高网络资源的利用效率的流量管理装置和流量管理方法。

[0025] 本发明的一个技术方案是一种流量管理装置,包括:应用请求接收单元,接收应用的流量优化请求,上述流量优化请求用于请求对网络中的该应用的流量进行优化;网络信息取得单元,取得与网络有关的网络信息;应用请求过滤单元,按照预定的准则,对上述应用请求接收单元接收到的上述流量优化请求进行过滤,判断是否保留上述流量优化请求;以及流量优化决策单元,针对经过上述应用请求过滤单元的过滤而保留的上述流量优化请求进行响应,利用上述网络信息,生成与经过过滤而保留的上述流量优化请求相对应的流量优化方案。

[0026] 本发明的另一个技术方案是一种流量管理方法,包括:应用请求接收步骤,接收应用的流量优化请求,上述流量优化请求用于请求对网络中的该应用的流量进行优化;网络信息取得步骤,取得与网络有关的网络信息;应用请求过滤步骤,按照预定的准则,对上述应用请求接收步骤中接收到的上述流量优化请求进行过滤,判断是否保留上述流量优化请求;以及流量优化决策步骤,针对经过上述应用请求过滤步骤的过滤而保留的上述流量优化请求进行响应,利用上述网络信息,生成与经过过滤而保留的上述流量优化请求相对应的流量优化方案。

[0027] 发明效果:

[0028] 本发明适用于承载第三方服务提供商流量的通信网络。根据本发明,使得网络在

通过定义的接口接收到应用侧发来的流量优化请求时,通过必要的请求过滤单元联合网络目前的资源现状,决策出适当的流量优化方案,从而能够更好地实现网络对流量的交付,更好地发挥网络资源的价值以满足服务提供商的需求,从而能够提高网络资源的利用效率。

## 附图说明

- [0029] 图 1 是用户通过移动网络连接到因特网或服务提供商的网络框架的示意图。
- [0030] 图 2 是 LTE 网络中建立承载所涉及的网络框架的示意图。
- [0031] 图 3 是包括第一实施方式涉及的流量管理装置的通信网络的结构框图。
- [0032] 图 4 是第一实施方式涉及的流量管理装置的应用请求控制的流程图。
- [0033] 图 5 是包括第二实施方式涉及的流量管理装置的通信网络的结构框图。
- [0034] 图 6 是第二实施方式涉及的流量管理装置的应用请求控制的流程图。
- [0035] 图 7 是基于第三方应用流量优化的流量控制的消息交互图。
- [0036] 图 8 是基于第三方应用流量优化的流量控制的另一消息交互图。
- [0037] 图 9 是包括第三实施方式涉及的流量管理装置的通信网络的结构框图。
- [0038] 图 10 是第三实施方式涉及的流量管理装置的应用请求控制的流程图。
- [0039] 图 11 是现有的非漫游状态下的策略和计费控制的架构图。

## 具体实施方式

[0040] 本发明适用于承载第三方服务提供商流量的通信网络。作为适用本发明的网络框架的示例,以图 1 为例进行说明。

[0041] 图 1 是用户通过移动网络连接到因特网或服务提供商的网络框架的示意图。

[0042] 图 1 所示的移动网络 102 例如是正在使用的最先进的移动网络技术 LTE “长期演进”,也称为第四代通信技术或 4G 技术。

[0043] 当用户终端 101 请求通过 LTE 网络连接至因特网 103 或者请求获得服务即访问服务提供商 103 以获取信息时,移动网络 102 会根据网络资源状况为用户终端 101 建立一条承载,以满足用户的需求。相应承载建立后,会在通信网络中产生相应的流量。流量的形式可以体现为通过相应网络节点如基站、数据网关等的网络中的流量 104 和接口上的流量 105。

[0044] 图 2 是 LTE 网络中建立承载所涉及的网络框架的示意图。

[0045] 在图 2 所示的移动网络中,通过在用户终端 201 与分组数据网 205 之间建立端到端的连接,来实现流量的传输。具体来说,由 EPS(Evolved Packet System,演进分组系统)承载 (Bearer) 实现 QoS(Quality of Service) 保证。

[0046] 其中,用户终端 201 与分组数据网 205 之间建立端到端的连接由无线接入网、核心网、以及因特网(互联网)构成。移动网络的主要构成单元在无线接入网侧包含基站 202,在核心网侧包含服务网关 203、数据网关 204、移动管理实体 206、归属用户服务器 207 以及策略与计费规则功能单元 208。当网络有充足的资源满足用户的传输需求时,会根据流量的类型为其建立相应的承载以提供适当的 QoS。

[0047] EPS 承载位于用户和数据网关之间,按照接口协议等不同可以分为承载 1(用户与基站之间的无线承载,由基站实现控制)、承载 2(基站与服务网关之间的 S1 承载)以及承

载 3 (服务网关与数据网关之间的 S5/S8 承载,由数据网关实现控制)。每段承载分别由相应的网络单元按照流量对应的 QoS 参数负责保证其传输需求、例如时延、丢包率等。对于一个用户的连接,全部承载 (承载 1、承载 2、承载 3) 由策略与计费规则功能单元 (PCRF) 联合移动管理实体 (MME) 中存储的用户信息确定。

[0048] 本发明的流量管理装置能够设置于在用户终端 201 与分组数据网 205 之间建立承载的各个网络设备中,或者作为独立的流量管理装置而与现有的网络设备相连接,从而对用户终端 201 与分组数据网 205 之间建立的承载的流量进行管理。

[0049] 此外,以上例举了 LTE 网络框架下的承载所涉及的网络设备,各个网络设备的具体功能可以参照现有的网络标准所赋予的功能。

[0050] 此外,网络类型也不限于 LTE 网络能够使用在用户终端与分组数据网之间建立承载的任意网络。

[0051] 也就是说,本发明涉及的流量管理装置能够应用于连接到应用服务商的网络中,从而对所建立的承载中的流量进行管理。

[0052] 在上述网络框架的背景下,以下,参照附图详细说明本发明的优选实施方式。

[0053] 此外,在不同实施方式的各个附图中,对于相同或类似的部分,赋予相同的附图标记,并且适当省略重复说明的部分。

[0054] (第一实施方式)

[0055] 图 3 是包括第一实施方式涉及的流量管理装置的通信网络的结构框图。

[0056] 如图 3 所示,本实施方式的流量管理装置 100 设置在通信网络 401 中,经由网络接口 41 而与通信网络中的设备 404 相连接。此处,设通信网络是 LTE 网络。通信网络中的设备 404 可以是图 2 中包含的 LTE 网络的设备中的某个。具体来说,是指基站、移动管理实体、服务网关、归属用户服务器、数据网关等中的某个。

[0057] 其中,现有通信网络中的设备 404 通过无线连接能够与移动终端 402 进行通信,并且通过有线连接能够与远端 403 进行通信,从而能够在移动终端 402 与远端 403 之间建立承载。

[0058] 流量管理装置 100 包括接收单元 10、应用请求过滤单元 20、流量优化决策单元 30、网络接口 41 以及网络设备交互单元 42。

[0059] 接收单元 10 从终端 402 或者远端 403 接收应用的流量优化请求,该流量优化请求用于请求对网络中的该应用的流量进行优化。这里,既可以从终端 402 侧发送应用的流量优化请求 407,也可以从终端 403 侧发送应用的流量优化请求 406。这里的接收单元 10 相当于“应用请求接收单元”。

[0060] 应用请求过滤单元 20 按照预定的准则,对接收单元 10 接收到的应用的流量优化请求进行过滤,判断是否保留流量优化请求。

[0061] 具体来说,应用请求过滤单元 20 按照被预先设定的判断准则,对流量优化请求进行过滤,忽略不符合准则要求的流量优化请求。应用请求过滤单元 20 进行过滤所依据的准则可以包括:应用提供商与网络运营商之间所签订的关于流量优化的签约信息、应用能够容忍的请求最长处理时间、或者网络运营商对于应用流量优化可能性的判断结果等。

[0062] 例如,设应用请求过滤单元 20 预先保存有应用提供商与网络运营商之间所签订的关于流量优化的签约信息或者从外部接收并保存该签约信息,从而过滤所依据的准则包



括应用提供商与网络运营商之间所签订的关于流量优化的签约信息。在应用提供商与网络运营商就应用 A 签订有协议,并且签约信息表示在网络拥塞时也尽力保障该应用 A 的流量交付的情况下,该应用 A 所对应的远端或移动终端侧向流量管理装置 100 发送了流量优化请求时,应用请求过滤单元 20 对该流量优化请求进行过滤处理,根据以上签约信息判断为尽力保障该应用 A 的流量交付,从而将应用 A 的流量优化请求判断为“通过”,即需要响应应用 A 的流量优化请求而对应用 A 的流量进行优化处理。

