

(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
2016 年 1 月 21 日 (21.01.2016)



(10) 国际公布号
WO 2016/008347 A1

- (51) 国际专利分类号:
G06F 17/24 (2006.01)
- (21) 国际申请号: PCT/CN2015/081626
- (22) 国际申请日: 2015 年 6 月 17 日 (17.06.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201410341665.9 2014 年 7 月 17 日 (17.07.2014) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [GB/CN]; 开曼群岛大开曼乔治城资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 及
- (71) 申请人 (仅对美国): 刘孙亮 (LIU, Sunliang) [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.); 中国北京市朝阳区西坝河西里 28 号英特公寓 C 座 104 室, Beijing 100028 (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)。

(54) Title: LAYOUT DOCUMENT REARRANGEMENT METHOD AND SYSTEM, AND ELECTRONIC READING TERMINAL

(54) 发明名称: 版式文档的重排方法、系统及电子阅读终端

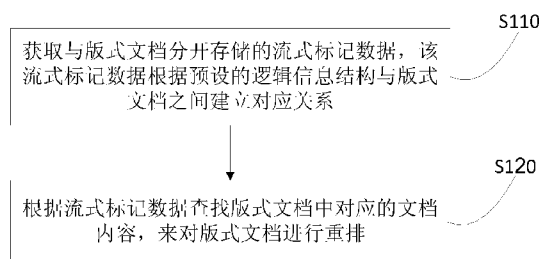


图 1 / FIG. 1

- S110 ACQUIRING STREAM-TYPE FLAG DATA STORED SEPARATELY FROM A LAYOUT DOCUMENT, WHEREIN A CORRELATION IS ESTABLISHED BETWEEN THE STREAM-TYPE FLAG DATA AND THE LAYOUT DOCUMENT ACCORDING TO A PRESET LOGIC INFORMATION STRUCTURE
- S120 ACCORDING TO THE STREAM-TYPE FLAG DATA, SEARCHING FOR THE CORRESPONDING DOCUMENT CONTENT IN THE LAYOUT DOCUMENT TO CONDUCT REARRANGEMENT ON THE LAYOUT DOCUMENT

(57) Abstract: A layout document rearrangement method, which comprises: acquiring stream-type flag data stored separately from a layout document, wherein a correlation is established between the stream-type flag data and the layout document according to a preset logic information structure (S110); and according to the stream-type flag data, searching for the corresponding document content in the layout document to conduct rearrangement on the layout document (S120). A layout document rearrangement system and an electronic reading terminal, wherein by means of the layout document rearrangement system, the flagged stream-type flag data is stored separately from the layout document, and the layout document is parsed according to the stream-type flag data during rearrangement, so that the stream-type flag data may not influence an original document, and therefore, the rearrangement effect and rearrangement efficiency of the layout document can be effectively improved in the case where the original document is not modified or damaged; and at the same time, the flagged stream-type flag data can be shared by multiple users and multiple terminals easily, thereby contributing to the technological upgrade of electronic devices.

(57) 摘要:

[见续页]



WO 2016/008347 A1

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种版式文档的重排方法，其包括：获取与版式文档分开存储的流式标记数据，该流式标记数据根据预设的逻辑信息结构与版式文档之间建立对应关系（S110）；根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排（S120）。一种版式文档的重排系统及电子阅读终端，其将标记后的流式标记数据与版式文档分开存储，在重排时根据流式标记数据来解析版式文档，这样流式标记数据就不会对原始文档造成影响，使得在不修改不破坏原始文档的情况下，可以有效地提高版式文档的重排效果和重排效率；同时，标记后的流式标记数据可以容易地被多用户多终端共享，有助于电子设备的技术升级。

版式文档的重排方法、系统及电子阅读终端

技术领域

本申请涉及数字阅读技术，尤其涉及一种版式文档的重排方法、系统及电子阅读终端。

背景技术

随着互联网的蓬勃发展及硬件水平的不断提升，电子文档正在逐步取代传统的图书和纸质文档。同时，人们的阅读习惯也不再局限于传统的纸质出版物，电子阅读（或称数字阅读）比重正在逐渐变大。由于各种手机、电子书等便携性电子设备的流行，使得人们可以利用生活中的碎片时间进行电子阅读，例如乘坐公交、地铁时就可以阅读电子图书。在巨大的市场需求下，对于电子阅读的信息提供及处理方式也提出了更高的要求。

