



(12) 发明专利申请

(10) 申请公布号 CN 102542019 A

(43) 申请公布日 2012. 07. 04

(21) 申请号 201110427621. 4

(22) 申请日 2011. 12. 19

(71) 申请人 北京地拓科技发展有限公司

地址 100084 北京市海淀区农大南路 1 号硅谷亮城 5 号楼 201

(72) 发明人 曹刚 黄兆伟 孙成宝 李团宏
李嵩 计文琦

(74) 专利代理机构 北京润泽恒知识产权代理有限公司 11319

代理人 苏培华

(51) Int. Cl.

G06F 17/30(2006. 01)

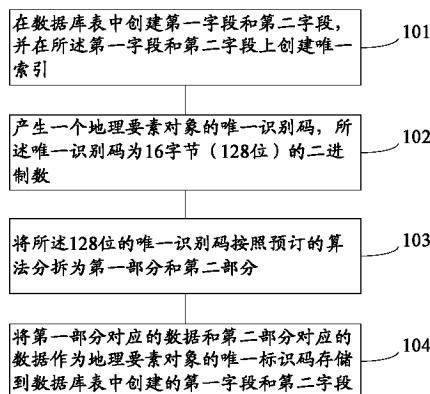
权利要求书 2 页 说明书 5 页 附图 2 页

(54) 发明名称

识别码存储方法及系统、识别码索引方法及系统

(57) 摘要

本发明提供了一种识别码存储方法,包括:在数据库表中创建第一字段和第二字段,并在该两个字段上创建唯一索引;产生一个地理要素对象的唯一识别码,唯一识别码为 16 字节(128 位)的二进制数;将 128 位的唯一识别码按照预订的算法分拆为长度为 64 位,数据类型为长整形的第一部分和第二部分;将第一部分对应的数据和第二部分对应的数据作为地理要素对象的唯一标识码分别存储到数据库表中创建的第一字段和第二字段。本发明还提供了一种实现前述方法的识别码存储系统、一种对前述方法存储的识别码进行索引的识别码索引方法及系统。本发明的识别码存储方法及系统、识别码索引方法及系统,可以降低识别码占用的存储空间,以及提高检索效率。



1. 一种识别码存储方法,其特征在于,包括以下步骤:
在数据库表中创建第一字段和第二字段,并在所述第一字段和第二字段上创建唯一索引;
产生一个地理要素对象的唯一识别码,所述唯一识别码为 16 字节(128 位)的二进制数;
将所述 128 位的唯一识别码按照预订的算法分拆为第一部分和第二部分,所述第一部分和第二部分的数据长度为 64 位,数据类型为长整形;
将所述第一部分对应的数据和第二部分对应的数据作为所述地理要素对象的唯一标识码分别存储到所述数据库表中创建的第一字段和第二字段。
2. 如权利要求 1 所述的识别码存储方法,其特征在于,将所述 128 位的唯一识别码按照预订的算法分拆为第一部分和第二部分包括:
将所述 128 位的唯一识别码的前 64 位分拆为第一部分,后 64 位分拆为第二部分。
3. 如权利要求 1 所述的识别码存储方法,其特征在于,所述方法还包括:
利用预定的型别标识码对所述地理要素对象的唯一识别码进行数据重构,得到新的唯一识别码,该新的唯一识别码携带有所述型别标识码。
4. 如权利要求 3 所述的识别码存储方法,其特征在于,所述数据重构包括:
按照预定的算法将所述型别标识码分别与所述唯一识别码上的变体位和 / 或版本位进行运算得到新的变体和 / 或版本;
将所述新的版本和变体写入所产生的唯一识别码对应的位置。
5. 一种识别码存储系统,其特征在于,包括:
字段创建模块,用于在数据库表中创建第一字段和第二字段,并在所述第一字段和第二字段上创建唯一索引;
识别码产生器,用于产生一个地理要素对象的唯一识别码,所述唯一识别码为 16 字节(128 位)的二进制数;
拆分模块,用于将所述 128 位的唯一识别码按照预订的算法分拆为第一部分和第二部分,所述第一部分和第二部分的数据长度为 64 位,数据类型为长整形;
存储模块,用于将所述第一部分对应的数据和第二部分对应的数据作为所述地理要素对象的唯一标识码分别存储到所述数据库表中创建的第一字段和第二字段。
6. 如权利要求 5 所述的识别码存储系统,其特征在于,所述系统还包括:
数据重构模块,用于利用预定的型别标识码对所述地理要素对象的唯一识别码进行数据重构,得到新的唯一识别码,该新的唯一识别码携带有所述型别标识码。
7. 如权利要求 6 所述的识别码存储系统,其特征在于,所述数据重构模块包括:
运算单元,用于按照预定的算法将所述型别标识码分别与所述唯一识别码上的变体位和 / 或版本位进行运算得到新的变体和 / 或版本;
写入单元,用于将所述新的版本和变体写入所产生的唯一识别码对应的位置。
8. 一种识别码索引方法,其特征在于,包括以下步骤:
获取需要检索的地理要素对象的唯一标识码;
获取所述数据库表中存储唯一标识码时的拆分算法,依据所述拆分算法将所述唯一标识码拆分为第一检索部分和第二检索部分;

