

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 17/30 (2006.01)



[12] 发明专利说明书

专利号 ZL 03814226.0

[45] 授权公告日 2008 年 11 月 26 日

[11] 授权公告号 CN 100437562C

[22] 申请日 2003.6.11 [21] 申请号 03814226.0

[30] 优先权

[32] 2002. 6. 20 [33] US [31] 10/176,232

[86] 国际申请 PCT/IB2003/002573 2003.6.11

[87] 国际公布 WO2004/001625 英 2003.12.31

[85] 进入国家阶段日期 2004.12.17

[73] 专利权人 皇家飞利浦电子股份有限公司

地址 荷兰艾恩德霍芬

[72] 发明人 D·J·皮奥特罗斯基

[56] 参考文献

CN1342368A 2002.3.27

Annotation - based Web content transcoding.
Masahiro Hori, Goh Kondoh, Kouichi Ono, Shin.
ichiHirose, Sandeep Singhal. Computer Networks,
Vol. 33 No. 1. 6. 2000

Efficient representation and streaming of XML
content over the Internet medium. Girardot. M.
Sundaresan. N. Multimedia and Expo, 0.1 CME2000
IEEE Internal Conference on, Vol. 1. 2000

审查员 李 圆

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 程天正 王 勇

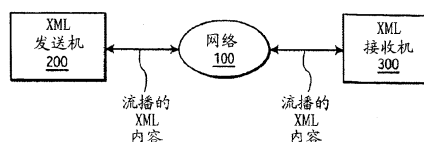
权利要求书 3 页 说明书 6 页 附图 3 页

[54] 发明名称

用于 XML 文档的结构化流播的方法和装置

[57] 摘要

公开了一种用于以允许接收机(300)译码 XML 文档的划分优先级的部分的结构化方式流播 XML 文档/内容的方法和装置。文档模型,例如 XML 大纲,被用于把 XML 文档转换成根据预先定义的方案被发送的、划分优先级的部分。因此,本发明允许 XML 接收机(300)首先以及在传输中开始处理 XML 流的最重要的 XML 部分并且允许用户在接收到较低优先级的 XML 部分之前结束传输。



1. 一种发送 XML 文档的方法, 包括步骤:
产生用于发送的该 XML 文档的 XML 部分;
根据预定的方案划分该 XML 部分的优先级, 其中每个预定的方案取决于一个特定的应用和用户定义的优先级; 以及
发送该 XML 的被划分优先级的部分。
2. 权利要求 1 所述的方法, 其中所述预定的方案是根据规定的 XML 大纲文档, 其中每个 XML 大纲元素被划分优先级。
3. 权利要求 2 所述的方法, 其中划分所述 XML 部分的优先级的步骤包括添加 XML 大纲标识符。
4. 权利要求 1 所述的方法, 还包括用户在第一 XML 部分被发送之后终止所述发送步骤的步骤。
5. 一种编码 XML 文档的方法, 包括步骤:
根据预先定义的划分优先级的方式, 把该 XML 文档分解成多个 XML 部分, 其中每个预先定义的方式取决于一个特定的应用和用户定义的优先级; 以及
独立地发送所述多个 XML 部分的每一 XML 部分连同指示该 XML 部分在所述较大的 XML 文档中如何被放置的信息。
6. 权利要求 5 所述的方法, 其中指示该 XML 部分在所述较大的 XML 文档中如何被放置的所述信息在一个 XML 部分中被发送。
7. 一种用于接收流播的 XML 文档的方法, 该 XML 文档包括多个 XML 部分, 其中每个 XML 部分包括基于一个特定的应用和用户定义的优先级的结构信息, 所述方法包括步骤:
从至少一个接收到的 XML 部分中提取预定的结构信息或者它的部分, 其中该预定的结构信息被用于划分所述 XML 部分的优先级;
直接处理所述 XML 部分; 以及
从所接收到的 XML 部分和结构信息中重新编译所述 XML 文档。
8. 权利要求 7 所述的方法, 其中所述处理步骤还包括显示该信息的步骤。
9. 权利要求 7 所述的方法, 其中所述处理步骤还包括存储该信息的步骤。
10. 权利要求 7 所述的方法, 还包括在接收到第一个 XML 部分之后

