

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06F 17/30 (2006.01)

G06Q 30/00 (2006.01)

H04L 29/06 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200710100736.6

[43] 公开日 2008 年 6 月 4 日

[11] 公开号 CN 101192235A

[22] 申请日 2007.4.11

[21] 申请号 200710100736.6

[71] 申请人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市高新科技园南区
高新南一道飞亚达高科技大厦 5-10 层

[72] 发明人 岳亚丁 廖焕华 宋大可 刘奕慧
龚磊 曾海涛 吴晓光

[74] 专利代理机构 深圳市顺天达专利商标代理有限公司

代理人 郭伟刚 蔡晓红

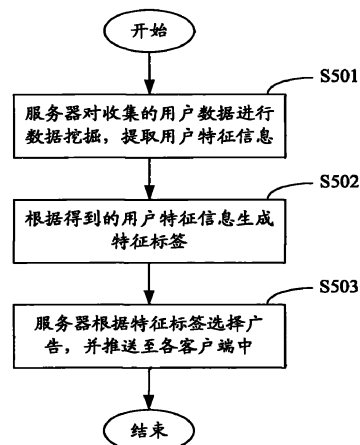
权利要求书 2 页 说明书 10 页 附图 4 页

[54] 发明名称

一种基于用户特征推送广告的方法、系统及设备

[57] 摘要

本发明涉及通信领域,提供了一种基于用户特征推送广告的方法、系统及设备。所述方法包括以下步骤:A.服务器对用户原始数据进行数据挖掘,并根据提取出的用户特征信息生成对应的特征标签;B.根据所述特征标签确定待投放的广告的属性,并将所述广告推送到客户端。本发明通过在服务器中收集存储大量用户原始数据,并对其进行数据挖掘,利用提取出的用户特征信息生成特征标签,再根据特征标签推送网络广告,提高了广告推送的针对性,进而提高了广告的命中率。



1、一种基于用户特征推送广告的系统，包括服务器和客户端，所述服务器包括用于存储用户原始数据的数据库，以及将广告发送至客户端的广告推送单元，其特征在于，所述服务器还包括一个特征挖掘单元；

所述特征挖掘单元与数据库及广告推送单元相连，用于对数据库中的用户原始数据进行数据挖掘，根据提取出的用户特征信息生成特征标签，并将所述特征标签送入广告推送单元以控制广告推送单元对广告的选择及推送。

2、根据权利要求1所述的基于用户特征推送广告的系统，其特征在于，所述特征挖掘单元进一步包括数据处理模块和特征标签模块；

所述数据处理模块用于对数据库中的用户原始数据进行数据挖掘，以提取出用户特征信息；

所述特征标签模块与数据处理模块相连，用于根据所述用户特征信息生成特征标签。

3、根据权利要求2所述的基于用户特征推送广告的系统，其特征在于，所述特征挖掘单元进一步包括一个数据分类模块，其与所述数据处理单元相连，用于对数据库中的用户原始数据进行分类，并将分类后的数据送入数据处理模块中。

4、根据权利要求3所述的基于用户特征推送广告的系统，其特征在于，所述特征挖掘单元进一步包括一个校验模块，其与数据处理模块相连，用于对数据处理模块的数据处理结果进行检验，以修正所述数据处理模块的处理精度。

5、一种基于用户特征推送广告的设备，即与客户端相连的服务器，所述服务器包括用于存储用户原始数据的数据库，以及将广告发送至客户端的广告推送单元，其特征在于，所述服务器还包括一个特征挖掘单元；

所述特征挖掘单元与数据库及广告推送单元相连，用于对数据库中的用户原始数据进行数据挖掘，根据提取出的用户特征信息生成特征标签，并将所述特征标签送入广告推送单元以控制广告推送单元对广告的选择及推送。

6、根据权利要求5所述的基于用户特征推送广告的设备，其特征在于，所述特征挖掘单元进一步包括数据处理模块和特征标签模块；

所述数据处理模块用于对数据库中的用户原始数据进行数据挖掘，以提取出用户特征信息；

所述特征标签模块与数据处理模块相连,用于根据所述用户特征信息生成特征标签。

7、一种基于用户特征推送广告的方法,其特征在于,所述方法包括以下步骤:

A.服务器对用户原始数据进行数据挖掘,并根据提取出的用户特征信息生成对应的特征标签;