将所述第一检索部分和第二检索部分作为索引在数据库表中检索,当所述第一检索部分与数据库表中的第一部分且第二检索部分与数据库表中的第二部分都能匹配时,则确定所述第一部分和第二部分对应的地理要素对象为所述需要检索的地理要素对象。

9. 如权利要求 8 所述的识别码索引方法,其特征在于,所述唯一标识码为利用预定的型别标识码对所述地理要素对象的唯一标识码进行数据重构,得到新的唯一标识码,该新的唯一标识码携带有所述型别标识码。

10. 一种识别码索引系统,其特征在于,包括:

唯一标识码获取模块,用于获取需要检索的地理要素对象的唯一标识码;

拆分模块,用于获取所述数据库表中存储唯一标识码时的拆分算法,依据所述拆分算法将所述唯一标识码拆分为第一检索部分和第二检索部分;

检索模块,用于将所述第一检索部分和第二检索部分作为索引在数据库表中检索,当所述第一检索部分与数据库表中的第一部分且第二检索部分与数据库表中的第二部分都能匹配时,则确定所述第一部分和第二部分对应的地理要素对象为所述需要检索的地理要素对象。

识别码存储方法及系统、识别码索引方法及系统

技术领域

[0001] 本申请涉及计算机数据处理技术领域,特别是涉及一种识别码存储方法及系统。

背景技术

[0002] 唯一识别码 (Universally Unique Identifier, UUID) 是一个软件建构的标准,亦为自由软件基金会 (Open Software Foundation, OSF) 的组织在分布式计算环境 (Distributed Computing Environment, DCE) 领域的一部份。UUID 的目的,是让分布式系统中的所有元素,都能有唯一的辨识信息,而不需要通过中央控制端来做辨识信息的指定。如此一来,每个人都可以创建不与其它人冲突的 UUID。在这样的情况下,就不需考虑数据库创建时的名称重复问题。目前最广泛应用的 GUID,即是微软的 Microsoft's Globally Unique Identifiers (GUIDs),而其他重要的应用,则有 Linuxext2/ext3 文件系统、LUKS 加密分区、GNOME、KDE、MacOSX 等等。另外,也可以在 e2fsprogs 包中的 UUID 库找到实现。

[0003] 一组 UUID,系由一串 16 位组 (亦称 16 字节,或 128 位) 的 16 进位数字所构成,是故 UUID 理论上的总数为 $2^{16 \times 8} = 2^{128}$,约等于 3.4×10^{38} 。也就是说若每秒产生 1 兆个 UUID,要花 100 亿年才会将所有 UUID 用完。使用 UUID 的好处在分布式的软件系统中就能体现出来,它能保证每个节点所生成的标识都不会重复,并且随着 WEB 服务等整合技术的发展,UUID 的优势将更加明显。通用唯一标识符还可以用来指向大多数的可能的物体。