终止其它的 XML 部分的接收的步骤。

11. 权利要求 7 所述的方法, 还包括即使所述 XML 部分的一个部分没有被正常接收, 也继续处理后续的 XML 部分的步骤。

12. 一种 XML 发送机 (200), 包括:

用于存储 XML 文档和计算机可读代码的存储器 (220); 以及

可操作地被耦合到所述存储器的处理器 (210), 所述处理器被配置成:

产生用于发送的所述 XML 文档的部分, 其中所述部分根据预定的方案被划分优先级, 其中每个预定的方案取决于一个特定的应用和用户定义的优先级; 以及

根据所述被划分优先级的部分各自的优先级来发送它们。

13. 权利要求 12 所述的 XML 发送机 (200), 其中所述预定的方案是根据规定的 XML 大纲文档的, 其中每个 XML 大纲元素被划分优先级。

14. 权利要求 13 所述的 XML 发送机 (200), 其中所述被划分优先级的部分包括 XML 大纲标识符。

15. 权利要求 12 所述的 XML 发送机 (200), 其中所述处理器可操作来允许用户在至少第一被划分优先级的部分被发送以后终止所述被划分优先级的部分的发送。

16. 一种 XML 发送机 (200) 包括:

用于存储 XML 文档和计算机可读代码的存储器 (220); 以及

可操作地被耦合到所述存储器的处理器 (210), 所述处理器被配置成:

根据预先定义的划分优先级的方式, 把所述 XML 文档分解成多个 XML 部分, 其中每个预先定义的方式取决于一个特定的应用和用户定义的优先级; 以及

独立地发送所述多个 XML 部分的每一 XML 部分连同指示所述 XML 部分在所述较大的 XML 文档中如何被放置的信息。

17. 一种用于接收流播的 XML 文档的 XML 接收机 (300), 所述 XML 文档包括多个 XML 部分, 其中每个 XML 部分包括基于一个特定的应用和用户定义的优先级的结构信息, 该接收机包括:

用于存储计算机可读代码的存储器 (320); 以及

可操作地被耦合到所述存储器的处理器 (310), 所述处理器被配置

成:

从至少接收到的 XML 部分提取预定的结构信息或者它的部分, 其中所述预定的结构信息被用于划分所述 XML 部分的优先级;

直接处理所述 XML 部分; 以及

从所接收到的 XML 部分和结构信息重新编译所述 XML 文档。

18. 权利要求 17 所述的 XML 接收机 (300), 其中所述处理器可操作来允许用户在至少第一个 XML 部分被发送之后终止所述 XML 部分的另外的传输。

19. 权利要求 17 所述的 XML 接收机 (300), 其中所述预定的结构信息是根据规定的 XML 大纲文档, 其中每个 XML 大纲元素被划分优先级。

20. 权利要求 19 所述的 XML 接收机 (300), 其中所接收到的 XML 部分包括 XML 大纲标识符。

用于 XML 文档的结构化流播的方法和装置

本发明涉及连续媒体的流播，并且更具体而言，涉及用于以结构化方式流播 XML 文档的方法及装置。

可扩展标示语言 (XML) 是已经由万维网协会 (W3C) 推荐的、用于编码文本信息的标准。对于 XML 标准的讨论，参见，例如可扩展标示语言 (XML) 1.0 W3C 建议，<http://www.w3.org/TR/1998/REC-xml-19980210>，该建议在此引入作为参考。XML 标准允许 XML 使能的应用程序与其它兼容系统相互协作用于被编码的信息的交换。

XML 文档使用了几个基本的概念。XML 文档在分层树结构中存储文本数据。每篇 XML 文档具有一个根节点，通常被称为根元素，在分层树中的其它节点被安排为根节点的子节点。XML 标准详细说明了四种类型的节点，即：字符节点，处理指令 (PI) 节点，注释节点以及元素节点。基于 XML 的应用程序能够在所有不同类型的节点中以及在每种节点类型的所有区域中存储数据。XML 也使用所谓的文档模型方法，例如文档类型定义或者 “DTD” 来定义文档的属性。