B.根据所述特征标签确定待投放的广告的属性,并将所述广告推送至客户端。

8、根据权利要求7所述的基于用户特征推送广告的方法,其特征在于,所述步骤A之前包括:服务器收集用户原始数据,并存储到数据库中;

所述用户原始数据包括:即时通信数据、网站数据、游戏数据、支付数据、场景数据以及广告点击数据。

9、根据权利要求8所述的基于用户特征推送广告的方法,其特征在于,所述步骤A进一步包括:

A1.服务器对数据库中存储的用户原始数据进行数据挖掘,提取出用户特征信息;

A2.根据用户特征信息生成对应的特征标签。

10、根据权利要求9所述的基于用户特征推送广告的方法,其特征在于,所述步骤A2进一步包括:将所述用户特征信息进行编码,并将编码结果作为特征标签。

一种基于用户特征推送广告的方法、系统及设备

技术领域

本发明涉及通信领域，更具体地说，涉及一种基于用户特征推送广告的方法、系统及设备。

背景技术

在这个以信息沟通为主导的经济时代，随着互联网技术的完善，网络智能广告也在迅猛发展。

网络智能广告的核心技术是进行受众分析。也即，通过对互联网用户的网络行为进行分析得出用户特征信息，比如该用户的年龄、性别、地理位置、收入状况以及其感兴趣的领域等，从而有针对性地投放用户感兴趣的个性化广告。

而目前最为普遍的受众分析，是通过对用户注册资料进行汇总，将其作为用户特征信息并推送广告。如图1所示，即为现有技术推送广告的系统结构，包括服务器100及与其相连的多个客户端（客户端200……客户端N）。其中服务器100包括数据库101和广告推送单元103：（1）数据库101用于存储收集到的用户原始数据，主要是用户在网络（各个网站或论坛等）中留下的注册资料等；（2）广告推送单元103利用数据库101中汇总的用户注册资料，确定广告属性并将广告推送至各客户端（客户端200……客户端N）中。

由上可知，该现有技术没有对用户原始数据进行深入挖掘，不能掌握完整精确的用户特征信息，因此广告推送的针对性较低，进一步导致命中率（也即广告的点击率）较低。

发明内容

本发明的目的在于提供一种基于用户特征推送广告的系统，旨在解决现有技术推送广告时针对性较低，导致广告命中率低的问题。

本发明的目的还在于提供一种基于用户特征推送广告的设备，以更好地解决现有技术中存在的上述问题。

本发明的目的还在于提供一种基于用户特征推送广告的方法,以更好地解决现有技术中存在的上述问题。

为了实现发明目的,所述基于用户特征推送广告的系统包括服务器和客户端,所述服务器包括用于存储用户原始数据的数据库,以及将广告发送至客户端的广告推送单元,所述服务器还包括一个特征挖掘单元;

所述特征挖掘单元与数据库及广告推送单元相连,用于对数据库中的用户原始数据进行数据挖掘,根据提取出的用户特征信息生成特征标签,并将所述特征标签送入广告推送单元以控制广告推送单元对广告的选择及投放。

优选地,所述特征挖掘单元进一步包括数据处理模块和特征标签模块;

所述数据处理模块用于对数据库中的用户原始数据进行数据挖掘,以提取出用户特征信息;

所述特征标签模块与数据处理模块相连,用于根据所述用户特征信息生成特征标签。

优选地,所述特征挖掘单元进一步包括一个数据分类模块,其与所述数据处理单元相连,用于对数据库中的用户原始数据进行分类,并将分类后的数据送入数据处理模块中。

优选地,所述特征挖掘单元进一步包括一个校验模块,其与数据处理模块相连,用于对数据处理模块的数据处理结果进行检验,以修正所述数据处理模块的处理精度。

为了更好地实现发明目的,所述设备基于用户特征推送广告的设备,即与客户端相连的服务器,包括用于存储用户原始数据的数据库,以及将广告发送至客户端的广告推送单元,所述服务器还包括一个特征挖掘单元;

所述特征挖掘单元与数据库及广告推送单元相连,用于对数据库中的用户原始数据进行数据挖掘,根据提取出的用户特征信息生成特征标签,并将所述特征标签送入广告推送单元以控制广告推送单元对广告的选择及投放。

优选地,所述特征挖掘单元进一步包括数据处理模块和特征标签模块;

所述数据处理模块用于对数据库中的用户原始数据进行数据挖掘,以提取出用户特征信息;