[0063] 又例如,设应用请求过滤单元 20 预先保存有网络运营商与使用的链路的对应关系等网络运营商信息,从而能够验证网络运营商对于应用流量优化可能性,因此,过滤所依据的准则包括网络运营商对于应用流量优化可能性的判断结果。远端 403 检测到应用 B 的性能下降,在判断为存在传输网络对流量没有可靠交付的可能性等情况下,向流量管理装置 100 发送流量优化请求。接收单元 10 接收到该应用 B 的流量优化请求后,应用请求过滤单元 20 对该流量优化请求进行过滤处理,根据网络运营商对于应用流量优化可能性的判断结果,判断出由于实际出现性能下降的链路不属于应用 B 所对应的运营商网络的范畴,运营商对于该应用 B 的流量优化请求没有实现可能性,这是因为流量优化请求所对应的运营商无法去改变其他运营商网络的配置来实现优化。因此,应用请求过滤单元 20 将应用 B 的流量优化请求判断为“不通过”,并将该流量优化请求丢弃。在将应用的流量优化请求判断为“不通过”的情况下,应用请求过滤单元 20 也可以将这一情况反馈给请求发送源的远端 403。

[0064] 又例如,设应用请求过滤单元 20 预先保存有或接收到应用能够容忍的请求最长处理时间,从而过滤所依据的准则包括应用能够容忍的请求最长处理时间。在远端或移动终端中某实时性较强的应用 C 检测到由于流量交付没有得到保证造成的应用性能下降的情况下,向流量管理装置 100 发送流量优化请求,并告知可容忍的请求最长处理时间为 3s(秒)。如应用请求过滤单元 20 判断为,接收到的该流量应用请求的经过时间超过 3s,或者应用请求过滤单元 20 根据现有的待处理流量优化请求和延迟时间等判定为不能在应用可容忍的最长时间内处理该流量优化处理,则应用请求过滤单元 20 将应用 C 的流量优化请求判断为“不通过”,并将该流量优化请求丢弃。在将应用的流量优化请求判断为“不通过”的情况下,应用请求过滤单元 20 也可以将这一情况反馈给请求发送源的远端 403。

[0065] 以上分别举例说明了应用请求过滤单元 20 进行过滤的多个准则,但应用请求过滤单元 20 所采用的准则并不限于以上所例示的内容,只要是能够对流量优化处理进行过滤的过滤条件即可,可以根据具体的网络环境进行设定。

[0066] 并且,以上多个准则也可以同时适用,或者分先后地适用来配合进行流量优化请求的判断。

[0067] 返回到图 3 的说明,流量优化决策单元 30 针对经过应用请求过滤单元 20 的过滤而保留(被判断为“通过”)的流量优化请求进行响应,利用网络信息,生成与经过过滤而保留的流量优化请求相对应的流量优化方案。

[0068] 具体来说,流量优化决策单元 30 通过连接到网络设备来获取网络信息的网络设备交互单元 42,对通过的流量优化请求进行优化,决策出适合具体应用请求的流量优化方案。此外,网络接口 41 通过连接到现有网络设备获知各设备处的资源占用情况、资源分配策略等信息,并提供给流量优化决策单元 30,由此用于流量优化决策单元 30 的流量优化决

策。例如,当无线网络部分出现拥塞时,网络设备交互单元 42 通过网络接口 41 获取到该拥塞信息,并且取得基站的流量分配策略,将这些网络信息提供给流量优化决策单元 30 来进行流量优化。例如通过网络信息了解到某应用由于无线网络部分拥塞而流量交付受到影响时,流量优化决策单元 30 根据该应用的具体信息结合网络设备交互单元 42 反馈的网络状态信息,做出决策,具体包括是否优化其流量以及分配出多少资源给应用所对应的承载等。此外,流量优化决策单元 30 具体的决策方法也可以采用现有的流量优化决策方法。这里不再赘述。

[0069] 网络接口 41 和网络设备交互单元 42 相当于“网络信息取得单元”,取得与网络有关的网络信息。

[0070] 图 4 是第一实施方式涉及的流量管理装置的应用请求控制的流程图。

[0071] 如图 4 所示,在步骤 S501,接收单元 10 从终端 402 或者远端 403 接收应用的流量优化请求的情况下,应用请求过滤单元 20 按照被预定的判断准则,例如应用提供商与网络运营商之间所签订的关于流量优化的签约信息、应用能够容忍的请求最长处理时间、或者网络运营商对于应用流量优化可能性的判断结果等,对所接收到流量优化请求进行过滤,判断是否接受该请求(步骤 S502)。

[0072] 在步骤 S502 判断为不接受时,即将应用的流量优化请求判断为“不通过”的情况下(步骤 S502 为“否”),应用请求过滤单元 20 将该流量优化请求丢弃,并通知给请求发送源的远端 403。

[0073] 在步骤 S502 判断为接受时,即将应用的流量优化请求判断为“通过”的情况下(步骤 S502 为“是”),应用请求过滤单元 20 将该流量优化请求发送给流量优化决策单元 30 加以保存(步骤 S504),等待流量优化决策单元 30 根据该请求而对相应的应用进行流量的优化。(步骤 S505)。

[0074] 通过以上流程,完成对流量优化请求的管理。

[0075] 通过对流量优化的请求进行过滤处理,能够预先过滤掉那些实行可能性低下或者占用过多资源而效率低的应用流量优化请求,从而能够合理配置流量优化资源,能够更好地实现网络对流量的交付,更好地发挥网络资源的价值以满足服务提供商的需求,从而能够提高网络资源的利用效率。

[0076] (第二实施方式)

[0077] 图 5 是包括第二实施方式涉及的流量管理装置的通信网络的结构框图。

[0078] 第二实施方式的流量管理装置与第一实施方式相比较,区别点在于还包括请求队列 50,其他部分的功能与第一实施方式相同。因此,对于相同或类似的部分,赋予相同的附图标记,并且适当省略重复说明的部分。

[0079] 如图 5 所示,本实施方式的流量管理装置 100 设置在通信网络 401 中,经由网络接口 41 而与通信网络中的设备 404 相连接。此处,设通信网络是 LTE 网络。通信网络中的设备 404 可以是图 2 中包含的 LTE 网络的设备中的某个。具体来说,是指基站、移动管理实体、服务网关、归属用户服务器、数据网关等中的某个。

[0080] 其中,现有通信网络中的设备 404 通过无线连接能够与移动终端 402 进行通信,并且通过有线连接能够与远端 403 进行通信,从而能够在移动终端 402 与远端 403 之间建立承载。

[0081] 流量管理装置 100 包括接收单元 10、应用请求过滤单元 20、流量优化决策单元 30、请求队列 50、网络接口 41 以及网络设备交互单元 42。

[0082] 接收单元 10 从终端 402 或者远端 403 接收应用的流量优化请求，相当于“应用请求接收单元”。

[0083] 应用请求过滤单元 20 按照预定的准则，对接收单元 10 接收到的应用的流量优化请求进行过滤，判断是否保留流量优化请求。其中，应用请求过滤单元 20 进行过滤所依据的准则可以包括：应用提供商与网络运营商之间所签订的关于流量优化的签约信息、应用能够容忍的请求最长处理时间、或者网络运营商对于应用流量优化可能性的判断结果等。

[0084] 应用请求过滤单元 20 进行过滤的多个准则是能够对流量优化处理进行过滤的过滤条件即可，可以根据具体的网络环境进行设定。并且，以上多个准则也可以同时适用，或者分先后地适用来配合进行流量优化请求的判断。

[0085] 此外，请求队列单元 50 将应用请求过滤单元 20 经过过滤而接受（判断为“通过”）的流量优化请求依次放入队列中加以排列，从而对多个流量优化请求赋予处理的顺序。

[0086] 在第二实施方式中，请求队列单元 50 通过先入先出的原则，即先收到的请求先处理的方式排列流量优化请求。只有通过应用请求过滤单元 20 的过滤而保留的请求才会进入到请求队列中。

[0087] 流量优化决策单元 30 按照请求队列单元 50 中的排列顺序，依次针对经过应用请求过滤单元 20 的过滤而保留（被判断为“通过”）的流量优化请求进行响应，利用网络信息，生成与经过过滤而保留的流量优化请求相对应的流量优化方案。

[0088] 网络接口 41 和网络设备交互单元 42 相当于“网络信息取得单元”，取得与网络有关的网络信息。

[0089] 在第二实施方式中，请求队列单元 50 的排列原则为先入先出的原则，但是，流量优化请求的排列顺序并不仅限于此，还可以是其他排列顺序。

[0090] 例如，请求队列单元 50 能够根据流量优化请求中所附带的信息对请求赋予优先级，将优先级高的请求排列在前而优先进行处理。例如在流量优化请求中附带了希望处理时间的信息等时效性要求时，使流量管理装置 100 能够优先处理时效性要求高的流量处理请求。

