

(12) 发明专利申请

(10) 申请公布号 CN 102118400 A

(43) 申请公布日 2011. 07. 06

(21) 申请号 200910206791. 2

(22) 申请日 2009. 12. 31

(71) 申请人 北京四维图新科技股份有限公司
地址 100028 北京市朝阳区曙光西里甲 5 号
凤凰置地广场 A 座写字楼 16 层

(72) 发明人 罗丽俊 陈文斌 汪洋 程鹏

(74) 专利代理机构 北京银龙知识产权代理有限公司 11243

代理人 许静

(51) Int. Cl.

H04L 29/08 (2006. 01)

G06F 17/30 (2006. 01)

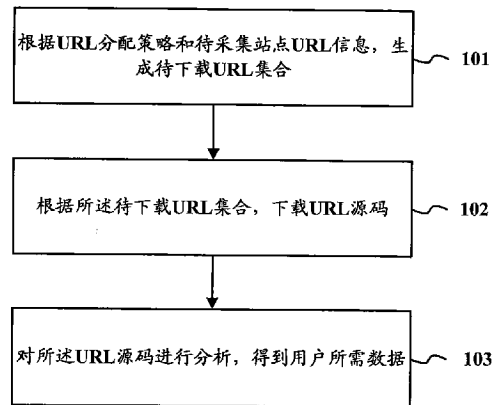
权利要求书 2 页 说明书 7 页 附图 6 页

(54) 发明名称

数据采集方法和数据采集系统

(57) 摘要

本发明提供一种数据采集方法和数据采集系统,所述数据采集方法包括以下步骤:根据 URL 分配策略和待采集站点 URL 信息,生成待下载 URL 集合;根据所述待下载 URL 集合,下载 URL 源码;对所述 URL 源码进行分析,得到用户所需数据。本发明能够自动从互联网上采集用户所需数据。



1. 一种数据采集方法,其特征在于,包括以下步骤:
根据统一资源定位符 URL 分配策略和待采集站点 URL 信息,生成待下载 URL 集合;
根据所述待下载 URL 集合,下载 URL 源码;
对所述 URL 源码进行分析,得到用户所需数据。
2. 根据权利要求 1 所述的数据采集方法,其特征在于,所述 URL 分配策略为:
预设待采集站点的 URL 在所述待下载 URL 集合中所占的条数;和 / 或
预设 URL 级别的 URL 在所述待下载 URL 集合中所占的条数。
3. 根据权利要求 2 所述的数据采集方法,其特征在于,所述对所述 URL 源码进行分析,得到用户所需数据,具体包括:
抽取所述 URL 源码中的情报数据;
从所述情报数据中,获取满足预设过滤规则的第一数据;
采用机器学习算法,将所述第一数据与语料库中的预设语料文件进行匹配,获取所述第一数据的信息价值;
将信息价值超过第一预设阈值的第一数据作为所述用户所需数据。
4. 根据权利要求 3 所述的数据采集方法,其特征在于,所述得到用户所需数据,之后还包括:
从所述用户所需数据中提取属于特定时间范围,且信息价值超过第二预设阈值的第二数据,所述第二预设阈值大于所述第一预设阈值;
将所述第二数据作为新的语料文件,更新所述语料库。
5. 根据权利要求 3 所述的数据采集方法,其特征在于,所述得到用户所需数据,之后还包括:
获取所述用户所需数据对应的 URL,作为有效 URL;
统计所述待采集站点中包含的有效 URL 的条数;
根据所述待采集站点中包含的有效 URL 的条数,更新所述待采集站点的 URL 在所述待下载 URL 集合中所占的条数。
6. 根据权利要求 1 所述的数据采集方法,其特征在于,所述根据所述待下载 URL 集合,下载 URL 源码,之后还包括:
从所述 URL 源码中,获取满足预设条件的 URL,所述预设条件为:所述 URL 不属于特定类型,且为所述待采集站点的 URL,且不属于所述待采集站点的站内 URL 黑名单;
根据所述满足预设条件的 URL,更新所述待采集站点 URL 信息。
7. 根据权利要求 6 所述的数据采集方法,其特征在于,所述从所述 URL 源码中,获取满足预设条件的 URL,之后还包括:
将所述满足预设条件的 URL 存储于内存映射文件中,且采用双层哈希表记录所述满足预设条件的 URL 的存储位置。
8. 根据权利要求 1 至 7 任一项所述的数据采集方法,其特征在于,所述用户所需数据为地理信息数据。
9. 一种数据采集系统,其特征在于,包括:
生成模块,用于根据 URL 分配策略和待采集站点 URL 信息,生成待下载 URL 集合;
下载模块,用于根据所述待下载 URL 集合,下载 URL 源码;

分析模块,用于对所述 URL 源码进行分析,得到用户所需数据。

10. 根据权利要求 9 所述的数据采集系统,其特征在于,所述生成模块、所述下载模块和所述分析模块设置于一台或多台计算机上。

数据采集方法和数据采集系统

技术领域

[0001] 本发明涉及互联网技术领域,尤其涉及一种数据采集方法和数据采集系统。

背景技术

[0002] 在现阶段的电子地图的地理信息数据的生产过程中,均是通过人工的方式进行地理信息数据的采集。在需要采集的地理信息数据规模较大时,通过人工方式采集的效率较低。另外,目前地理信息数据的更新速度也急剧加快,通过人工方式也无法做到地理信息数据的及时更新,难以满足大规模地理信息数据生产的需求。