所述特征标签模块与数据处理模块相连,用于根据所述用户特征信息生成特征标签。

为了更好地实现发明目的,所述基于用户特征推送广告的方法包括以下步

骤:

A.服务器对用户原始数据进行数据挖掘,并根据提取出的用户特征信息生成对应的特征标签;

B.根据所述特征标签确定待投放的广告的属性,并将所述广告推送到客户端。

优选地,所述步骤A之前包括:服务器收集用户原始数据,并存储到数据库中;

所述用户原始数据包括:即时通信数据、网站数据、游戏数据、支付数据、场景数据以及广告点击数据。

优选地,所述步骤A进一步包括:

A1.服务器对数据库中存储的用户原始数据进行数据挖掘,提取出用户特征信息;

A2.根据用户特征信息生成对应的特征标签。

优选地,所述步骤A1中的数据挖掘包括以下方式:归纳、计算及预测。

优选地,所述步骤A1中的用户特征信息包括:个人属性、家庭属性、网络行为、兴趣爱好。

优选地,所述步骤A2进一步包括:将所述用户特征信息进行编码,并将编码结果作为特征标签。

本发明通过在服务器中收集存储大量用户原始数据,并对其进行数据挖掘,利用提取出的用户特征信息生成特征标签,再根据特征标签推送网络广告,提高了广告推送的针对性,进而提高了广告的命中率。

附图说明

图1是现有技术中基于用户特征推送广告的系统结构图;

图2是本发明中基于用户特征推送广告的系统结构图;

图3是图2所示系统中特征挖掘单元的内部结构图;

图4是本发明中基于用户特征推送广告的另一系统结构图;

图5是本发明中基于用户特征推送广告的方法流程图;

图6是本发明中基于用户特征推送广告的另一方法流程图。

具体实施方式

为了使本发明的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本发明进行进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本发明，并不用于限定本发明。

本发明中，服务器通过各种渠道收集存储大量的用户原始数据，并利用建立的数据挖掘模型对用户原始数据进行数据挖掘，提取出有效的用户特征信息并生成对应的特征标签，再根据特征标签推送网络广告，因此提高了广告推送的针对性。

图2示出了本发明中基于用户特征推送广告的系统结构，该系统包括服务器100，以及与其相连的多个客户端（客户端200、客户端300...客户端N）。应当说明的是，本发明所有图示中各设备之间的连接关系是为了清楚阐释其信息交互及控制过程的需要，因此应当视为逻辑上的连接关系，而不应仅限于物理连接。

各客户端（客户端200、客户端300...客户端N）典型的可为各种能够...的终端设备，例如个人计算机（Personal Computer, PC）、个人数字助理（Personal Digital Assistant, PDA）、移动电话（Mobile Phone, MP）等，因此本发明的保护范围不应限定为某种特定类型的客户端。

服务器100用于收集和存储用户原始数据，并从中提取用户特征信息，以及根据用户特征信息进行网络广告的针对性推送。该服务器100典型的可为专门的广告伺服器，或者具有广告伺服功能的大型网站服务器等，因此本发明的保护范围不应限定为某种特定类型的服务器。

在本发明中，服务器100包括数据库101、特征挖掘单元102和广告推送单元103，其中：

（1）数据库101用于存储所收集到的用户原始数据，本发明中用户原始数据的种类有多种，可通过多种方式和渠道收集用户原始数据。

在本发明的一个示例方案中，用户原始数据可包括：即时通信（Instant Message, IM）数据、网站数据、游戏数据、支付数据、场景数据、广告点击数据等等。而收集上述用户原始数据的方式可以从网站提取用户注册信息、跟踪用户在网站的网络行为，以及进行调研，等等。