众所周知，电子文档分为流式文档和版式文档。流式文档的基本单位是字符，它是有序字符的集合，长度为该文件所包含的字符个数。如Word文件就是一种流式文档，其主要记录流式信息，其中也可以加入某些定版的对象（如图像漂浮等）。版式文档作为一种绝对描述方式，在自定义的坐标系中，明确记录每个文档的位置和尺寸等，从而使得文档打印出来的结果和计算机上浏览的结构一致，而且在任何计算机环境下具有显示一致性的特点，由此可以保证真实地重现文档的原貌。如pdf、xps、ceb等文件就是较为典型的版式文档，它们具有版面固定、所见即所得（What you see is what you get, WYSIWYG）等特点，因而非常适合于成文后的文件发布、传播和存档。

流式文档不存在电子阅读的排版障碍，目前已经有成熟的排版引擎可以适用。对于版式文档而言，则经常会因为版面固定而在小屏幕设备下导致阅读不便。如果版式文档的一个页面的内容在设备的屏幕上显示，将受到文字、图像等太小而看不

清楚等诸多限制；若对页面进行放大或缩小操作，将势必影响用户的阅读体验。这就要求电子阅读终端能够突破版式文档显示固定的局限性，以便可以根据版式文档的内容重新进行排版，最终保证用户具有较好的阅读体验。

对于版式文档的重排问题，业界纷纷推出了各种解决方案。实现版式文档重排的现有方案主要有以下两种：

一种现有版式文档的重排方案是：为了补救在各种电子设备上阅读电子文档的需要，在制作相应的版式文档时标记版面的流式显示信息，并把标记数据存储于原始文档中一起发布。这种重排方案中，以版式文档中精确定位的版面描述为基础，在其中附加足够的流式逻辑结构信息用来支持重排、抽取表格结构等流式应用。例如，Adobe在1999年推出的PDF 1.3规范中引入了Logical Structure（逻辑结构），并在2001年推出的PDF 1.4产品中引入了tagged PDF来完善流式信息的表达，之后又在其发布的MARS文档格式中使用XML对这部分信息进行结构化的描述。这种xml标记语言，在理论上可以描述一切格式，如Word新版本的Docx就是基于xml进行描述的。此外，方正阿帕比在2010年发布的CEBXv1.1规范中定义了包含文章、章节、段落、片段、块的多层可嵌套的树状逻辑结构，其中的块通过直接引用版式页面上的版面块或图元（v1.2）来实现数据共享，由此可以支持在移动终端等电子阅读设备上的实时排版和屏幕自适应显示，具体的标准手册及软件可参考方正阿帕比的官方网站（<http://www.apabi.cn/download/index.html>）上的有关介绍。

另一种现有版式文档的重排方案是：在打开一个版式文档时，通过某些预设的算法和规则，对版面信息进行解析，并根据解析的结果，交给排版引擎来进行实时重排，即通过实时排版来进行屏幕自适应显示。这种版式文档的实时重排方法，目前在各种电子阅读终端上得到广泛使用。

上述两种方案均可对版式文档进行重排显示，但它们均存在一定的问题，简述如下：

第一种重排方案中，文档内容、标记数据数据位于同一个文件内，没有标记过流式显示信息的版式电子文档的数据同步可能会存在困难。如果发现原始文档标记有错误，需要再次修改文档，而修改文档时有可能对原始文档造成破坏。尤其在大量文档已经归档的情况下，采用这种方式对文档进行同步可能会引发更多不良后果。

第二种重排方案在打开文档的时候实时地解析该版式文档，电子阅读终端在每次阅读时通过算法实时分析、标记、重排等，因此比较耗时耗电。此外，该重排方案依赖某种算法的可靠性，因而可能存在重排效果不好的问题。