发明内容

[0003] 有鉴于此,本发明提供一种数据采集方法和数据采集系统,能够自动从互联网上采集用户所需数据。

[0004] 为解决上述问题,本发明提供一种数据采集方法,包括以下步骤:

[0005] 根据 URL 分配策略和待采集站点 URL 信息,生成待下载 URL 集合;

[0006] 根据所述待下载 URL 集合,下载 URL 源码;

[0007] 对所述 URL 源码进行分析,得到用户所需数据。

[0008] 所述 URL 分配策略为:

[0009] 预设待采集站点的 URL 在所述待下载 URL 集合中所占的条数;和/或

[0010] 预设 URL 级别的 URL 在所述待下载 URL 集合中所占的条数。

[0011] 所述对所述 URL 源码进行分析,得到用户所需数据,具体包括:

[0012] 抽取所述 URL 源码中的情报数据;

[0013] 从所述情报数据中,获取满足预设过滤规则的第一数据;

[0014] 采用机器学习算法,将所述第一数据与语料库中的预设语料文件进行匹配,获取所述第一数据的信息价值;

[0015] 将信息价值超过第一预设阈值的第一数据作为所述用户所需数据。

[0016] 所述得到用户所需数据,之后还包括:

[0017] 从所述用户所需数据中提取属于特定时间范围,且信息价值超过第二预设阈值的第二数据,所述第二预设阈值大于所述第一预设阈值;

[0018] 将所述第二数据作为新的语料文件,更新所述语料库。

[0019] 所述得到用户所需数据,之后还包括:

[0020] 获取所述用户所需数据对应的 URL,作为有效 URL;

[0021] 统计所述待采集站点中包含的有效 URL 的条数;

[0022] 根据所述待采集站点中包含的有效 URL 的条数,更新所述待采集站点的 URL 在所述待下载 URL 集合中所占的条数。

[0023] 所述根据所述待下载 URL 集合,下载 URL 源码,之后还包括:

[0024] 从所述 URL 源码中,获取满足预设条件的 URL,所述预设条件为:所述 URL 不属于

特定类型,且为所述待采集站点的 URL,且不属于所述待采集站点的站内 URL 黑名单;

[0025] 根据所述满足预设条件的 URL,更新所述待采集站点 URL 信息。

[0026] 所述从所述 URL 源码中,获取满足预设条件的 URL,之后还包括:

[0027] 将所述满足预设条件的 URL 存储于内存映射文件中,且采用双层哈希表记录所述满足预设条件的 URL 的存储位置。

[0028] 所述用户所需数据为地理信息数据。

[0029] 本发明还提供一种数据采集系统,包括:

[0030] 生成模块,用于根据 URL 分配策略和待采集站点 URL 信息,生成待下载在 URL 集合;

[0031] 下载模块,用于根据所述待下载在 URL 集合,下载 URL 源码;

[0032] 分析模块,用于对所述 URL 源码进行分析,得到用户所需数据。

[0033] 所述生成模块、所述下载模块和所述分析模块设置于一台或多台计算机上。

[0034] 本发明具有以下有益效果:

[0035] 系统的自动化程度高。能够自动从互联网上获取用户所需数据,有效提高了数据采集速度,从而能够快速增加数据采集总量,满足大规模数据生产的需求,相对人工方式采集数据来说,既极大地提高了数据生产效率,又节省了人力成本;

[0036] 系统学习能力强。通过对得到的用户所需数据进行分析,能够自动调整 URL 分配策略和语料库,从而能够满足数据更新速度急剧加快的需求。

[0037] 系统并发性高。系统各模块之间互相对立,能够分布式部署,从而提高了系统的并发能力。

[0038] 系统处理能力高。采用内存映射机制,能够支持大批量的 URL 分析处理,同时采用自动修复映射文件机制,能够有效避免服务器突然中断带来的损失。

附图说明

[0039] 图 1 为本发明实施例的数据采集方法的一流程示意图;

[0040] 图 2 为本发明实施例的数据采集方法的另一流程示意图;

[0041] 图 3 为本发明实施例的数据采集系统的一结构示意图;

[0042] 图 4 为本发明实施例的数据采集系统的另一结构示意图;

[0043] 图 5 为本发明实施例的 URL 下载模块的示意图;

[0044] 图 6 为本发明实施例的信息价值分析模块的示意图;

[0045] 图 7 为本发明实施例的 URL 分析模块的示意图;

[0046] 图 8 为本发明实施例的日志分析模块的示意图;

[0047] 图 9 为本发明实施例的 URL 分配策略模块的示意图。

具体实施方式

[0048] 下面结合附图和实施例,对本发明的具体实施方式作进一步详细描述。

[0049] 如图 1 所示为本发明实施例的数据采集方法的一流程示意图,所述数据采集方法包括以下步骤:

[0050] 步骤 101,根据 URL(Uniform/Universal Resource Locator,统一资源定位符)分

配策略和待采集站点 URL 信息,生成待下载 URL 集合。

[0051] 所谓 URL 也称成网页地址,是因特网(Internet)上标准的资源地址(Address)。所述 URL 分配策略可以为包括:预设待采集站点的 URL 在所述待下载 URL 集合中所占的条数。预设待采集站点的 URL 在所述待下载 URL 集合中所占的条数,也称为预设待采集站点的下载权重,例如,URL 分配策略可以设置为:待采集站点 A 的下载权重为 30%,即每 100 条待下载 URL 集合中属于待采集站点 A 的 URL 占 30 条。所述 URL 分配策略还可以为包括:预设 URL 级别的 URL 在所述待下载 URL 集合中所占的条数。本发明实施例中,可以将 URL 划分为四个级别,其中,级别 1:通常需要快速下载,例如,下载期限为 2 天(即需要在 2 天内下载完成),级别 2:一般为列表页,其下载期限略长于级别 1 的下载期限,例如,下载期限为 5 天,级别 3:一般为详细页,其下载期限略长于级别 2 的下载期限,例如,下载期限为一周,级别 4:其下载期限最长,例如,下载期限为一个月。此时,URL 分配策略可以设置为:级别 1 的 URL 的下载权重为 50%,即每 100 条待下载 URL 集合中级别 1 的 URL 占 50 条。

[0052] 所述待采集站点 URL 信息包括:待采集站点的 URL 和待采集站点的站内 URL 黑名单(即不需要访问的站内 URL)。所述待采集站点的 URL 可以为待采集站点的首页 URL(也称为站点种子),也可以为待采集站点的子页面 URL。

[0053] 在系统初始化时,通常设定 URL 分配策略为空,且将所述待采集站点的首页 URL 作为最初的待下载 URL,生成待下载 URL 集合。另外,为了获取大量的用户所需数据,在系统初始化时,待采集站点的个数通常设置的较多,例如为 1000 个。

[0054] 步骤 102,根据所述待下载 URL 集合,下载 URL 源码;

[0055] 所述待下载 URL 集合可以存储于一“URL 下载队列”中,并按照先进先出的顺序从所述“URL 下载队列”中获取待下载 URL,根据所述待下载 URL 下载 URL 源码。

[0056] 具体的,系统通过域名服务器解析所述待下载 URL,将所述待下载 URL 转换成待 IP 地址,从而根据 IP 地址,下载 URL 源码。

[0057] 步骤 103,对所述 URL 源码进行分析,得到用户所需数据。

[0058] 所述用户所需数据可以为地理信息数据、图书数据、产品数据等。

[0059] 通过上述实施例提供的数据采集方法,能够自动从互联网上获取用户所需数据,有效提高了数据采集速度,从而能够快速增加数据采集总量,满足大规模数据生产的需求,相对人工方式采集数据来说,既极大地提高了数据生产效率,又节省了人力成本。

[0060] 下面对 URL 源码的分析过程进行详细描述。

[0061] 如图 2 所示为本发明实施例的数据采集方法的另一流程示意图,所述数据采集方法包括以下步骤:

[0062] 步骤 201,根据 URL 分配策略和待采集站点 URL 信息,获取待下载 URL 集合;

[0063] 步骤 202,根据所述待下载 URL 集合,下载 URL 源码;

[0064] 步骤 203,抽取所述 URL 源码中的情报数据;

[0065] 所述情报数据可以包括:当前页面的 URL、页面标题、页面内容、页面创建时间、所属站内 URL 目录等。

[0066] 步骤 204,从所述情报数据中,获取满足预设过滤规则的第一数据;

[0067] 所述预设过滤规则可以为时间过滤规则,例如,如果只需获取某一时间段内待采集站点的更新数据,则可以通过时间过滤规则以及情报数据中的页面创建时间,仅获取某

一时间段内创建的情报数据。

[0068] 步骤 205, 采用机器学习算法, 将所述第一数据与语料库中的预设语料文件进行匹配, 获取所述第一数据的信息价值;

[0069] 所述语料库中记载有与用户所需数据相关的多个语料文件, 所述语料文件中记载有用户所需数据的数据格式、字符编码或文本描述等信息。假设用户所需数据为交通类数据, 则所述语料库中则存储有多个交通类数据相关的语料文件, 所述语料文件中记载有交通类数据的特定规则。可以通过机器学习算法, 将所述第一数据与语料库中的预设语料文件进行匹配, 判断所述第一数据与所述预设语料文件的匹配度, 从而得到所述第一数据的信息价值。

[0070] 步骤 206, 将信息价值超过第一预设阈值的所述第一数据作为所述用户所需数据。

[0071] 所述信息价值可以设定为 0 至 100 分, 例如, 通过与预设语料文件匹配, 得到所述第一数据的信息价值为 70 分。假设所述第一预设阈值为 80 分, 则所述第一数据则不是用户所需数据。

[0072] 步骤 207, 从所述用户所需数据中提取属于特定时间范围, 且信息价值超过第二预设阈值的第二数据, 所述第二预设阈值大于所述第一预设阈值;

