

(12) 发明专利

(10) 授权公告号 CN 101222348 B

(45) 授权公告日 2011.05.11

(21) 申请号 200710000085.3

(22) 申请日 2007.01.10

(73) 专利权人 阿里巴巴集团控股有限公司
地址 开曼群岛大开曼乔治敦

(72) 发明人 董启江 王超 孙永明 谢瑶
庄奇 李晔 朱振华

(74) 专利代理机构 北京挺立专利事务所 11265
代理人 崔吉甫

(51) Int. Cl.

H04L 12/24(2006.01)

H04L 12/28(2006.01)

H04L 29/06(2006.01)

G06F 17/30(2006.01)

H04L 9/32(2006.01)

G06Q 10/00(2006.01)

(56) 对比文件

US 7137009 B1, 2006.11.14, 摘要, 说明书第2栏15行—第10栏20行, 图1-6.

US 7137009 B1, 2006.11.14, 摘要, 说明书第2栏15行—第10栏20行, 图1-6.

US 2003/0005046 A1, 2003.01.02, 说明书第1页[0007]—第7页[0073], 图1A、5、7.

CN 1878096 A, 2006.12.13, 全文.

审查员 刘锐

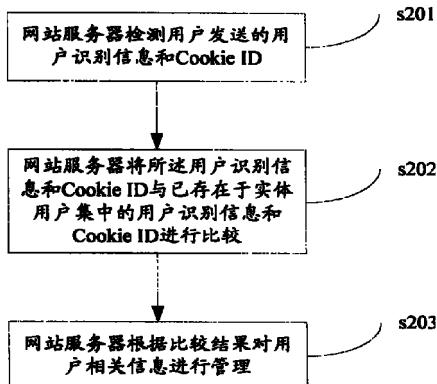
权利要求书 2 页 说明书 6 页 附图 2 页

(54) 发明名称

统计网站真实用户的方法及系统

(57) 摘要

本发明公开了一种统计网站真实用户的方法,包括以下步骤:网站服务器检测用户发送的用户识别信息和Cookie ID;网站服务器将用户识别信息和Cookie ID与已存在于实体用户集中的用户识别信息和Cookie ID进行比较;网站服务器根据比较结果对用户相关信息进行管理。本发明还提供了一种统计网站真实用户的系统。本发明解决了对于申请多个账号的真实用户进行管理的问题,网站服务器不再片面分析单个虚拟用户的信息、行为和属性,而是将网络上的虚拟用户和现实的真实用户对应,收集到一个真实用户在网站服务器上注册的所有用户信息;以及以未注册用户身份在网站服务器上的行为信息,从而加以利用,全面分析。



1. 一种统计网站真实用户的方法,其特征在于,包括以下步骤:

网站服务器检测用户发送的用户识别信息和Cookie ID;

所述网站服务器将所述用户识别信息和Cookie ID与已存在于实体用户集中的用户识别信息和Cookie ID进行比较;

所述网站服务器根据比较结果对所述用户相关信息进行管理;

所述根据比较结果对所述用户相关信息进行管理具体包括:

如果所述用户识别信息和所述Cookie ID都属于所述实体用户集,则判断所述用户识别信息和所述Cookie ID是否属于同一个实体用户,如果不属于同一个实体用户,则合并为一个实体用户。

2. 如权利要求1所述统计网站真实用户的方法,其特征在于,所述根据比较结果对所述用户相关信息进行管理具体还包括:

如果所述用户识别信息属于实体用户集中的实体用户,而所述实体用户中不存在所述Cookie ID,则所述网站服务器将所述Cookie ID加入所述实体用户;

如果所述Cookie ID属于所述实体用户集中的实体用户,而所述实体用户中不存在所述用户识别信息,则所述网站服务器将所述用户识别信息加入所述实体用户;

如果所述用户识别信息和所述Cookie ID都不属于所述实体用户集中的实体用户,则所述网站服务器将所述用户识别信息和所述Cookie ID加入到所述实体用户集中的新实体用户。

3. 如权利要求1所述统计网站真实用户的方法,其特征在于,在网站服务器检测之前还包括:

用户通过用户设备向所述网站服务器发送登录消息;

所述网站服务器判断所述登录消息中是否包括Cookie ID,如果没有,则为所述用户设备创建一个Cookie ID,记录到实体用户集的实体用户中,并通知所述用户设备。

4. 如权利要求3所述统计网站真实用户的方法,其特征在于,所述网站服务器为所述用户设备分配Cookie ID后还包括:所述用户设备删除所述Cookie ID。

5. 如权利要求1至4中任一项所述统计网站真实用户的方法,其特征在于,所述用户识别信息包括用户ID。

6. 一种统计网站真实用户的系统,包括网站服务器和用户设备,其特征在于,所述网站服务器进一步包括检测单元、比较单元和管理单元;

所述检测单元,用于检测用户通过用户设备发送的用户识别信息和Cookie ID;

所述比较单元,与所述检测单元连接,用于将所述用户识别信息和Cookie ID与已存在于网站服务器的实体用户集中的用户识别信息和Cookie ID进行比较;

所述管理单元,与所述比较单元连接,用于根据所述比较结果对所述用户相关信息进行管理;

所述管理单元根据所述比较结果对所述用户相关信息进行管理具体包括:

如果所述用户识别信息和所述Cookie ID都属于所述实体用户集,则判断所述用户识别信息和所述Cookie ID是否属于同一个实体用户,如果不属于同一个实体用户,则合并为一个实体用户。

7. 如权利要求6所述统计网站真实用户的系统,其特征在于,所述管理单元根据所述

比较结果对所述用户相关信息进行管理具体还包括：

如果所述用户识别信息属于实体用户集中的实体用户，而所述实体用户中不存在所述Cookie ID，则所述网站服务器将所述Cookie ID 加入所述实体用户；

如果所述Cookie ID 属于所述实体用户集中的实体用户，而所述实体用户中不存在所述用户识别信息，则所述网站服务器将所述用户识别信息加入所述实体用户；

如果所述用户识别信息和所述Cookie ID 都不属于所述实体用户集中的实体用户，则所述网站服务器将所述用户识别信息和所述Cookie ID 加入到所述实体用户集中的新实体用户。

8. 如权利要求 6 所述统计网站真实用户的系统，其特征在于，还包括存储单元，与所述比较单元连接，用于存储用户设备的实体用户集。

9. 如权利要求 8 所述统计网站真实用户的系统，其特征在于，所述实体用户集包括用户识别信息和网站服务器为用户设备分配的 Cookie ID。

10. 如权利要求 8 所述统计网站真实用户的系统，其特征在于，还包括 Cookie ID 分配单元，用于为第一次登录的用户设备分配 Cookie ID，并将所述Cookie ID 存储于所述存储单元。

11. 如权利要求 6 至 10 中任一项所述统计网站真实用户的系统，其特征在于，所述用户识别信息包括用户 ID。

统计网站真实用户的方法及系统

技术领域

[0001] 本发明涉及网络用户注册信息管理技术领域,尤其涉及一种统计网站真实用户的方法及系统。

背景技术

[0002] 用户登录网站时,需要输入用户名和密码等信息,网站通过上述信息识别该用户后,才能向该用户提供相应的服务。用户输入的用户名和密码通常是自己设定的,具有很大随机性,即一个真实的用户可以设置多个不同的用户名和密码。网站中存储的多个用户名信息可能都属于同一个真实的用户,然而,该网站无法区分哪些用户名是哪个真实用户的,只能分别将这些用户名看作属于不同的真实用户。

[0003] 网站上存在一个真实用户注册多个账号(如用户 ID),然后以不同的账号在网站上活动,网站只能单独的跟踪到单个用户 ID 的行为,只能从局部掌握用户的行为和信息,收集到一个实体用户的部分数据,不能从整体上了解用户真正的行为和信息。

[0004] 因此,在某些领域需要对一个真实用户的所有用户名进行关联,进行统一管理,例如,在银行业务中使用真实用户的登录信息中的身份证号码作为关联,即同一个身份证号码所开的账户都是同一个实体用户的账户,该账户的所有记录都被记录后,作为一个整体来应用,比如判断该用户的信用等级,查找该用户的不良纪录,判断该用户的还款能力等。