[0004] 用 UUID 来唯一标识一个要素类、要素或者要素上的一个属性方面,是已知的公认的做法。就是把每一个实体序列化时都用一个 UUID 来标识,同时用一个约定的方式规定为具体的类型 (上述所说的要素类、要素、要素属性等),然后再根据不同的类型指定需要序列化的数据。但是现有的 UUID 在使用的过程中比较占存储空间,简单的和一个整型相比存储一个 UUID 要花费更多的空间,二是在网络资源有限的情况下额外的增加了传输的负担。

发明内容

[0005] 本申请提供一种识别码存储方法及系统、识别码索引方法及系统,以解决识别码占用过多存储空间,以及检索效率低下的问题。

[0006] 为了解决上述问题,本申请公开了一种识别码存储方法,包括以下步骤:

[0007] 在数据库表中创建第一字段和第二字段,并在所述第一字段和第二字段上创建唯一索引;

[0008] 产生一个地理要素对象的唯一识别码,所述唯一识别码为 16 字节 (128 位) 的二进制数;

[0009] 将所述 128 位的唯一识别码按照预订的算法分拆为第一部分和第二部分,所述第一部分和第二部分的数据长度为 64 位,数据类型为长整形;

[0010] 将所述第一部分对应的数据和第二部分对应的数据作为所述地理要素对象的唯一标识码分别存储到所述数据库表中创建的第一字段和第二字段。

[0011] 进一步地,将所述 128 位的唯一识别码按照预订的算法分拆为第一部分和第二部

分包括：

[0012] 将所述 128 的唯一识别码的前 64 位分拆为第一部分,后 64 位分拆为第二部分。

[0013] 进一步地,所述方法还包括：

[0014] 利用预定的型别标识码对所述地理要素对象的唯一识别码进行数据重构,得到新的唯一识别码,该新的唯一识别码携带有所述型别标识码。

[0015] 进一步地,所述数据重构包括：

[0016] 按照预定的算法将所述型别标识码分别与所述唯一识别码上的变体位和 / 或版本位进行运算得到新的变体和 / 或版本；

[0017] 将所述新的版本和变体写入所产生的唯一识别码对应的位置。

[0018] 为了解决上述问题,本申请还公开了一种识别码存储系统,包括：

[0019] 字段创建模块,用于在数据库表中创建第一字段和第二字段,并在所述第一字段和第二字段上创建唯一索引；

[0020] 识别码产生器,用于产生一个地理要素对象的唯一识别码,所述唯一识别码为 16 字节 (128 位) 的二进制数；

[0021] 拆分模块,用于将所述 128 位的唯一识别码按照预订的算法分拆为第一部分和第二部分,所述第一部分和第二部分的数据长度为 64 位,数据类型为长整形；

[0022] 存储模块,用于将所述第一部分对应的数据和第二部分对应的数据作为所述地理要素对象的唯一标识码分别存储到所述数据库表中创建的第一字段和第二字段。

[0023] 进一步地,所述系统还包括：

[0024] 数据重构模块,用于利用预定的型别标识码对所述地理要素对象的唯一识别码进行数据重构,得到新的唯一识别码,该新的唯一识别码携带有所述型别标识码。

[0025] 进一步地,所述数据重构模块包括：

[0026] 运算单元,用于按照预定的算法将所述型别标识码分别与所述唯一识别码上的变体位和 / 或版本位进行运算得到新的变体和 / 或版本；

[0027] 写入单元,用于将所述新的版本和变体写入所产生的唯一识别码对应的位置。

[0028] 为了解决上述问题,本申请还公开了一种识别码索引方法,包括以下步骤：

[0029] 获取需要检索的地理要素对象的唯一标识码；

[0030] 获取所述数据库表中存储唯一标识码时的拆分算法,依据所述拆分算法将所述唯一标识码拆分为第一检索部分和第二检索部分；

[0031] 将所述第一检索部分和第二检索部分作为索引在数据库表中检索,当所述第一检索部分与数据库表中的第一部分且第二检索部分与数据库表中的第二部分都能匹配时,则确定所述第一部分和第二部分对应的地理要素对象为所述需要检索的地理要素对象。

[0032] 进一步地,所述唯一标识码为利用预定的型别标识码对所述地理要素对象的唯一识别码进行数据重构,得到新的唯一识别码,该新的唯一识别码携带有所述型别标识码。