已经出现了诸如视频点播以及其它连续媒体应用的、用于编码并发送连续的媒体流的多种应用。例如，从运动图像组建议的 MPEG-7 标准提供了用于编码视频信息以及与视频源有关的文本信息的技术规范。连续的媒体流通常是利用基于分组的通信系统而被发送的。然而由于基于分组的通信系统的不可靠性，接收到的流的质量可能会受到分组丢失的影响。因此，这样的连续媒体传输系统通常必须包括允许接收机适应于分组丢失的机制。已经提出或建议了许多种技术来解决在连续的媒体传输系统中的分组丢失，包括冗余传输，重发，交织以及前向纠错技术。对于用于解决在连续的媒体系统中的分组丢失的这种技术的总的讨论，参见，例如 “Options for Repair of Streaming Media (关于流播媒体的修复的选项)，” 网络工作组，请求注释 No. 2354 (1998 年六月)，该文档可以从 <ftp://isi.edu/in-notes/rfc2354.txt> 下载，在此引入作为参考。

XMLNet 是用于流播 XML 文档的应用程序接口 (API)。XMLNet 允许信息作为 XML 文档系列在因特网或另一个网络上实时地被发送。XML 文档

以系列方式被递送给接收机。然而，在接收机能对包含在 XML 文档中的任意 XML 内容进行译码和处理之前，接收机必须接收到完整的 XML 文档。对于 XMLNet 的讨论，参见，例如可以从 home.earthlink.net/%7Earabbit/xmlnet (1998 年 12 月 9 日) 下载的“XMLNet”。

因此存在对这样一种方法和装置的需要，其允许用户对用于传输的部分 XML 编码文档划分优先级，并且即使没有接收到完整的 XML 文档的其它的或优先级较低的部分时，例如如果在接收到完整的 XML 文档之前发生分组丢失时，也允许接收机译码接收到的部分。还存在对如下方法和装置的需要，即：如果用户期望的话，准许以允许只有部分 XML 文档被发送并被接收机译码的划分优先级的方式来进行 XML 文档流播。

通常，公开的方法和装置是用于以结构化方式进行 XML 文档/内容的流播，该结构化方式允许接收机以划分优先级的方式来译码 XML 文档的部分。文档模型被用于把 XML 文档转换成根据预先定义的方案被发送的、划分优先级的部分，元素，段或子树（在下文中被称为“XML 部分”）。因此，本发明允许 XML 接收机首先或者在传输中开始处理 XML 流的最重要的 XML 部分并且允许用户在接收到较低优先级的 XML 部分之前结束传输。

根据本发明的一个实施例，XML 文档根据基于特定应用（例如：购货单，发货单等等）的特定 XML 大纲而被分割成 XML 部分。例如，用于在因特网上的电子商务交易的购货单的 XML 文档可被分割成不同的部分，例如，姓名，地址，产品编号，价格等等。然后，这些部分基于用户的准则被划分优先级。这些 XML 部分然后被分组和或被发送给接收机以首先译码 XML 文档的最重要的 XML 部分。具体地，每个 XML 文档被分解并编码成 XML 大纲元素的集合，这些元素根据具体的参数组（取决于应用或者预先定义的用户参数）被划分优先级。较大的 XML 文档的每个被划分优先级的 XML 部分被 XML 接收机解析并验证它是不是 XML 文档。

根据本发明的另一个方面，用户可以在接收到 XML 文档的几个 XML 部分之后或者在接收到完整的 XML 文档之前终止请求。这样，用户（1）可以终止 XML 部分的其它传输并且能够只译码他或她认为在 XML 文档中重要的 XML 部分，以及（2）倘若接收到了足够多的流播的 XML 部分，即使丢失了某些较低优先级的 XML 部分，利用已知的 XML 大纲也可重建整