（2）特征挖掘单元102与数据库101及广告推送单元103相连，用于对数据库101中的用户原始数据进行数据挖掘，根据提取出的用户特征信息生成特征

标签，并将特征标签送入广告推送单元103中。该特征挖掘单元102的内部结构将在其后进行详细阐述。

本发明中用户特征信息包括多种，例如个人属性、家庭属性、网络行为、兴趣爱好等等。在一个示例方案中，用户特征信息以如下表格的形式展现：

特征属性类别	特征属性	取值
个人属性	年龄	6岁以下，6-12，13-15，15-18，19，23，24-30，31-35，36-40，41-50，51岁以上
	性别	男、女
	婚姻状况	已婚、未婚
	民族	汉族、或56个少数民族之一
	国家	100多个国家
	省	24个省，5个自治区，4个直辖市，2个特别行政区
	地区	各省拥有的行政地区
	教育程度	高中（中专）以下、高中（中专）、大专、本科、硕士、博士及以上
	身份类型	无业、学生、公司职员、工人、个体工商业、企业主、农民、军人、其他
	所处行业	农、林、牧、渔业 地质勘查业、水利管理业 社会服务业 房地产业 金融、保险业 卫生、体育和社会福利业 制造业 批发和零售贸易、餐饮业 教育、文化艺术及广播电影电视业 电力、蒸汽及水的生产和供应业 交通运输、仓储、及邮电通信业 科学研究和综合技术服务业 建筑业 采掘业 国家机关、政党机关和社会团体 其他行业

	个人月收入	无收入、500 以下、501-1000、1001-1500、1501-2000、2501-3000、3001-4000、4001-5000、5001-8000、8001-10000、10000 以上
家庭属性	有无子女	无子女、有子女
	家庭月收入	无收入、1000 以下、1001 - 3000、3001 - 6000、6001 - 8000、8001 - 10000、10001 - 15000、150001 - 30000、30000 以上
	家庭成员人数	1, 2 - 3, 3 以上
	居住房屋性质	自己所有、租住
	居住房屋面积	50 以下, 51 - 100, 101 - 150, 151 - 300, 300 以上
	居住区域类型	农村、郊区、城市
	家庭主要交通工具 (可多选)	无、自行车、汽车 (分自有小汽车、出租车、公共交通)
兴趣爱好	兴趣爱好 (可多选)	汽车、房产、旅游、数码、音乐、动漫、游戏、体育、交友、读书、军事、财经、文学、美食、
网络行为	上网地点 (可多选)	家里、工作场所、网吧、学校、公共场所、其它
	上网设备 (可多选)	台式计算机、笔记本电脑、手机
	上网方式	专线、拨号、宽带
	上网时段	0 - 1 点、1 - 2 点、.....、23 - 24 点
	平均每周上网时间	(以小时计)
	平均每月实际花费的上网费用	(以元计)
	最近三个月是否变更上网接入城市	是、否
	经常使用的网络服务 (可多选)	浏览新闻、搜索引擎、收发邮件、论坛/BBS/讨论组等、即时通信、获取信息、在线影视收看及下载、在线音乐收听及下载、文件上传下载、网上游戏、网上校友录、网上购物、个人主页空间、博客、网上招聘、网络聊天室、网上金融、电子杂志、网上教育、网上销售、短信/彩信服务、网络电话、网上预订、电子政务、征婚/交友/社区俱乐部、其它

特征挖掘单元102可采取多种方式,例如归纳、计算、预测等,从数据库101存储的用户原始数据中提取出上述表格中的各种用户特征信息。

(3) 广告推送单元103与特征挖掘单元102相连,用于根据特征挖掘单元102发送的特征标签,确定待推送的广告的属性,并将所确定的广告推送至各客户端(客户端200、客户端300……客户端N)中。

图3示出了图2的系统中特征挖掘单元102的内部结构,包括数据分类模块1021、数据处理模块1022、特征标签模块1023和校验模块1024,其中:

(1) 数据分类模块1021用于对数据库101中存储的大量的用户原始数据进行分类,也即,将用户分成多个群体,再将分类后的数据输入数据处理模块1022中。当然本发明中此模块并非必要,也可不进行数据分类,而利用数据处理模块1022直接对数据库101中的用户原始数据进行处理。

(2) 数据处理模块1022用于对数据库101中的用户原始数据进行数据挖掘,以提取出用户特征信息。本发明中数据处理模块1022对用户特征信息的提取可有多种方式,包括归纳、计算、预测等,分别针对不同种类的用户特征信息。

例如,可通过归纳方式获得用户的兴趣爱好相关的特征信息,包括汽车、房产、旅游、数码、音乐、动漫、游戏、体育、交友、读书、军事、财经、文学、美食等等;可通过计算方式获取用户使用某企业服务的忠诚度,包括用户的注册时间、登录频率、使用项目、累计消费额等等;可通过调研结果和数据筛选预测用户的其他特征信息。