[0005] 然而,网站可以对付付费用户利用唯一的身份证明信息关联,例如在登录网站时要求用户提供身份证明,公司营业执照等真实用户唯一的信息,但是网站上有大量的免费会员和未注册的用户,要求他们提供这样的信息显然门槛太高,对用户来说难以接受,对网站的运营成本来说也是非常高的。

[0006] 另一种方式是利用客户端收集用户的信息作为关联,用硬件或者软件的方式来搜集同一台计算机上的用户信息,通过唯一的标识作为关联,从而判定唯一的实体用户。例如,银行系统的 U 盾,即一种网络环境里工行用来识别客户身份的数字证书,一种带智能芯片且形状类似于闪存的实物硬件,拥有 U 盾的客户在使用网上银行时,除了需要提交网上银行用户名和密码外,还需要将 U 盾插入计算机的 U S B 接口,银行经过多重验证无误后,才允许进入网上银行操作自己的账户;又如财务系统的软件狗,开发商程序通过调用软件狗开发套件提供的接口模块对软件狗操作,软件狗响应应该操作并通过接口模块将相应数据返回给开发商程序;或者收集用户计算机的机器码作为标识。

[0007] 但是以上方法都需要客户付出一定的成本,对广大的免费会员和未注册用户来说基本上是不可行的;且由于用户可以选择是否安装,如果不安装客户端,数据就没有办法收集到,如果强制安装,强行或者暗中收集用户的信息,法律上也不允许。

发明内容

[0008] 本发明要解决的问题是提供一种统计网站真实用户的方法及系统,以解决现有技术中不能将一个真实用户注册的多个用户 ID,以及该真实用户在网站上行为的未注册用户

相关联,从而真正从整体上判断一个真实用户的行为和信息的缺陷。

[0009] 为了实现以上目的,本发明提供了一种统计网站真实用户的方法,包括以下步骤:

[0010] 网站服务器检测用户发送的用户识别信息和Cookie ID;

[0011] 所述网站服务器将所述用户识别信息和Cookie ID与已存在于实体用户集中的用户识别信息和Cookie ID进行比较;

[0012] 所述网站服务器根据比较结果对所述用户相关信息进行管理。

[0013] 优选的,所述根据比较结果对所述用户相关信息进行管理具体包括:

[0014] 如果所述用户识别信息属于实体用户集中的实体用户,而所述实体用户中不存在所述Cookie ID,则所述网站服务器将所述Cookie ID加入所述实体用户;

[0015] 如果所述Cookie ID属于所述实体用户集中的实体用户,而所述实体用户中不存在所述用户识别信息,则所述网站服务器将所述用户识别信息加入所述实体用户;

[0016] 如果所述用户识别信息和所述Cookie ID都不属于所述实体用户集中的实体用户,则所述网站服务器将所述用户识别信息和所述Cookie ID加入到所述实体用户集中的新实体用户;

[0017] 如果所述用户识别信息和所述Cookie ID都属于所述实体用户集,则判断所述用户识别信息和所述Cookie ID是否属于同一个实体用户,如果不属于同一个实体用户,则合并为一个实体用户。

[0018] 优选的,在网站检测之前还包括:

[0019] 用户通过用户设备向所述网站服务器发送登录消息;

[0020] 所述网站服务器判断所述登录消息中是否包括Cookie ID,如果没有,则为所述用户设备创建一个Cookie ID,记录到实体用户集的实体用户中,并通知所述用户设备。

[0021] 优选的,所述网站服务器为所述用户设备分配Cookie ID后还包括:用户设备删除所述Cookie ID。

[0022] 优选的,所述用户识别信息包括用户ID。

[0023] 本发明还提供了一种统计网站真实用户的系统,包括网站服务器和用户设备,所述网站服务器进一步包括检测单元、比较单元和管理单元;

[0024] 所述检测单元,用于检测用户通过用户设备发送的用户识别信息和Cookie ID;

