(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2017 年 3 月 9 日 (09.03.2017)

WIPO | PCT

(10) 国际公布号
WO 2017/036348 A1

(51) 国际专利分类号:
G06F 17/30 (2006.01)

(21) 国际申请号: PCT/CN2016/096790

(22) 国际申请日: 2016 年 8 月 25 日 (25.08.2016)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201510561440.9 2015 年 9 月 6 日 (06.09.2015) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 英属开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人; 及

(71) 申请人 (仅对美国): 魏强 (WEI, Qiang) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京国昊天诚知识产权代理有限公司 (CO-HORIZON INTELLECTUAL PROPERTY INC.); 中国北京市朝阳区小关北里甲 2 号渔阳置业大厦 B 座 605, Beijing 100029 (CN)。

(81) 指定国 (除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明，要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

(54) Title: METHOD AND DEVICE FOR COMPRESSING AND DECOMPRESSING EXTENSIBLE MARKUP LANGUAGE DOCUMENT

(54) 发明名称 : 一种可扩展标记语言 XML 文档的压缩、解压方法和装置

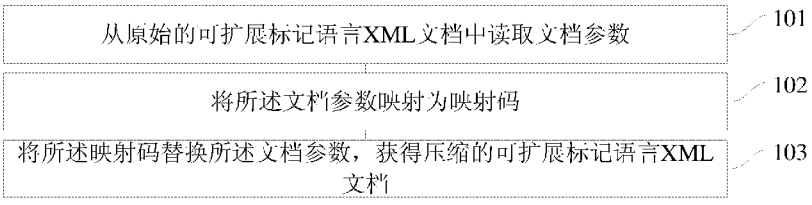


图 1

101 READ DOCUMENT PARAMETER IN SOURCE XML DOCUMENT
102 MAP DOCUMENT PARAMETER INTO MAPPING CODE
103 REPLACE DOCUMENT PARAMETER WITH MAPPING CODE TO OBTAIN COMPRESSED XML DOCUMENT

(57) Abstract: A method and device for compressing and decompressing an extensible markup language (XML) document. The compressing method comprises: reading a document parameter in a source XML document (101); mapping the document parameter into a mapping code (102); and replacing the document parameter with the mapping code to obtain a compressed XML document (103). Since the document parameter has duplications in many situations, the replacement can greatly reduce the volume of data stored in the XML document and increase the compression efficiency. Also, the mapping is easy to operate and thus reduces the compression time.

(57) 摘要: 一种可扩展标记语言 XML 文档的压缩、解压方法和装置，该压缩方法包括：从原始的可扩展标记语言 XML 文档中读取文档参数（101）；将所述文档参数映射为映射码（102）；将所述映射码替换所述文档参数，获得压缩的可扩展标记语言 XML 文档（103）。由于文档参数很多情况下存在重复，因此，替换之后可以大大减少 XML 文档存储的数据量，提高了压缩效率，并且，映射操作简单，减少了压缩的耗时。

WO 2017/036348 A1

一种可扩展标记语言XML文档的压缩、解压方法和装置

技术领域

本申请涉及计算机处理的技术领域，特别是涉及一种可扩展标记语言XML文档的压缩方法、一种可扩展标记语言XML文档的解压方法、一种可扩展标记语言XML文档的压缩装置和一种可扩展标记语言XML文档的解压装置。

背景技术

XML（Extensible Markup Language，可扩展标记语言）文档可以对数据进行结构化处理，从而能够在部门、客户和供应商之间进行交换，实现动态内容生成，企业集成和应用开发。

XML文档可以使用户能够更准确的搜索，更方便的传送应用组件，更好的描述一些事物，如电子商务交易等。

为方便传输或存储，通常对XML文档进行压缩。

由于XML文档是一个文本文档，通常压缩文本文档的方法主要是GZIP压缩，GZIP对于需要压缩的文档，首先使用LZ77算法的一个变种进行压缩，对得到的结果再使用Huffman编码的方法进行压缩。

可见，GZIP压缩是一个通用的文本压缩方法，因此普遍性高，也就是任意一个文本文档都可以使用GZIP压缩，但是，这种压缩方法压缩操作复杂，耗时较高，应用在XML文档压缩时，压缩后的XML文档的体积依然较大，压缩效率较低。

由于XML文档的体积较大，因此，在传输或存储压缩后的XML文档时，会消耗较大的带宽或占用较多的存储空间。

在解压缩时，需要对Huffman编码和LZ77算法进行反向解压缩，解压操作复杂，耗时也较高。

发明内容

鉴于上述问题，提出了本申请实施例以便提供一种克服上述问题或者至

少部分地解决上述问题的一种针对可扩展标记语言 XML 文档的压缩方法、一种针对可扩展标记语言 XML 文档的解压方法和相应的一种针对可扩展标记语言 XML 文档的压缩装置、一种针对可扩展标记语言 XML 文档的解压装置。

为了解决上述问题，本申请实施例公开了一种可扩展标记语言 XML 文档的压缩方法，包括:

从原始的可扩展标记语言 XML 文档中读取文档参数;

将所述文档参数映射为映射码;

将所述映射码替换所述文档参数，获得压缩的可扩展标记语言 XML 文档。

可选地，还包括:

将所述文档参数与所述映射码之间的映射关系嵌入所述可扩展标记语言 XML 文档中。

可选地，还包括:

传输和/或存储压缩的可扩展标记语言 XML 文档。

可选地，所述文档参数包括元素和/或属性。

可选地，所述将所述文档参数映射为映射码的步骤包括:

对所述文档参数进行去重处理;

将去重处理之后的文档参数映射为唯一的映射码，所述映射码的字符串长度小于或等于所述文档参数的字符串长度。

可选地，所述将去重处理之后的文档参数映射为唯一的映射码的步骤包括:

从去重处理之后的文档参数提取候选字符串;

判断所述候选字符串是否与已映射的映射码相同;

当不相同时，确认所述候选字符串为所述文档参数的映射码;