由此可见，现有版式文档的重排技术仍然存在着较大的改进空间，这就有必要提出一种有效提高重排效果和重排效率的版式文档重排的技术方案。

发明内容

针对现有技术存在的缺陷，本申请的目的在于提供一种版式文档的重排方法、系统和电子阅读终端，可以有效地改善重排效果和重排效率。

为解决以上技术问题，本申请提供一种版式文档的重排方法，该方法包括：

获取与版式文档分开存储的流式标记数据，该流式标记数据根据预设的逻辑信息结构与版式文档之间建立对应关系；

根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。

可选地，流式标记数据包括与版式文档的文档内容相对应的逻辑信息，未包括版式文档的实质内容。

可选地，流式标记数据包括版式文档的摘要内容。

可选地，预先查找是否存在与版式文档对应的预处理的流式标记数据；

若是，获取该流式标记数据；

若否，按照预设的流式逻辑信息结构对版式文档进行标记，以获取流式标记数据并进行存储。

可选地，通过算法分析或人工分析或算法分析与人工分析相结合的方式对版式文档进行版面解析，在按照预设的流式逻辑信息结构进行标记后获得相应的流式标记数据。

可选地，流式标记数据以文件或数据库记录的形式外置存储于服务器端或本地。

可选地，通过本地选定的流式逻辑信息结构，根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。

可选地，本地选定的流式逻辑信息结构对应于本地算法实现的、本地预处理的、用户指定的、或最新标记技术标记的流式逻辑信息结构。

可选地，通过本地选定的流式逻辑信息结构确定的流式标记数据与版式文档的对应关系，从流式标记数据中获取全部或部分的流式标记，针对每一流式标记查找到版式文档中的对应文档内容，交由排版引擎重新排版及显示。

与此对应地，本申请同时提供一种版式文档的重排系统，该系统包括：

流式标记析取器，被配置为获取与流式标记数据，该流式标记数据根据预设的逻辑信息结构与版式文档之间建立对应关系；

存储器，被配置为存储流式标记数据，该流式标记数据与版式文档分开存储；

排版引擎，被配置为根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。

此外，本申请还相应提供一种电子阅读终端，可对版式文档进行重排，该电子阅读终端被配置为：

获取与版式文档分开存储的流式标记数据，该流式标记数据根据预设的逻辑信息结构与版式文档之间建立对应关系；以及

根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。

与现有技术相比，本申请利用将流式标记数据外置存储的方式，在不修改不破

坏原始文档的情况下,可以实现有效地提高版式文档的重排效果和重排效率,具体而言:本申请通过对版式文档的实时流式逻辑标记及预处理标记,可以适应版面大小而进行重排显示,这既可以获得较好的排版效果,又可以很好地缩短重排时间;同时,通过版面分析并把版式文档的流式逻辑信息标记外置化,可以解决大量已有缺少流式标记数据的版式文档的重排问题,不需要担心修改对原始文档造成的破坏及其后续文档泛滥不统一的问题;此外,本申请中版式文档只需标记一次,即可供多用户多终端共享,从整个系统来讲,其不仅耗电耗时少,也有助于进行技术升级。

附图说明

通过阅读下文优选实施方式的详细描述,各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的,而并不认为是对本申请的限制。而且在整个附图中,用相同的参考符号来表示相同的部件。在以下附图中:

图 1 示出了根据本申请一个实施例的版式文档的重排方法的流程图;

图 2 为根据图 1 所示版式文档的重排方法的较为完整的实例;

图 3 示出了根据本申请一个实施例的版式文档的重排系统的组成框图;

图 4 示出了根据本申请一个实施例的电子阅读终端的组成框图。

具体实施方式

本申请的以下实施例阐述了一种版式文档的重排方法、系统和电子阅读终端,其利用流式标记数据外置存储的方式,在不修改不破坏原始文档的情况下,可以提高版式文档的重排效果和重排效率,以下结合附图与具体实施例来进行详细描述。