[0025] 所述比较单元,与所述检测单元连接,用于将所述用户识别信息和Cookie

[0026] ID与已存在于网站服务器的实体用户集中的用户识别信息和Cookie ID进行比较;

[0027] 所述管理单元,与所述比较单元连接,用于根据所述比较结果对所述用户相关信息进行管理。

[0028] 优选的,所述管理单元根据所述比较结果对所述用户相关信息进行管理具体包括:

[0029] 如果所述用户识别信息属于实体用户集中的实体用户,而所述实体用户中不存在所述Cookie ID,则所述网站服务器将所述Cookie ID加入所述实体用户;

[0030] 如果所述Cookie ID属于所述实体用户集中的实体用户,而所述实体用户中不存在所述用户识别信息,则所述网站服务器将所述用户识别信息加入所述实体用户;

[0031] 如果所述用户识别信息和所述 Cookie ID 都不属于所述实体用户集中的实体用户，则所述网站服务器将所述用户识别信息和所述 Cookie ID 加入到所述实体用户集中的新实体用户；

[0032] 如果所述用户识别信息和所述 Cookie ID 都属于所述实体用户集，则判断所述用户识别信息和所述 Cookie ID 是否属于同一个实体用户，如果不属于同一个实体用户，则合并为一个实体用户。

[0033] 优选的，还包括存储单元，与所述比较单元连接，用于存储用户设备的实体用户集。

[0034] 优选的，所述实体用户集包括用户识别信息和网站服务器为用户设备分配的 Cookie ID。

[0035] 优选的，还包括 Cookie ID 分配单元，用于为第一次登录的用户设备分配 Cookie ID，并将所述 Cookie ID 存储于所述存储单元。

[0036] 优选的，所述用户识别信息包括用户 ID。

[0037] 与现有技术相比，本发明具有以下优点：

[0038] 本发明解决了网站服务器对于申请多个账号的真实用户进行管理的问题，网站服务器不再片面的分析单个虚拟用户的信息、行为和属性，而是将网络上的虚拟用户和现实中的真实用户对应，收集到一个真实用户在网站服务器上注册的所有用户信息，例如用户名、用户密码等；以及以未注册用户身份在网站服务器上的行为信息，例如用户邮箱地址、电话号码等；从而可以加以利用，全面分析。

[0039] 另外，本发明利用 Cookie ID 和用户 ID 作为关联，应用成本很低，对用户不会有任何需要付出的成本，对用户的应用没有影响，也没有法律上的风险。

附图说明

[0040] 图 1 是本发明一种统计网站真实用户的系统结构图；

[0041] 图 2 是本发明一种统计网站真实用户的方法流程图；

[0042] 图 3 是图 2 的一具体实施例流程图。

具体实施方式

[0043] 本发明的核心思想是：在同一台计算机上登录网站服务器的用户，对应同一个 Cookie ID。对于未注册为网站会员的用户，网站服务器只记录其 Cookie ID；对于注册为网站会员的用户，网站服务器记录其用户 ID 和 Cookie ID。网站服务器根据用户 ID 和 Cookie ID 的对应关系，判定同一个用户 ID 对应的 Cookie ID 属于一个实体用户，同一个 Cookie ID 对应的用户 ID 属于同一个实体用户，同时这对用户 ID 和 Cookie ID 属于同一个实体用户，从而系统将一组用户 ID 和 Cookie ID 指定对应某一个实体用户 (User)，并给予一个编号 (User ID)。网站服务器可以根据一个实体用户的信息全面分析单个虚拟用户的信息、行为和属性。

[0044] 本发明引入企业用户和个人用户两个概念，一般情况下，使用国际网站进行商业行为的用户大多是企业用户，例如大中小企业或采购商等。由于商业行为在网吧进行就像在网吧使用网上银行一样，用户信息存在被盗用的风险，安全很难保证，因此企业用户通常

使用企业内部的计算机,使用公共场所计算机的概率比较小,多个用户同时使用同台计算机的概率就更小。而且,本发明主要是统计企业用户的信息,当然不排除有误关联的情况,如果个人用户在登录时能提供企业用户的合法证明材料,网站也会象对待企业用户一样统计相关信息。