[0033] 为了解决上述问题,本申请还公开了一种识别码索引系统,包括：

[0034] 唯一标识码获取模块,用于获取需要检索的地理要素对象的唯一标识码；

[0035] 拆分模块,用于获取所述数据库表中存储唯一标识码时的拆分算法,依据所述拆分算法将所述唯一标识码拆分为第一检索部分和第二检索部分；

[0036] 检索模块,用于将所述第一检索部分和第二检索部分作为索引在数据库表中检

索,当所述第一检索部分与数据库表中的第一部分且第二检索部分与数据库表中的第二部分都能匹配时,则确定所述第一部分和第二部分对应的地理要素对象为所述需要检索的地理要素对象。

[0037] 与现有技术相比,本申请包括以下优点:

[0038] 本申请中通过唯一标识码来作为地理要素对象的标识,即 128 位的数据长度,可以通过索引的方法来实现检索。在对唯一标识码进行存储的时候采用拆分成两个字段,每一个字段采用 64 位的长整形的方式进行存储,从而可以减少占用存储空间。

[0039] 在检索时,可以分别对两个字段同时进行检索,因为拆分后的字段数据长度小于拆分前,从而可以减少检索的数据量,提高检索效率。拆分后的字段分成两个 64 位的数据,并采用长整形结构存储也可以帮助提高检索效率。特别是在海量的地理要素中检索的时候,此种方式可以提高检索效率,快速响应结果。

[0040] 另外,利用型别标识码对产生的唯一识别码进行数据重构,使得重构后的唯一识别码携带型别标识信息,从而实现对象或物体类别的识别,可以在不破坏唯一识别码的特性基础上注入了新的数据变量,减少对物体的描述数据资源量和网络传输量,有力地支撑了在地理环境中海量数据的标识、存储、传输的一些问题。使用该识别码就能快速识别出该识别码所代表的物体类型,并加快了信息检索的速度。

[0041] 当然,实施本申请的任一产品不一定需要同时达到以上所述的所有优点。

附图说明

[0042] 图 1 是本申请的标识码存储方法实施例一的流程图;

[0043] 图 2 是本申请的标识码存储系统实施例一的结构示意图;

[0044] 图 3 是本申请的标识码索引方法实施例一的流程图;

[0045] 图 4 是本申请的标识码索引系统实施例一的结构示意图。

具体实施方式

[0046] 为使本申请的上述目的、特征和优点能够更加明显易懂,下面结合附图和具体实施方式对本申请作进一步详细的说明。

[0047] 参照图 1,示出本申请的一种标识码存储方法实施例一,包括以下步骤:

[0048] 步骤 101,在数据库表中创建第一字段和第二字段,并在所述第一字段和第二字段上创建唯一索引。

[0049] 步骤 102,产生一个地理要素对象的唯一识别码,所述唯一识别码为 16 字节(128 位)的二进制数。

[0050] 该唯一识别码通过通常的 UUID 发生器产生,每一个唯一识别码作为一个地理要素对象的唯一标识码。

[0051] 步骤 103,将所述 128 位的唯一识别码按照预订的算法分拆为第一部分和第二部分,所述第一部分和第二部分的数据长度为 64 位,数据类型为长整形。

[0052] 所述 128 位的唯一识别码的分拆可以根据需要设定,可以预先在系统中存储分拆的算法,在需要分拆时,则直接获取对应的算法进行分拆。优选地,为了减少拆分时的计算量,可以直接将 128 的唯一识别码的前 64 位分拆为第一部分,后 64 位分拆为第二部分。可

以理解,还可以通过其他方式对 128 的唯一识别码进行分拆,只要能将该唯一识别码分成两个 64 位的数据,并且这两个 64 位的数据能够唯一还原为拆分前的 128 位数据即可,本申请对此并不限制。

[0053] 第一部分和第二部分的数据类型为长整形。因为,长整型的数据长度也是 64 位,能够满足拆分后的单个数据的最大要求,可以避免数据的溢出,同时可以减少对存储空间的占用。

[0054] 步骤 104,将所述第一部分对应的数据和第二部分对应的数据作为所述地理要素对象的唯一标识码分别存储到所述数据库表中创建的第一字段和第二字段。

[0055] 优选地,在步骤 102 和步骤 103 之间还可以包括以下步骤:

[0056] 利用预定的型别标识码对所述地理要素对象的唯一识别码进行数据重构,得到新的唯一识别码,该新的唯一识别码携带有所述型别标识码。

[0057] 具体过程如下:按照预定的算法将所述型别标识码分别与所述唯一识别码上的变体位和 / 或版本位进行运算得到新的变体和 / 或版本;将所述新的版本和变体写入所产生的唯一识别码对应的位置。

[0058] 其中,型别标识码优选采用 8 位二进制数,将该型别标识码拆分为高 4 位和低 4 位,然后分别与唯一识别码上的变体和版本位进行运算得到新的变体和版本位。采用的运算方法,如,将 8 位的版本和型别标识码的高 4 位进行“按位与”或“按位或”的位运算,得到一个新的 8 位的版本,同理也可对变体做类似的处理。

[0059] 利用型别标识码对产生的唯一识别码进行数据重构,使得重构后的唯一识别码携带型别标识信息,从而实现对象或物体类别的识别,可以在不破坏唯一识别码的特性基础上注入了新的数据变量,减少对物体的描述数据资源量和网络传输量,有力地支撑了在地理环境中海量数据的标识、存储、传输的一些问题。使用该识别码就能快速识别出该识别码所代表的物体类型,并加快了信息检索的速度。

[0060] 参照图 2,示出本申请的一种识别码存储系统实施例一,包括字段创建模块 10、识别码产生器 20、拆分模块 30 和存储模块 40。

[0061] 字段创建模块 10,用于在数据库表中创建第一字段和第二字段,并在所述第一字段和第二字段上创建唯一索引。

[0062] 识别码产生器 20,用于产生一个地理要素对象的唯一识别码,所述唯一识别码为 16 字节(128 位)的二进制数。

[0063] 拆分模块 30,用于将所述 128 位的唯一识别码按照预订的算法分拆为第一部分和第二部分,所述第一部分和第二部分的数据长度为 64 位,数据类型为长整形。优选地,拆分模块 30 采用将所述 128 的唯一识别码的前 64 位分拆为第一部分,后 64 位分拆为第二部分来对唯一识别码进行分拆。

[0064] 存储模块 40,用于将所述第一部分对应的数据和第二部分对应的数据作为所述地理要素对象的唯一标识码分别存储到所述数据库表中创建的第一字段和第二字段。

[0065] 优选地,该系统还包括数据重构模块,用于利用预定的型别标识码对所述地理要素对象的唯一识别码进行数据重构,得到新的唯一识别码,该新的唯一识别码携带有所述型别标识码。其中,数据重构模块包括运算单元和写入单元。运算单元,用于按照预定的算法将所述型别标识码分别与所述唯一识别码上的变体位和 / 或版本位进行运算得到新的

变体和 / 或版本。写入单元,用于将所述新的版本和变体写入所产生的唯一识别码对应的位置。

[0066] 参照图 3,示出本申请的识别码索引方法实施例一,包括以下步骤:

[0067] 步骤 301,获取需要检索的地理要素对象的唯一标识码。

[0068] 其中,若唯一标识码在产生之后采用如前识别码存储方法中描述的方法进行过数据重构,此处获取的唯一标识码则是经过数据重构后得到的唯一标识码,若没有进行过数据重构,则是产生器直接产生得到的唯一标识码。

[0069] 步骤 302,获取所述数据库表中存储唯一标识码时的拆分算法,依据所述拆分算法将所述唯一标识码拆分为第一检索部分和第二检索部分。

[0070] 步骤 303,将所述第一检索部分和第二检索部分作为索引在数据库表中检索,当所述第一检索部分与数据库表中的第一部分且第二检索部分与数据库表中的第二部分都能匹配时,则确定所述第一部分和第二部分对应的地理要素对象为所述需要检索的地理要素对象。

