



[12] 发明专利申请公开说明书

[21] 申请号 01810733.8

[43] 公开日 2003 年 7 月 30 日

[11] 公开号 CN 1433615A

[22] 申请日 2001.5.4 [21] 申请号 01810733.8

[30] 优先权

[32] 2000.5.5 [33] US [31] 60/202,326

[86] 国际申请 PCT/US01/14493 2001.5.4

[87] 国际公布 WO01/86877 英 2001.11.15

[85] 进入国家阶段日期 2002.12.5

[71] 申请人 诺玛迪克斯公司

地址 美国加利福尼亚州

[72] 发明人 乔尔·E·肖特

阿纽雷格·K·加格

维杰·K·巴哈格瓦斯

[74] 专利代理机构 中国国际贸易促进委员会专利
商标事务所

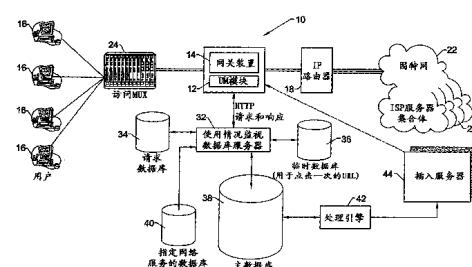
代理人 董 萍

权利要求书 6 页 说明书 18 页 附图 5 页

[54] 发明名称 网络使用情况监视设备和相关方法

[57] 摘要

本发明提供一种用于在网络接入点，即网络通信集合点，一般在网关装置或者类似的网络接口装置监视网络使用情况的网络使用情况监视模块。这样，网络使用情况监视模块能够监视试图通过网关装置访问提供的各种网络服务的若干网络用户的使用情况。从而，和由常规的监视技术提供的使用情况信息相比，使用情况监视模块收集的使用情况信息更全面。从而，对于网络服务提供商、网络用户和网络受益者等来说，该信息价值更高。另外，使用情况监视方法和设备提供许多特殊的功能，以便改进监视过程以及提高收集的使用情况信息的价值。



1、一种监视通信网络使用情况的设备，包括：

在网络接入点工作的使用情况监视模块，该模块捕获从若干网络用户传来的所有请求分组和从若干网络服务传来的所有响应分组；其中使用情况监视模块过滤捕获的分组，以便从中抽取使用情况监视数据；和

与所述使用情况监视模块通信的使用情况监视数据库，该数据库从使用情况监视模块接收过滤后的使用情况监视数据，并保存和用户请求分组及网络响应分组相关的数据。

2、按照权利要求1所述的设备，还包括实现使用情况监视模块的网关装置，其中网关装置在网络内被布置在网络通信集合点，从而它接收来自于若干网络用户的所有请求，并且接收来自于网络服务的所有响应。

3、按照权利要求1所述的设备，其中使用情况监视模块过滤捕获的分组，以便抽取网络地址，使用情况监视数据库保存该网络地址。

4、按照权利要求3所述的设备，其中网络地址还包括URL（统一资源定位符）。

5、按照权利要求1所述的设备，其中使用情况监视模块过滤捕获的分组，以便抽取使用情况监视数据，所述使用情况监视数据包括选自用户标识符、网络地址、分组时间戳、参阅网络地址、内容类型、内容长度、响应状态码和用户查询串的至少一种数据。

6、按照权利要求1所述的设备，其中使用情况监视模块对捕获的分组进行浏览排序，以便监视用户所访问的一系列网络地址。

7、按照权利要求1所述的设备，其中使用情况监视模块对网络服务响应分组进行状态码检查，以便证实用户已访问请求的网络服务地址，以致只有与可访问的网络服务相关的数据才被保存在使用情况监视数据库中。

8、按照权利要求1所述的设备，其中使用情况监视数据库还包括

在从请求的网络服务接收响应之前，临时保存用户请求数据的用户请求数据库。

9、按照权利要求 1 所述的设备，其中使用情况监视数据库还包括临时数据库，所述临时数据临时保存关于用户对其的访问次数不大于预定最小次数的网络地址的用户监视信息。

10、按照权利要求 9 所述的设备，其中预定最小次数为 1。

11、按照权利要求 1 所述的设备，其中使用情况监视数据库还包括保存对于浏览排序所指定的网络地址的指定网络地址数据库。

12、按照权利要求 11 所述的设备，其中使用情况监视模块适合于从使用情况监视客户接收已关于浏览排序指定的网络地址。

13、按照权利要求 11 所述的设备，其中使用情况监视模块根据用户访问网络地址的频率，确定已对于浏览排序所指定的网络地址。

14、按照权利要求 1 所述的设备，其中使用情况监视数据库还包括保存和用户请求分组及网络响应分组相关的数据的主数据库。

15、按照权利要求 1 所述的设备，还包括与使用情况监视数据库通信的插入服务器，该服务器根据网络用户使用情况监视数据向网络用户提供针对性信息。

16、按照权利要求 15 所述的设备，其中插入服务器还包括根据网络用户使用情况监视数据，向网络用户提供针对性广告的广告插入服务器。

17、按照权利要求 15 所述的设备，其中插入服务器还包括根据网络用户使用情况监视数据，向网络用户提供针对性调查的调查插入服务器。

18、一种在通信网络中提供使用情况监视的方法，所述方法包括：

在网络接入点捕获从若干用户和若干网络服务传来的所有数据分组；

过滤捕获的分组，提供使用情况监视数据；和

把使用情况监视数据保存在使用情况监视数据库中。

19、按照权利要求 18 所述的方法，其中在网络接入点捕获从若干用户和若干网络服务传来的所有数据分组还包括在网关装置捕获从若干用户和若干网络服务传来的所有数据分组。

20、按照权利要求 18 所述的方法，其中过滤捕获的分组以提供使用情况监视数据还包括过滤捕获的分组，以便抽取使用情况监视数据，所述使用情况监视数据包括选自用户标识符、网络地址、分组时间戳、参阅网络地址、内容类型、内容长度、响应状态码和用户查询串的至少一种数据。

21、按照权利要求 18 所述的方法，其中把使用情况监视数据保存在使用情况监视数据库中还包括在把使用情况监视数据保存在主数据库中之前，把使用情况监视数据保存在临时数据库中。

22、一种在通信网络中提供使用情况监视的方法，所述方法包括：
在网络接入点捕获传送的数据分组；
确定传送的数据分组是否是用户产生的请求数据分组； 和
如果确定传送的数据分组是用户产生的请求数据分组，则把传送的数据分组保存在第一数据库中。

23、按照权利要求 22 所述的方法，其中在网络接入点捕获传送的数据分组还包括在网络网关装置捕获传送的数据分组。

24、按照权利要求 22 所述的方法，其中如果确定传送的数据分组是用户产生的请求数据分组，则把传送的数据分组保存在第一数据库中还包括如果传送的数据分组被确定为用户产生的请求数据分组，则把传送的数据分组保存在第一临时数据库中。

25、按照权利要求 22 所述的方法，还包括：
如果传送的数据分组和可访问的网络服务响应数据分组相匹配，则把传送的数据分组从第一数据库转移到第二数据库。

26、按照权利要求 25 所述的方法，其中第二数据库包括如果确定用户还没有超过预定的网络服务地址访问次数，则保存传送的数据分组和匹配的网络服务响应数据分组的临时数据库。

27、按照权利要求 25 所述的方法，其中第二数据库包括如果确定用户已超过预定的网络服务地址访问次数，则保存传送的数据分组和匹配的网络服务响应数据分组的主数据库。

28、按照权利要求 22 所述的方法，还包括：

确定传送的数据分组是否超出话路流限度，以便保证把第一传送数据分组保存在第一数据库中。

29、一种在通信网络中提供使用情况监视的方法，所述方法包括：

在网络接入点捕获传送的数据分组，传送的数据分组被确定为网络服务产生的响应分组；

证实第二传送数据分组的用户网络可访问性；

匹配传送的数据分组和用户产生的请求数据分组；和

把来自于匹配的传送数据分组和用户产生的请求数据分组的使用情况监视信息保存在数据库中。

30、按照权利要求 29 所述的方法，其中在网络接入点捕获传送的数据分组还包括在网络网关装置捕获传送的数据分组。

31、按照权利要求 29 所述的方法，其中证实传送的数据分组的网络可访问性还包括检查传送的数据分组的状态码，以便证实用户网络可访问性。

32、按照权利要求 29 所述的方法，其中把来自匹配的传送数据分组和用户产生的请求数据分组的使用情况监视信息保存在数据库中还包括如果确定用户还没有超过预定数目的网络服务地址访问，则把来自匹配的传送数据分组和用户产生的数据分组的使用情况监视信息保存在临时数据库中。