为了更好地理解本申请实施例的技术方案,先对有关的术语进行阐述。

1、逻辑信息结构

逻辑信息结构指文档组织信息的逻辑描述关系,比如对标题、段落、公式、表格、注释等结构化的信息来指定这些要素之间的逻辑结构关系(如一个图是居中的,

它的图题是什么内容等), 这些逻辑结构关系构成一个个的有序排列。

本申请实施例中的逻辑信息结构, 特别地指定外置化存储文档和原始文档二者之间的关系, 例如: 它可以在外置化存储文档中指定段落, 段落里面有多少span(不可拆分文字, 比如不希望在显示的时候被断行的字符串), 每个span有什么文字。当然也可以没有span, 而直接说明这个段落里面都有什么文字, 每个文字对应着原始版式文档的里面坐标或者文档二进制流偏移位置。

由此可见, 本申请实施例的逻辑信息结构与纯版式文档信息逻辑结构有所不同。在纯版式文档逻辑信息结构中, 它仅描述一个个字符、图像、图形有多大, 在页面的哪个坐标位置显示。这种纯版式文档逻辑信息结构对于整篇文档而言, 因为强调的是版面上的呈现而不是逻辑信息, 由此可能导致文档各部分之间的顺序关系可能是全乱的, 也可能是部分乱的。

2、自适应呈现版面信息

上述的逻辑信息结构描述的是文档结构及版面的逻辑信息, 按照该逻辑信息结构来识别并标记文档就可得到相应的流式标记数据。换言之, 流式标记数据中包括版式文档中的文档结构信息和/或版式文档中文档版面信息进行标记的结果, 其中的文档版面信息是一种自适应呈现版面信息。

根据这种自适应呈现版面信息, 电子阅读终端可以重构整个文档的版面, 最终使得版式文档重排显示的结果与电子阅读终端屏幕的大小相匹配。例如, 自适应呈现版面信息中描述版式文档的这部分是标题, 那部分是段落(段落里面可能有1000个文字), 等等。根据这种描述, 在不同阅读设备上可以根据屏幕的大小调整显示的相应内容: 例如, 在电脑上可能一屏显示900个文字, 即前述段落自适应显示为一屏多一点; 在手机上可能一屏只显示100个文字, 即前述段落自适应显示为10屏。但无论怎么自适应显示, 某些内容不会显示到不同屏里面去, 例如“and”这个词可以是一个span, 它无论如何自适应都不能被断行。

按照上述的逻辑信息结构可以解析版式文档而得到流式标记数据,也可以识别流式标记数据来重构版式文档。由此得到本申请实施例的构思:获取与版式文档分开存储的流式标记数据,该流式标记数据根据预设的逻辑信息结构与版式文档之间建立对应关系;根据流式标记数据查找版式文档中对应的文档内容,来对版式文档进行重排。以下进一步结合附图来进行详细阐述。

在下面的描述中阐述了很多具体细节以便于充分理解本申请。但是本申请能够以很多不同于在此描述的其它方式来实施,本领域技术人员可以在不违背本申请内涵的情况下做类似推广,因此本申请不受下面公开的具体实施例的限制。参见图 1,其示出根据本申请版式文档的重排方法的一个实施例。该方法的构思侧重于流式标记数据的存储外置化,以便根据流式标记数据与版式文档之间的对应关系,来确定版式文档中的文档结构信息以及文档版面信息,从而较好地实现版式文档的重排显示。这一实施例采用了流式标记数据外置存储的方式,可以在不修改不破坏原始文档的前提下,有效改善版式文档的重排效率和重排效果,以下进一步详细描述。

本申请实施例的版式文档可以指整个的版式文档,也可以指版式文档中的一页或几页。这种版式文档采用绝对描述方式,其在自定义的坐标系中明确记录每个文档的位置和尺寸等,从而可使得文档打印出来的结果和计算机上浏览的结构一致,实现所见即所得的效果。

如上所述,本申请实施例的流式标记数据包括版式文档中文档结构信息和/或版式文档中文档版面信息。文档结构信息包括文档的章节信息、各章节内部内容顺序以及内容块中各图元顺序等。文档版面信息包括版式文档相应版面重排时决定图元及其它要素最终呈现效果的信息、图元自身或者内容块自身的版面信息、以及同一内容块中各图元之间或者各内容块之间的关系,例如指定图片的文字衬托方式或者多个内容块的分栏信息等。此处的版面重排指的是由于版面大小或版面内容发生变化时,根据一定规则重新组织版面