(3) 特征标签模块1023与数据处理模块1022相连,用于根据数据处理模块1022提取出的用户特征信息生成对应的特征标签,并送入广告推送单元103中。

本发明中特征标签的生成可包括多种方式。在一个典型的实施例中,特征标签模块1023通过对所提取的用户特征信息进行编码处理,将所生成的编码作为特征标签。

(4) 校验模块1024与数据处理模块1022相连,用于对数据处理模块1022的数据处理结果进行检验,以修正数据处理模块1022的处理精度。

图4示出了本发明中基于用户特征推送广告的另一系统结构,该系统包括

服务器100及与其相连的多个客户端（客户端200、客户端300……客户端N）。与图2所示系统结构相比，服务器100中除包括数据库101、特征挖掘单元102和广告推送单元103外，还包括一个效果分析单元104。

该效果分析单元104根据各个客户端（客户端200、客户端300……客户端N）反馈的结果，对广告推送效果进行分析，即计算广告的曝光率、命中率（即点击率）等，并将所得数据反馈给特征挖掘单元102，以判定其数据挖掘的成效，从而可在之后进一步进行性能优化。

本发明中，曝光率和命中率的计算方法可有多种。在一个示例方案中，效果分析单元104计算曝光率的公式如下：曝光率 = 覆盖用户数 / 总用户数；其计算命中率的公式如下：命中率 = 点击数 / 曝光数。

在上述示例方案的一个实施例中，所得曝光率和命中率如下表：

属性类别	用户数	广告名称	曝光次数	曝光率	点击次数	命中率
汽车群	100 万	宝马 S 系列	90 万	90 %	30 万	33.3 %
女性用户	500 万	力士香皂	300 万	60 %	240 万	80 %

在上述示例方案的一个实施例中，所得曝光率和命中率如下表：

属性类别	用户数	广告名称	曝光次数	曝光率	点击次数	命中率
深圳 25 ~ 30 岁的男性白领	100 万	南山小区销售广告	40 万	40 %	10 万	25 %
家庭收入超过五十万的 30 岁以上的女性	50 万	新品服装	40 万	80 %	30 万	75 %

当然，本发明中效果分析单元104还可通过其他方式计算广告的曝光率和命中率，并不限于以上所述的方法。

图5示出了本发明中基于用户特征推送广告的方法流程，该方法流程基于图2、图3、图4所示的系统结构，具体过程如下：

在执行本发明的所有步骤之前，服务器100通过各种渠道或方式收集用户原始数据并存储在数据库101中，这些用户原始数据包括IM数据、网站数据、游戏数据、支付数据、场景数据、广告点击数据等等。而收集上述用户原始数据的方式可以从网站提取用户注册信息、跟踪用户在网站的网络行为，以及进行调研，等等。

在步骤S501中，服务器100对收集的用户原始数据进行数据挖掘，从中提取出用户特征信息。本发明中用户特征信息包括多种，例如个人属性、家庭属性、网络行为、兴趣爱好等等，在前述图2的一个示例方案中用表格形式对用户特征信息进行了展示。在此步骤中，服务器100可利用其特征挖掘单元102从数据库101中存储的用户原始数据中提取用户特征信息，针对不同类型的用户特征信息的提取方式可有多种，例如归纳、计算、预测等。

例如，可通过归纳方式获得用户的兴趣爱好相关的特征信息，包括汽车、房产、旅游、数码、音乐、动漫、游戏、体育、交友、读书、军事、财经、文学、美食等等；可通过计算方式获取用户使用某企业服务的忠诚度，包括用户的注册时间、登录频率、使用项目、累计消费额等等；可通过调研结果和数据筛选预测用户的其他特征信息。

此外，步骤S502可先对用户原始数据进行分类，然后再从分类后的数据中提取用户特征信息。

在步骤S502中，服务器100根据得到的用户特征信息生成特征标签。此步骤中，可采取多种方式生成特征标签。在一个典型的实施例中，特征标签模块1023通过对数据处理模块1022所提取的用户特征信息进行编码处理，将所生成的编码作为特征标签。

在步骤S503中，服务器100根据特征标签选择要投放的广告，并将所选择的广告推送至各客户端（客户端200、客户端300……客户端N）中。

如前所述，该特征标签中包含了用户特征信息，即用户的个人属性、家庭属性、网络行为、兴趣爱好等，因此服务器100的广告推送单元103可根据上述用户特征信息针对性地选择要投放的广告，并进行推送。