33、按照权利要求 29 所述的方法，其中把来自匹配的传送数据分组和用户产生的请求数据分组的使用情况监视信息保存在数据库中还包括如果确定用户已超过预定数目的网络服务地址访问，则把来自匹配的传送数据分组和用户产生的数据分组的使用情况监视信息保存在主数据库中。

34、一种通信网络中浏览顺序使用情况监视方法，所述方法包括：

指定需要浏览排序的网络地址；

在网络接入点捕获与网络地址相关的用户传送数据分组；

确定与捕获的用户传送数据分组相关的网络地址是否是指定的网络地址；和

如果确定与捕获的用户传送数据分组相关的网络地址是指定的网络地址，则把紧随指定网络地址被访问的预定数目的网络地址保存为浏览顺序。

35、按照权利要求34所述的方法，其中把紧随指定网络地址访问的预定数目的网络地址保存为浏览顺序还包括把在访问指定网络地址之后访问的预定数目的网络地址保存为浏览顺序。

36、按照权利要求34所述的方法，其中把紧随指定网络地址访问的预定数目的网络地址保存为浏览顺序还包括把在访问指定网络地址之前访问的预定数目的网络地址保存为浏览顺序。

37、按照权利要求34所述的方法，其中把紧随指定网络地址访问的预定数目的网络地址保存为浏览顺序还包括把在访问指定网络地址之前访问的预定数目的网络地址保存为浏览顺序。

38、按照权利要求34所述的方法，其中把紧随指定网络地址访问的预定数目的网络地址保存为浏览顺序还包括在把访问指定网络地址之后和之前访问的预定数目的网络地址保存为浏览顺序。

39、按照权利要求34所述的方法，还包括：

确定发送用户传送数据分组的用户是否已在预定数目的在先网络地址访问中访问了指定网络地址；和

如果确定用户已在预定数目的在先网络地址访问中访问了指定地址，则把来自用户传送的数据分组的网络地址添加到浏览排序列表中。

40、按照权利要求34所述的方法，还包括把指定的网络地址保存在数据库中。

41、按照权利要求34所述的方法，其中根据最频繁受访的网络地

址指定需要浏览排序的网络地址。

42、按照权利要求 34 所述的方法，其中根据被访问预定次数的网络地址，指定需要浏览排序的网络地址。

43、按照权利要求 34 所述的方法，其中根据使用情况监视客户选择的网络地址，指定需要浏览排序的网络地址。

网络使用情况监视设备和相关方法

技术领域

本发明涉及网络使用情况监视，更具体地说，涉及网络使用情况监视设备和在网络接入点捕捉使用情况数据的相关使用情况监视方法。

背景技术

诸如因特网、公司内部网之类的通信网络已成为工作场所及家庭中信息分发的选择形式。随着从这种网络获取信息的需求的逐步增加，越来越需要提供一种用于监视由这种通信网络分发的信息的装置。至此称为“使用情况监视”的这种性质的监视可向网络服务提供商、网络用户或者诸如网络广告商之类的网络受益者提供有价值的统计信息或者其它信息。

在多种不同的应用中，网络服务提供商可受益于使用情况监视；例如，它将提供根据用户已获取的信息和用户已提供给网络的信息产生网络用户简表的能力。在多种应用中，对服务提供商来说，用户简表是有益的，例如智能高速缓存数据（即用户访问的网页的存储副本）和/或预取数据（即预测未来的网页访问并保存这样的网页访问）。另外，在网络安全破坏方面，使用情况监视信息是有作用的。它提供可用于追踪破坏方历史的有价值的破坏后（post-breach）信息。

诸如因特网之类通信网络提供了另一种媒介，广告商可通过所述媒介分发信息和诱惑消费者。由于频繁（即使不是每天）依赖于因特网的人们很多，因此广告商已投入大量的资金以便通过因特网传送广告。但是对于广告商来说不幸的是，大批分发的广告通常费效比不高，因为许多（即使不是多数）受众并不对正在做广告的特殊产品或服务感兴趣。这样，广告商通常努力使广告以据认为对正在做广告的产品

或者服务最感兴趣的一部分人们作为目标。例如，新的私房屋主可能对安全系统最感兴趣，从而住宅安全公司可使他们的广告针对新的私房屋主。为了能够以他们潜在的消费群为目标，广告商必须能够实时地获取使用情况监视数据，并且能够获得格式化的数据或者易于格式化的数据，以便致力于人口统计，他们的特定产品针对所述人口统计结果而言。

另外，当内容分发网络（CDN）用于根据使用情况信息优化它们的分发网络，并且向它们的客户和订户提供较好的服务时，对于内容分发网络（CDN）来说，使用情况监视信息是重要的。

诸如家庭或公司之类的网络用户通过监视家庭或者公司内个人的使用情况，可受益于使用情况监视信息。这种监视确保个人没有获取不适宜的信息或者花费过多的时间浏览因特网或者企业内部网。

虽然诸如因特网之类通信网络提供了一种许多人可通过其通信的媒介，但是网络服务提供商、网络受益者和用户他们自己一直不能确定用户的确切本性，用户获取信息的内容，用户的人口统计细分（breakdown）和其它与使用情况相关的信息。例如，网络广告商一直难以实时确定关于用户的的具体人口统计结果和与用户所请求的实际内容相关的信息。网络服务提供商一直难以实时跟踪任意一个用户所访问的信息，以便提供更智能的高速缓存和预取能力、安全破坏研究能力等等。另外，网络用户（即家庭或公司）不具备就所访问内容或访问话路的持续时间而论监视个人（即家庭成员或雇员）的使用情况的适当手段。

在这方面，存在在通信网络内的不同离散位置监视网络使用情况的现有技术。例如，位于网络通信流内的网络服务器和/或路由器通过保存关于服务器提供的内容的访问请求的日志，可监视使用情况。这些日志提供和访问特定网页或者一组网页的用户的数目以及浏览这些网页的浏览顺序有关的信息。还提供了用于进行服务器日志的统计处理的软件。同样地，通过在存储器中保存用户所访问内容的日志，可在用户级，即在个人的计算机监视使用情况。该日志提供和单个用户

的使用模式相关的信息，包括访问的内容，内容的浏览顺序和每页上所花的时间量。此外，某些因特网服务提供商（ISP）可包括保存最频繁请求的信息的代理-高速缓存（proxy-caching）服务器。这些代理-高速缓存服务器可被设计成监视通过代理-高速缓存服务器被送到网络的那些用户的 web 内容使用情况。不幸的是，为了恰当地配置代理高速缓存，代理-高速缓存服务器需要网络管理人员等的用户干预。此外，不是几乎所有的客户请求都被引向代理-高速缓存服务器，从而使使用情况数据不完整，价值大大降低。

如同将指出的一样，这些现有技术均和环球网的离散装置或者部分的使用情况相关，并且监视所述使用情况。例如，在服务器级和在用户级进行的监视显然只分别确定在服务器产生的使用和用户所进行的使用。另外，通过代理-高速缓存服务器进行的监视只确定关于特定 ISP 网络的一些客户请求。虽然这些各种类型的使用情况监视对网络服务提供商、服务受益者和网络用户有一定帮助，但是这些实体需要在包括试图访问由多个不同服务器（可能位于多个不同服务提供商的网络上）提供的信息的许多不同用户的更大规模上监视使用情况。

发明内容

于是提供一种在网络接入点，即网络通信集合点，一般在网关装置或者类似的网络接口装置监视网络使用情况的网络使用情况监视模块。这样，本发明的网络使用情况监视模块可监视正在试图通过网关装置访问提供的各种网络服务的若干网络用户的使用情况。从而，和由常规的监视技术提供的使用情况信息相比，由本发明的使用情况监视模块收集的使用情况信息更全面。这样，对于网络服务提供商、网络用户、网络受益者等来说，该信息更有价值。另外，本发明的使用情况监视方法和设备提供许多特殊的特征，以便改进监视过程以及收集的使用情况信息的价值。