[0091] 此外，请求队列单元 50 还能够根据所请求流量优化的应用的类别进行排序，以便优先处理重要度高的应用的流量优化请求。

[0092] 图 6 是第二实施方式涉及的流量管理装置的应用请求控制的流程图。

[0093] 如图 6 所示，在步骤 S801，接收单元 10 从终端 402 或者远端 403 接收应用的流量优化请求的情况下，应用请求过滤单元 20 按照被预定的判断准则，例如应用提供商与网络运营商之间所签订的关于流量优化的签约信息、应用能够容忍的请求最长处理时间、或者网络运营商对于应用流量优化可能性的判断结果等，对所接收到流量优化请求进行过滤，判断是否接受该请求（步骤 S802）。

[0094] 在步骤 S802 判断为不接受时，即将应用的流量优化请求判断为“不通过”的情况下（步骤 S802 为“否”），应用请求过滤单元 20 将该流量优化请求丢弃，并通知给请求发送源的远端 403。

[0095] 在步骤 S802 判断为接受时，即将应用的流量优化请求判断为“通过”的情况下（步

骤 S802 为“是”),应用请求过滤单元 20 将该流量优化请求发送给请求队列单元 50,请求队列单元 50 根据预定的排列原则,将该流量优化请求加入到队列中,等待进行处理(步骤 S804)。

[0096] 流量优化请求在队列中处于最前列时,流量优化决策单元 30 根据该流量优化请求而对相应的应用进行流量的优化(步骤 S805)。

[0097] 通过以上流程,完成对流量优化请求的管理。

[0098] 通过对流量优化的请求进行过滤处理,能够预先过滤掉那些实行可能性低下或者占用过多资源而效率低的应用流量优化请求,从而能够合理配置流量优化资源,实现与第一实施方式相同的效果。并且,通过对过滤后的流量优化请求进行排列,能够按照适当的顺序更加有效地进行流量优化处理,从而进一步提高网络资源的利用效率。

[0099] 此外,流量优化请求不但能够由远端发起,还可以由移动终端发起,以下分别说明这两种发起模式下的消息交互顺序。

[0100] 图 7 是基于第三方应用流量优化的流量控制的消息交互图。图 7 表现出移动终端发起请求并且流量优化请求在过滤中被接收的情形。

[0101] 如图 7 所示,当移动终端 402 希望通过移动网络实现某种应用时,会依次通过无线接入网 200、核心网 300 连接到远端 403,具体的承载从移动终端 402 连接到分组数据网关(认为在核心网 300 的范围内),且应用产生的流量传输发生在移动终端 402 与远端 403 之间(步骤 605)。

[0102] 承载信息包含承载类型为 GBR 或 Non-GBR、QCI 以及网络对应保证的流量交付时延等。当移动终端 402 观测到应用性能下降(步骤 606),且判断可能由于网络问题导致,而向流量管理装置 100 发送请求网络进行流量优化的流量优化请求(步骤 608)。

[0103] 流量管理装置 100 接收应用侧(承载连接的双方,即移动终端 402 或远端 403)发起的流量优化请求,并且通过应用请求过滤单元 20 对接收到的应用请求进行过滤,得到“通过”的过滤结果,从而接受该流量优化请求(步骤 609)。过滤后的流量优化请求进入请求队列单元 50,由流量优化决策单元根据该流量优化请求得到相应的流量优化方案。该流量优化方案将被告知给网络侧的设备按照所决策的方案具体实施流量优化(步骤 610),且会涉及到移动终端和网络之间的 QoS 信令交互(步骤 611)。在步骤 611,根据传统的 QoS 建立或改变承载的信令交互来进行。

[0104] 最终,得到被优化了的承载(步骤 612:例如由 Non-GBR 改为 GBR),从而流量优化过程结束。

[0105] 图 8 是基于第三方应用流量优化的流量控制的另一消息交互图。图 8 表现出远端发起请求并且流量优化请求在过滤中被丢弃的情形。

[0106] 如图 8 所示,当移动终端 402 希望通过移动网络实现某种应用时,会依次通过无线接入网 200、核心网 300 连接到远端 403,具体的承载从移动终端 402 连接到分组数据网关(认为在核心网 300 的范围内),且应用产生的流量传输发生在移动终端侧和远端侧之间(步骤 705)。当远端 403 观测到应用性能下降(步骤 706)且判断可能由于网络问题导致,则向流量管理装置 100 发送请求网络进行流量优化的流量优化请求(步骤 708)。

[0107] 流量管理装置 100 接收应用侧发起的流量优化请求,并且通过应用请求过滤单元 20 对接收到的流量优化请求进行过滤,得到“不通过”的过滤结果,而丢弃该流量优化请求

(步骤 709)。

[0108] 被丢弃的流量优化请求不会进入请求队列单元 50,且不会得到流量管理装置 100 所决策的流量优化方案。流量管理装置 100 将过滤结果告知给应用侧(远端或者移动终端)(在步骤 710 为告知给移动终端 402),由应用侧自行决定是否自身可进行流量优化,这种自身的流量优化包括:在流量为上传的视频时可在移动终端侧选择降低视频清晰度或减少流量传输需要的数据速率等)。

[0109] 最终,之前建立的承载将继续被保持(步骤 711),由网络实施的流量优化过程结束,应用侧可自行选择其他不涉及网络的流量优化方案。

[0110] (第三实施方式)

[0111] 图 9 是包括第三实施方式涉及的流量管理装置的通信网络的结构框图。

[0112] 第三实施方式的流量管理装置与第二实施方式相比较,区别点在于还包括定时器 60,其他部分的功能与第二实施方式相同。因此,对于相同或类似的部分,赋予相同的附图标记,并且适当省略重复说明的部分。

[0113] 如图 9 所示,本实施方式的流量管理装置 100 设置在通信网络 401 中,经由网络接口 41 而与通信网络中的设备 404 相连接。此处,设通信网络是 LTE 网络。通信网络中的设备 404 可以是图 2 中包含的 LTE 网络的设备中的某个。具体来说,是指基站、移动管理实体、服务网关、归属用户服务器、数据网关等中的某个。

[0114] 其中,现有通信网络中的设备 404 通过无线连接能够与移动终端 402 进行通信,并且通过有线连接能够与远端 403 进行通信,从而能够在移动终端 402 与远端 403 之间建立承载。

[0115] 流量管理装置 100 包括接收单元 10、应用请求过滤单元 20、流量优化决策单元 30、请求队列 50、定时器 60、网络接口 41 以及网络设备交互单元 42。

[0116] 接收单元 10 从终端 402 或者远端 403 接收应用的流量优化请求,相当于“应用请求接收单元”。

[0117] 应用请求过滤单元 20 按照预定的准则,对接收单元 10 接收到的应用的流量优化请求进行过滤,判断是否保留流量优化请求。其中,应用请求过滤单元 20 进行过滤所依据的准则可以包括:应用提供商与网络运营商之间所签订的关于流量优化的签约信息、应用能够容忍的请求最长处理时间、或者网络运营商对于应用流量优化可能性的判断结果等。

[0118] 此外,请求队列单元 50 将应用请求过滤单元 20 经过过滤而接受(判断为“通过”)的流量优化请求依次放入队列中加以排列,从而对多个流量优化请求赋予处理的顺序。

[0119] 定时器 60 用于设定定时。具体来说,定时器 60 按照应用请求中包含的可容忍的请求最长处理时间来设置定时,当从接收到某个流量优化请求并进行过滤起经过了该流量优化请求的最长处理时间时,定时器 60 告知请求队列单元 50 查找是否包含此流量优化请求,如该流量优化请求仍在队列中,则删除相应的流量优化请求且不进行流量优化。定时器 60 相当于“定时器单元”。

[0120] 例如,请求最长处理时间可以包含在流量优化请求中,从而通过提取来获得,也可以预先以协议信息等类型存储在流量管理装置中。

[0121] 例如,在请求最长处理时间为 3s(秒)且包含该请求最长处理时间的流量优化请求通过了应用请求过滤单元 20 的过滤而进入队列中的情况下,定时器 60 启动,在定时结束

时,查找该流量优化请求,如果该流量优化请求仍在队列中未被处理,则丢弃该流量优化请求。这是因为,超过应用可忍受的最长处理时间后,由于应用性能可能已经随着网络拥塞状况的变化而有所改善等原因,而不再需要流量优化,从而处理该流量优化请求变得无意义。

[0122] 此外,若所接收到的流量优化请求并不具有时效性要求,则定时器 60 不需要启动,或者也可以省略该定时器 60。

[0123] 此外,以上定时器 60 根据请求最长处理时间进行定时,但是这仅是一个例子,定时器 60 也可以被设定为其他定时,例如能够对所有进入流量管理装置的流量优化请求设定一个最长处理时间,以避免存在一直在队列中排队而长期未被处理的流量优化请求。