[0045] 根据某大型中文站的统计数据,2002年只有10%的中文网站会员(企业用户及非企业用户)利用网吧上网,到4年后的今天这个比例应该已经降到更低;其中,企业用户使用网吧等公共场所的计算机登录中文网站的比例更是可以忽略。另外,对于外文网站,由于大多涉及外贸商务,其进入的门槛比中文网站更高,因此,企业用户利用网吧等个人计算机登录的概率更低。因此,上述数据可以说明不同的企业用户使用同一台电脑的概率很低,而使用企业自有计算机的概率很高。另外,即使有不同的企业用户使用同一台计算机的情况发生,他们访问同一家网站的概率也会很低。

[0046] 因此,本发明可以认为同一个用户ID对应现实中的一个真实用户,不论这个用户ID在哪台计算机上登录,都认为这是同一个真实用户的行为;并且认为使用同一台计算机登录网站服务器的用户有很高的概率是一个真实用户或者是同一家公司的行为。

[0047] 本发明提供一种统计网站真实用户系统,如图1所示,包括网站服务器100和多个用户设备200,网站服务器100进一步包括检测单元101、比较单元102、管理单元103、存储单元104和Cookie ID分配单元105。

[0048] 其中,检测单元101用于检测用户通过用户设备200发送的用户识别信息和Cookie ID;比较单元102与检测单元101连接,用于将用户识别信息和Cookie ID与已存在于网站服务器100的实体用户集中的用户识别信息和Cookie ID进行比较;管理单元103与比较单元102连接,用于根据比较结果对用户相关信息进行管理;存储单元104与比较单元102连接,用于存储实体用户集,实体用户集中可能包括多个实体用户,每个实体用户包括用户识别信息和网站服务器为用户设备分配的Cookie ID;Cookie ID分配单元105用于为第一次登录的用户设备分配Cookie ID,并将Cookie ID存储于存储单元104。

[0049] 管理单元103根据比较结果对用户相关信息进行管理具体包括:如果用户识别信息属于实体用户集中的一个实体用户,而该实体用户中不存在Cookie ID,则管理单元103将Cookie ID加入该实体用户;如果Cookie ID属于实体用户集中的一个实体用户,而该实体用户中不存在用户识别信息,该用户识别信息为同一实体用户的概率很高,则管理单元103将该用户识别信息加入该实体用户;如果用户识别信息和Cookie ID都不属于实体用户集中的实体用户,则管理单元103将用户识别信息和Cookie ID加入到实体用户集中的一个新实体用户;如果用户识别信息和Cookie ID都属于实体用户集,则管理单元103判断用户识别信息和Cookie ID是否属于同一个实体用户,如果不属于同一个实体用户,则合并为一个实体用户,如果属于同一个实体用户,则不操作。

[0050] 网站服务器可以根据上述实体用户的信息对登录该网站的真实用户的信息进行统计,并根据一个真实用户的信息全面分析单个虚拟用户的信息、行为和属性,实现对真实用户的行为进行统一管理。

[0051] 本发明还提供了一种统计网站真实用户的方法,如图2所示,包括以下步骤:

[0052] 步骤s201,网站服务器检测用户通过用户设备发送的用户识别信息和Cookie ID。其中,用户识别信息对于已注册的用户来说,包括用户ID、用户密码、邮件地址等信息,对于

未注册的用户来说,不包括用户 ID ;用户设备通常为用户使用的计算机 ;Cookie 是网站服务器在用户使用的计算机上创建的一些小文本文件,这些文件用于存储搜集有关用户的站点访问信息,其中包括访问的站点、执行的操作以及提供的任何个人信息,网站服务器使用 Cookie 中的信息来提供个性化的内容、完成交易及搜集统计数据,只要用户浏览了网站,就可以根据用户登录、注册等动作留下对应的 Cookie,只要用户不清除该 Cookie,则下次用户再次访问时,网站就可以根据该 Cookie 中保存的信息获知用户的部分信息,例如 Email 地址、用户名等。