在一个实施例中，使用情况监视方法和设备捕获在用户和网络服务之间传送的所有字节，并且记录和网络地址（即统一资源定位符

(URL)) 相关的使用情况信息。通常，使用情况监视方法和设备将把信息的存储限制为用户获得网络服务，并且丢弃和被确定为不可访问的地址相关的任意使用情况信息的情况，从而减小最终数据库的大小，并且改进数据库的随后搜索和数据库的高速缓存。使用情况监视方法和设备还允许实时监视流式内容，并且简化基于 web 的用户调查的实现。本发明的使用情况监视方法和设备还便于监视用户查询串以及用户输入 web 表格中的信息。

重要的是，本发明的使用情况监视方法和设备捕获包括一个或多个指定网络地址的浏览顺序。在这方面，指定地址可以由使用情况监视模块预先指定或者临时确定为最受欢迎的地址。另一方面，使用情况监视方法和设备可监视位于指定网络地址的邻域中的地址，而不考虑地址受访的顺序。按照和本发明的使用情况监视方法和设备依据其监视指定地址邻域中的地址的技术相同的方式，使用情况监视方法和设备可监视指定域之前和/或之后的域。此外，使用情况监视方法和设备可监视位于指定地址邻域中的相关地址，即在指定地址之前或者之后不超过预定数目地址的相关 URL，而不考虑特定的浏览顺序。可以多种方式确定该相关地址，并且该相关地址一般由用户或者网络管理人员确定，以便识别其它相关地址。

通过捕获和通信网络的使用情况相关的一组更全面的信息，本发明的使用情况监视方法和设备能够向使用情况监视信息客户提供非常有用的信息。例如，使用情况监视方法和设备可包括详细的用户人口统计信息和与受访站点或者本地流式媒体的内容相关的信息。另外，通过在去往因特网的网关，即通信集合点监视使用情况，使用情况监视方法和设备能够监视请求访问许多不同服务提供商的大量用户的使用情况。

附图说明

图 1 是根据本发明一个实施例的使用情况监视设备的方框图。

图 2 是根据本发明另一实施例的使用情况电视设备的方框图。

图 3 是图解说明由本发明一个实施例的使用情况监视方法和设备执行的一般操作的流程图。

图 4 是图解说明由根据本发明另一实施例的使用情况监视方法和设备执行的浏览排序操作的流程图。

图 5 图解说明了可被本发明的使用情况监视方法和设备的一个实施例利用的关系数据库。

具体实施方式

下面将参考附图更充分地说明本发明，附图中表示了本发明的优选实施例。但是，本发明也可以多种形式体现，不应被认为局限于这里陈述的实施例；相反，提供这些实施例的目的是使本公开更充分更完整，并且向本领域的技术人员全面告知本发明的范围。附图中相同的标记代表相似的部件。

图 1 中描述了根据本发明实现使用情况监视的网络 10。使用监视模块 12 最好实现于网关装置 14 中或者被布置成与网关装置 14 通信。例如，在网关装置中实现使用监视模块的实施例中，网关可以是由 California, Westlake Village 的 Nomadix Inc. 提供，并在美国专利申请序列号 No.08/816174、09/458602、09/458569、09/541877 中描述的通用用户网关，上述专利申请的内容作为参考整体包含于此。另一方面，网关装置可以是本领域普通技术人员已知的能够实现编程模块的许多其它网关装置中的任意一种。应注意虽然优选和网关装置一起实现，不过使用监视模块也可和其它装置一起实现，或者可以是独立的装置，只要实现该模块的模块或装置被布置在网络中，以致它接收来自若干用户的所有请求/询问，并且接收来自于构成该网络的所有网络服务（即网络通信聚合点）的所有响应即可。在本发明的一个实施例中，实现使用情况监视模块的网关装置将在若干用户装置 16 和通信网络的剩余部分（即路由器 18、服务提供商集合体 20 和因特网 22）之间被安置在网络内。从而，在用作包含在通信网络内的所有网络服务的通信聚合点的网络接入点实现使用情况监视模块。

网络用户可通过若干不同的用户装置 6，例如手持式通信装置、便携式通信装置（膝上型计算机）、个人计算机等等与网关装置通信，所述用户装置 6 通过访问多路复用器 24 与网关装置通信。另外，根据用户和网络之间的通信装置，网络体系结构可包括用于链接各种通信装置，例如 DSL、电缆、LMDS（本地多点分发服务）、拨号、租用线路、无线等等的网络接口装置（NID）（图 1 中未示出）。虽然图 1 描述通过单个访问多路复用器与网络通信的通信网络用户，不过在典型的通信网络中，为了便于扩展用户能力，会实现多个访问多路复用器以及多个网络接口装置。

网关装置 14 一般允许用户与由 ISP（因特网服务提供商）或者由企业网络（即公司内部网等）提供的各种各样 IP 网络服务通信。如图 1 中所示，网关装置可与用于把网络通信送给服务提供商服务器集合体 20、因特网 22 或者其它任意网络服务（图 1 中未示出）的路由器 18 通信。虽然图 1 中只描述了单个路由器，但是在典型的通信网络中，多个路由器和/或交换装置将与网关装置通信，以便恰当地把网络通信送给它们的寻址目的地的。

根据本发明，使用情况监视模块捕获在用户和网络服务之间传送的所有分组。进行过滤以便抽取使用情况监视客户希望监视的所有相关分组（例如 HTTP 请求和响应，DNS 请求和响应等等），并把这些相关分组转发给相关的使用情况监视数据库 30。用户发出的关于网络服务的请求和业自网络服务的响应由使用情况监视数据库捕获。使用情况监视数据可在实现使用情况监视模块的装置之内或者可在该装置之外，只要使用情况监视数据库可通过通信与使用情况监视模块相连即可。在本发明的一个实施例中，使用情况监视数据库如后所述保存和 HTTP 请求及响应相关的信息，并处理所述信息。

图 2 根据本发明的一个实施例，示意表示了实现具有一系列数据库的使用情况监视的网络。由于网关装置 14 位于网络体系结构中通信聚合点的缘故，网关装置 14 从若干用户接收关于因特网或者其它网络服务的请求。在这方面，在网关装置中实现的使用情况监视模块能够

捕获所有的分组，因为网关装置按照混杂模式工作。在图 2 的实施例中，使用情况监视数据库由可以访问与使用情况监视过程相关的辅助数据库的使用情况监视数据库服务器 32 实现。

根据本发明的一个方面，使用情况监视模块 12 将捕获分组，并把它们送给它们相应的数据库。如果捕获的分组代表来自用户的请求或询问，则该请求保存在请求数据库 34 中。除了捕获用户请求分组之外，使用情况监视模块将捕获代表由网络服务传送的响应和其它消息的分组。在这方面，使用情况监视模块将把响应和它们相应的请求进行匹配。如果用户请求或询问成功，即网络服务批准访问或者响应询问，随后把响应和用户请求进行匹配，并把属于该组匹配的请求/响应的信息转发给临时数据库 36 或者主数据库 38。如果来自网络服务的响应指示没有出现有效响应，例如访问被拒绝，访问无效，产生服务错误等等，或者如果在接收响应之前发出超时，则请求数据库从数据库删除相关的请求。按照类似的方式，如果来自网络服务的响应不能和用户请求匹配，或者如果在预定的许可时间窗口之外产生响应，则丢弃来自网络服务的响应。本发明这方面，即提供丢弃不能和有效响应相符的请求，以及丢弃不能请求相符的响应的能力允许使用情况监视设备节省存储量，从而使支持使用情况监视设备所需的存储空间降至最小。通过非永久地保存不能产生成功响应的请求/询问，主数据库 38 可在搜索/高速缓存操作中智能提供优良性能。

一旦使用情况监视模块已使用户请求和网络服务响应匹配，则在请求和响应中找到的信息被保存在临时数据库 36 或者主数据库 38 中。使用情况监视模块将确定用户请求是否是相对于特定网络服务，例如网页的预定数目的初始请求之一。如果确定用户请求是初始请求或者在预定数目的初始请求之内，则在该请求和响应中找到的信息被保存在临时数据库。一旦该网络服务被访问次数超过预定的次数，则该信息被传送给并保存在主数据库 38 中，后来的请求/响应被委托给并保存在主数据库中。临时数据库的实现考虑了许多网络服务只被请求一次或者只在极少量的时间被请求，从而当处理使用情况数据信息时，