图6示出了本发明基于用户特征推送广告的另一方法流程，该方法流程基

于图4所示的系统结构，具体过程如下：

在执行本发明的所有步骤之前，服务器100通过各种渠道或方式收集用户原始数据并存储在数据库101中，这些用户原始数据包括IM数据、网站数据、游戏数据、支付数据、场景数据、广告点击数据等等。而收集上述用户原始数据的方式可以从网站提取用户注册信息、跟踪用户在网站的网络行为，以及进行调研，等等。

在步骤S601中，服务器100对收集的用户原始数据进行数据挖掘，从中提取出用户特征信息，其具体过程与前述步骤S501一致。

在步骤S602中，服务器100根据得到的用户特征信息生成特征标签。此步骤中，可采取多种方式生成特征标签，具体过程与前述步骤S502一致。

在步骤S603中，服务器100根据特征标签选择要投放的广告，并将所选择的广告推送至各客户端（客户端200、客户端300……客户端N）中，具体过程与前述步骤S503一致。

在步骤S604中，根据服务器100的推送数据以及各客户端（客户端200、客户端300……客户端N）反馈的点击数据，计算广告的曝光率和命中率，并将计算结果发送至特征挖掘单元102中，转步骤S601，从而利用曝光率和命中率对数据挖掘进行进一步的优化。