[0124] 在定时器 60 的定时未结束的情况下,流量优化决策单元 30 按照请求队列单元 50 中的排列顺序,依次针对经过应用请求过滤单元 20 的过滤而保留(被判断为“通过”)的流量优化请求进行响应,利用网络信息,生成与经过过滤而保留的流量优化请求相对应的流量优化方案。

[0125] 网络接口 41 和网络设备交互单元 42 相当于“网络信息取得单元”,取得与网络有关的网络信息。

[0126] 此外,以上说明中,在定时结束时,在队列查找该到期的流量优化请求,如果还在队列中则进行丢弃。但是,也可以在对某个流量优化请求进行处理时,判断其对应的定时器 60 是否过期,在定时器 60 过期时不进行该流量优化请求所对应的流量优化决策,而丢弃该流量优化请求,进行针对队列中的下个流量优化请求的处理。

[0127] 图 10 是第三实施方式涉及的流量管理装置的应用请求控制的流程图。

[0128] 如图 10 所示,在步骤 S901,接收单元 10 从终端 402 或者远端 403 接收应用的流量优化请求的情况下,应用请求过滤单元 20 按照被预定的判断准则,判断是否接受该请求(步骤 S902)。

[0129] 在步骤 S902 判断为不接受时,即将应用的流量优化请求判断为“不通过”的情况下(步骤 S902 为“否”),应用请求过滤单元 20 将该流量优化请求丢弃,并通知给请求发送源的远端 403。

[0130] 在步骤 S902 判断为接受时,即将应用的流量优化请求判断为“通过”的情况下(步骤 S902 为“是”),应用请求过滤单元 20 将该流量优化请求发送给请求队列单元 50,并且提取该流量优化请求中包含的请求处理最长时间参数,根据请求处理最长时间参数设定并启动定时器 60(设置为  $T_n$ )。请求队列单元 50 根据预定的排列原则,将该流量优化请求加入到队列中,等待进行处理(步骤 S904)。

[0131] 接着,在步骤 S905,进行定时器 60 是否过期的判断。当队列中进行到的流量优化请求所对应的定时器被判断为过期,或者在  $T_n$  过期时在请求队列中查找到所对应的流量优化请求时(步骤 S905 为“是”),进入步骤 S903,不进行流量优化决策。

[0132] 例如,流量优化请求中包括的请求最长处理时间为 3s,且此流量优化请求通过应用请求过滤单元 20 进入队列中,则定时器 60 启动且设置  $T_n = 3$ ,在定时 3s 结束时,如果所对应的流量优化请求仍未能被处理,则被队列丢弃。

[0133] 另一方面,当队列中进行到的流量优化请求所对应的定时器被判断为未过期,或者在  $T_n$  过期时在请求队列中没有查找到所对应的流量优化请求时(步骤 S905 为“否”),流量优化决策单元 30 根据该流量优化请求而对相应的应用进行流量的优化(步骤 S906)。

[0134] 通过以上流程,完成对流量优化请求的管理。

[0135] 通过对流量优化的请求进行过滤处理,能够预先过滤掉那些实行可能性低下或者占用过多资源而效率低的应用流量优化请求,从而能够合理配置流量优化资源,实现与第一实施方式和第二实施方式相同的效果。并且,通过设定定时器来控制对某个流量优化请求的处理时限,能够进一步提高网络资源的利用效率。

[0136] 下面例举出几个具体应用的实例,来说明如何具体实施本发明涉及的流量管理方法。

[0137] (实施例 1)

[0138] 随着移动网络技术的发展,3G 以及 LTE 网络已经能够支持较大的上行速率。加上中国为了更好地实现对公共安全的保护,正在对安防领域逐步升级,尤其基于智能安防的应用得到了较快发展。而且基于移动网络实现智能安防已经变得可能,3G 或者 LTE 已经能够支持视频这样数据量的流量通过网络上行传输到监控中心。

[0139] 并且,一些外出的巡警车也配备有移动摄像头,这一类的摄像头无法通过有线方式连接到监控中心,最可能的网络接入方式是通过网络运营商的 3G/LTE 网络,或者 Wifi 接入。而由于所述巡警车常常停靠在有警情的地段,此地很可能没有 Wifi 接入,而通常网络运营商的信号覆盖已经非常全面,故配有移动摄像头的巡警车可将 3G/LTE 接入作为主要的接入方式,其拍摄的视频将通过上传的方式发送到远处的监控中心,即移动摄像头作为移动终端,其产生的流量通过移动网络上行发送至远端的监控中心。

[0140] 实施例 1 中,以智能安防应用展开对本发明的流量管理方法的说明,其中,利用 LTE 网络,但是在 3G 网络中实施也相同。

[0141] 在实施例 1 中,移动摄像头作为移动终端通过 LTE 网络或 Wifi 向远端的监控中心不断上传实时拍摄的视频画面,以智能安防为例,说明流量优化装置拒绝应用优化请求的过程。具体网络接入方式的选择可能是在某些已设置好 Wifi 的地段(如人群密集的大型购物中心附近、机场航站楼附近等),摄像头通过 Wifi 接入,但在其他未设置好 Wifi 的地段(可包含其他任意有 LTE 网络覆盖的地段)摄像头通过 LTE 网络接入。当监控中心发现某个移动摄像头拍摄的画面出现了重要目标或异常事件时,会严格需要此时的拍摄画面清晰流畅,使得监控中心能够根据现场情况判断是否需要出警以及需要调派何种程度的警力。如果此时在远端的监控中心观测到出现重要目标或异常事件的视频画面出现花屏、卡顿等影响判断现场的情况,远端判断可能是网络出现了问题导致视频流量传输出现了较大的丢包(会导致花屏,看不清画面)或很大的时延(会导致卡顿,画面可能很久才恢复,失去了追踪重要目标的良机),从而向网络侧发送“流量优化请求”。网络侧的流量管理装置(例如第一实施方式~第三实施方式中的流量管理装置 100)会依据收到的请求进行一系列流量优化的步骤。

[0142] 首先,应用请求过滤单元处理该流量优化请求,所依据的准则包括服务提供商是否与网络运营商依据流量优化达成合约以及运营商优化流量的可能性等。在本实例中,假定智能安防的服务提供商(通常是公安局)与网络运营商已约定对于此应用的流量尽力优化。

[0143] 此外,假定出现应用性能下降时,移动摄像头的连接方式是 Wifi 连接。在应用请求过滤单元处的判断时,虽然服务提供商与网络运营商已签约,但是,该承载的连接并非所

属此网络运营商运营的网络,而是 Wifi 网络,则接收到流量优化请求的运营商无法去改变其他运营商网络的配置(包括 Wifi 网络)来实现优化,因此,应用请求过滤单元判断为不接受该流量优化请求。

[0144] 接着,流量管理装置将该结果告知给移动终端侧,移动终端侧可自行降低清晰度以换取低清晰度并流畅的视频效果,或者告知远端当前网络不属于运营商管辖范围,没有优化的可能性。应用将依然维持之前的承载。

[0145] (实施例 2)

[0146] 在实施例 2 中,依然与实施例 1 同样地以智能安防应用为例。

[0147] 移动摄像头作为移动终端通过 LTE 网络向远端的监控中心不断上传实时拍摄的视频画面。在本实例中,假设监控中心刚刚发现犯罪嫌疑人并希望追踪,而且期望网络以很快速度响应,否则犯罪嫌疑人已经离开某个摄像头的拍摄区域,则流量优化变得意义不大。故本实施例 2 中假设应用远端发送的流量优化请求携带“可容忍请求被处理的最长时间为 3s”这一信息,且智能安防的服务提供商(通常是公安局)与网络运营商已约定对于此应用的流量尽力优化。网络侧的流量管理装置依据收到的请求进行一系列流量优化的步骤。

[0148] 首先,应用请求过滤单元会处理该流量优化请求,所依据的准则包括服务提供商是否与网络运营商依据流量优化达成合约以及运营商优化流量的可能性等。在应用请求过滤单元进行判断时,由于服务提供商与网络运营商已签约而且应用使用的是 LTE 网络(有流量优化的可能性),从而应用请求过滤单元接受该流量优化请求,并将其加入到队列中。

[0149] 但是,假定流量管理装置的请求队列单元中排队的请求较多,由于定时器过期时此请求仍未被处理,则结果是请求队列将请求丢弃,并告知应用的远端或移动终端侧“未作流量优化”的决定。移动终端或远端可通过自行降低视频清晰度(视频码率)等自适应方式去适应网络,以获得较满意的应用性能。

[0150] (实施例 3)

[0151] 在实施例 3 中,依然以智能安防应用为例。

[0152] 移动摄像头作为移动终端通过 LTE 网络向远端的监控中心不断上传实时拍摄的视频画面,且一些公安人员携带有可移动的视频终端(可认为是移动终端的另一种实际形式),可以通过访问远端的监控中心观看视频,并实施对犯罪嫌疑人在安全距离内的追踪。