[0073] 为了使得语料库中的语料文件的信息价值较高, 因此, 设定的第二预设阈值要远大于第一预设阈值, 例如, 所述第一预设阈值可以设定为 80 分, 所述第二预设阈值可以设定为 95 分。

[0074] 步骤 208, 将所述第二数据作为新的语料文件, 更新所述语料库。

[0075] 另外, 还可以对得到的用户所需数据进行分析, 自动调整 URL 分配策略, 从而能够满足数据更新速度急剧加快的需求。此时, 上述方法还包括:

[0076] 步骤 209, 获取所述用户所需数据对应的 URL, 作为有效 URL;

[0077] 步骤 210, 统计所述待采集站点中包含的有效 URL 的条数;

[0078] 步骤 211, 根据所述待采集站点中包含的有效 URL 的条数, 更新所述待采集站点的 URL 在所述待下载 URL 中所占的条数 (即所述待采集站点的下载权重)。

[0079] 另外, 还可以统计待采集站点与数据类别的比重, 调整数据类别在待采集站点中的阈值。

[0080] 以及, 统计预设 URL 目录中包含的有效 URL 的条数, 将不包含有效 URL 或者包含很少个有效 URL 的 URL 目录列入 URL 黑名单。

[0081] 上述实施例中, 在抽取所述 URL 源码中的情报数据时, 还可以从所述 URL 源码中抽取 URL, 从而更新所述待采集站点 URL 信息。因此, 上述方法还包括:

[0082] 步骤 212, 从所述 URL 源码中, 抽取 URL;

[0083] 步骤 213, 从所述抽取的 URL 中, 获取满足预设条件的 URL, 所述预设条件为: 所述 URL 不属于特定类型 (例如: 图片类型 (URL 的后缀为 *.jpg 等)、特定文档类型 (URL 后缀为 *.ppt 等)), 且属于所述待采集站点 (本发明实施例中也将待采集站点称为白名单) 的 URL, 且不属于所述待采集站点的站内 URL 黑名单。

[0084] 步骤 214, 根据所述满足预设条件的 URL, 更新所述待采集站点 URL 信息。

[0085] 另外, 为了防止随着 URL 的增加而使得系统负担增大, 可以将满足预设条件的 URL 存储于内存映射文件中, 通过内存映射机制, 在单个服务器 (4G 内存) 能够支持一亿个 URL

的分析处理。另外,为了方便查找,可以采用双层哈希表(hash)记录所述满足预设条件的URL的存储位置。采用双层哈希表可以确定URL的唯一存储位置,其中,第一层hash为固定值(例如1~500),可用于分布式扩展,第二层可自动扩展(即大小不限)。另外,由于内存映射文件时时变动,为了防止由于服务器突然中断而带来的影响,需对具有一定机制对内存映射文件进行还原。例如,可以每隔一定时间将内存映射文件保存一次,然而由于内存映射文件通常较大(1千万URL的信息大概有2G大小),因此,每次使用拷贝将占用了大量的时间与空间。本发明实施例中,可以仅保存内存hash的信息(最多为500M)为hash文件,通过hash文件,可快速还原内存映射文件,通常情况下,在3分钟之内可还原1千万的URL信息。

[0086] 对应于上述数据采集方法,本发明实施例还提供一种数据采集系统,如图3所示,所述数据采集系统包括:

[0087] 生成模块301,用于根据URL分配策略和待采集站点URL信息,生成待下载URL集合;

[0088] 下载模块302,用于根据所述下下载URL集合,下载URL源码;

[0089] 分析模块303,用于对所述URL源码进行分析,得到用户所需数据。

[0090] 所述生成模块301、所述下载模块302和所述分析模块304可以设置于一台计算机上,当然,为了提高系统的并发能力,所述生成模块301、所述下载模块302和所述分析模块304还可以设置于多台计算机上。

[0091] 下面举例对上述实施例的数据采集系统进行详细说明。

[0092] 如图4所示为本发明实施例的数据采集系统的另一结构示意图,所述数据采集系统包括:WEB管理系统模块、URL下载模块、信息抽取模块、信息价值分析模块、URL分析模块、日志分析模块和URL分配策略模块,上述各模块可以分布在一台或多台计算机上,由于每台计算机的处理能力有限,通常将上述模块设置于多台计算机上,各模块通过指定类型的信息包与管道进行通信。管道是一种通信机制,由计算机的操作系统提供。信息包可以采用轮寻的方式或广播的方式在管道中传送,每个模块只处理(接收和或删除)属于该模块的指定规定类型的信息包,处理后按约定的方式向管道中发送新的信息包。

[0093] 各模块的如下所示:

[0094] 1、WEB管理系统模块:

[0095] WEB管理系统模块在系统初始化时使用,用于输入待采集站点URL信息(例如,待采集站点的首页URL、待采集站点的站内URL黑名单等),得到待下载URL集合,并根据所述待下载URL集合,生成“URL下载包”发送给管道。

[0096] 所述待下载URL集合中包括多个URL,例如1000个,根据所述待下载URL集合生成的“URL下载包”可以为多个,每一“URL下载包”中可以包括若干个待下载URL,例如10个。