[0053] 步骤 s202,网站服务器将用户识别信息和 Cookie ID 与已存在于实体用户集中的用户识别信息和 Cookie ID 进行比较。

[0054] 步骤 s203,网站服务器根据比较结果对用户相关信息进行管理,如果用户识别信息属于实体用户集中的实体用户,而实体用户中不存在 Cookie ID,则网站服务器将 Cookie ID 加入实体用户 ;如果 Cookie ID 属于实体用户集中的实体用户,而实体用户中不存在用户识别信息,该用户识别信息为同一实体用户的概率很高,则网站服务器将用户识别信息加入实体用户 ;如果用户识别信息和 Cookie ID 都不属于实体用户集中的实体用户,则网站服务器将用户识别信息和 Cookie ID 加入到实体用户集中的新实体用户 ;如果用户识别信息和 Cookie ID 都属于实体用户集,则判断用户识别信息和 Cookie ID 是否属于同一个实体用户,如果不属于同一个实体用户,则合并为一个实体用户。

[0055] 对于第一次在网站服务器上登录的用户设备,在步骤 s201 之前还包括 :

[0056] 网站服务器生成一个唯一的、不重复的 Cookie ID(即 IP 地址加序列号) : 用户通过用户设备向网站服务器发送登录消息,网站服务器判断登录消息中是否包括 Cookie ID,如果没有,则为用户设备创建一个 Cookie ID,记录到实体用户集的实体用户中,并通知用户设备。用户登录的计算机中会保存该 Cookie ID,该 Cookie ID 是永久有效的,除非用户自行删除,否则在用户登录网站服务器时,网站服务器会判断,如果该用户所在的计算机中所存的 CookieID 已经存在于网站服务器,则该计算机已经登录过该网站服务器,该网站服务器不再生成新的 Cookie。

[0057] 结合图 2 的本发明基本原理流程图,本发明一个具体应用实施例如图 3 所示,网站服务器收集用户 ID 和 Cookie ID 信息 ; 网站服务器将用户 ID 和 Cookie

[0058] ID 信息归入实体用户集 ; 网站服务器判断新接收的用户 ID 和 Cookie ID 信息是否已存在于实体用户集,如果用户识别信息属于实体用户集中的实体用户,而实体用户中不存在 Cookie ID,则网站服务器将 Cookie ID 加入实体用户 ; 如果 Cookie ID 属于实体用户集中的实体用户,而实体用户中不存在用户识别信息,则网站服务器将用户识别信息加入实体用户 ; 如果用户识别信息和 CookieID 都不属于实体用户集中的实体用户,则网站服务器将用户识别信息和 Cookie ID 加入到实体用户集中的新实体用户 ; 如果用户识别信息和 Cookie ID 都属于实体用户集,则判断用户识别信息和 Cookie ID 是否属于同一个实体用户,如果不属于同一个实体用户,则合并为一个实体用户。

[0059] 以下实例是本发明在几种实际场景中的具体应用 :

[0060] 真实用户第一次登录网站服务器,先以未注册用户的身份在网站上活动,而后注册为会员。当真实用户以未注册用户的身份登录网站服务器时,网站服务器会在其计算机上设置一个 Cookie,同时系统记录下这个 Cookie ID,并对应到一个实体用户 ; 而后该真实

用户注册成会员，网站服务器记录下其用户 ID，并根据同一个 Cookie ID，将该用户 ID 也对应到同一个实体用户。

[0061] 由于用户 ID 不同而 Cookie ID 相同时，该用户 ID 为同一真实用户的概率很高，因此，同一个真实用户在同一台计算机上，用多个用户 ID 登录网站服务器，网站服务器记录下这些用户 ID，并根据同一个 Cookie ID，将这些用户 ID 也对应到同一个实体用户。网站服务器判断如果这些用户 ID 曾经在其他计算机上登录过该网站服务器，会将这些用户 ID 原来对应的实体用户和这台计算机上的 Cookie ID 对应的实体用户合并成一个实体用户。