当相同时，从所述文档参数中提取包含所述候选字符串的目标字符串，作为新的候选字符串，返回执行所述判断所述候选字符串是否与已映射的映

射码相同的步骤。

可选地，所述将去重处理之后的文档参数映射为唯一的映射码的步骤还包括:

按照字符串长度和/或字符顺序对去重处理之后的文档参数进行排序。

本申请实施例还公开了一种可扩展标记语言 XML 文档的解压方法，包括:

获取压缩的可扩展标记语言 XML 文档，压缩的可扩展标记语言 XML 文档中包括映射码;

查找压缩的可扩展标记语言 XML 文档的、映射码与文档参数之间的映射关系;

按照所述映射关系将所述文档参数替换所述映射码，获得原始的可扩展标记语言 XML 文档。

可选地，还包括:

当所述映射关系嵌在压缩的可扩展标记语言 XML 文档中时，删除所述映射关系。

可选地，所述获取压缩的可扩展标记语言 XML 文档的步骤包括:

读取在先存储的压缩的可扩展标记语言 XML 文档;

或者，

接收传输到来的压缩的可扩展标记语言 XML 文档。

本申请实施例还公开了一种可扩展标记语言 XML 文档的压缩装置，包括:

文档参数读取模块，用于从原始的可扩展标记语言 XML 文档中读取文档参数;

映射模块，用于将所述文档参数映射为映射码;

文档参数替换模块，用于将所述映射码替换所述文档参数，获得压缩的可扩展标记语言 XML 文档。

可选地，还包括:

映射关系嵌入模块，用于将所述文档参数与所述映射码之间的映射关系嵌入所述可扩展标记语言 XML 文档中。

可选地，还包括:

传输模块，用于传输压缩的可扩展标记语言 XML 文档;

和/或，

存储模块，用于存储压缩的可扩展标记语言 XML 文档。

可选地，所述文档参数包括元素和/或属性。

可选地，所述映射模块包括:

去重子模块，用于对所述文档参数进行去重处理;

去重映射子模块，用于将去重处理之后的文档参数映射为唯一的映射码，所述映射码的字符串长度小于或等于所述文档参数的字符串长度。

可选地，所述去重映射子模块包括:

候选字符串提取单元，用于从去重处理之后的文档参数提取候选字符串;

映射码判断单元，用于判断所述候选字符串是否与已映射的映射码相同；当不相同时，调用映射码确认单元，当相同时，调用目标字符串提取单元，返回调用映射码判断单元;

映射码确认单元，用于确认所述候选字符串为所述文档参数的映射码;

目标字符串提取单元，用于从所述文档参数中提取包含所述候选字符串的目标字符串，作为新的候选字符串。

可选地，所述去重映射子模块还包括:

排序单元，用于按照字符串长度和/或字符顺序对去重处理之后的文档参数进行排序。

本申请实施例还公开了一种可扩展标记语言 XML 文档的解压装置，包括:

XML 文档获取模块，用于获取压缩的可扩展标记语言 XML 文档，压缩

的可扩展标记语言 XML 文档中包括映射码;

映射关系查找模块，用于查找压缩的可扩展标记语言 XML 文档的、映射码与文档参数之间的映射关系;

映射码替换模块，用于按照所述映射关系将所述文档参数替换所述映射码，获得原始的可扩展标记语言 XML 文档。

可选地，还包括:

映射关系删除模块，用于在所述映射关系嵌在压缩的可扩展标记语言 XML 文档中时，删除所述映射关系。

可选地，所述 XML 文档获取模块包括:

XML 文档读取子模块，用于读取在先存储的压缩的可扩展标记语言 XML 文档;

或者，

XML 文档接收子模块，用于接收传输到来的压缩的可扩展标记语言 XML 文档。

本申请实施例包括以下优点:

本申请实施例通过对原始的 XML 文档的文档参数进行映射，将映射码替换文档参数，实现压缩，由于文档参数很多情况下存在重复，因此，替换之后可以大大减少 XML 文档存储的数据量，提高了压缩效率，并且，映射操作简单，减少了压缩的耗时。

进而，在传输、存储压缩的 XML 文档时，可以减少带宽的消耗，或者，较少存储空间的占用。

本申请实施例对文档参数进行去重处理之后映射为唯一的映射码，文档参数的唯一性保证了映射码的唯一性，从而保证两者可以相互转换，保证了压缩之后恢复操作的准确性。

本申请实施例依次增大候选字符串的长度进行映射检测，在保证映射码的唯一性的同时，减少了映射码的长度。

本申请实施例通过映射关系，将原始的文档参数替换映射码，映射操作

简单，减少了解压的耗时。

附图说明

图 1 是本申请的一种可扩展标记语言 XML 文档的压缩方法实施例的步骤流程图；

图 2 是本申请实施例的一种 XML 文档的结构示例图；

图 3 是本申请的一种可扩展标记语言 XML 文档的解压方法实施例的步骤流程图；

图 4 是本申请的一种可扩展标记语言 XML 文档的压缩装置实施例的结构框图；

图 5 是本申请的一种可扩展标记语言 XML 文档的解压装置实施例的结构框图。

具体实施方式

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。

参照图 1，示出了本申请的一种可扩展标记语言 XML 文档的压缩方法实施例的步骤流程图，具体可以包括如下步骤：

步骤 101，从原始的可扩展标记语言 XML 文档中读取文档参数；

在本申请实施例中，针对 XML 文档的特性，设计了一套 XML 文档通用的压缩方法，可以较大幅度的压缩 XML 文档。

需要说明的是，XML 文档是一个文本文档。

以下是一本书的 XML 文档的示例：

```
<bookstore>
<book category="COOKING">
  <title lang="en">Everyday Italian</title>
  <author>Giada De Laurentiis</author>
  <year>2005</year>
```

```
    <price>30.00</price>
  </book>
  <book category="CHILDREN">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
  <book category="WEB">
    <title lang="en">Learning XML</title>
    <author>Erik T. Ray</author>
    <year>2003</year>
    <price>39.95</price>
  </book>
  </bookstore>
```