[0097] 由于是数据采集的初始化阶段,因此该步骤中URL分配策略设置为空。

[0098] 2、URL下载模块:

[0099] 所述URL下载模块主要用于下载URL,其主要包括以下步骤:

[0100] 1) 在管道中获取“URL下载包”,把待下载URL存到URL下载对列中;所述URL下载对列的容量有限,例如长度为100,即仅能存储100个待下载URL。当URL下载对列中的URL小于某固定值(例如90)时,所述URL下载模块可以主动向管道发送“请求URL包”,以

请求新的待下载 URL。

[0101] 2) 从 URL 下载队列中获取待下载 URL, 把待下载 URL 发送给客户端。

[0102] 如图 5 所示为本发明实施例的 URL 下载模块的示意图, 所述 URL 下载队列可以位于一服务器上, 而具体的 URL 下载任务可以由一个或多个客户端执行, 服务器与客户端之间可以通过 socket (套接字) 通信, 在待下载 URL 数据较大时, 通常采用多个客户端同时执行下载任务。

[0103] 3) 获取客户端下载的 URL 源码, 并生成“URL 源码包”向管道发送。

[0104] 3、信息抽取模块:

[0105] 1) 在管道中获取“URL 源码包”, 对所述“URL 源码包”进行分析, 从 URL 源码中抽取 URL 列表, 生成“URL 列表包”, 所述“URL 列表包”用于待采集站点 URL 信息的更新;

[0106] 2) 从 URL 源码中抽取出情报信息, 生成“信息分析包”;

[0107] 3) 向管道发送所述“URL 列表包”和“信息分析包”。

[0108] 4、信息价值分析模块:

[0109] 所述信息价值分析模块主要用于获取用户所需数据, 其可以包括以下步骤:

[0110] 1) 在管道中获取“信息分析包”, 根据时间过滤规则以及语料库, 对所述

[0111] “信息分析包”中的情报信息进行分析, 得到用户所需数据。如 6 图所示为本发明实施例的信息价值分析模块的示意图。

[0112] 2) 将分析得到的用户所需数据的相关信息 (信息价值、所属类别等) 以及情报信息整合在一起, 生成“日志分析包”, 并向管道发送。

[0113] 5、URL 分析模块:

[0114] 如 7 图所示为本发明实施例的 URL 分析模块的示意图, 所述 URL 分析模块的工作流程主要包括以下步骤:

[0115] 1) 在管道中获取“URL 列表包”, 获取满足预设条件的 URL, 所述满足预设条件的 URL 为不属于特定类型的 URL, 且属于待采集站点的 URL, 且不属于站内 URL 黑名单;

[0116] 2) URL 排重:

[0117] 判断内存映射文件中是否存在重复的 URL, 如果存在重复的 URL, 则判断 URL 的更新时间是否发生变化, 如果 URL 的更新时间发生变化, 则修改内存映射文件中存储的该 URL 的相关信息, 例如: 修改 URL 更新时间、URL 下载时间 (可以将下载时间置为空); 如果 URL 的更新时间未发生变化, 则结束;

[0118] 如果不存在重复的 URL, 则对 URL 的级别进行计算, 得到 URL 的级别信息, 并将 URL 信息插入内存映射文件。

[0119] 6、日志分析模块:

[0120] 在管道中获取“日志分析包”, 将所述“日志分析包”中的数据存入数据库中的两个表格中, 其中第一个表格主要包含的字段为: URL、信息价值、类别 ID、标题、内容、记录插入时间 (即数据插入数据库的时间) 等, 第二个表格主要包含的字段为: URL、信息价值、类别 ID (用户所需数据所属的类别的编号)、站点 ID (用户所需数据对应的待采集站点的编号)、站内目录 ID (用户所需数据对应的待采集站点的站内 URL 目录的编号)、记录插入时间等。

[0121] 如图 8 所示为本发明实施例的日志分析模块的示意图, 所述日志分析模块的工作流程主要包括以下步骤:

[0122] 1) 定期对第一个表格进行分析,提取特定时间范围内的数据,把信息价值度高的信息,推送给对应类别的语料文件,更新语料库,推送方式由机器自动推到和人工审核两种方式。

[0123] 2) 定期对第二个表格进行分析,提取特定时间范围内的数据(将用户所需数据对应的 URL 称为有效 URL);

[0124] A:统计待采集站点中有效 URL 所占的比重,对于比重高的待采集站点,赋予更高的下载权重;

[0125] B:统计待采集站点的站内 URL 目录中有效 URL 所占的比重,把比重低的站内 URL 目录,列入站内 URL 黑名单;此时,为了提高准确度,还可以由人工来审核系统推送的 URL 是否需要列入 URL 黑名单;

[0126] C:统计待采集站点与数据类别的比重,调整数据类别在待采集站点中的阈值。

[0127] 根据上述分析结果,对 URL 分配策略进行更新。

[0128] 7、URL 分配策略模块:

[0129] 如图 9 所示为本发明实施例的 URL 分配策略模块的示意图。

[0130] 1) 当“URL 调度队列”中的 URL 的个数小于一定值时,例如小于 100 时,扫描 URL 分析模块产生的映射文件,以及日志分析模块产生的 URL 分配策略,生成“URL 调度队列”;所述“URL 调度队列”用于存储通过 URL 分配策略从所述内存映射文件中获取的待下载 URL 集合。

[0131] 或者,所述“URL 调度队列”也可以定时(例如每天零点)扫描 URL 分析模块产生的映射文件,以及日志分析模块产生的 URL 分配策略,生成“URL 调度队列”;

[0132] 2) 在管道中获取“请求 URL 包”,从“URL 调度对列”中,向管道发送“URL 下载包”。

[0133] 以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以作出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

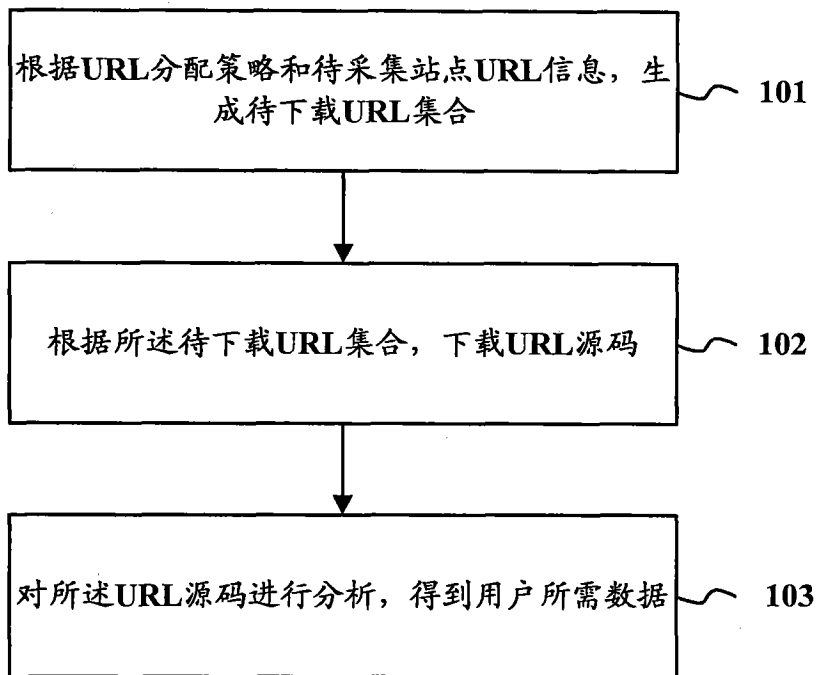


图 1

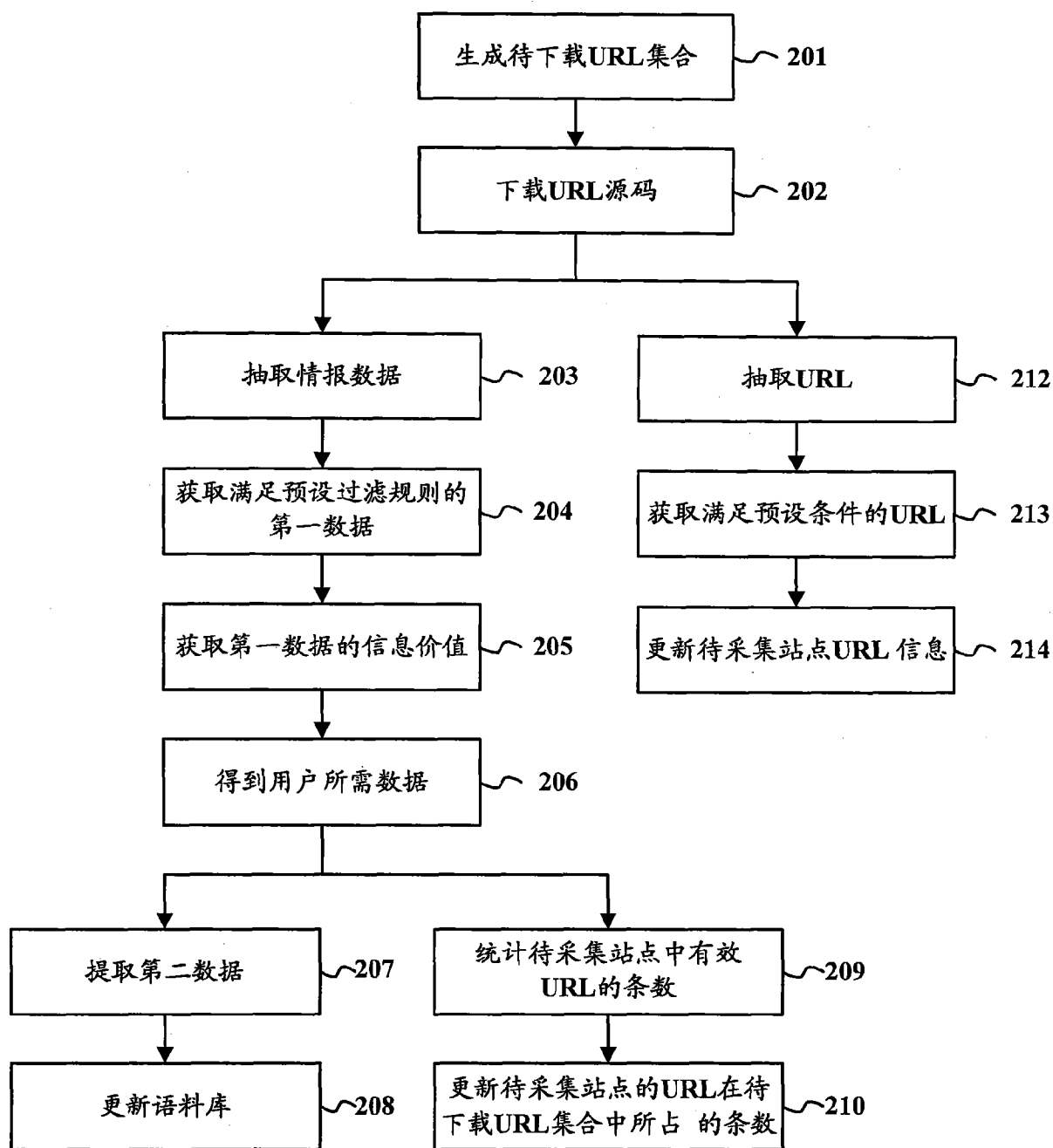


图 2

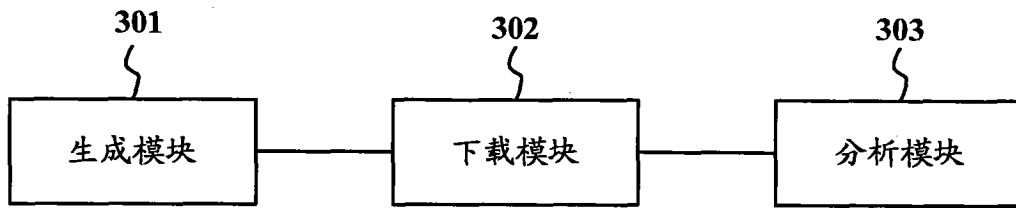