中的各图元及其它要素，以形成版面展现的结果。此外，本申请实施例的流式标记数据还可以包括阅读线索信息。除了上述文档结构信息提供的阅读顺序外，所述流式标记数据还可以根据具体需要提供的额外的阅读顺序信息。所述阅读线索信息是提供给用户的可选的阅读顺序信息。

值得注意的是，本申请实施例的流式标记数据包括与版式文档的文档内容相对应的逻辑信息，而未包括版式文档的实质内容。特别地，这种流式标记数据可以包括版式文档的摘要信息，例如基于 MD5 或 SHA 算法得到的版式文档的摘要信息。按照这种方式以预定的流式逻辑信息结构来对版式文档进行标记，由此获得的流式标记数据可以实现流式标记数据与版式文档之间的强关联。

本申请实施例通过以一定的逻辑标记算法来分析需要重排的版式文档，可以有效地析取版式文档中那些字可以组成词，哪些词可以组成段落，哪些字是上标或者下标，哪些对象是图，哪些文字是图题等等，由此就能够更好地对版式文档进行充分、有效的描述，最终有利于版式文档的重排显示。与此不同，通常的版式文档中仅描述页面中每个文字、图形或图像的位置，而没有从逻辑上描述这些对象之间的关系，由此会影响版式文档的重排效率及显示效果。

本申请实施例的流式标记数据的具体逻辑信息结构可以参照某些现有技术标准确定，如方正阿帕比的有关技术手册确定；也可以重新自定义有关的逻辑信息结构，以保证具有较好的兼容性为好，不再赘述。

请参见图 1，为本申请版式文档的重排方法的实施例。该实施例通过利用外置存储的流式标记数据对版式文档进行重排的方式，可以在不修改不破坏原始文档的情况下，有效地提高版式文档的重排效果和重排效率。下面对本申请实施例所述版式文档的重排方法的具体技术方案进行详细的描述。

在步骤 S110 处，获取与版式文档分开存储的流式标记数据。该流式标记

数据根据预设的逻辑信息结构与版式文档之间建立对应关系，如采用原始文档摘要方式及其他数据库键值对来建立对应关系。换言之，这些流式标记数据是按照预设的流式逻辑信息结构对版式文档进行版面解析的结果；通过这些流式标记数据，可以查找版式文档中的相应文档内容的结构信息及版面信息。由于该流式标记数据记录有丰富的版式文档中的文档结构信息、文档版面信息等，因而可以较好地与原始的版式文档（简称原始文档）之间建立强关联的对应关系。这样，根据这些优化的流式标记数据不仅可以查找原始文档中的对应文档内容，同时可以确定这些文档内容的结构信息及版面信息，由此有助于方便地对整个版式文档进行重排显示。

在步骤 S120 处，根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。通过该步骤可以完成版式文档的重排，最终得到与电子阅读终端等显示界面相适应的、重排的版式文档（简称重排文档）。由于本申请实施例中的流式标记数据对版式文档进行了更充分、有效的描述，因而有助于提高版式文档的重排显示效果。

为了深入理解本申请实施例的版式文档的重排方法的技术方案，进一步对步骤 S110 和步骤 S120 详细叙述如下。

在步骤 S110，获取以流式逻辑信息结构与版式文档相关联的流式标记数据，该流式标记数据与版式文档分开存储。本申请实施例通过预处理方式或实时标记版式文档的方式来获得流式标记数据，但无论是预处理的流式标记数据还是实时标记的流式标记数据，都可以与版式文档分开进行存储，之后通过查找与该流式标记数据相对应的版式文档内容来对原始文档进行重排。

该步骤 S110 可以包含两个具体步骤：一是查找是否有与版式文档关联的预处理过的流式标记数据，二是在没有预处理的流式标记数据的情况下对版式文档进行实时的标记。具体而言，该步骤 S110 的基本流程是：预先查找是否存在与版式文档对应的预处理的流式标记数据；存在的话，获取该流式