[0153] 在本实例中,假设公安人员刚刚从携带的移动视频终端的视频画面中发现犯罪嫌疑人并希望追踪,而且期望网络以很快速度响应,否则犯罪嫌疑人的行动轨迹很难继续被相应公安人员获知,则实施追踪变得十分困难,流量优化变得意义不大。故本实例假设应用远端发送的流量优化请求携带“可容忍请求被处理的最长时间为 3s”这一信息,且智能安防的服务提供商(通常是公安局)与网络运营商已约定对于此应用的流量尽力优化。

[0154] 网络在应用的普通情况(公安人员未从移动终端的视频画面发现犯罪嫌疑人时)为此应用连接建立了 Non-GBR(尽力而为的)承载,此承载可能是从连接最初开始时就建立,公安人员观察到的视频质量尚可。

[0155] 假定在应用的特殊情况时(公安人员在携带的移动视频终端的视频画面发现犯罪嫌疑人并希望追踪)突发网络拥塞,导致移动终端侧的视频画面出现花屏、卡顿等影响判断现场的情况。移动终端侧判断可能是网络出现了问题导致视频流量传输出现了较大的丢包(会导致花屏,看不清画面)或很大的时延(会导致卡顿,画面可能很久才恢复,失去



了追踪重要目标的良机),从而向网络侧发送“流量优化请求”。

[0156] 网络侧的流量管理装置依据收到的流量优化请求,首先,应用请求过滤单元根据应用的签约信息(已签约)、优化可能性(使用此运营商的LTE网络,可优化)并且记录了应用可容忍的请求最长处理时间(3s),判断为接受该流量优化请求,从而将该流量优化请求放入请求队列中。

[0157] 假定在定时器(3s)过期前,此请求已经被受理,流量管理装置针对该流量优化请求制作适当的流量优化方案。具体来说,流量优化决策单元根据通过网络设备交互单元得到网络中各设备处的信息,这些信息进一步由网络接口从现有网络的各设备处获取。流量优化决策单元根据网络中的流量数量及各流量承载对应的类型、QCI等信息,判断为将另一未签约的FTP应用(承载也为Non-GBR类型)速率进行限制,以节约出更多的资源给此智能安防应用。则网络通过必要的QoS信令将承载修改为保证速率的GBR承载,此后在发起流量优化的移动终端侧观察到视频质量改善。

[0158] 又或者流量优化决策单元根据网络中的流量数量及各流量承载对应的类型、QCI等信息,判断为网络中已经没有足够的资源分配给此智能安防应用,但由于网络运营商与服务提供商已签约“尽力”优化其流量,故网络分析自身的负荷后,估计出拥塞可以缓解的时间(假定为5min),告知给移动终端:“未作流量优化,建议降低速率,大致5min后可恢复高清晰度传输”。

[0159] 以上的第二种流量优化决策方案也可视为网络在应用接收的时间内处理其流量优化请求,并在某种程度上提供了流量优化(或性能提升)的可能。从而应用获得到更好的流量交付体验。

[0160] 通过应用请求过滤单元对流量优化请求的过滤,更好地实现了网络对流量的交付,更好地发挥网络资源的价值以满足服务提供商的需求。

[0161] 流量优化设备处理不带有时效性要求的流量优化请求的过程更为简单,但也具有很强的现实性。例如,一些与网络运营商签约的第三方应用服务提供商,希望在指定时间段内尽量优化其流量(与前面实例“希望立即优化”的请求相比,有很弱的时效性要求,甚至没有时效性要求)。以一些餐饮类点评应用为例,应用服务提供商希望在每天的晚餐时间段(例如16点至20点)网络优先保证其流量的交付,这样用户使用此应用时体验更佳,应用会获得更好的用户评价。当网络状况不佳导致流量交付体验差时,此时应用侧发起的流量请求的时效性并不强(应用侧的请求是“希望网络对其流量做优化,但可以排队等待网络的处理”),从而流量优化设备在处理流量优化请求的流程将如图6所示,同时流量优化网络设备起作用的功能模块如图5所示的结构图(不需要定时功能)。

[0162] 此外,请求被接受或被拒绝的消息交互过程在第一实施方式至第三实施方式中并没有区别,都是如图7、8所示。

[0163] (变形例)

[0164] 在以上各个实施方式中,流量管理装置100可以将流量优化请求所得到的结果告知请求源或者应用侧的移动终端或远端,以便应用侧设备进行其他流量优化方案,但是也可以不告知流量优化请求所得到的结果,同样能够实施本发明。

[0165] 此外,实施方式中的各设备与网络的联系方式并没有特别限制,例如能够使用无线连接或者有线连接。

[0166] 此外,在以上各个实施方式中,流量管理装置作为独立的装置,与网络中的已有设备连接进行动作,但是也可以安装到网络中的某个设备中而作为网络设备的一部分进行动作。该流量管理装置能够通过软件的功能模块来实现,也能够通过具有相应功能的电路来实现。

[0167] 以上对本发明的几个实施方式进行了说明,但这些实施方式是作为例子而提示的,并非旨在限定发明的范围。这些新的实施方式能够以其他各种方式来实施,在不脱离发明的主旨的范围内能够进行各种省略、替换、变更。这些实施方式、其变形包含于发明的范围、主旨内,并且包含于权利要求所记载的发明及其等同的范围内。