如图 2 所示，XML 文档的文档参数包括元素，其可以形成了一棵文档树，这棵树从根部开始，并扩展到树的最底端。

其中，父、子以及同胞等术语用于描述元素之间的关系。

具体而言，该元素包括根元素（如 book store）是所有其他元素（如 book、title、author、year、price）的父元素，相对而言，其他元素为该根元素的子元素。

相同层级上的子元素成为同胞（兄弟或姐妹，如 title、author、year、price），所有元素均可拥有文本（如 Harry Potter、J K.Rowling、2005、29.99）和属性（如 lang、category）。

在实际应用中，若应用在 java 语言编程的计算机中，则可以使用自带的 DOM 和 SAX 技术可以解析 XML 文档，读取元素、属性等文档参数。

步骤 102，将所述文档参数映射为映射码；

在 XML 文档里，元素、属性等文档参数用于描述数据，而它们是可以

自定义的，通常为了便于理解，都是使用相应的英文。

例如，用 book 元素代表书，用 title 元素代表标题等等，一般无特别的限制，这也是 XML 文档的特点之一。

但对于计算机来说，计算机不关心元素、属性等文档参数的具体意义，在传输、存储等过程中可以将元素、属性等文档参数用其他更简单的映射码来表示，缩小元素、属性等文档参数的字符串数量，减小传输量、存储量等等，而在使用的时候，再恢复回来供计算机使用即可。

需要说明的是，XML 文档的文本是用户具体表达的数据，一般不能进行篡改，本申请实施例可以不对其进行压缩。

元素、属性等文档参数的压缩，即单词的压缩，是用尽可能少的字符代表原始的元素、属性等文档参数，即映射码的字符串长度小于或等于文档参数的字符串长度。

并且，可以恢复原始的元素、属性等文档参数，通常满足以下几个要求：

1、代表元素、属性等文档参数的单词（后文称为映射码）尽可能的短，占用的字符数量少，以降低压缩率；

2、映射码互相唯一，不重复，避免混乱；

3、可以双向选择，即可以通过映射码找到原始的元素、属性等文档参数，也可以通过原始的元素、属性等文档参数找到映射码。

因此，本申请实施例可以对文档参数进行去重处理，将去重处理之后的文档参数映射为唯一的映射码。

在具体实现中，可以设置一个集合，判断提取的元素、属性等文档参数是否在该集合中，若是，则忽略，若否，则放入该集合中，使得该集合中的元素、属性等文档参数具有唯一性，这样就可以去除重复的元素、属性等文档参数。

例如，属性是 title，元素也是 title，那么就会只有一个 title 在集合中。

本申请实施例对文档参数进行去重处理之后映射为唯一的映射码，文档参数的唯一性保证了映射码的唯一性，从而保证两者可以相互转换，保证了压缩之后恢复操作的准确性。

进一步而言，可以按照字符串长度和/或字符顺序对去重处理之后的文档参数进行排序，字符串长度越短，排序越前，而排序越前，越优先进行映射。

例如，book 和 year 的字符串长度为 4，title 的字符串长度为 5，因此，book 和 year 可以排在 title 之前。

对同字符串长度的文档参数，可以按字符的顺序进行排序，

例如，book 和 year 的字符串长度相同，而 book 的第一个字符 b 比 year 的第一个字符 y 的顺序要前，因此，book 可以排在 year 之前。

对于排序好的的元素、属性等文档参数，则可以逐个进行映射，获得映射码。

在一种映射方式中，从去重处理之后的文档参数提取候选字符串；

判断候选字符串是否与已映射的映射码相同；

当不相同时，确认候选字符串为文档参数的映射码；

当相同时，从文档参数中提取包含候选字符串的目标字符串，作为新的候选字符串，返回判断候选字符串是否与已映射的映射码相同。

此种映射方式，在每次取文档参数的子串（候选字符串）时，都会与之前生成的映射码进行比较，相同时会继续取子串（目标字符串），直到取得唯一的子串为止，获得唯一的映射码。

并且，由于去重处理之后的文档参数是唯一的，映射码也是唯一的，就可以互相转换了，是一个双向的引用集合。

为了尽可能减少映射码的字符的数量，通常，候选字符串初始为第一个字符，目标字符串为由候选字符串与相邻的一个字符组成的字符串。

假设已经生成的映射码集合为 M，文档参数为 A，那么 A 的映射码 F 为：

F(A) = Min(A.substr(0, [1-A.length]) not in M)

即 A 的映射码为，由第一个字符开始的所有子串中，在已经生成的映射码中不存在的最短子串。

例如，对于属性 book，从第一个字符 b（即初始的候选字符串）开始，检查是否已经有这个映射码，如果不存在映射码 b，就将 b 与 book 进行映射，

即 b 作为 book 的映射码。

如果存在映射码 b，就取下一个子串 bo（即目标字符串），作为新的候选字符串，同样检查是否已经有这个映射码，直至生成映射码为止。

理想情况下，第一个字符即可作为映射码，最差的情况，是使用整个字符串作为映射码。

由于去重处理已经使得文档参数本身具有唯一性，因此，其映射码也具有唯一性。

应用上述方式进行映射，按照字符串长度和/或字符顺序对去重处理之后的文档参数进行排序，可以在保证映射码的唯一性的情况下，使得映射码的字符尽量少。

例如，若同时对 b 与 book 进行映射，若先映射 book，则会获得映射码 b，再对 b 进行映射时，无法映射出唯一的映射码；反之，若先映射 b，则会获得映射码 b，再对 book 进行映射时，获得映射码 bo，保证两者的唯一性。

对于上述书的 XML 文档的示例，提取的元素、属性如下：

bookstore

book

category

title

lang

author

year

price

排序之后的元素、属性如下：

book

lang

year

price

title

author

category

bookstore

按照取最短子串的规则取映射码如下:

book <-> b

lang <-> l

year <-> y

price <->p

title<-> t

author <->a

category <->c

bookstore <->bo

本申请实施例依次增大候选字符串的长度进行映射检测，在保证映射码的唯一性的同时，减少了映射码的长度。

当然，上述映射方法只是作为示例，在实施本申请实施例时，可以根据实际情况设置其他映射方法，可以保证映射关系的唯一性，以及映射码的合法性即可，本发明实施例对此不加以限制。