以上所述仅为本发明的较佳实施例而已，并不用以限制本发明，凡在本发明的精神和原则之内所作的任何修改、等同替换和改进等，均应包含在本发明的保护范围之内。

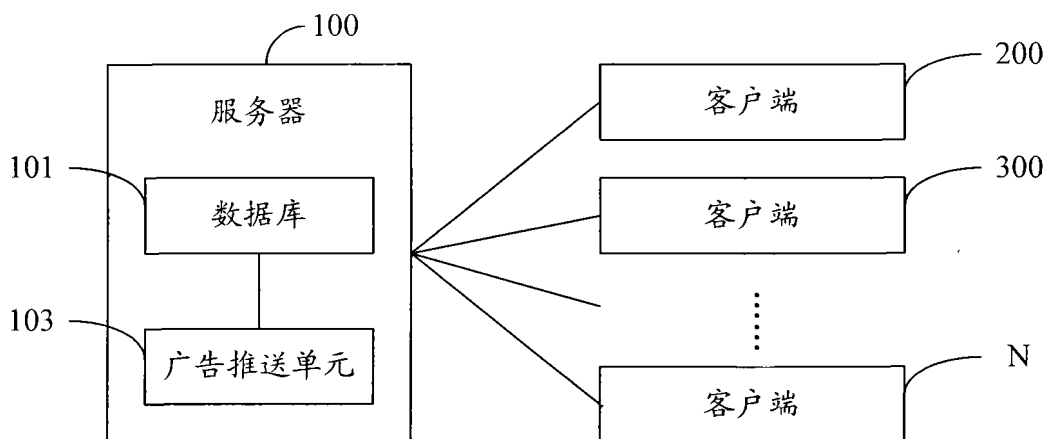


图 1

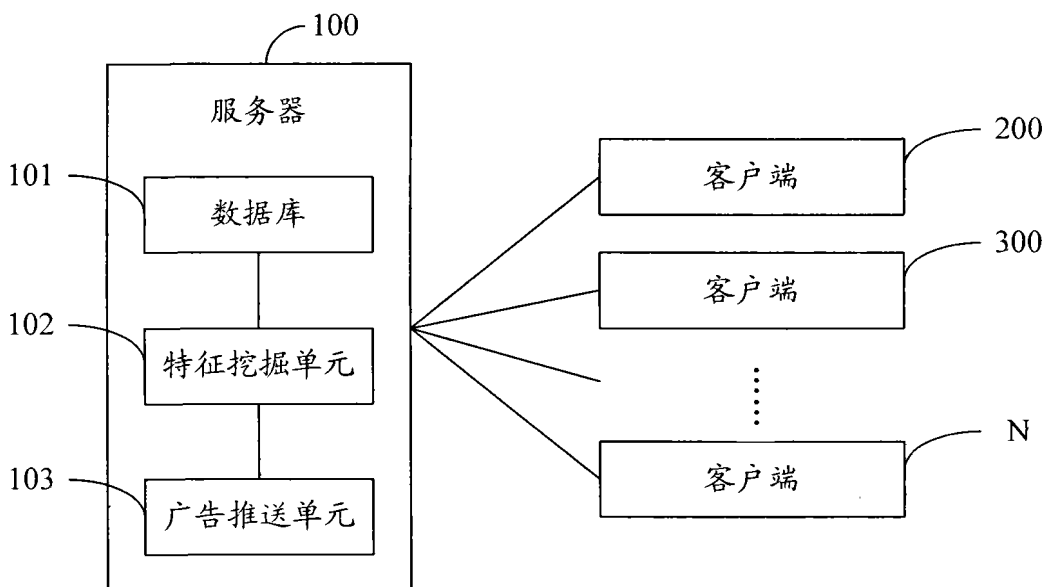


图 2