不需要这些极少使用请求/响应的统计存在的事实。在这方面，主数据库被最小化，以便更好地适应最后的使用情况监视信息客户（即网络服务提供商、网络受益者或网络用户）对数据库的后续处理。使用情况监视模块可被配置成选择预定数目的初始请求，以平衡记录用户的所有活动和简化数据库并使之更有效率的竞争目标。但是在某些情况下，预定数目被设置为 1。

另外，本发明的网络使用情况监视设备可包括指定网络地址的数据库 40。指定网络地址是由网络使用情况监视管理人员，一般是网关装置管理人员等预定的频繁受访网络地址。指定网络地址大多数时候是使用情况监视信息客户认为特别重要的地址。使用情况监视模块允许把预定为“指定”的那些网络地址的浏览顺序保存在指定网络地址的数据库中。浏览顺序一般包括预定数目的紧接在前和/或紧接在后的用户所访问网络地址。例如，诸如商业广告商之类的网络受益者可能希望为使用情况监视目的指定特定的因特网站点。这样，当网络用户访问指定的因特网站点，使用情况监视模块将把该网络地址识别为指定的网络地址，并且将记录和保存访问指定网络地址之前和之后的在前和在后浏览顺序。应注意的是浏览顺序可超出单个用户话路，只要话路之间的间隔在预定的话路窗口之内即可。

在本发明的一个实施例中，使用情况监视设备与诸如 PC 之类的处理引擎 42（即数据库前端）通信。处理引擎与使用情况监视数据库通信，并且和在处理引擎上实现的 GUI（图形用户界面）一起访问数据库。处理引擎根据使用情况监视客户的命令访问数据库，以便按照特定的用户要求处理使用情况监视数据。在本发明的一个实施例中，处理引擎与插入服务器 44（图 2 中所示）通信。例如，插入服务器可包括目标广告插入服务器和目标调查插入服务器。处理引擎执行根据收集的使用情况监视数据，确定哪些广告或调查有助于网络提供商或者网络受益者的必要命令。一旦确定目标广告或调查，则向插入服务器发送命令，并且通过与网关装置 14 的通信，把恰当的广告或者调查转发给网络用户。广告或调查一般呈发送给用户的通信装置的弹出式

控制画面的形式。

图 3 根据本发明的一个实施例图解说明了使用情况监视方法的流程图。在步骤 100，通常实现于网关装置上的使用情况监视模块读取正从通过网关装置访问网络的若干用户传来的信息的所有字节和来自于网络服务的字节。网关装置一般按照混杂模式工作，以便确保信息的所有字节（不考虑目的地地址）都被使用情况监视模块处理。在步骤 110，使用情况监视模块检测 TCP（传输控制协议）流中 HTTP（超文本传送协议）报头的开始。

可选地，使用情况监视模块可被设计成只监视在一个话路中访问多于预定数目的不同网络地址，例如因特网 URL（统一资源定位符）的网络用户的使用情况。这可在不会显著损害数据完整性的情况下减小数据库，因为只访问极少量地址的用户不会向使用情况客户提供重要的统计数据。于是，在可选的步骤 120，确定关于该用户的当前话路流是否已超过预定的临界最小值。如果在一个话路中，用户还没有访问多于预定数目的 URL，则使用情况监视模块在步骤 110 继续监视用户传送的请求，但是在请求的数目超过预定阈值之前，并不进一步分析和保存所述请求。

一旦在一个话路内用户发出多于预定数目的请求或者如果使用情况监视模块不实现该可选的筛选程序，则在步骤 130，使用情况监视模块一般通过分析分组的 HTTP 报头，确定该分组是否代表请求或者响应。如果捕获的分组代表请求，则在步骤 140，该请求被保存在请求数据库中。如果捕获的分组代表响应，则在步骤 150，检查该响应以确定该响应中的状态代码是有效的状态代码还是无效的状态代码。

通常，虽然监视网络服务访问或者其它请求，但是不考虑出于某些原因（所述原因可以是停用链路或者服务器可能已被临时停机），用户可能一直不能访问该服务的事实。但是在本发明的一个方面，还监视来自被请求网络服务的响应状态码，以便只记录在从数据库中删除其它请求的情况下，用户实际访问的网络服务的地址。这可显著节省存储量和减小处理计算量。

状态码是被请求服务器返回的三位整数。状态码的第一位代表响应的类别。返回的某些最常见的状态码是：

- 1) 200 OK
- 2) 301 永久移动
- 3) 304 未修改
- 4) 401 未经许可
- 5) 403 禁用
- 6) 404 未找到
- 7) 500 因特网服务器错误
- 8) 501 未实现

状态码 200 指示该网页可被访问，而其它状态码指示网页不可访问。如果确定响应状态码无效，即除 200 之外的状态码，则在步骤 160，丢弃该响应，并且从请求数据库删除匹配的请求。另外，如果在接收相关的网络响应之前，超过预定的超时时段，则从请求数据库删除用户请求。如果确定响应状态码有效，则在步骤 170，使网络响应和请求数据库中的相关用户请求进行匹配。如果对于返回的网络响应没有找到任何匹配，例如在请求和响应之间过去过多的时间的情况下，则在步骤 180，丢弃网络响应。

可选地，在步骤 190，使用情况监视方法和设备确定匹配的请求/响应是在主数据库中还是在临时数据库中。虽然所有的请求/响应和与之相关的信息可保存在主数据库中，但是如果被请求的网络服务首次被请求，或者如果被配置成这样，被请求的网络服务的被请求次数少于预定次数，则本发明一方面的使用情况监视方法和设备一开始（并且可选地）把数据保存在临时数据库中。之后，一旦网络服务被请求的次数超过预定次数，则请求/响应信息被传送给主数据库和/或记录在主数据库中。如前所述，由于许多网络地址，例如因特网 URL 只被请求一次，因此在继续保存任何用户访问一次以上的所有网络服务请求/响应的时候，可使主数据库的大小降至最小。通过为很少被请求的，请求次数少于预定最小次数的网络服务建立单独的但是临时的数据

库，可实现这一点。

从而，如果在永久数据库和临时数据库中都没有找到任何入口（或者如果对于临时数据库中的入口来说还没有超过最小的阈值），则在步骤 200，请求/响应信息被保存在临时数据库中。如果在永久数据库中找到一个入口，则在步骤 210，更新永久数据库以指示新的请求/响应入口。另外，如果在临时数据库中找到一个入口，并且当前的访问使之有资格包含在永久数据库中（即正在超过预定的最小阈值），则在步骤 210，来自临时数据库的入口被转移到永久数据库。

在本发明的一个实施例中，使用情况监视数据库保存请求和相关响应的 UR。一般来说，URL 将包括协议前缀，端口号，域名，子目录名称和文件名称。在本发明的一个备选实施例中，使用情况监视数据库最好在主数据库中保存和请求及响应相关的各种其它数据。在这方面，通常从请求/响应收集下述数据，并将其保存在适当的数据库中：

- a) 用户标识符（通过将其转换成随机数，在提供给使用情况监视客户之前，该标识符可被匿名化，并且将相对于提供和用户的年龄、性别、收入等有关信息的用户简表编制索引）
- b) URL（通常以群的方式逻辑排列 URL）
- c) 时间戳-指示分组被使用情况监视模块处理的时刻
- d) 参阅地址（从 HTTP 报头获得）-通常，该地址提供关于受访地址的超链接。
- e) 内容类型（一般从 HTTP 报头获得）
- f) 内容长度
- g) 响应状态码
- h) 用户询问串（通常以变元的形式发送给 CGI（公共网关接口）应用程序或者作为网络用户登记表格的结果）。

上面的列表不应被认为是穷举的，如同信息最终应用于的应用程序指示的那样，对于请求/响应也可抽取其它信息。

虽然可以不同的方式保存包含在用户请求和网络响应中的信息，不过本发明的一个实施例的主数据库是如图 5 中描述的关系数据库，

所述关系数据库使用户和相应的地址/URL 联系起来。关系数据库是一种合适的保存彼此相关数据集的既定方式。这种情况下，每个网络用户和每个受访地址只被保存一次，而不是每次网络用户访问地址时，都要保存这两个项目。随后页面的各个访问可被表现为把访问该地址的用户的入口和正被访问的地址的入口连接在一起的节点（在关系数据库的环境中称为元组）。该节点将包括和访问相关的附加信息，例如访问的时间、访问的频率，内容长度和类型等等。例如，在用户 11 访问地址/URL 25 的情况下，关系数据库把用户 11 保存在表 1 中，把地址/URL 保存在表 2 中，同时一个节点链接来自表 1 的入口和来自表 2 的入口。这种形式的数据存储便于数据检索，从而为了编辑信息，使用情况监视客户可访问某一用户与之链接的所有节点。