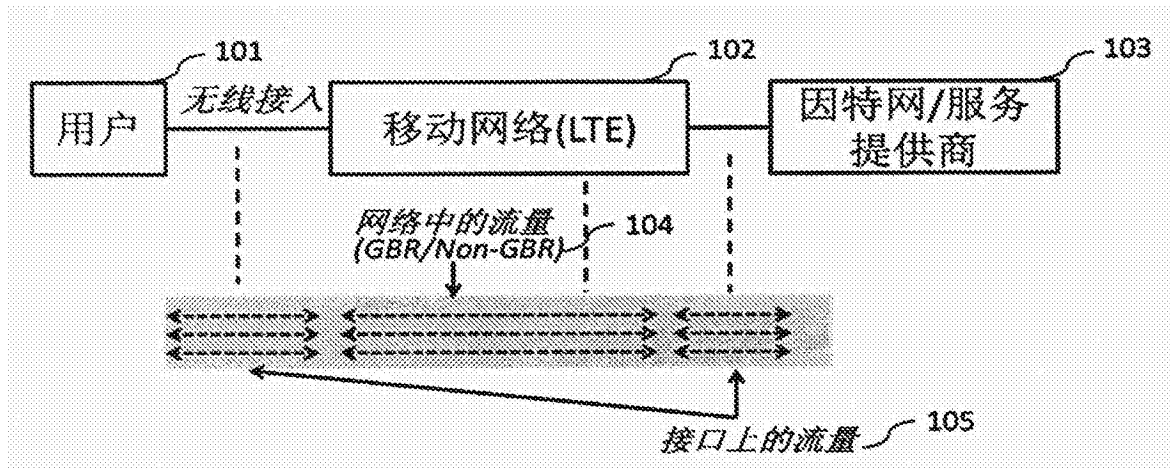


图 1

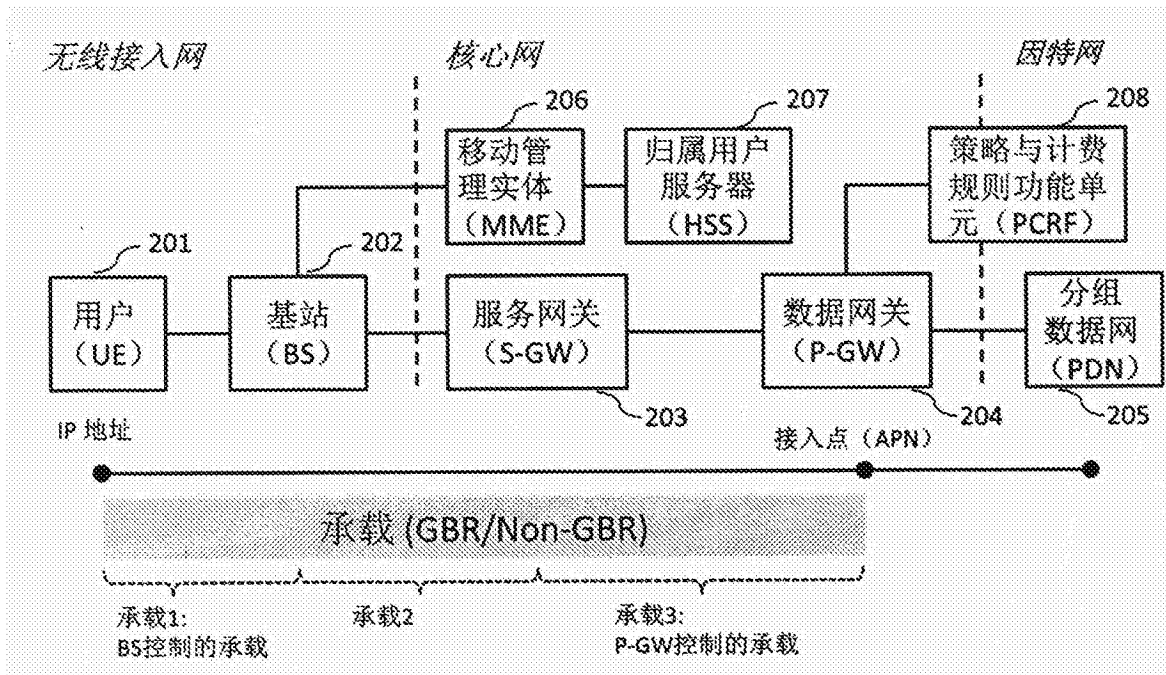


图 2

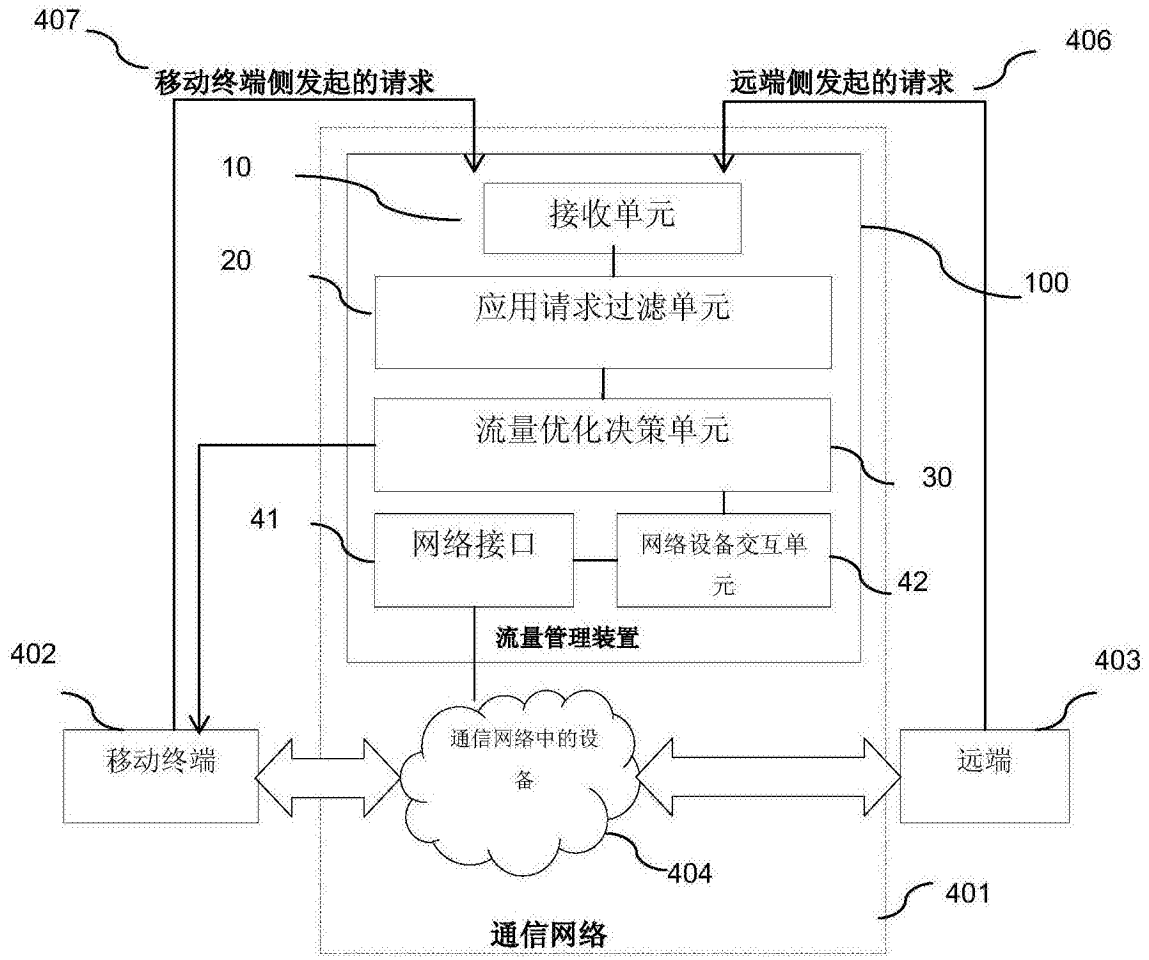


图 3

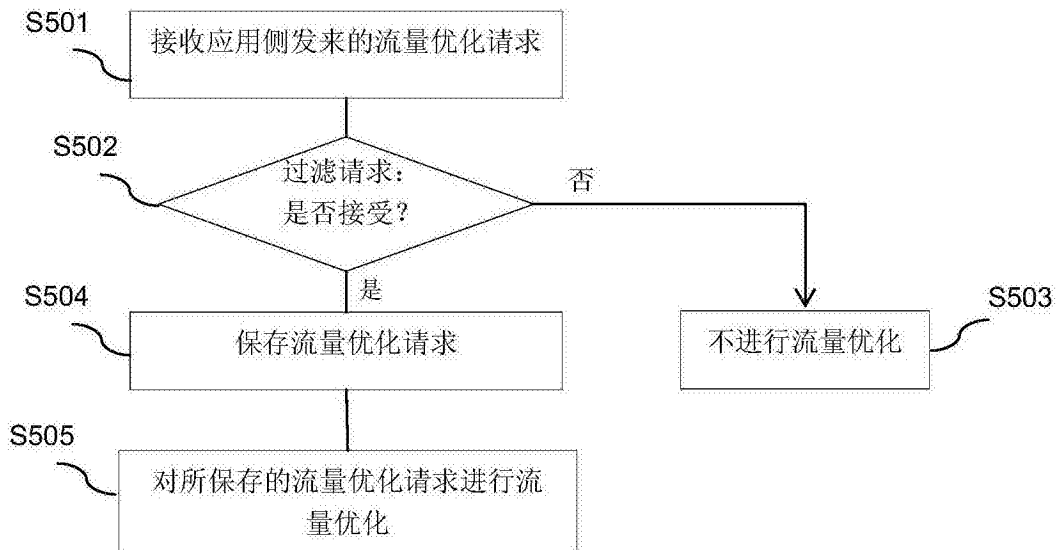