例如，预先设定映射码，如 a-z、A-Z、0-9 及其组合等等，对去重处理之后的文档参数直接配置映射码，为保证唯一性，映射码仅配置一次，为保证映射码的字符串长度尽可能少，优先配置字符串长度为 1 的映射码，如对排序第一的文档参数配置映射码 a，对排序第二的文档参数配置映射码 b，等等，配置完字符串长度为 1 映射码的之后，配置字符串长度为 2 的映射码，如此类推，如对排序第 N 的文档参数配置映射码 a0、对排序第 N +1 的文档参数配置映射码 a1 等等。

另外，除了上述判断处理方法外，本领域技术人员还可以根据实际需要采用其它判断处理方法，本发明实施例对此也不加以限制。

步骤 103，将所述映射码替换所述文档参数，获得压缩的可扩展标记语言 XML 文档。

在本申请实施例中，将映射码逐个进行替换，实现 XML 文档的压缩。

对于上述书的 XML 文档的示例，将映射码替换原始的文档参数，获得的压缩的 XML 文档如下：

```
<bo>
<b c="COOKING">
  <t l="en">Everyday Italian</t>
  <a>Giada De Laurentiis</a>
  <y>2005</y>
  <p>30.00</p>
</b>
<b c="CHILDREN">
  <t l="en">Harry Potter</t>
  <a>J K. Rowling</a>
  <y>2005</u>
  <p>29.99</p>
</b>
<b c="WEB">
  <t lang="en">Learning XML</t>
  <a>Erik T. Ray</a>
  <y>2003</y>
  <p>39.95</p>
</b>
</bo>
```

可以看到，XML 文档有大幅的压缩，尤其当 XML 很大时，压缩效果就更为明显了，并且，XML 结构本身没有被破坏，且元素、属性等文档参数与映射码都具有唯一性，在应用时可以进行恢复。

在实际应用中，很多情况下需要重复对相同的对象进行定义，XML 文档会有大量重复的元素、属性等文档参数。

例如，云盘的接口基于 XML 文档进行传输，该 XML 文档通常需要定义存储的文件，会存在大量的 field（表征文件）、name（表征名称）等元素、属性等文档参数。

又例如，在图书馆的数据库中，该 XML 文档通常需要定义存储的书本，会存在大量的 book（表征书本）、title（表征标题）、author（表征作者）等元素、属性等文档参数。

此外，应用 XML 文档也会有大量重复的元素。

例如，基于 XML 的一些规范，如 BPEL，BPMN，都有大量重复的元素。

基于 XML 的定义的通常都是元素的规范，而实际使用中，符合规范的元素都是很多的。

例如，电子商务的某个接口用于获取商品列表，每个商品的描述是一样的，如 item（表征项目），有时候多达数百个。

应用本申请实施例进行压缩，假设原始的文档参数的字符串长度为 X，数量为 Y，其映射码的字符串长度为 Z，则减少的数据量为（X-Z）*Y，XML 文档存储的数据量就大大减少了，在 XML 文档较大、文档参数的重复数量较多时尤为明显。

在某项实验数据表明，平均 20M 的 XML 文档，只经过 GZIP 压缩，可以压缩到 5M，但经过本申请实施例的压缩方法进行压缩，可以压缩到 1M 左右，文档的体积有较大幅度的降低。

除此之外，文档参数与映射码之间的映射关系需要作为元信息，用于恢复。

在一种情况中，可以将文档参数与映射码之间的映射关系嵌入可扩展标记语言 XML 文档中。

对于上述书的 XML 文档的示例，映射关系嵌入 XML 文档如下：

<!—

book <-> b

lang <-> l

year <-> y

```
price <->p
title <-> t
author <-> a
category <->c
bookstore <->bo
-->
<bo>
<b c="COOKING">
  <t l="en">Everyday Italian</t>
  <a>Giada De Laurentiis</a>
  <y>2005</y>
  <p>30.00</p>
</b>
<b c="CHILDREN">
  <t l="en">Harry Potter</t>
  <a>J K. Rowling</a>
  <y>2005</u>
  <p>29.99</p>
</b>
<b c="WEB">
  <t lang="en">Learning XML</t>
  <a>Erik T. Ray</a>
  <y>2003</y>
  <p>39.95</p>
</b>
</bo>
```

在此示例中，映射关系以“key <-> value”（key 为文档参数，value 为映射码）的形式进行记录，以注解的形式嵌入 XML 文档的头部。

当然，映射关系也可以以其他形式进行记录，嵌入 XML 文档的其他位置，甚至，可以以独立的文件记录，本申请实施例对此不加以限制。

在实际应用中，很多开源的数据库都是用 XML 文档进行存储的，某些电子商务平台上接口的返回请求参数也是基于 XML 文档的，等等。

在压缩 XML 文档之后，为保证压缩有效，则可以对比原始的可扩展标记语言 XML 文档与压缩的可扩展标记语言 XML 文档的体积大小。

若原始的可扩展标记语言 XML 文档比压缩的可扩展标记语言 XML 文档的体积小，则可以认为压缩有效，可以传输和/或存储压缩的可扩展标记语言 XML 文档。

若原始的可扩展标记语言 XML 文档比压缩的可扩展标记语言 XML 文档的体积大，则可以认为压缩无效，针对此可扩展标记语言 XML 文档，可以不应用本申请实施例的压缩方法进行压缩。

需要说明的是，在传输和/或存储之前，还可以采用预设的文本压缩方式（如 GZIP）对压缩的可扩展标记语言 XML 文档进一步进行压缩，进一步提高压缩率。

本申请实施例通过对原始的 XML 文档的文档参数进行映射，将映射码替换文档参数，实现压缩，由于文档参数很多情况下存在重复，因此，替换之后可以大大减少 XML 文档存储的数据量，提高了压缩效率，并且，映射操作简单，减少了压缩的耗时。