根据图 4 的流程图，图解说明了本发明的必然伴有呈浏览顺序捕获形式的使用情况监视的备选实施例。网络服务提供商和诸如网络广告商之类的网络受益者不仅对了解哪些网络服务经常被访问和花费多少时间访问这些网络服务感兴趣，而且对包括这些频繁受访地址的浏览顺序感兴趣。应服务提供商或网络受益者的要求，网络管理人员和/或网关管理人员可为浏览排序监视预定或者“指定”具体的网络地址。管理人员将把指定网络地址保存在指定网络地址数据库中。涉及指定网络地址的浏览顺序将包括访问指定网络地址之前和/或之后，用户所访问的预定数目的网络地址。应注意虽然典型的浏览顺序产生于单个用户话路之内，但是浏览顺序也可超出单个用户话路，只要两个话路之间的间隔在规定的阈值之内即可。另外，指定的网络地址可是以一个以上的浏览顺序的一部分；于是，本发明的使用情况监视模块可捕获并保存下述信息：

■ 指定的网络地址为其一部分的所有浏览顺序。

■ 只是网络地址为其一部分的最普及的浏览顺序（它超过了一定的访问阈值）。

■ 指定的网络地址为其一部分的特定浏览顺序的频率。

通常包含在主数据库中的浏览顺序数据库的配置将由网络管理人

员或者网关管理人员任意选择。

如同本领域的普通技术人员已知的那样，Zipf 定律陈述某一单词在足够长的一段文本中的出现次数是频率发生顺序的倒数。例如，出现频率为第 10 的单词的出现次数比最频繁出现的单词的出现次数小 10 倍。虽然不易于产生纯粹的 ZIPF 定律关系，不过在诸如网络服务访问频率之类的各种各样情况中存在幂定律（出现频率的顺序和出现频率之间的关系不是线性的）。其结果之一是少数网络地址构成绝大多数的用户访问（例如 5% 的适用网络地址构成 95 的网络访问）。这些假定意味只需对较小百分比的网络地址（例如 5%）捕获统计数据，即浏览顺序等等，并将其保存在使用情况监视数据库中即可。在这方面，网络服务提供商和/或网络受益者被赋予预定将受到浏览顺序监视的指定网络地址，或者允许使用情况监视模块监视由数据库确定的最频繁受访网络地址的浏览顺序的选择。

重新参见图 4 的流程图，本发明的使用情况监视模块实时监视指定网络地址的浏览顺序。应注意图 4 流程图的步骤 100-170 和图 3 流程图的步骤相同，于是，不需要对这些步骤进行更多的说明。根据本发明，通过持续跟踪在各个用户的当前话路中，各个用户所访问的前 k 个网络地址来完成浏览顺序监视。数字 k 规定在指定的网络地址之前和之后的网络地址的数目。k 值可由网络或网关管理人员指定，并且对于不同的指定网络地址可不同。虽然本说明书指出在指定的网络地址之前和之后监视相同数目的网络地址，但是如果需要，可为在前的网络地址和在后的网络地址分配不同的数值。

在步骤 220，使用情况监视模块确定网络用户是否已在前面的 k 次网络服务访问中访问了指定的网络地址。如果确定网络用户在前面的 k 次访问中访问了指定的网络地址，则在步骤 230，当前被访问的网络地址被添加到正在构造的浏览顺序中。如果确定网络用户在前面的 k 次访问中没有访问指定的网络地址，则在步骤 240，使用情况监视模块通过尝试在指定网络地址数据库中定位某一入口，确定当前受访的网络地址是否是指定的网络地址。如果使用情况监视模块确定当

前受访的网络地址是指定的网络地址，则在步骤 250，发送把用户访问的前 k 个网络地址保存为浏览序列的命令。如果使用情况监视模块确定当前受访的网络地址不是指定的网络地址，则在步骤 260，把当前受访的网络地址添加到该特定用户的 k_{max} 网络地址的列表中，并且如果该 k_{max} 列表已具有 k 个网络地址，则丢弃该 k_{max} 列表中最陈旧的网络地址。保存 k_{max} 列表，因为预先并不知道网络用户将访问哪些网络地址，从而如果网络用户访问指定的网络地址并且需要浏览顺序，则 k_{max} 列表提供一系列的在先网络地址。

根据本发明的又一实施例，除了监视用户访问请求和网络服务响应之外，本发明的使用情况监视方法和设备可以监视流式内容的使用情况。流式内容可以是实况转播的流式内容也可以是录制的流式内容，并且当通过网关装置从本地流式媒体服务器向网络用户传送流式内容时进行捕获。流式内容一般是高容量内容，因此不可能或者不适用于保存和所有流式内容数据分组相关的信息。于是，在使用情况监视模块捕获所有流式内容分组的时候，考虑到高通信容量的缘故，相当大的一部分分组将被滤出。持续监视流式内容分组便于进行效果显著的滤除，并且将提供和流式连接过程中传送的数据的数量，流式连接的持续时间等有关的信息。为了实现这种持续监视，实现使用情况监视模块的装置，一般为网关装置，需要在存储器中分配一定数量的状态，以便持续跟踪流式内容的当前或者最新已知状态。另外，使用情况监视模块或者后处理应用程序将跟踪当前正在访问流式内容的用户，流式内容来自的站点，在连接中传送的数据的数量以及关于特定用户/流式内容对观察到流式内容分组的最后时间。由于在流式内容的情况下不存在明确的连接终止，因此使用情况监视模块需要实现超时，以便确定流式内容何时被终止。这些跟踪功能可在使用情况监视模块完成，或者可以每个流式分组为基础，向后处理应用程序发送未格式化的数据，以便确定跟踪参数。

在本发明的又一实施例中，使用情况监视设备可被用于向网络用户提供针对性调查。如前所述，可实现插入服务器，以便根据使用情

况监视数据向网络用户发送针对性调查。调查可被主动发送给网络用户或者网络用户可选择参与所述调查。在这方面，可随机选择网络用户，并且在随机选择的时间，在这些网络用户的通信装置上出现弹出式屏幕。例如，参见美国专利申请序列号 No.09/541877 “Information And Control Console For Use With A Network Gateway Interface”（申请日 2000 年 4 月 3 日，发明人 Short 等），该专利申请和本发明一样被转让给相同的受让人。美国专利申请序列号 No.09/541877 的内容作为参考整体包含于此。

在本发明的一个备选实施例中，使用情况监视方法和设备还监视网络用户向网络服务应用程序，例如搜索引擎等等提交的查询。通过监视网络用户向诸如搜索引擎之类的网络服务应用程序提交的查询，可获得和用户的喜好有关的辅助了解。现有的使用情况监视方法不提供实时监视网络查询的能力。网络用户查询串的实时监视向诸如网络广告商之类的网络受益者提供简要描述网络用户的能力，并且向用户提供更有针对性的信息。例如，如果用户在 FordTM网站上搜索型号 TaurusTM，则网络广告商可利用该信息把竞争产品的广告引向该用户。

使用情况监视方法和设备可监视的另一种用户输入是用户投寄的 web 表格，例如网络服务登记表、申请表等等。目前的方法不能捕获环球网上用户提交的表格中的信息。

如前所述，本发明的网络使用情况监视设备和方法提供监视并保存各种各样和网络用户相关的信息的能力。这些信息包括（但不限于）被访问的网络服务的内容，用户群体的人口统计信息（即用户群体年龄、性别、地理位置、服务计划等等）和网络服务受访的持续时间。在这方面，网络服务提供商或者受益者将能够确定在任意特定时间访问任意特定网络服务的用户的实时统计性质。例如，“有多少年龄在 34-55 的妇女正在访问特定的网站？”或者“收听特定古典音乐台的用户在 35 岁以下的百分比为多少？”这些信息随后可被提供给网络受益者，所述网络受益者可利用该信息有效地有益于他们的核心听众。