标记数据；不存在的话，则按照预设的流式逻辑信息结构识别版式文档并进行标记，以获取流式标记数据并进行存储。具体地，这些流式标记数据可以以文件或数据库记录的形式外置存储于服务器端（例如，云端服务器）或本地，这样可以方便地做到与版式文档分开存储。

在步骤 S110 中，以预处理方式或实时标记方式获取流式标记数据的基本过程是：按照预设的流式逻辑信息结构对版式文档进行版面解析并进行标记，得到的所有标记信息的集合构成流式标记数据。可以理解的是，在根据上述方式标记版式文档时，可通过算法分析、人工分析、或算法分析与人工分析相结合的方式对版式文档进行版面解析，并按照预设的流式逻辑信息结构进行标记，以获得相应的流式标记数据。这些流式标记数据与原始文档之间按照预设的流式逻辑信息结构建立了强关联的对应关系。

下面列出一个 pdf 版式文档及其流式标记数据，来作为流式标记数据外置化的具体例子来阐述本申请所述方法的技术方案。

(1) 原始的版式文档如下：

a.pdf

2 0 obj<< /Type /Page

/Contents 30R

... >>

endobj

3 0 obj<< /Length ... >>

stream ...

...(Here is some text1) ...//对应一个词，坐标x = 100, y = 100, 内容是：你好，

...(Here is some text2) ...//对应一个词，坐标x = 110, y = 200, 内容是：标题

...(Here is some text3) ...//对应一个词，坐标x = 130, y = 100, 内容是：中国。

... endstream

endobj

(2) 流式标记数据如下:

a.mark,

<SrcDoc> "xxx"</SrcDoc> //xxx 对应 a.pdf 整个文档的摘要

<Head> // 一个标题

<obj = 3, offset = xxx, length = xxx> //对应内容是: 标题

</Head>

<P> // 一个段落

<obj = 3, offset = xxx, length = xxx> //对应内容是: 你好,

<obj = 3, offset = xxx, length = xxx> //对应内容是, 中国。

</P>

由此可以看出, 该版式文档的标记实例中, 按照预设的流式逻辑信息结构对版式文档进行版面解析, 并将解析结果的标记数据集合作为版式文档的流式标记数据。其中对各部分文档内容标记有丰富的流式结构信息及版面信息, 因而可以较好地对应于原始的版式文档, 最终可以方便地用来进行重排显示。

需要指出的是, 本实施例所述的流式标记数据也可不局限于上述标记实例的描述方式, 它完全可以采用二进制的描述、xml 描述等等。实际上, 本申请实施例并不侧重某种文件格式的具体描述标准, 因而对如何形成所述的流式标记数据不再进行详细说明。

在步骤 S120 中, 根据流式标记数据查找版式文档中对应的文档内容, 并识别所述文档内容的结构信息及版面信息 (例如确定某些文档内容是文字、图形或表格, 并确定它们之间的关系, 据此确定相应的排版方案), 来对版式文档进行重排。经过重排的版式文档可以在移动终端等电子设备上进行实时排版和屏幕自适应显示, 由此可以有效提高用户的阅读体验。这里的屏幕

自适应显示包括获取设备的屏幕尺寸信息，并根据屏幕尺寸信息自适应对文档内容进行排版。

此处对版式文档的重排包括：在版面大小或版面内容发生变化时，根据一定规则重新组织版面中的各图元及其它要素，形成版面展现结果的过程。本申请实施例对排版引擎可不做特定要求，现在市面上的成熟的排版引擎（如webkit）均可作为选择对象，当然用户也可自主开发其它合适的排版引擎，在此不再展开说明。