图 4

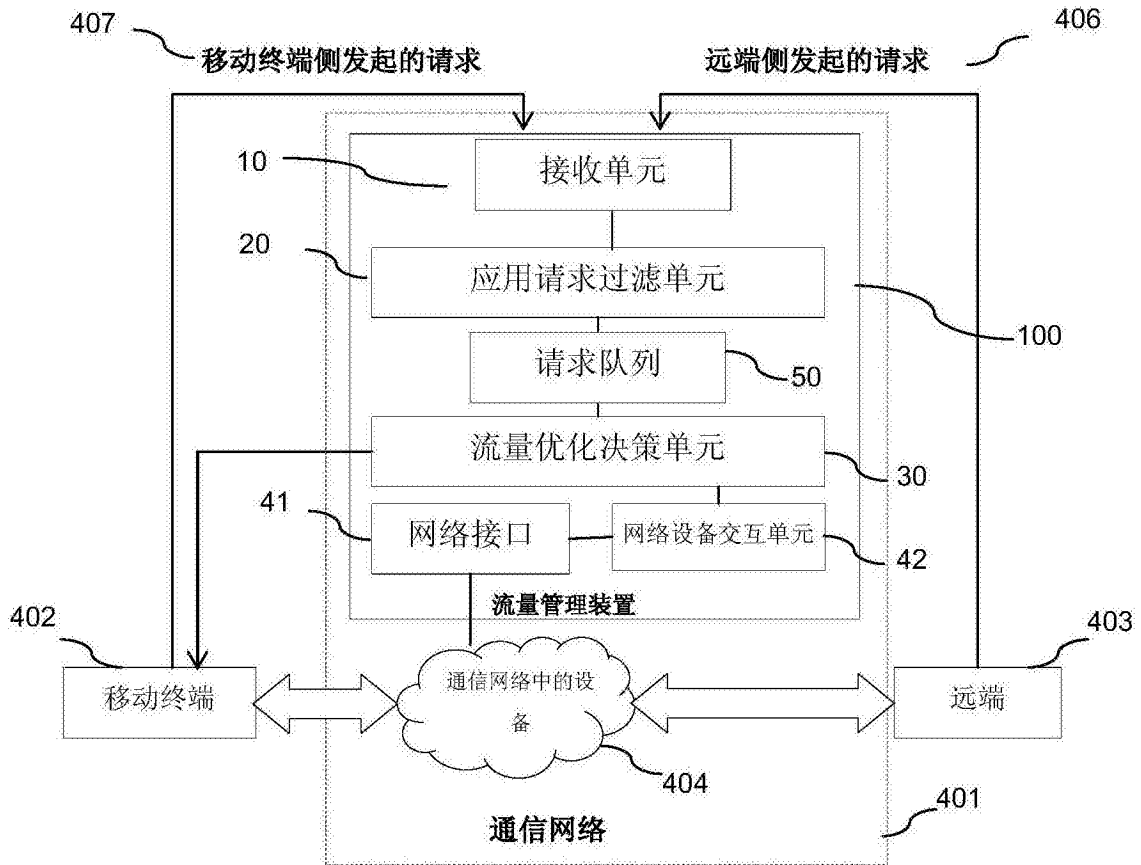


图 5

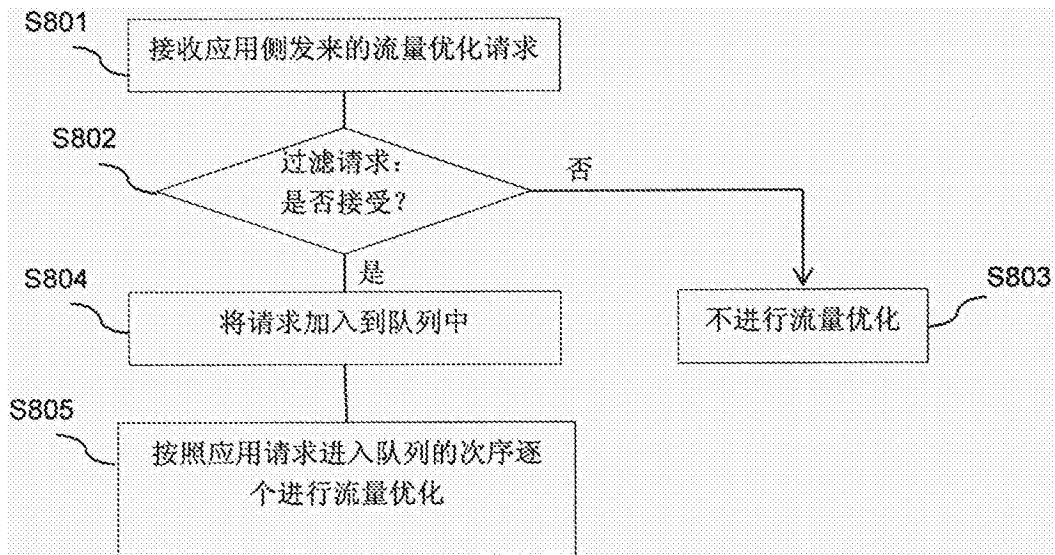


图 6

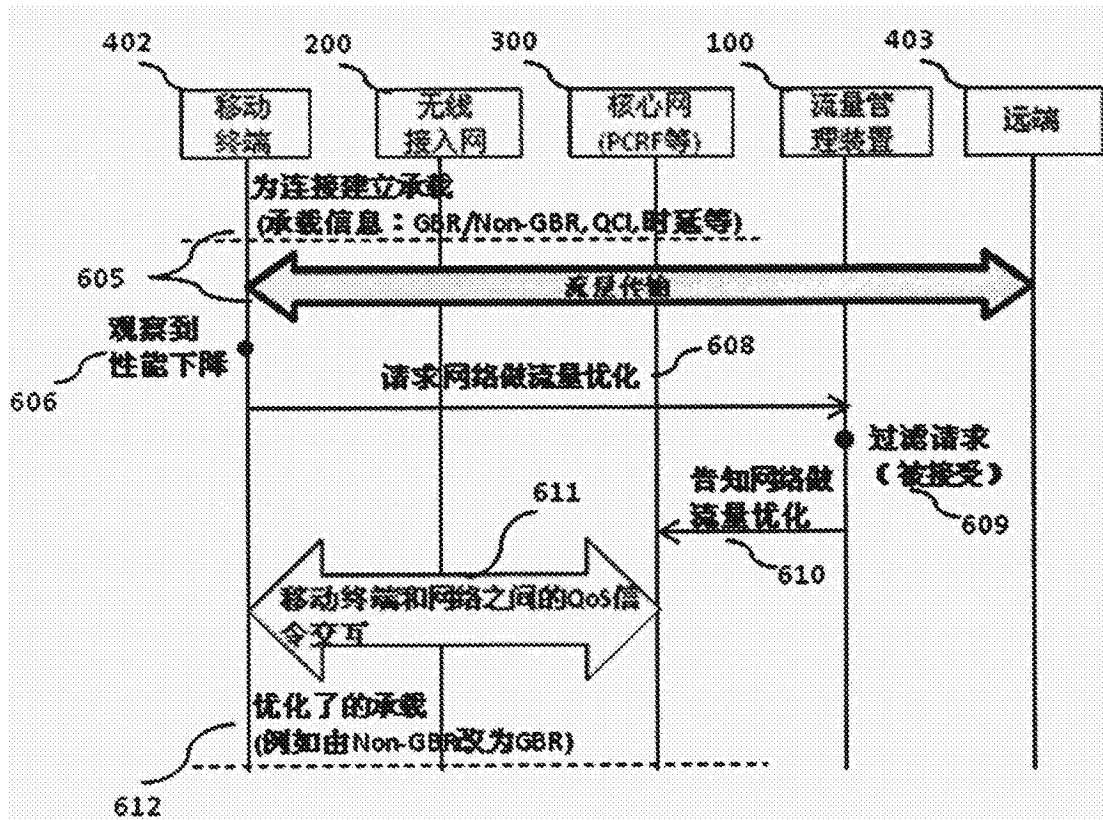


图 7