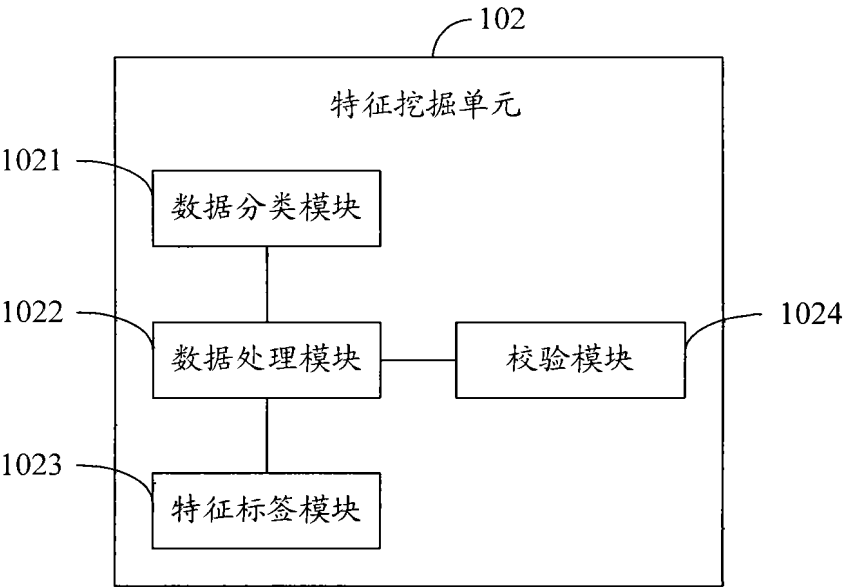


图 3

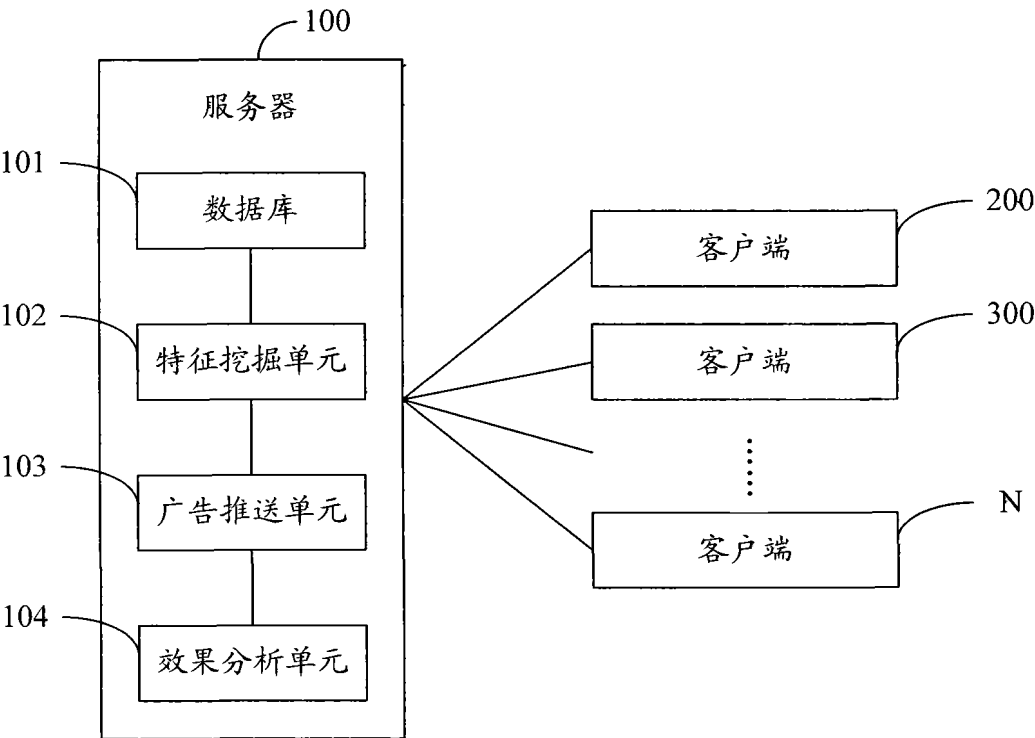


图 4

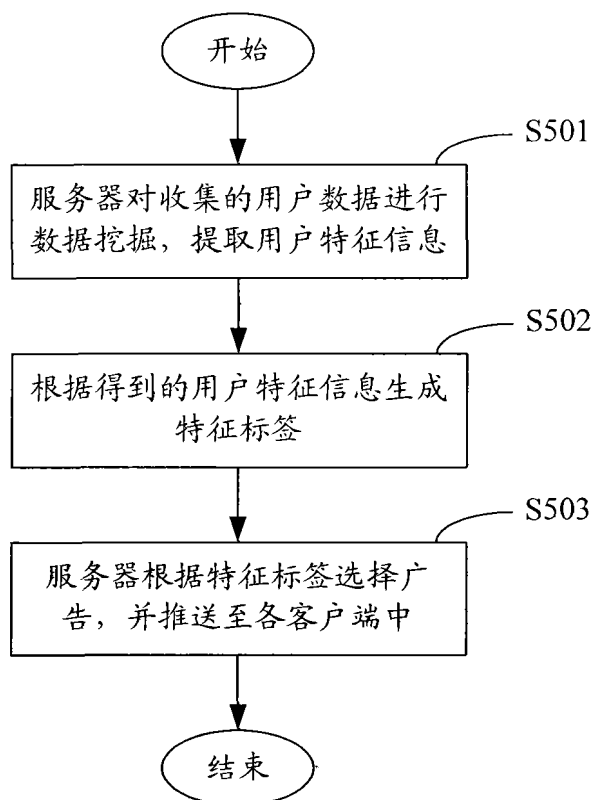


图 5

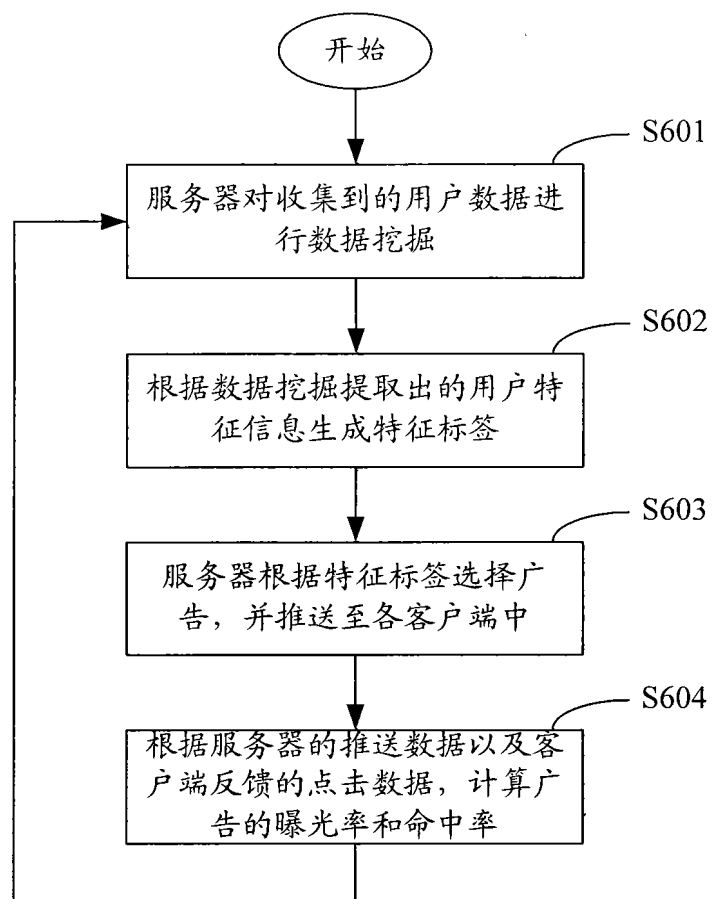


图 6