个的 XML 文档。

通过参考以下的详细描述和附图，将得到本发明的更彻底的理解以及本发明的另外的特征和优点。

图 1 说明了其中本发明可以运行的典型的网络环境；

图 2 是显示根据本发明的示例性 XML 发送机的构造的框图；

图 3 是显示根据本发明的示例性 XML 接收机的构造的框图；以及

图 4 是描述由图 3 的 XML 接收机执行的示范流播的 XML 部分处理的流程图。

图 1 说明了其中本发明可以运行的网络环境 100。如图 1 中所示，XML 发送机 200 发送被流播的 XML 文档/内容到 XML 接收机 300。根据下面进一步讨论的本发明的特性，XML 发送机 200 以允许 XML 接收机 300 只译码接收到的被发送的 XML 文档的划分优先级的 XML 部分的这样方式编码并发送 XML 内容。例如，如果用户终止其它的传输或者部分 XML 数据出现分组丢失或者如果整个的 XML 文档尚未到达，XML 接收机 300 可以只译码部分被流播的 XML 文档。因此，根据本发明，XML 接收机 300 能在传输中截取 XML 流并基于 XML 编码的文档的接收到的 XML 部分仍能完成有用的任务。以这种方式，用户规定 XML 文档的最相关的 XML 部分并且一旦接收到这些 XML 部分就可以停止传输。

每个 XML 文档被编码成段（例如，XML 部分）的集合，这使得接收机 300 在处理信息之前可以不用等待去接收整个的 XML 文档。根据转让给本发明的受让人并在此引入作为参考的、在 1999 年 4 月 10 日提交的、序列号为 09/411, 756 的、题目为“Method and Apparatus for Streaming XML Content（用于流播 XML 内容的方法和装置）”的美国专利申请的学说，子树被使用。

例如，XML 文档被分解并被编码成部分/段的集合或所谓的子树。子树被设置在给定的节点，并且包括这个给定的节点以及分层树结构中的在这个给定的节点以下的所有节点。较大的 XML 文档树中的每个子树被接收机 300 分解并验证它是不是独立的树。给定的子树可以包括其它子树。流播的 XML 文档中的每个子树使用在每个独立的子树周围用作子树外包功能的结构节点。该结构节点指示了这个子树跟其它子树之间的关系。以这种方式，倘若接收到足够多的流播的 XML 内容，XML 接收机 300 能重建全部树的结构。

如上面所指出的，XML 使用文档模型方法，例如文档类型定义或者“DTD”。特别地，XML 大纲表示共享的词表并且提供用于规定 XML 文档的结构，内容和语义的方法。XML 大纲被使用来规定 XML 文档类或者在应用之间发送的字节流，作为在数据库记录中的域，或者作为 XML 信息项的集合。这些模型通常提前提供，但是也可以按需要来任意地创建。XML 使用 DTD 或者其它结构化文档模型来把内容与适当的标识指令相联系以使得内容能够以预期的表现形式或风格显示。为了识别的目的，标示语言为文档的每个“元素”或部分添加了标识符。例如，DTD 可以规定文档模型为具有标题，主要段落和几个次要段落。然后标示语言添加被称作“标签”的标识符来指明特定的元素的开始和结束。

根据一个实施例，已知的 XML 大纲被修正和使用以规定 XML 部分的优先级。例如，用户在 XML 大纲中预先定义它何时被创建、其中的哪部分具有最高的优先级。此外，用户可以，例如在用于特定的 XML 文档的请求消息中通过许多常规方法中的任意方法修改 XML 大纲的优先级。

DTD 也可以被修改以允许 DTD (DTD 段) 通过 XML 部分被呈现。DTD 段根据本发明被用于检验正确性并且也可以规定 XML 部分的优先级。替换地，DTD 段如果不被包括在单个的 XML 部分中，可以作为参考被提供并包括在被给定最高优先级的全部的 DTD 中。

图 2 是显示根据本发明的示例性 XML 发送机 200 的构造的框图。XML 发送机 200 可以具体是通用计算系统，例如图 2 中所示的通用计算系统。如图 2 中所示，XML 发送机 200 优选地包括处理器 210 以及诸如数据存储设备 220 的相关存储器，该存储器可以是分布式的或者局部的。处理器 210 可以具体是单个的处理器，或者并行运行的多个局部的或分布式的处理器。数据存储设备 220 和/或只读存储器 (ROM) (未示出) 可操作来存储一条或多条指令，处理器 210 可操作来取回、解释并执行这些指令。

数据存储设备 220 包括可以从存储器取回或者实时产生的文本源 250。因此，文本源 250 可以是诸如数据库或其它文档的预先录入的文本文件，或者实时产生的文档，例如，通过用户从键盘 (未示出) 键入文本信息或者通过语音识别系统 (未示出)。数据存储设备 220 也包括一个或者多个 XML 部分处理 260 以及指示在构建 XML 部分中，文本信息应该如何被分解的划分优先级的 XML 大纲 (或者划分优先级的 DTD 信息) 270。例如，XML 部分处理 260 使用以上所描述的子树方法以及 XML 大纲