进而，在传输、存储压缩的 XML 文档时，可以减少带宽的消耗，或者，较少存储空间的占用。

参照图 3，示出了本申请的一种可扩展标记语言 XML 文档的解压方法实施例的步骤流程图，具体可以包括如下步骤：

步骤 301，获取压缩的可扩展标记语言 XML 文档；

在具体实现中，若压缩的 XML 文档在先存储在数据库，则可以从数据库读取在先存储的压缩的可扩展标记语言 XML 文档；

若压缩的 XML 文档在先进行传输，则可以接收传输到来的压缩的可扩

展标记语言 XML 文档，等等。

在本申请实施例中，压缩的可扩展标记语言 XML 文档中包括映射码，该映射码为由原始的文档参数映射获得。

例如，某本书的压缩的 XML 文档的示例如下：

```
<bo>
<b c="COOKING">
  <t l="en">Everyday Italian</t>
  <a>Giada De Laurentiis</a>
  <y>2005</y>
  <p>30.00</p>
</b>
<b c="CHILDREN">
  <t l="en">Harry Potter</t>
  <a>J K. Rowling</a>
  <y>2005</u>
  <p>29.99</p>
</b>
<b c="WEB">
  <t lang="en">Learning XML</t>
  <a>Erik T. Ray</a>
  <y>2003</y>
  <p>39.95</p>
</b>
</bo>
```

其中，b、l、y、p、t、a、c、bo 均为映射码。

需要说明的是，在传输和/或存储之前，采用了预设的文本压缩方式（如 GZIP）对压缩的可扩展标记语言 XML 文档进一步进行压缩，则在获取压缩的可扩展标记语言 XML 文档之后，可以按照该文本压缩方式（如 GZIP）对

压缩的可扩展标记语言 XML 文档进进行解压。

步骤 302，查找压缩的可扩展标记语言 XML 文档的、映射码与文档参数之间的映射关系；

若该映射关系可以嵌入 XML 文档中（如头部），则可以从 XML 文档中（如头部）读取该映射关系。

例如，对于上述书的压缩的 XML 文档，映射关系嵌入 XML 文档的示例如下：

```
<!—
book <-> b
lang <-> l
year <-> y
price <->p
title <-> t
author <-> a
category <->c
bookstore <->bo
-->
<bo>
<b c="COOKING">
  <t l="en">Everyday Italian</t>
  <a>Giada De Laurentiis</a>
  <y>2005</y>
  <p>30.00</p>
</b>
<b c="CHILDREN">
  <t l="en">Harry Potter</t>
  <a>J K. Rowling</a>
  <y>2005</u>
```

```
    <p>29.99</p>
  </b>
  <b c="WEB">
    <t lang="en">Learning XML</t>
    <a>Erik T. Ray</a>
    <y>2003</y>
    <p>39.95</p>
  </b>
  </bo>
```

其中，映射关系以注解的形式嵌入 XML 文档的头部，包括:

book <-> b

lang <-> l

year <-> y

price <->p

title <-> t

author <-> a

category <->c

bookstore <->bo

当然，若该映射关系采用其他方式存储，如应用独立的文件记录，则可以采用相应的方式进行读取，如查找该独立的文件，本申请实施例对此不加以限制。

步骤 303，按照所述映射关系将所述文档参数替换所述映射码，获得原始的可扩展标记语言 XML 文档。

在具体实现中，可以按照映射关系的记录形式，查找压缩 XML 文档中映射码的文档参数，将该文档参数替换该映射码，进行恢复。

例如，对于上述书的压缩的 XML 文档的示例，映射关系以“key <-> value”的形式进行记录，其中，key 为文档参数，value 为映射码。

若需要对映射码 b 恢复，则可以查找到 value 为 b 的映射关系，即 book

<-> b，将 key，即 book 替换映射码 b。

则恢复后的书的 XML 的示例如下:

```
<bookstore>
<book category="COOKING">
  <title lang="en">Everyday Italian</title>
  <author>Giada De Laurentiis</author>
  <year>2005</year>
  <price>30.00</price>
</book>
<book category="CHILDREN">
  <title lang="en">Harry Potter</title>
  <author>J K. Rowling</author>
  <year>2005</year>
  <price>29.99</price>
</book>
<book category="WEB">
  <title lang="en">Learning XML</title>
  <author>Erik T. Ray</author>
  <year>2003</year>
  <price>39.95</price>
</book>
</bookstore>
```

需要说明的是，当映射关系嵌在压缩的可扩展标记语言 XML 文档中时，删除映射关系，最终获得原始的 XML 文档，进行正常使用。

本申请实施例通过映射关系，将原始的文档参数替换映射码，映射操作简单，减少了解压的耗时。

需要说明的是，对于方法实施例，为了简单描述，故将其都表述为一系

列的动作组合，但是本领域技术人员应该知悉，本申请实施例并不受所描述的动作顺序的限制，因为依据本申请实施例，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作并不一定是本申请实施例所必须的。