如前所述，本申请实施例通过预设的流式逻辑信息结构建立起了流式标记数据与版式文件之间的对应关系。根据该流式逻辑信息结构，可以对版式文档进行预先标记或实时标记，由此得到相应的流式标记数据。对版式文档进行预先标记或实时标记可以理解成是对版式文档进行解析的过程。根据该流式逻辑信息结构，也可以根据标记好的流式标记数据来重构版式文档，具体流式标记数据中的文档结构信息及版面信息来查找版式文档中的对应文档内容，并根据当前的显示要求（例如字体大小的要求、根据屏幕尺寸自适应显示的要求）对这些文档内容进行排版显示。简单地讲，重构版式文档可以理解成是一个对流式标记数据进行反解析的过程。

由于本申请实施例通过某种流式逻辑信息结构在版式文档与流式标记数据之间建立了对应关系，因此在重新排版时应当保证标记时的流式逻辑信息结构与重排时的流式逻辑信息结构保持匹配。可以理解的是，标记版式文档时的预设流式逻辑信息结构与重排时的流式逻辑信息结构实际上可能存在并不匹配的情况，因此排版引擎在选定流式逻辑信息结构时，一般应优先考虑对应于本地算法实现的、本地预处理的、用户指定的、或最新标记技术标记的流式逻辑信息结构。

本实施例中通过上述规则在本地选定某一流式逻辑信息结构之后，如果该流式逻辑信息结构与标记版式文档时的逻辑信息结构能够匹配，则在版式

文档重排时就可通过该本地选定的流式逻辑信息结构，来对版式文档进行解析，即根据流式标记数据查找版式文档中的对应文档内容，并进一步识别这些文档内容的结构信息及版面信息，最终来实现对版式文档进行重排。

由此可见，对于预先处理的流式标记数据，如果本地选定的流式逻辑信息结构与标记时的流式逻辑信息结构相匹配，则重排时流式标记数据与版式文档之间可以建立起有效的对应关系，它与标记时流式标记数据与版式文档的对应关系一致。这样，在版式文档重排时就可从流式标记数据中获取全部或部分的流式标记（记录），由此可针对每一流式标记查找到版式文档中的对应文档内容，并识别这些文档内容的结构信息及版面信息，之后就可交由排版引擎重新排版及显示。

可以理解的是，针对本地选定的流式逻辑信息结构一般需要确定某种解析算法对版式文档进行重排。这些解析算法可以有不同的方案，但由于本申请并不侧重如何实时解析某种系统算法，故此对相应的解析算法亦不具体展开描述。

参见图 2，为根据图 1 所示版式文档的重排方法的一个较完整的实例。该实例主要包括以下步骤 210~步骤 250，以下简要进行描述。

步骤S210，接收版式文档。所述版式文档可以根据当前的显示条件（例如根据显示屏幕的大小等因素）进行重排。

步骤S220，查找是否存在与版式文档对应的流式标记数据。

查找是否存在与版式文档对应的流式标记数据也就是判断是否存在预处理的流式标记数据，该流式标记数据为对版式文档进行流式标记预处理而得到的。得到的流式标记数据与版式文档可以分开进行存储。若存在预处理的流式标记数据，则进入步骤S230，若不存在预处理的流式标记数据进入步骤S240。

步骤 S230，获取该预处理的流式标记数据，作为版式文档的重排时的解

析要素，来实现对版式文档的重排。

步骤 S240，实时标记该版式文档，以获取该流式标记数据并进行存储，实现对版式文档的流式标记数据进行更新。

步骤 S250，根据获取的流式标记数据，来查找版式文档中的对应文档内容，并识别所述文档内容的结构信息及版面信息，来实现对版式数据进行重排。

图 2 作为图 1 所示版式文档的重排方法的一个完整实例，可以清晰地显示本申请所阐述技术方案的基本脉络，其中的大多数细节已在图 1 中进行阐述。如图 2 中所描述的内容有不详尽之处，请进一步参见对图 1 的描述部分。