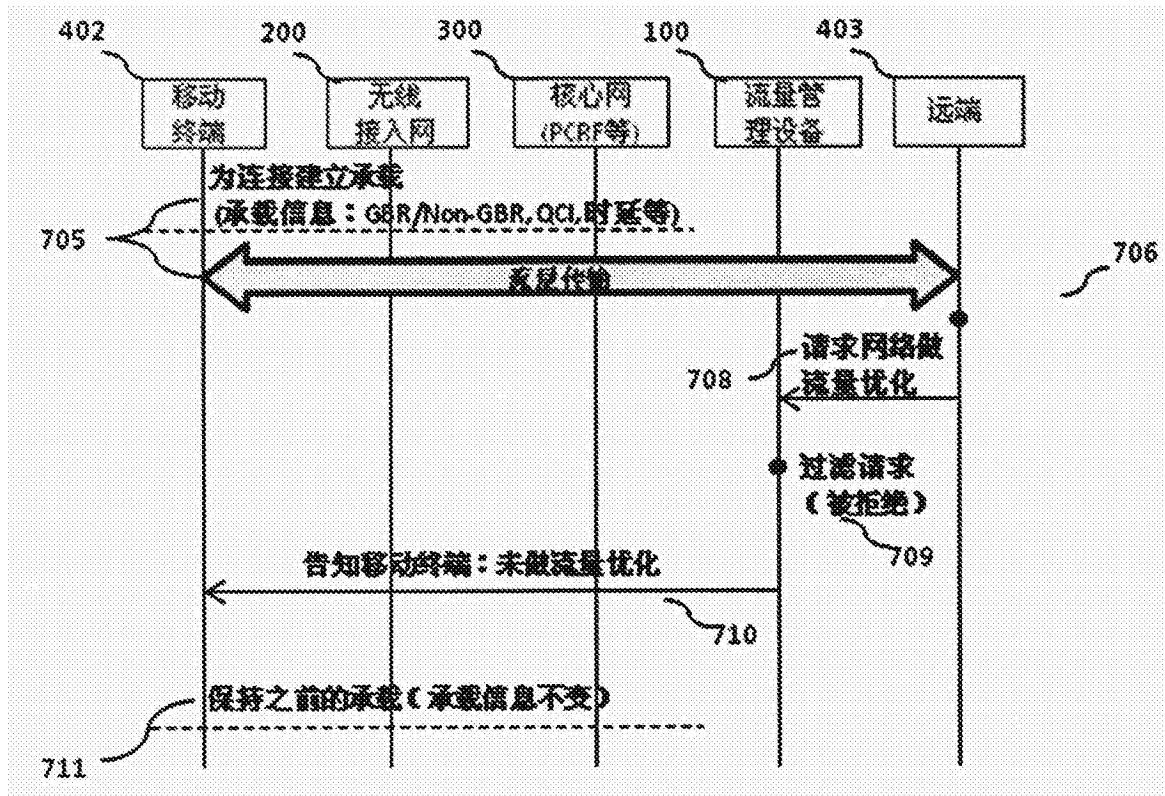


图 8

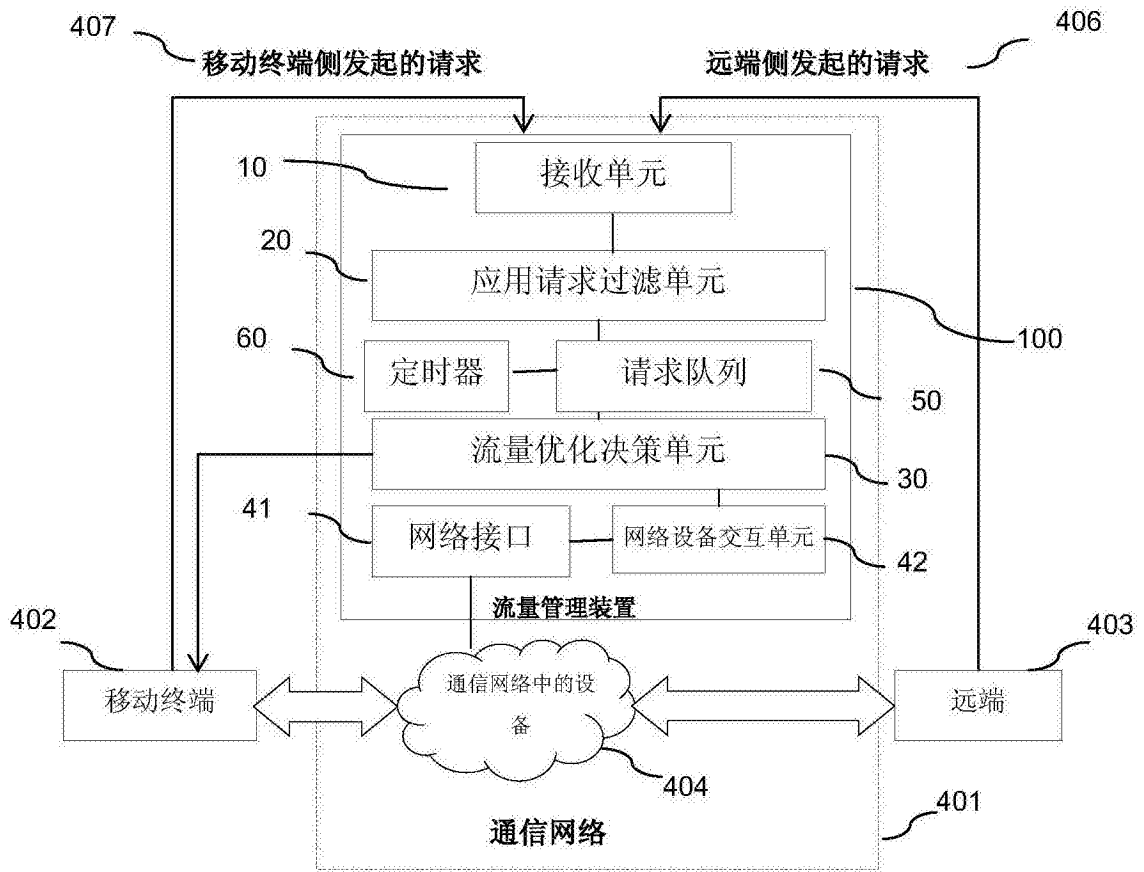


图 9



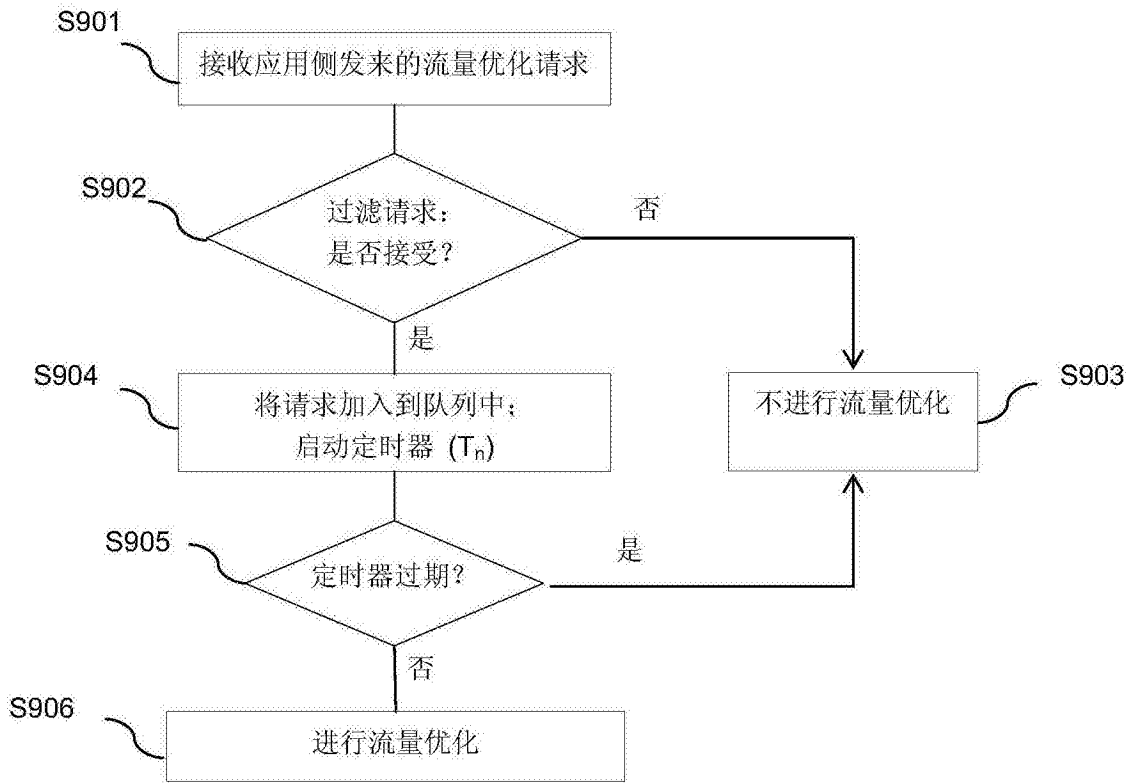


图 10

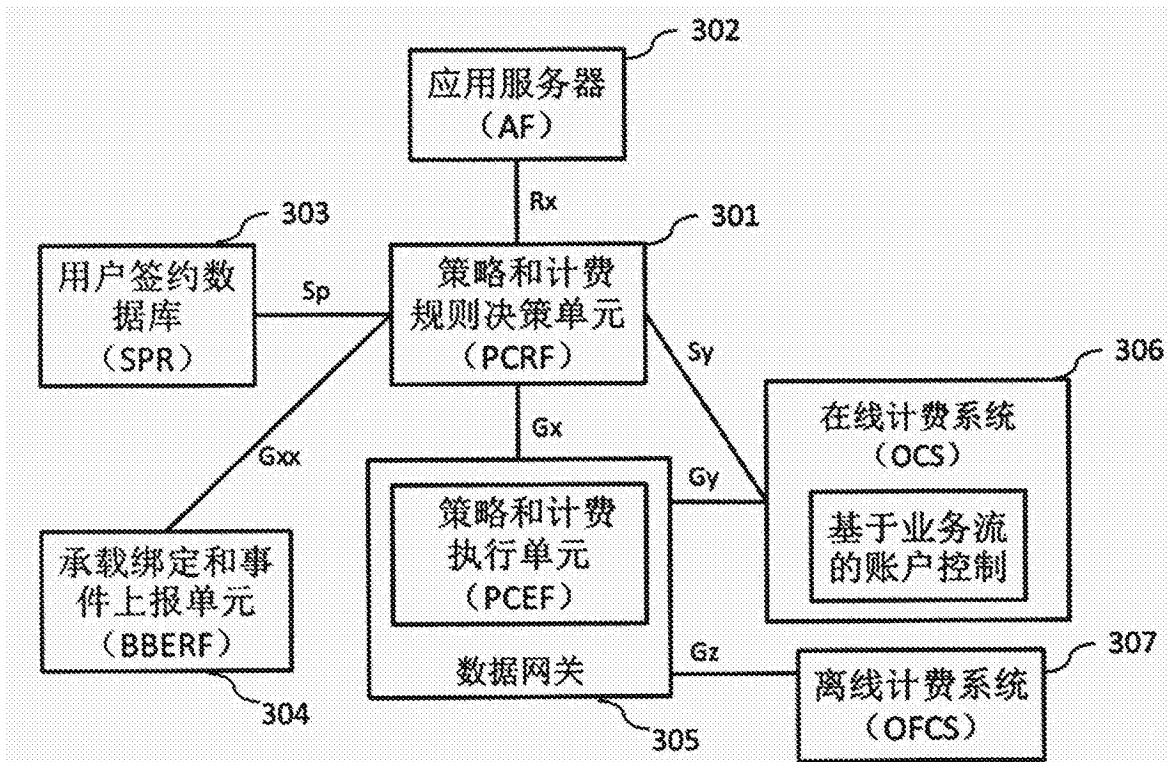


图 11