主存储器要求

下面的讨论用于假想地估计为操纵根据本发明的使用情况监视模块所需的主存储器的数量。

假定地址的平均长度为 l 字节，有效用户的数目为 n ，则存储量可被计算为

$$S_n \text{ (以字节表示的存储量)} = l * 2k_{max} * n$$

最坏的情况是存储量要求取决于 N ，即在特定访问点访问网络的用户的总数。

$$S_n \text{ (以字节表示的存储量)} = l * 2k_{max} * N$$

ISP（因特网服务提供商）研究指出在任意时候，只有 10-25% 的用户处于活动状态。从而，典型的存储量要求 S_n 仅约为最坏情况下存储量要求 S_N 的 25%。

典型分组的长度一般小于 100 字节，虽然一些 URL 确实扩大到该长度的两倍以上。较长的 URL 一般起因于不常见的查询，从而不会频繁发生。于是，对目前实际应用来说，可假定 l 的平均值为 100 字节。

用户话路的典型长度从拨号用户的 55 分钟到 DSL 用户的 95 分钟不等。并且由于用户花在网页上的平均时间为 2 分钟，估计在一个网络话路中任意用户访问的网络地址的数目一般约为 100 左右是合理的。这向我们提供了 k 值的上限。

从而，如果假定主数据库中 128MB 的 RAM 专用于保存根据本方法识别的浏览顺序，并且 k 值为 100， l 为 100 字节，则使用情况监视模块能够监视 12800 个用户的近邻。可推断出在所需的状态的数量，即保存的在先和在后网络地址的数目以及主存储器要求方面，该方法将按比例增加。

辅助存储器要求

当用户通过不同的浏览顺序访问更多的指定网络地址时，保存在主数据库中的浏览顺序的数目将随着时间而增加。当浏览顺序本身变得更长时，该数目还随着 k 值的增大而增加。在不损失和浏览顺序相关的有价值信息的同时，减少所需存储量的途径有两种。

- 陈化浏览顺序：和每个浏览顺序一起保存一个时间戳。该时间

截指示该顺序被访问的最后时间。如果在规定的一段时间内一直没有访问该顺序，则该浏览顺序超时，并被删除。

· 消除罕见的顺序：Zipf 定律再次指出只有少部分的顺序负责绝大多数的访问。从而能够消除访问次数少于规定次数的浏览顺序。

根据本发明的一个备选实施例，使用情况监视设备和方法可用于捕获并保存指定网络地址的邻域。该实施例中，用户访问网络地址的顺序并不重要。保存在数据库中的唯一信息是在指定网络地址的规定数目的页面（或者时间）内发生的网络地址。当用户访问更多的网络地址时，定义为 k 的邻域中的网络地址的数目通常增大。该数目还随着 k 值的增大而增大。类似于上面定义的辅助存储器要求，在不损失和邻域相关的有用信息的同时，减少所需存储量的途径有两种。

· 陈化邻域：和为邻域一部分的每个网络地址入口一起保存一个时间截。该时间截指示该网络地址被访问的最后时间。如果在规定的一段时间内一直没有访问该网络服务，则该网络服务超时，并从所述邻域中删除。

· 消除少见的相邻网络地址：Zipf 定律再次指出邻域中只有少部分的网络地址负责绝大多数的访问。从而能够从邻域内删除访问次数低于规定次数的网络地址。

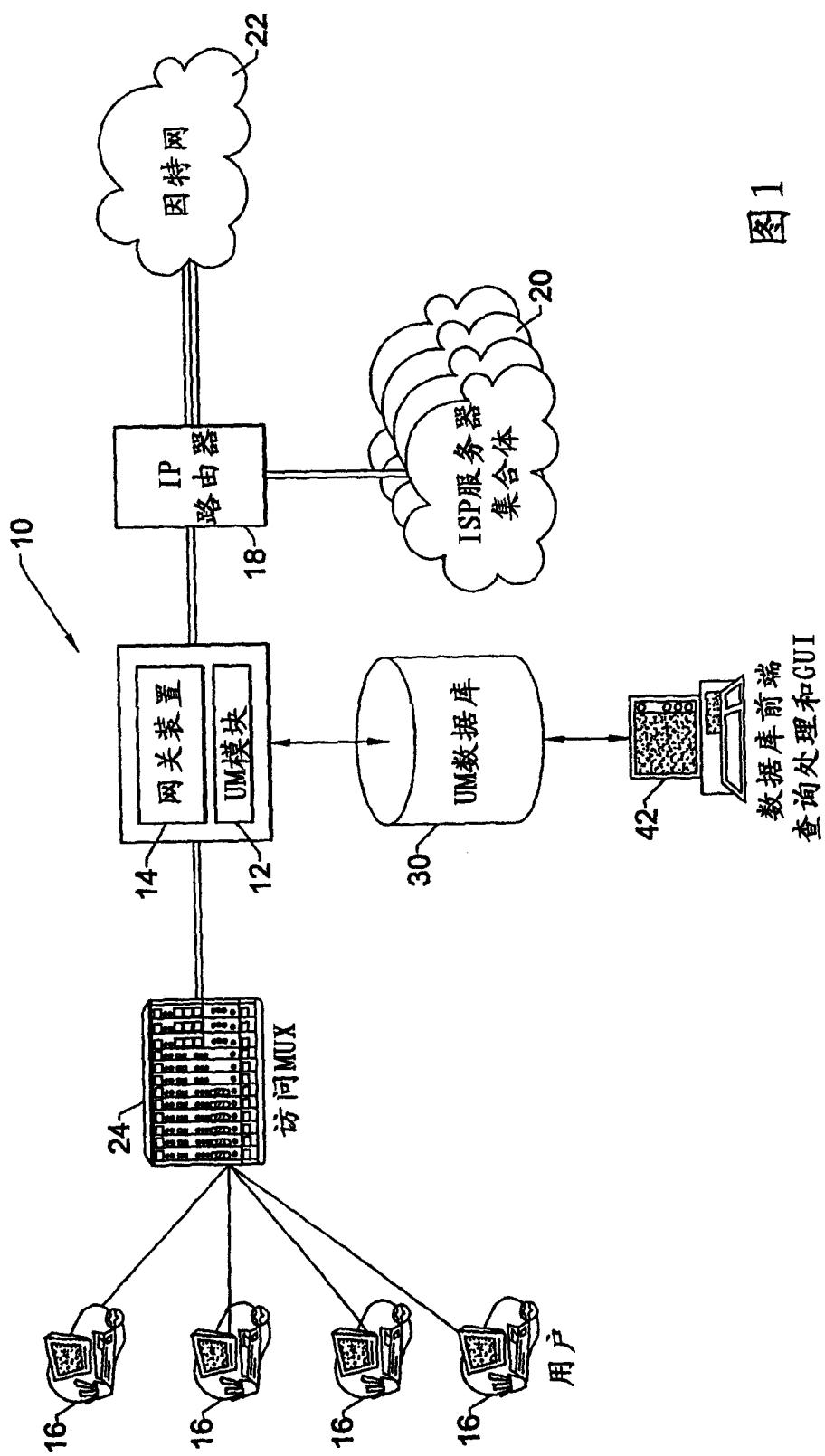
和前面的方法描述捕获指定网络地址的网络地址邻域的方式相同；还能够捕获指定域的域邻域。广告商可指定希望使用情况监视模块监视的域。使用情况监视模块随后始终监视用户在任意时刻访问的前 k 个域。当用户访问指定域时，使用情况监视模块保存用户访问的前 k 个域以及用户将访问的后 k 个域。通过陈化域以及消除很少受访的域，同样能够减少所需的存储量。