270 来确定 XML 文档将如何被分解。然而，其它常规的分段处理也可以被使用，倘若它们能够被修改来结合 XML 大纲和用户设立的优先级使用的话。如上面所解释的，XML 大纲 270 包括用户给每一个 XML 部分规定的优先级。重要地，由 DTD 或 XML 大纲所规定的每个“元素”被用户指配一个优先级。XML 大纲 270 可以被包括在 XML 部分处理 260 中或者可以被处理器 210 调用来帮助确定 XML 部分。因此，XML 发送机 200 将利用识别的 XML 部分处理 260 处理文本源 250 以产生流播的 XML 格式的发送内容（例如，XML 部分）。

图 3 是显示根据本发明的示例性 XML 接收机 300 的构造的框图。XML 接收机 300 可以具体是通用计算系统，例如图 3 中所示的通用计算系统，或者 XML 接收机 300 可以与诸如数字电视（DTV）的另外的设备相集成。XML 接收机 300 包括诸如处理器 310 的某些标准硬件以及诸如数据存储设备 320 的相关的存储器，如以上结合 XML 发送机 200 所讨论的。

数据存储设备 320 包括下面结合图 4 讨论的流播的 XML 部分处理 400 以及划分优先级的 XML 大纲（或者划分优先级的 DTD 信息）270。通常，即使没有接收到较大的 XML 文档的其它 XML 部分，流播 XML 部分处理 400 处理从 XML 发送机 200 接收到的每一个 XML 部分。

图 4 是描述由图 3 的 XML 接收机 300 执行的示范流播 XML 部分处理 400 的流程图。流播 XML 部分处理 400 在步骤 410 和 420 中提取在 XML 部分中接收到的结构（例如 DTD 或者 XML 大纲）信息，并且确定相应的在 XML 接收机 300 中存储的划分优先级的 XML 大纲 270。这个处理一直重复直到确定了特定的优先级的 XML 大纲。在步骤 430 中，接收到的 XML 部分被直接处理，例如，通过显示内容或者将内容存在特定的位置。此后，在步骤 440 中执行测试来确定是否已经接收到了与当前的 XML 文档相关的附加 XML 部分。如果它被确定，在将被处理的当前的 XML 文档中存在附加的 XML 部分的步骤 420 中，使用确定的划分优先级的 XML 大纲，程序控制返回步骤 410 并且以上面所描述的方式继续处理下一个 XML 部分。此外要确定传输是否已经被终止，例如，在步骤 450 中由用户通过终止消息或者其它常规方法。同样，终止可以是强制性的，例如由于噪声，其它干扰或者网络拥塞造成的传输错误/故障。如果确定在正被构造的当前的 XML 文档中没有要被处理的附加 XML 部分，或者传输已经被终止，那么在步骤 460 中，如果可能的话，全部的 XML 文档可以被装配在