参照图 4，示出了本申请的一种可扩展标记语言 XML 文档的压缩装置实施例的结构框图，具体可以包括如下模块：

文档参数读取模块 401，用于从原始的可扩展标记语言 XML 文档中读取文档参数；

映射模块 402，用于将所述文档参数映射为映射码；

文档参数替换模块 403，用于将所述映射码替换所述文档参数，获得压缩的可扩展标记语言 XML 文档。

在本申请的一个实施例中，该装置还可以包括如下模块：

映射关系嵌入模块，用于将所述文档参数与所述映射码之间的映射关系嵌入所述可扩展标记语言 XML 文档中。

在本申请的一个实施例中，该装置还可以包括如下模块：

传输模块，用于传输压缩的可扩展标记语言 XML 文档；

和/或，

存储模块，用于存储压缩的可扩展标记语言 XML 文档。

在具体实现中，所述文档参数可以包括元素和/或属性。

在本申请的一个实施例中，所述映射模块 402 可以包括如下子模块：

去重子模块，用于对所述文档参数进行去重处理；

去重映射子模块，用于将去重处理之后的文档参数映射为唯一的映射码，所述映射码的字符串长度小于或等于所述文档参数的字符串长度。

在本申请的一个示例中，所述去重映射子模块可以包括如下单元：

候选字符串提取单元，用于从去重处理之后的文档参数提取候选字符串；

映射码判断单元，用于判断所述候选字符串是否与已映射的映射码相

同；当不相同时，调用映射码确认单元，当相同时，调用目标字符串提取单元，返回调用映射码判断单元；

映射码确认单元，用于确认所述候选字符串为所述文档参数的映射码；

目标字符串提取单元，用于从所述文档参数中提取包含所述候选字符串的目标字符串，作为新的候选字符串。

在本申请的另一个示例中，所述去重映射子模块还可以包括如下单元：

排序单元，用于按照字符串长度和/或字符顺序对去重处理之后的文档参数进行排序。

参照图 5，示出了本申请的一种可扩展标记语言 XML 文档的解压装置实施例的结构框图，具体可以包括如下模块：

XML 文档获取模块 501，用于获取压缩的可扩展标记语言 XML 文档，压缩的可扩展标记语言 XML 文档中包括映射码；

映射关系查找模块 502，用于查找压缩的可扩展标记语言 XML 文档的、映射码与文档参数之间的映射关系；

映射码替换模块 503，用于按照所述映射关系将所述文档参数替换所述映射码，获得原始的可扩展标记语言 XML 文档。

在本申请的一个实施例中，该装置还可以包括如下模块：

映射关系删除模块，用于在所述映射关系嵌在压缩的可扩展标记语言 XML 文档中时，删除所述映射关系。

在本申请的一个实施例中，所述 XML 文档获取模块 501 可以包括如下子模块：

XML 文档读取子模块，用于读取在先存储的压缩的可扩展标记语言 XML 文档；

或者，

XML 文档接收子模块，用于接收传输到来的压缩的可扩展标记语言 XML 文档。

对于装置实施例而言，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。

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。

本领域内的技术人员应明白，本申请实施例的实施例可提供为方法、装置、或计算机程序产品。因此，本申请实施例可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本申请实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

在一个典型的配置中，所述计算机设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器(RAM) 和/或非易失性内存等形式，如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存 (PRAM)、静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、其他类型的随机存取存储器 (RAM)、只读存储器(ROM)、电可擦除可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括非持续性的电脑可读媒体(transitory media)，如调制的数据信号和载波。

本申请实施例是参照根据本申请实施例的方法、终端设备(系统)、和计

算机程序产品的流程图和／或方框图来描述的。应理解可由计算机程序指令实现流程图和／或方框图中的每一流程和／或方框、以及流程图和／或方框图中的流程和／或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理终端设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理终端设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理终端设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理终端设备上，使得在计算机或其他可编程终端设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程终端设备上执行的指令提供用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的步骤。

尽管已描述了本申请实施例的优选实施例，但本领域内的技术人员一旦得知了基本创造性概念，则可对这些实施例做出另外的变更和修改。所以，所附权利要求意欲解释为包括优选实施例以及落入本申请实施例范围的所有变更和修改。

最后，还需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者终端设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者终端设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者终

端设备中还存在另外的相同要素。

以上对本申请所提供的一种可扩展标记语言 XML 文档的压缩方法、一种可扩展标记语言 XML 文档的解压方法、一种可扩展标记语言 XML 文档的压缩装置和一种可扩展标记语言 XML 文档的解压装置，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。

权 利 要 求 书

1、一种可扩展标记语言 XML 文档的压缩方法，其特征在于，包括:

从原始的可扩展标记语言 XML 文档中读取文档参数;

将所述文档参数映射为映射码;

将所述映射码替换所述文档参数，获得压缩的可扩展标记语言 XML 文档。

2、根据权利要求 1 所述的方法，其特征在于，还包括:

将所述文档参数与所述映射码之间的映射关系嵌入所述可扩展标记语言 XML 文档中。

3、根据权利要求 1 或 2 所述的方法，其特征在于，还包括:

传输和/或存储压缩的可扩展标记语言 XML 文档。

4、根据权利要求 1 所述的方法，其特征在于，所述文档参数包括元素和/或属性。

5、根据权利要求 1 或 2 或 4 所述的方法，其特征在于，所述将所述文档参数映射为映射码的步骤包括:

对所述文档参数进行去重处理;

将去重处理之后的文档参数映射为唯一的映射码，所述映射码的字符串长度小于或等于所述文档参数的字符串长度。

6、根据权利要求 5 所述的方法，其特征在于，所述将去重处理之后的文档参数映射为唯一的映射码的步骤包括:

从去重处理之后的文档参数提取候选字符串;

判断所述候选字符串是否与已映射的映射码相同;

当不相同时，确认所述候选字符串为所述文档参数的映射码;

当相同时，从所述文档参数中提取包含所述候选字符串的目标字符串，作为新的候选字符串，返回执行所述判断所述候选字符串是否与已映射的映射码相同的步骤。

7、根据权利要求 6 所述的方法，其特征在于，所述将去重处理之后的

文档参数映射为唯一的映射码的步骤还包括:

按照字符串长度和/或字符顺序对去重处理之后的文档参数进行排序。

8、一种可扩展标记语言 XML 文档的解压方法，其特征在于，包括:

获取压缩的可扩展标记语言 XML 文档，压缩的可扩展标记语言 XML 文档中包括映射码;

查找压缩的可扩展标记语言 XML 文档的、映射码与文档参数之间的映射关系;

按照所述映射关系将所述文档参数替换所述映射码，获得原始的可扩展标记语言 XML 文档。

9、根据权利要求 8 所述的方法，其特征在于，还包括:

当所述映射关系嵌在压缩的可扩展标记语言 XML 文档中时，删除所述映射关系。

10、根据权利要求 8 或 9 所述的方法，其特征在于，所述获取压缩的可扩展标记语言 XML 文档的步骤包括:

读取在先存储的压缩的可扩展标记语言 XML 文档;

或者，

接收传输到来的压缩的可扩展标记语言 XML 文档。

11、一种可扩展标记语言 XML 文档的压缩装置，其特征在于，包括:

文档参数读取模块，用于从原始的可扩展标记语言 XML 文档中读取文档参数;

映射模块，用于将所述文档参数映射为映射码;