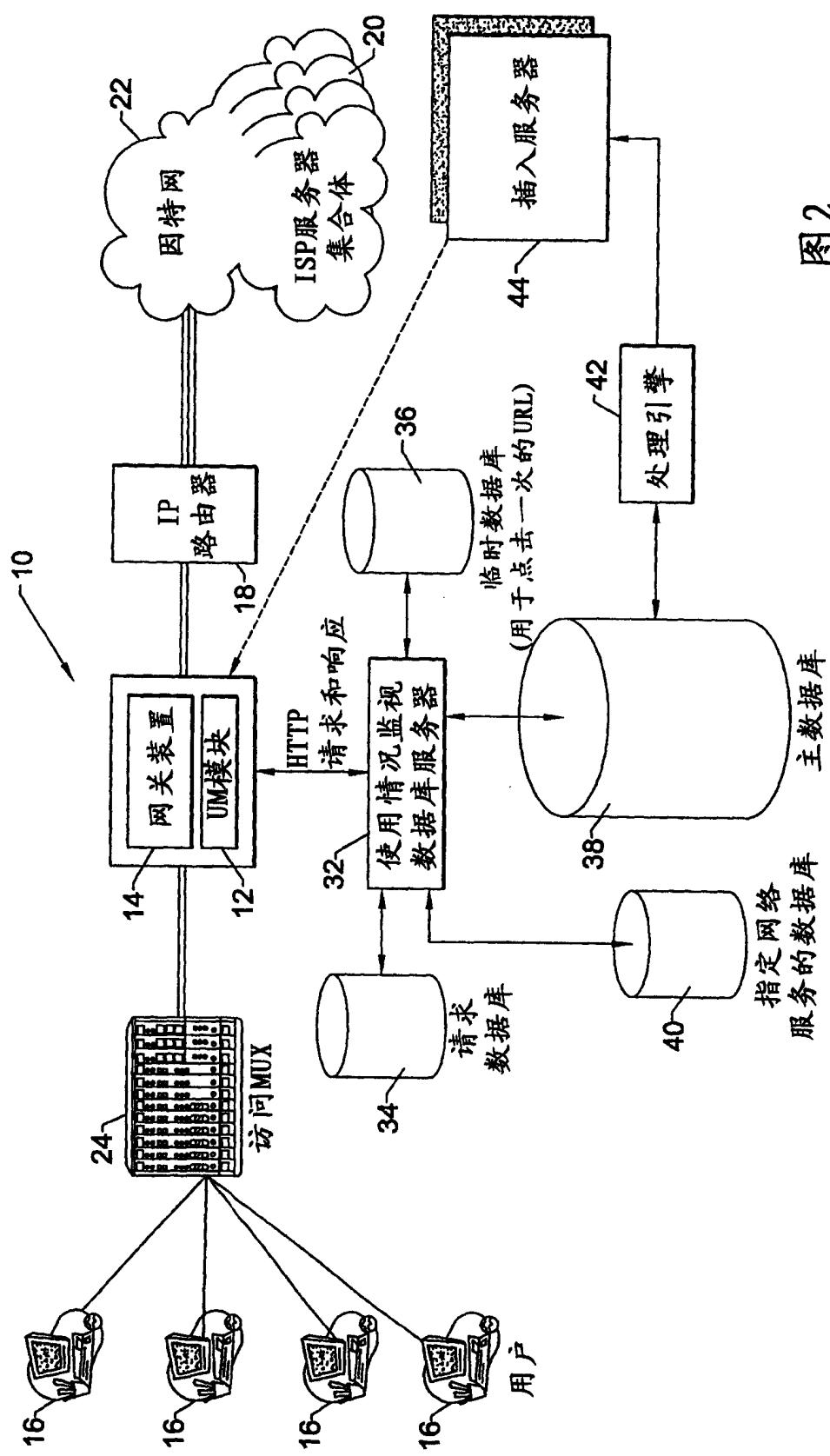
捕获指定网络地址和确定的一系列相关网络地址为其一部分的整个浏览顺序可被认为既需要很大的存储量又需要很大的计算量。在这方面，提供一种在较大的超时时段 T （这是为了确保存储量需求和状态信息合理并且能够实现）内，跟踪指定网络地址的邻域中的一组相关网络地址的备选方法，所述一组相关网络地址一般由网关或者网络

管理人员确定。就网络广告商来说，所述一组相关网络地址可能属于指定网络地址的竞争者。网络广告商可能希望监视未来用户在指定网络地址（例如福特网站）的邻域中浏览时，是否正在访问确定的一系列相关地址（例如通用汽车、沃尔沃、BMW 或梅赛德斯网站）。使用情况监视模块将提供搜索包括指定网络地址和相关网络地址的浏览顺序的能力，并向广告商提供访问相关网站的网络地址的用户的统计结果。

由于根据本发明的使用情况监视模块位于通信网络边缘的缘故，它能够监视由若干用户发出的请求，和由若干不同网络服务提供的响应，以便收集对网络服务提供商、网络用户和网络受益者有用的一组规范量度。该数据相当全面，可包括环球网和本地流式内容以及用户人口统计信息和浏览顺序。根据从正在通过因特网访问各种各样网络服务和其它实体的若干用户收集的该组全面数据，本发明的使用情况监视方法和设备能够向网络服务提供商、网络用户、网络受益者等等提供大量的有用数据。

根据上面的说明和相关附图中给出的教导，本发明所属领域中的技术人员将想起本发明的多种修改和其它实施例。于是，本发明并不局限于公开的具体实施例，并且这些修改和其它实施例包含在附加的权利要求的范围之内。虽然这里采用了特定的术语，不过只是在类属和说明的意义上使用这些特定术语，而不是对本发明的限制。