一起。此后，在步骤 470 中程序控制终止直到接收到用于处理的附加的 XML 部分。

应当理解，这里所示出和描述的实施例以及变动只是本发明的原理的说明，而且，本领域的技术人员可以不偏离本发明的范围和精神来实现不同的修改。

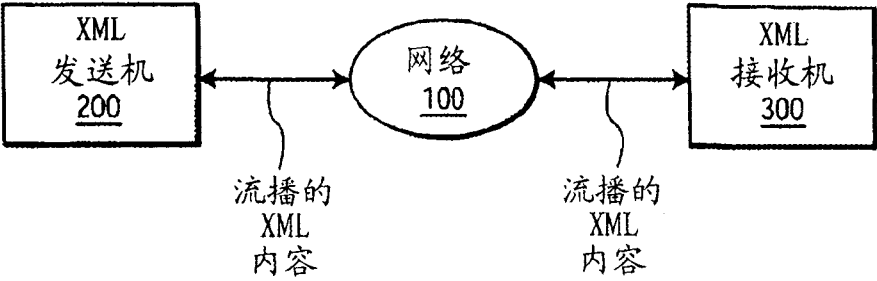


图 1

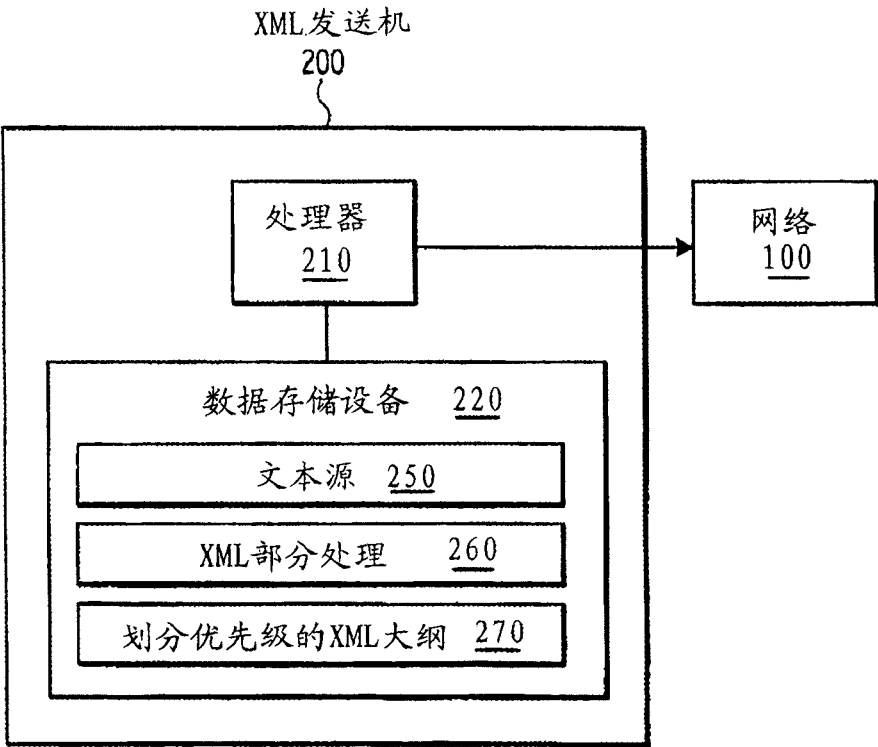


图 2

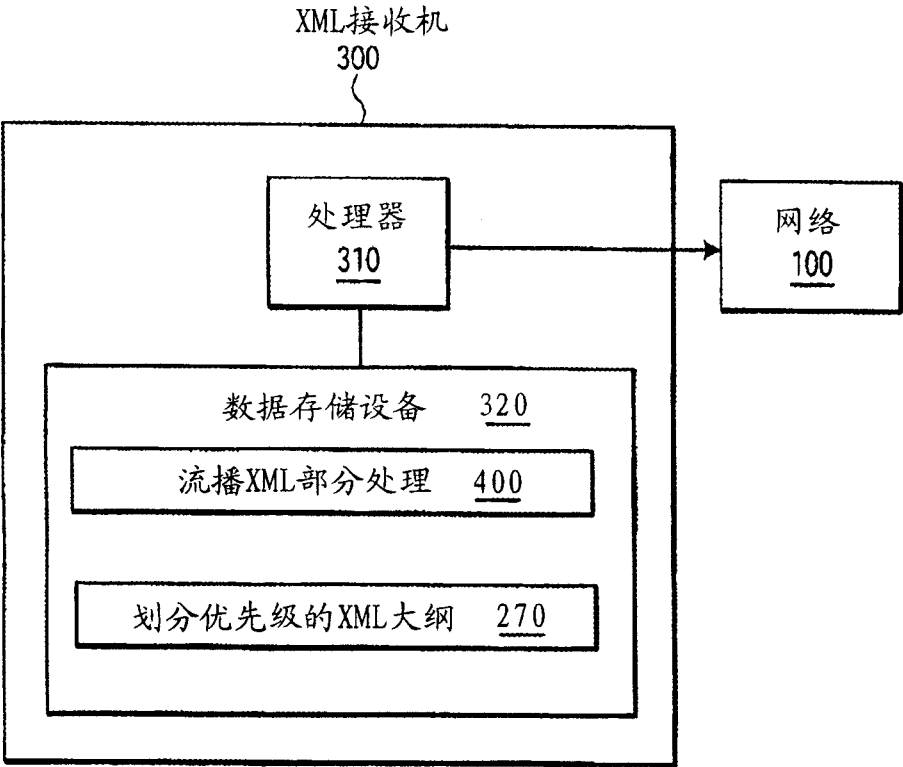


图 3

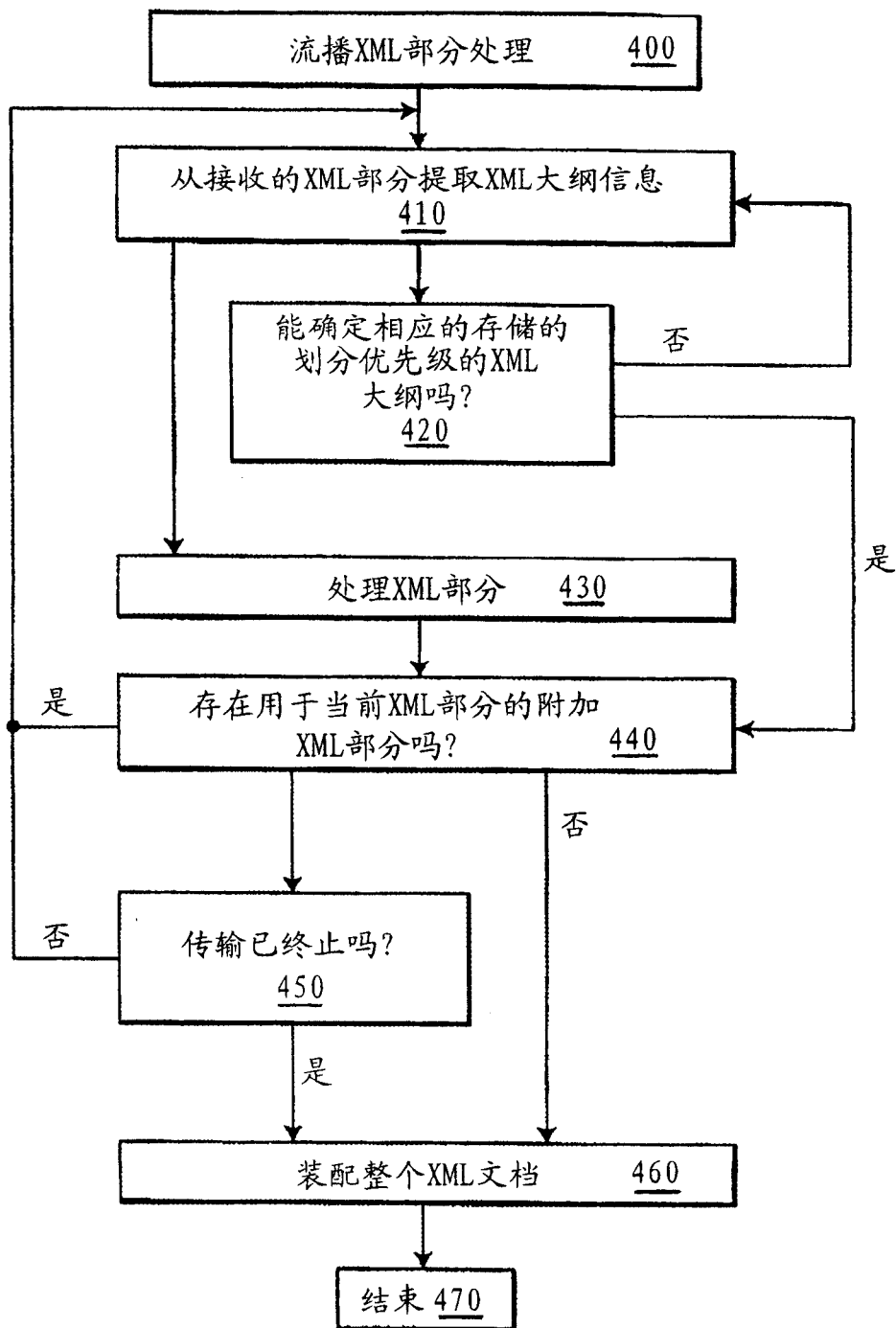


图 4