文档参数替换模块，用于将所述映射码替换所述文档参数，获得压缩的可扩展标记语言 XML 文档。

12、根据权利要求 11 所述的装置，其特征在于，还包括:

映射关系嵌入模块，用于将所述文档参数与所述映射码之间的映射关系嵌入所述可扩展标记语言 XML 文档中。

13、根据权利要求 1 或 2 所述的装置，其特征在于，所述映射模块包括:

去重子模块，用于对所述文档参数进行去重处理；

去重映射子模块，用于将去重处理之后的文档参数映射为唯一的映射码，所述映射码的字符串长度小于或等于所述文档参数的字符串长度。

14、根据权利要求 13 所述的装置，其特征在于，所述去重映射子模块包括：

候选字符串提取单元，用于从去重处理之后的文档参数提取候选字符串；

映射码判断单元，用于判断所述候选字符串是否与已映射的映射码相同；当不相同时，调用映射码确认单元，当相同时，调用目标字符串提取单元，返回调用映射码判断单元；

映射码确认单元，用于确认所述候选字符串为所述文档参数的映射码；

目标字符串提取单元，用于从所述文档参数中提取包含所述候选字符串的目标字符串，作为新的候选字符串。

15、根据权利要求 14 所述的装置，其特征在于，所述去重映射子模块还包括：

排序单元，用于按照字符串长度和/或字符顺序对去重处理之后的文档参数进行排序。

16、一种可扩展标记语言 XML 文档的解压装置，其特征在于，包括：

XML 文档获取模块，用于获取压缩的可扩展标记语言 XML 文档，压缩的可扩展标记语言 XML 文档中包括映射码；

映射关系查找模块，用于查找压缩的可扩展标记语言 XML 文档的、映射码与文档参数之间的映射关系；

映射码替换模块，用于按照所述映射关系将所述文档参数替换所述映射码，获得原始的可扩展标记语言 XML 文档。

17、根据权利要求 16 所述的装置，其特征在于，还包括：

映射关系删除模块，用于在所述映射关系嵌在压缩的可扩展标记语言 XML 文档中时，删除所述映射关系。

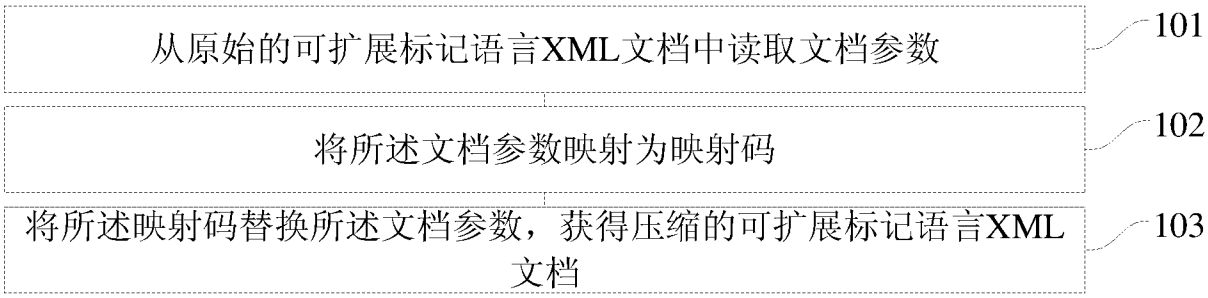

从原始的可扩展标记语言XML文档中读取文档参数
101
将所述文档参数映射为映射码
102
将所述映射码替换所述文档参数，获得压缩的可扩展标记语言XML文档
103

图 1

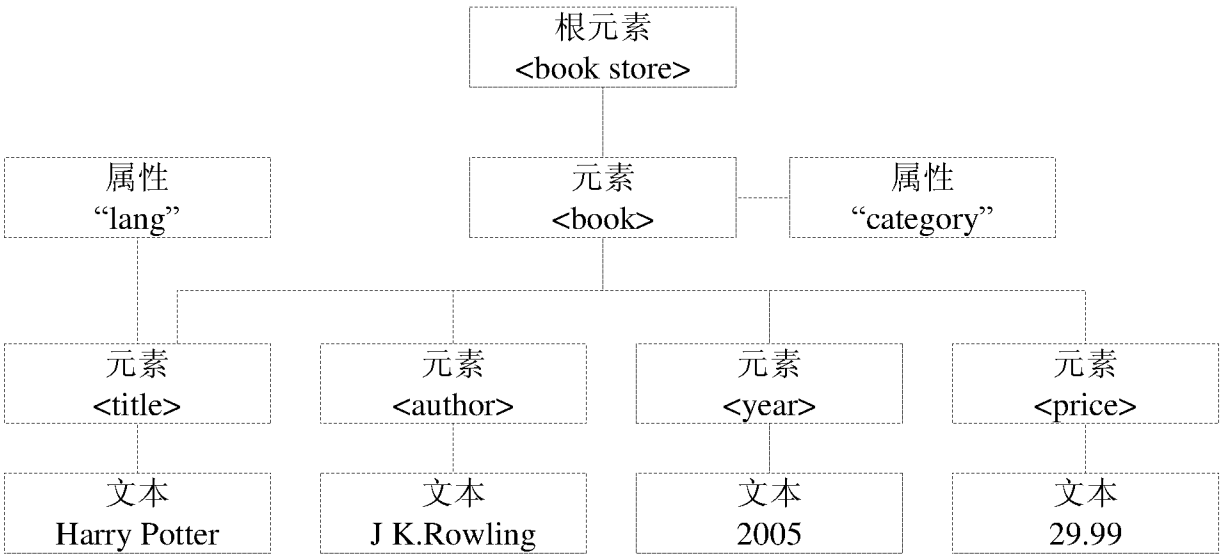

根元素
<book store>
属性
“lang”
元素
<book>
属性
“category”
元素
<title>
元素
<author>
元素
<year>
元素
<price>
文本
Harry Potter
文本
J K.Rowling
文本
2005
文本
29.99