图 3

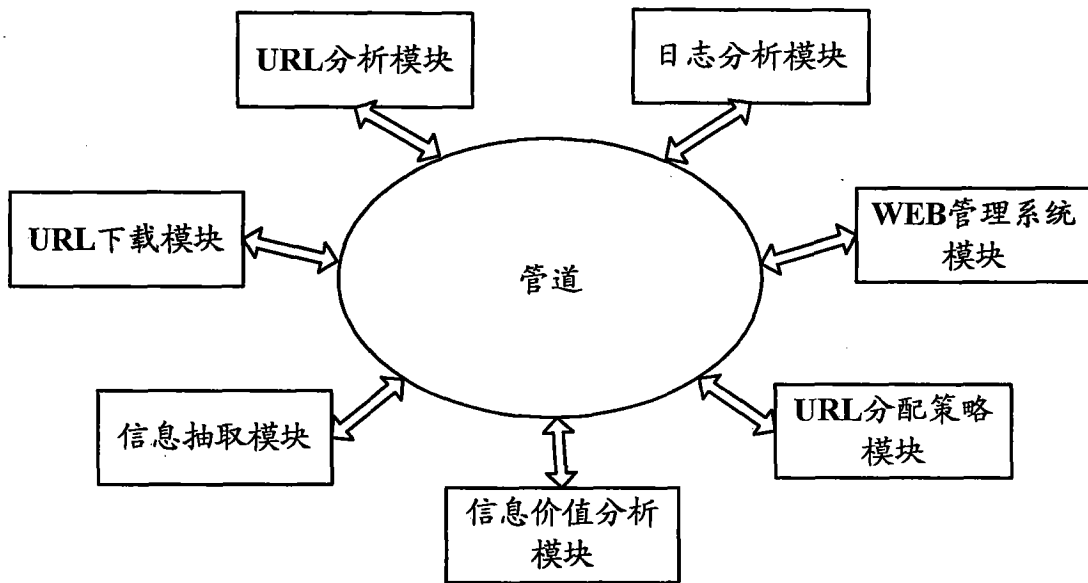


图 4

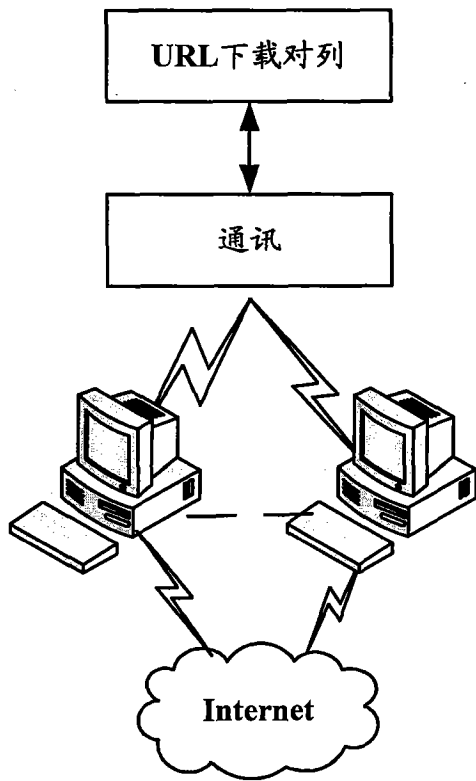


图 5

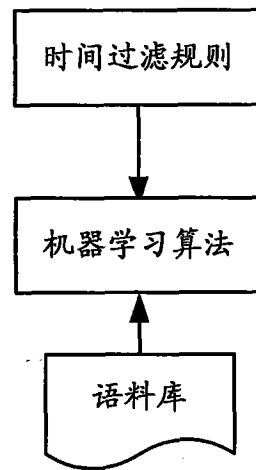


图 6

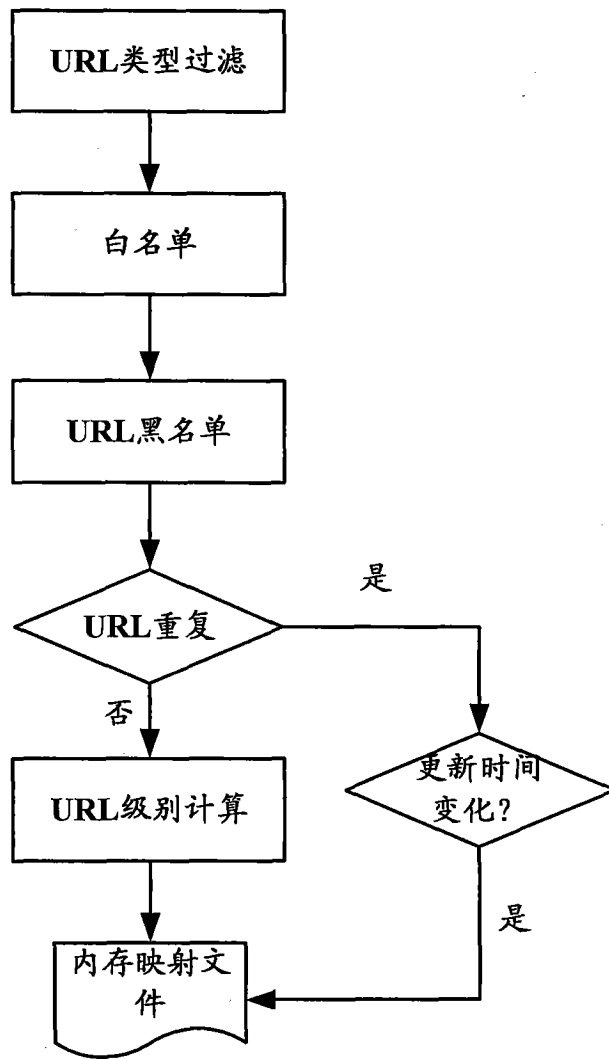


图 7

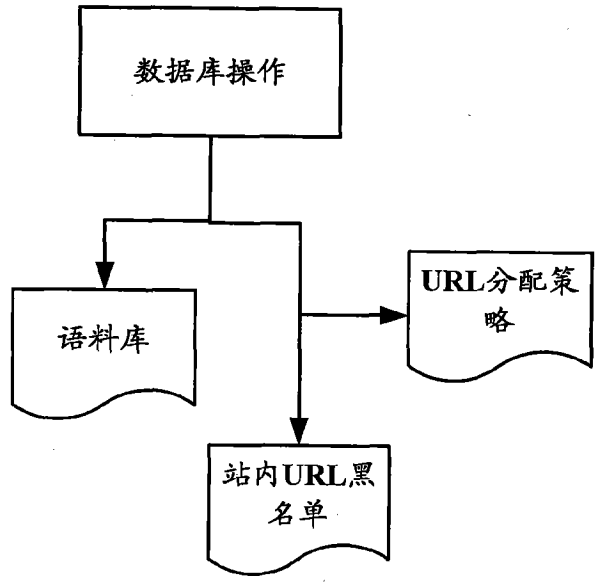


图 8

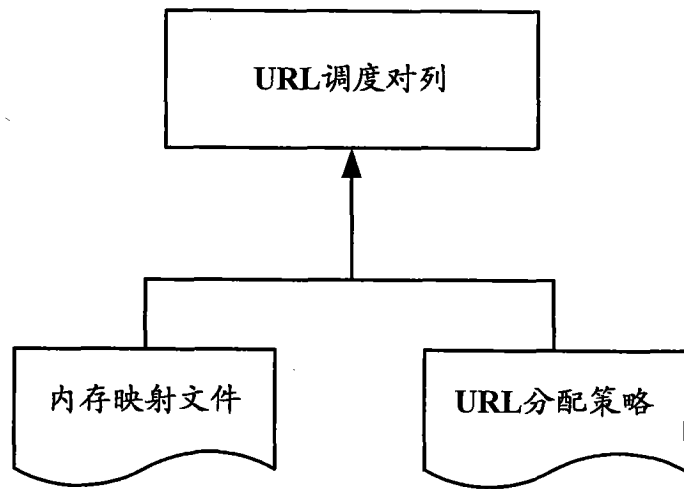


图 9