[0062] 真实用户使用相同的用户 ID 在不同计算机上登录网站。网站服务器判断这台计算机是否曾经登录过网站，如果没有登录过，网站服务器会在其计算机上设置一个 Cookie，同时系统记录下这个 Cookie ID，并根据同一个用户 ID，将这个 Cookie ID 对应到同一个实体用户；如果已经登录过，网站服务器会将这台计算机上的 Cookie ID 原本对应的实体用户和该登录用户 ID 原来对应的实体用户合并成一个实体用户。

[0063] 真实用户将 Cookie 删除后登录网站，如果真实用户以未注册用户活动，或者注册新的用户 ID 后在网站上活动，网站服务器将这个 Cookie ID 或新的用户 ID 对应到一个新实体用户；如果登录的帐号是原来登录过的帐号，实体用户用原有的用户 ID 登录，网站服务器会将这个新的 Cookie ID 对应到该用户 ID 对应的实体用户。

[0064] 以上公开的仅为本发明的几个具体实施例，但是，本发明并非局限于此，任何本领域的技术人员能思之的变化都应落入本发明的保护范围。

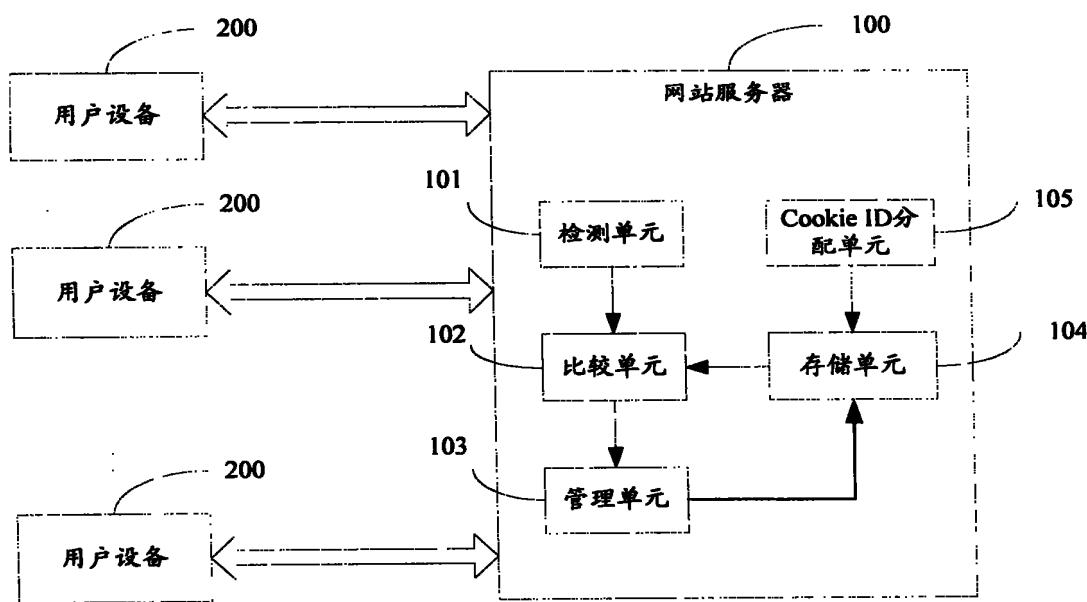


图 1

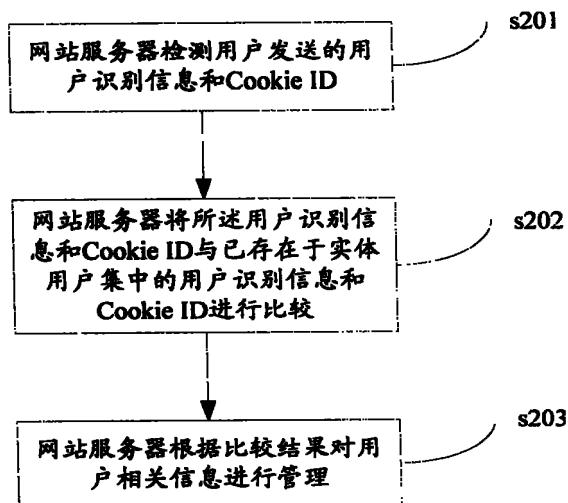


图 2

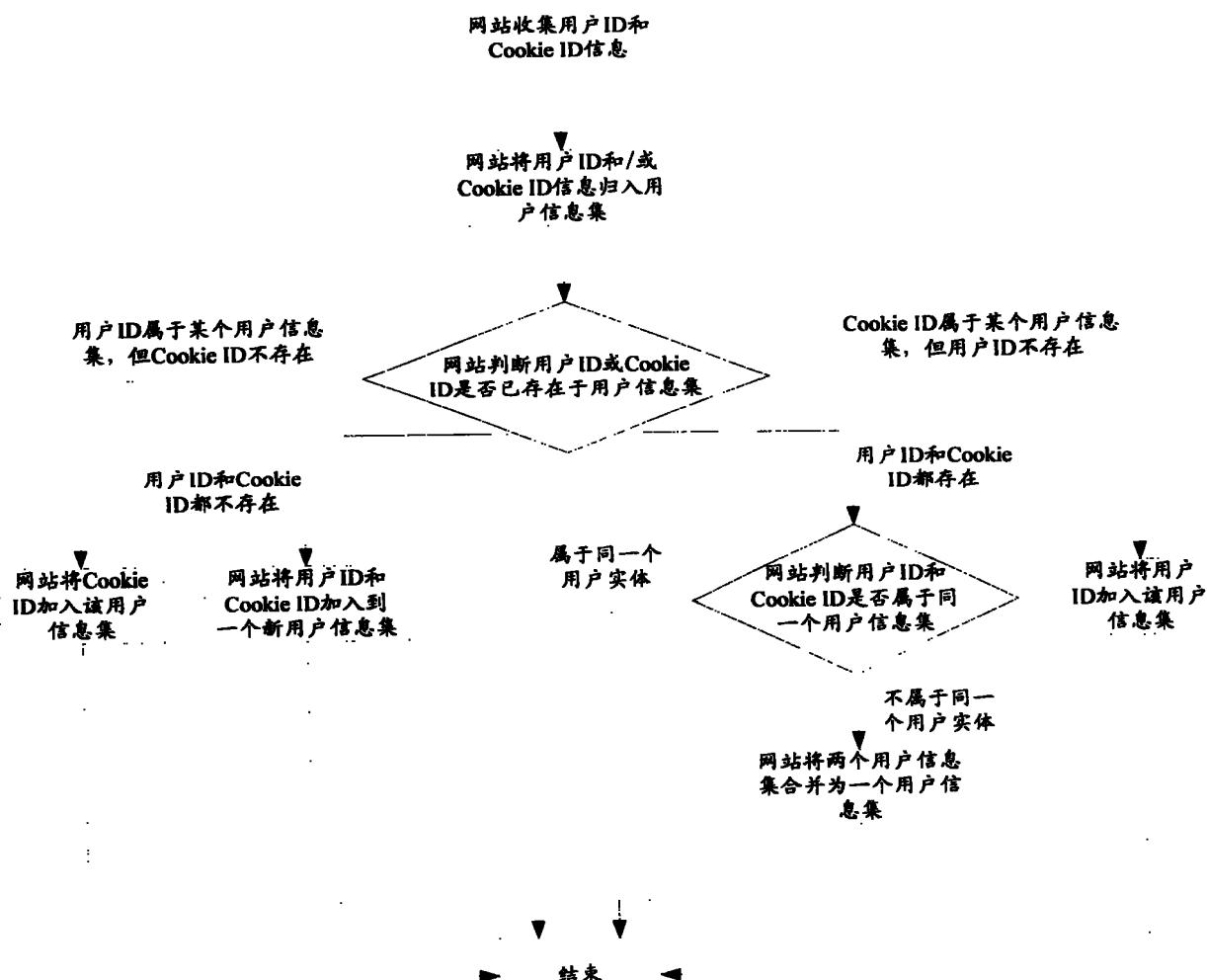


图 3