[0071] 参照图 4,示出本申请的识别码索引系统实施例一,包括唯一标识码获取模块 70、拆分模块 80 和检索模块 90。

[0072] 唯一标识码获取模块 70,用于获取需要检索的地理要素对象的唯一标识码。

[0073] 拆分模块 80,用于获取所述数据库表中存储唯一标识码时的拆分算法,依据所述拆分算法将所述唯一标识码拆分为第一检索部分和第二检索部分。

[0074] 检索模块 90,用于将所述第一检索部分和第二检索部分作为索引在数据库表中检索,当所述第一检索部分与数据库表中的第一部分且第二检索部分与数据库表中的第二部分都能匹配时,则确定所述第一部分和第二部分对应的地理要素对象为所述需要检索的地理要素对象。

[0075] 本申请中通过唯一标识码来作为地理要素对象的标识,即 128 位的数据长度,可以通过索引的方法来实现检索。在对唯一标识码进行存储的时候采用拆分成两个字段,每一个字段采用 64 位的长整形的方式进行存储,从而可以减少占用存储空间。

[0076] 在检索时,可以分别对两个字段同时进行检索,因为拆分后的字段数据长度小于拆分前,从而可以减少检索的数据量,提高检索效率。拆分后的字段分成两个 64 位的数据,并采用长整形结构存储也可以帮助提高检索效率。特别是在海量的地理要素中检索的时候,此种方式可以提高检索效率,快速响应结果。

[0077] 本说明书中的各个实施例均采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似的部分互相参见即可。对于系统实施例而言,由于其与方法实施例基本相似,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0078] 以上对本申请所提供的识别码存储方法及系统、识别码索引方法及系统进行了详细介绍,本文中应用了具体个例对本申请的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本申请的方法及其核心思想;同时,对于本领域的一般技术人员,依据本申请的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本申请的限制。