图 2

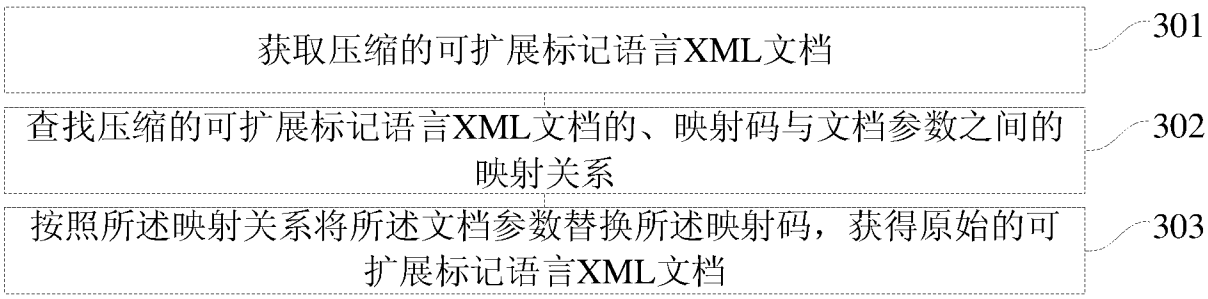

获取压缩的可扩展标记语言XML文档
301
查找压缩的可扩展标记语言XML文档的、映射码与文档参数之间的映射关系
302
按照所述映射关系将所述文档参数替换所述映射码，获得原始的可扩展标记语言XML文档
303

图 3

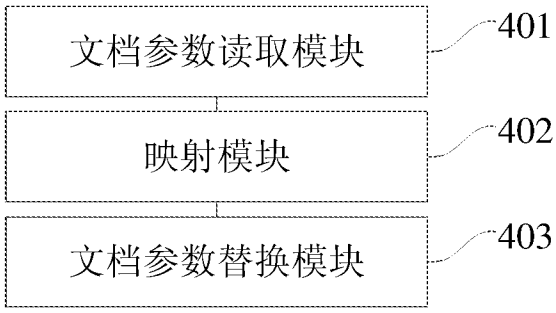

文档参数读取模块
401
映射模块
402
文档参数替换模块
403

图 4

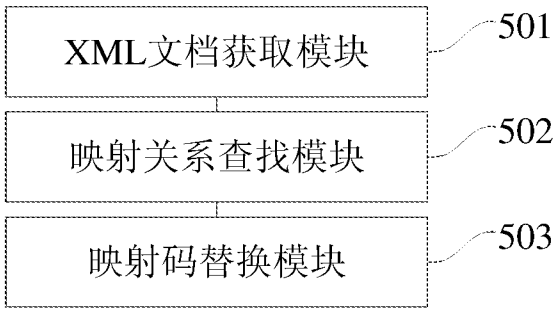

XML文档获取模块
501
映射关系查找模块
502
映射码替换模块
503

图 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/096790

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPODOC, WPI, CNPAT, CNKI, IEEE, GOOGLE: extensible markup language, XML, compress+, decompress+, map+ table, map+ code, document, element, attribute

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103605730 A (SHANXI SANHENG AUTOMATION EQUIPMENT CO., LTD. et al.), 26 February 2014 (26.02.2014), description, paragraphs [0012] and [0018]-[0044], and figures 1-3	1-17
X	CN 102096704 A (BEIJING FEINNO COMMUNICATION TECH CO., LTD.), 15 June 2011 (15.06.2011), description, paragraphs [0006]-[0019] and [0071]	1-17
A	CN 104484337 A (XIDIAN UNIVERSITY), 01 April 2015 (01.04.2015), the whole document	1-17
A	CN 1635492 A (ROYAL DUTCH PHILIPS ELECTRONICS LTD.), 06 July 2005 (06.07.2005), the whole document	1-17
A	US 6883137 B1 (INTERNATIONAL BUSINESS MACHINES CORPORATION), 19 April 2005 (19.04.2005), the whole document	1-17

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
03 November 2016 (03.11.2016)	**30 November 2016 (30.11.2016)**
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer **WU, Qing** Telephone No.: (86-10) **61648111**

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/096790

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103605730 A	26 February 2014	None	
CN 102096704 A	15 June 2011	None	
CN 104484337 A	01 April 2015	None	
CN 1635492 A	06 July 2005	US 2007273564 A1	29 November 2007
		WO 2005067153 A1	21 July 2005
		JP 2007520112 A	19 July 2007
		EP 1702412 A1	20 September 2006
		CN 1902827 A	24 January 2007
US 6883137 B1	19 April 2005	None	

国际检索报告

国际申请号
PCT/CN2016/096790

A. 主题的分类

G06F 17/30(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

EPODOC, WPI, CNPAT, CNKI, IEEE, GOOGLE：可扩展标记语言，压缩，解压，映射表，映射码，文档，元素，属性，extensible markup language, XML, compress+, decompress+, map+ table, map+ code, document, element, attribute

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
X	CN 103605730 A (山西三恒自动化设备有限公司 等) 2014年 2月 26日 (2014-02-26) 说明书第[0012], [0018]-[0044]段、附图1-3	1-17
X	CN 102096704 A (北京新媒传信科技有限公司) 2011年 6月 15日 (2011-06-15) 说明书第[0006]-[0019], [0071]段	1-17
A	CN 104484337 A (西安电子科技大学) 2015年 4月 1日 (2015-04-01) 全文	1-17
A	CN 1635492 A (皇家飞利浦电子股份有限公司) 2005年 7月 6日 (2005-07-06) 全文	1-17
A	US 6883137 B1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2005年 4月 19日 (2005-04-19) 全文	1-17

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2016年 11月 3日	2016年 11月 30日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 吴卿 电话号码 (86-10)61648111

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2016/096790

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	103605730	A	2014年 2月 26日	无			
CN	102096704	A	2011年 6月 15日	无			
CN	104484337	A	2015年 4月 1日	无			
CN	1635492	A	2005年 7月 6日	US	2007273564	A1	2007年 11月 29日
				WO	2005067153	A1	2005年 7月 21日
				JP	2007520112	A	2007年 7月 19日
				EP	1702412	A1	2006年 9月 20日
				CN	1902827	A	2007年 1月 24日
US	6883137	B1	2005年 4月 19日	无			