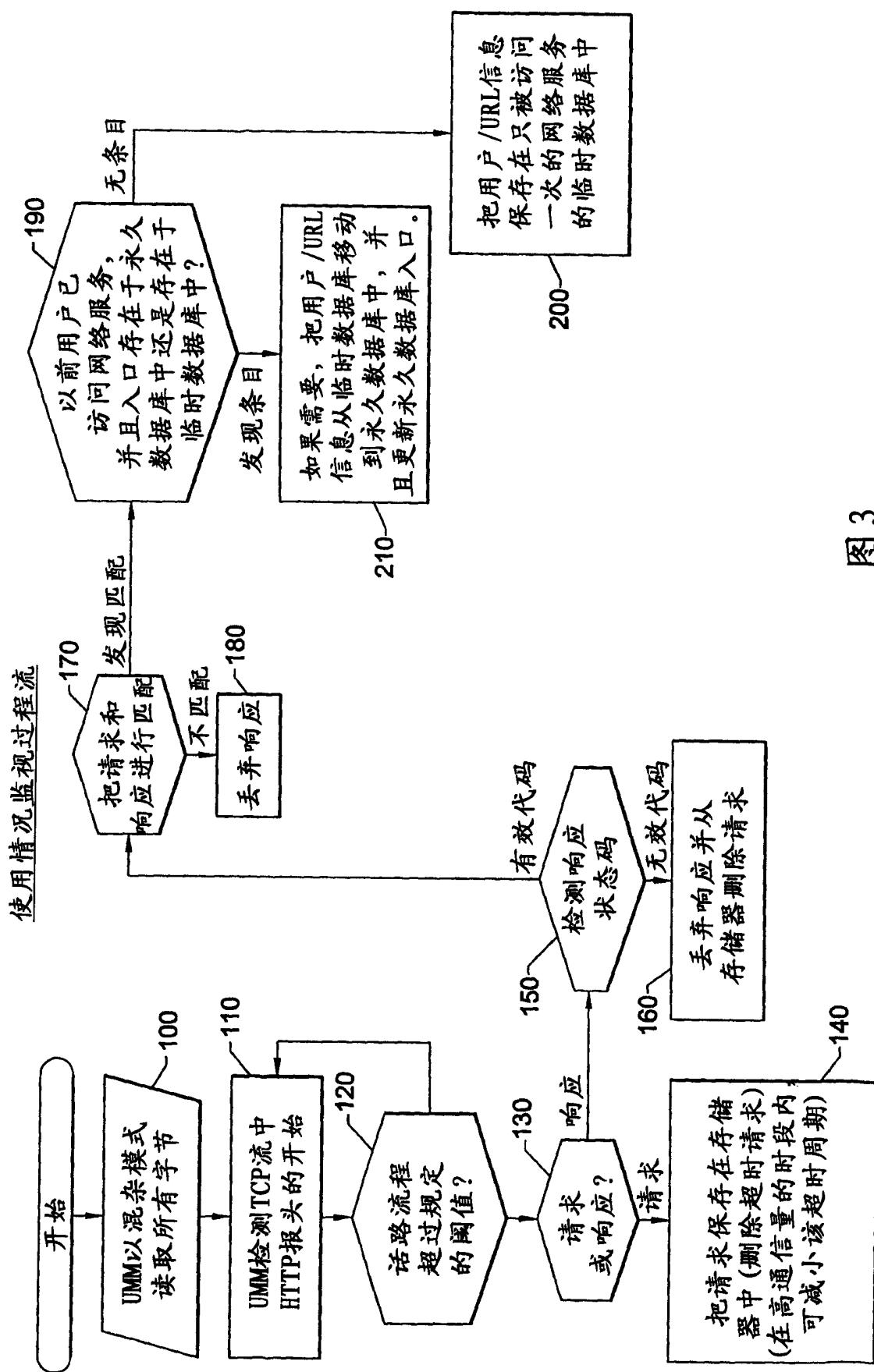


图 3

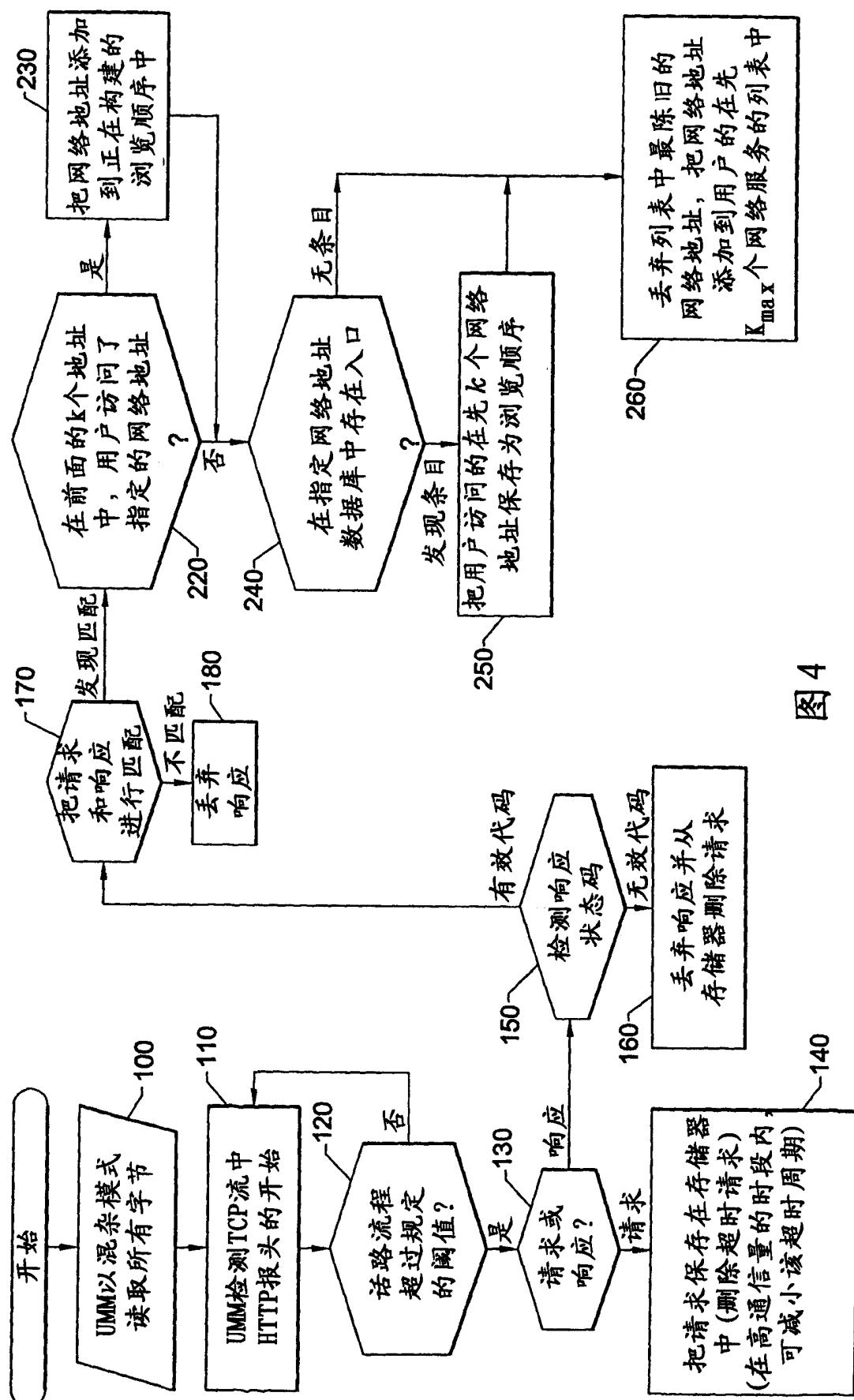


表 1

	用户 1
	用户 2
	•
	•
	用户 11
	•
	•
	用户 N

关于用户和URL索引的关系数据库

表 2

URL 1																							
URL 2																							
URL 3		•																					
•																							
•																							
•																							
•																							
•																							
•																							
•																							
•																							
•																							
•																							
•																							
•																							
•																							
•																							
•																							
URL 25		•																					
•																							
•																							
•																							
•																							
•																							
•																							
URL M																							

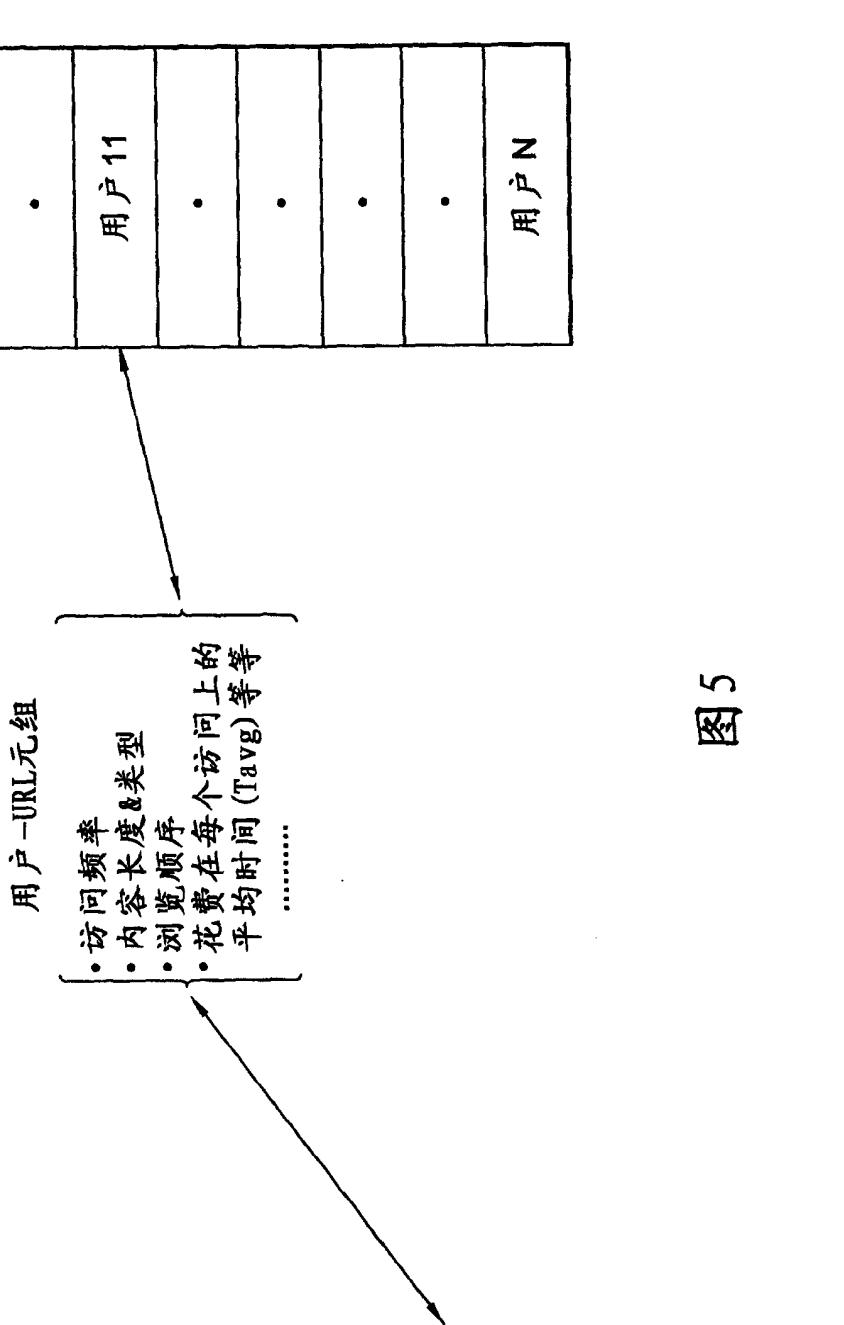


图 5