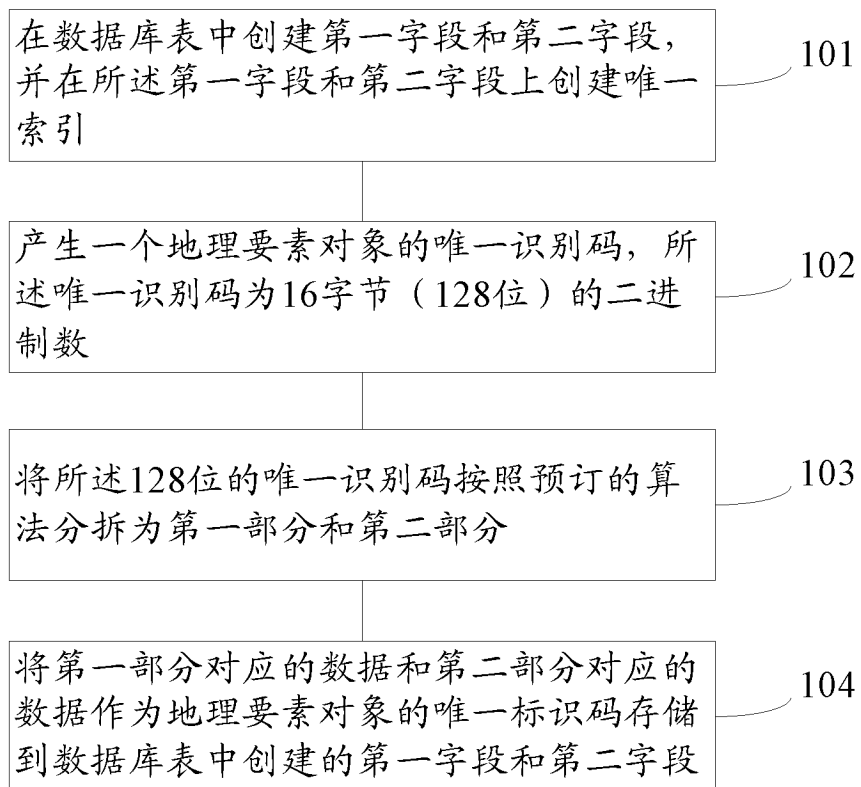


图 1

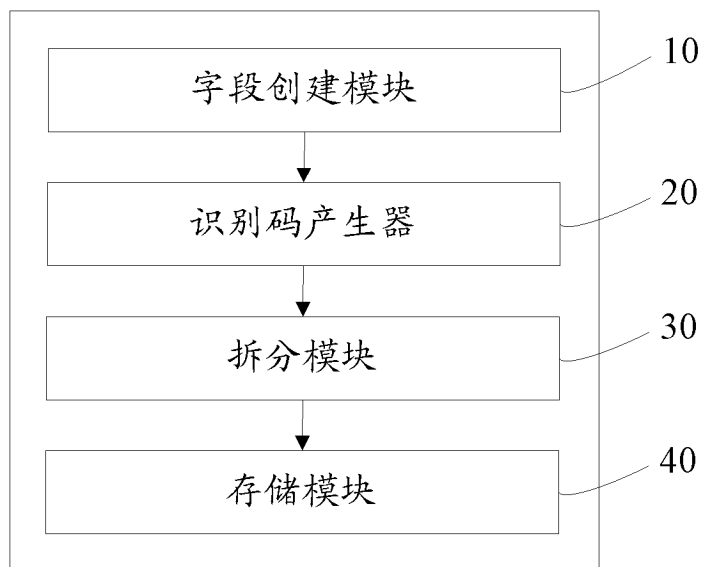


图 2

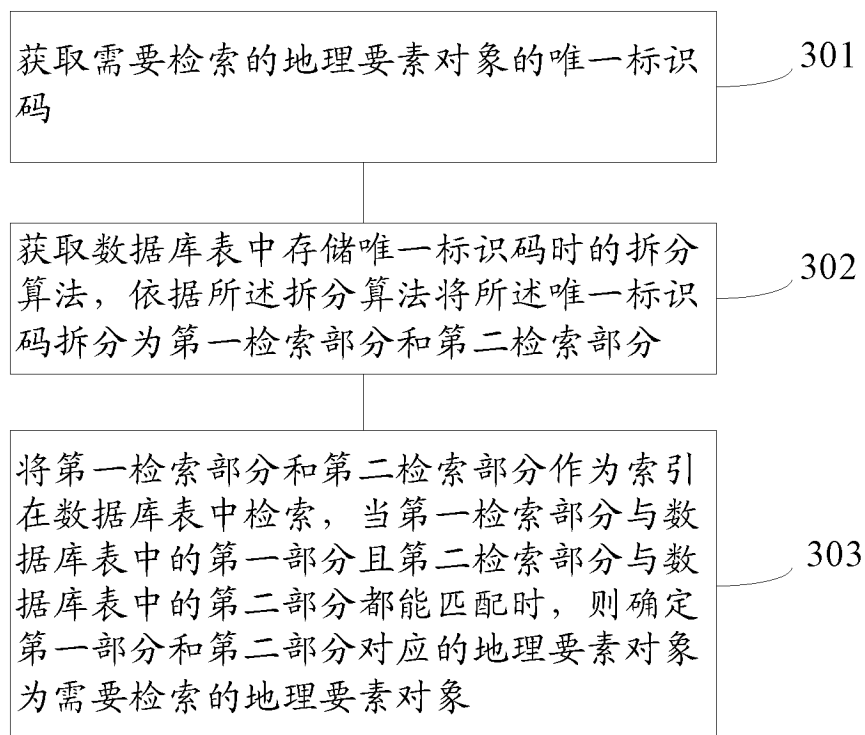


图 3

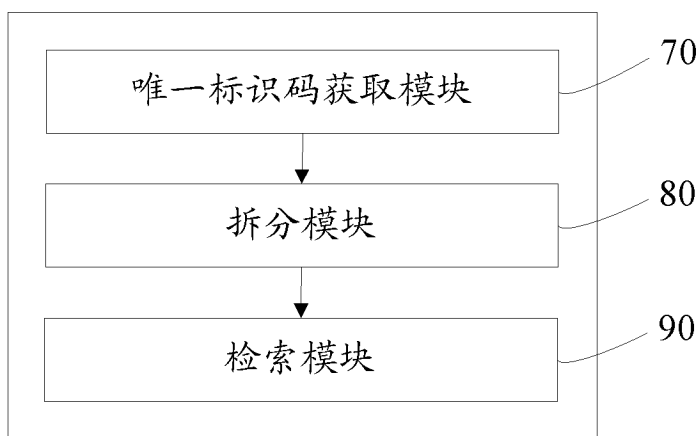


图 4