通过以上对图 1 和图 2 的描述可知，本申请针对现有版式文档重排技术的缺点，采用外部存储流式标记数据的方式，即通过分析并把版面的流式逻辑信息标记外置化，可以解决大量已有缺少流式标记数据的重排问题，不需要担心修改对原始文档造成的破坏及其后续文档泛滥不统一的问题。同时，本申请通过对版面的实时流式逻辑标记及预处理标记，对版式文档进行了更充分、有效的描述，因而既可以获得较好的排版效果，也可以很好地缩短重排时间。此外，本申请采用流式标记数据外置存储的方式，在流式标记数据中记录标记类型、电子阅读系统版本、服务器识别系统版本、人工识别版本等内容，可使得版式文档只需标记一次，即可供多用户多终端共享，因而也有助于对电子阅读系统进行技术升级。

需要指出的是，本申请中的流式标记数据一般可通过某种算法来标记，标记后需再将标记结果外置存储，以方便下一次使用。当然，这种标记过程也可以通过人工方式或者算法与人工相结合的方式进行标记。但不管是采用算法标记，还是采用人工标记，抑或人工与算法相结合的方式来标记版式文档，本申请实施例都应该按照某种指定的标准来获得这些流式标记数据。但是，本申请实施例并不拘泥于某种特定的标准，本申请实施例中的流式标记数据可以采用很多种不同

的逻辑信息描述标准，它们既可以用 xml 描述，也可以用二进制描述，而这些标记结果还可以直接存储在数据库或云端服务器，在此不展开说明。

以上对版式文档的重排方法进行了详细的描述。在此基础上，本申请还相应地构设了版式文档的重排系统（以下简称系统），以下进行详细的描述。

顺便指出的是，本实施例系统中如有描述不尽之处，请参见前文方法部分的描述内容；同样地，前述方法部分中如涉及到系统，也可引见以下系统部分的描述内容。

参见图 3，示出了根据本申请一个实施例的版式文档的重排系统的组成框图。该版式文档的重排系统（简称系统）300 由流式标记析取器 310、存储器 320、排版引擎 330 及流式标记预处理器 340 等部分构成，其通过流式标记数据外置存储的方式，在不修改不破坏原始文档的情况下，可以实现有效地提高版式文档的重排效果和重排效率，以下进一步对该系统 300 的各部分结构及功能进行描述。

如图 3 所示，该 300 具有流式标记析取器 310，其可以获取与流式标记数据，该流式标记数据根据预设的逻辑信息结构与版式文档之间建立对应关系，即该流式标记数据是按照预设的流式逻辑信息结构标记对版式文档进行版面解析的结果。具体地，流式标记析取器 310 包括流式标记查找模块 311、流式标记读取模块 312 及实时标记引擎模块 313 等，其中：流式标记查找模块 311 被配置为预先查找是否存在与版式文档对应的流式标记数据；流式标记读取模块 312，被配置为存在与版式文档对应的流式标记数据时，获取该流式标记数据；实时标记引擎模块 313，被配置为不存在与版式文档对应的流式标记数据时，按照预设的流式逻辑信息结构对版式文档进行标记，以获得流式标记数据并进行存储。所述的实时标记引擎模块 313 具体可以配置在本地或服务器端，其可以通过算法分析或人工分析或算法与人工相结合的方式对版式文档进行版面解析，在按照预设的流式逻辑信息结构进行标记后，

获得相应的流式标记数据。

如图 3 所示，该系统 300 还具有存储器 320，它可以是云存储器或本地存储器，可以以文件或数据库记录的形式存储流式标记数据，这些流式标记数据与版式文档分开存储。在本实施例中，流式标记数据是按照预设的流式逻辑信息结构对版式文档进行版面解析的标记结果，其中记录有丰富的流式文档结构信息及相应文档内容的结构信息与版面信息，因而可以较好地对应于原始的版式文档，有利于方便地重构版式文档，即对版式文档进行重排。

该系统 300 还具有排版引擎 330，根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。具体地，排版引擎 330 可通过本地选定的流式逻辑信息结构，根据流式标记数据查找版式文档中对应的文档内容，以便来对版式文档进行重排。重排的基本过程是，通过本地选定的流式逻辑信息结构确定的流式标记数据与版式文档之间的对应关系，从流式标记数据中获取和版式文档对应的流式标记（记录），针对每一流式标记查找到版式文档中的对应内容，之后交由排版引擎重新排版。

进一步地，该系统 300 还可以具有流式标记预处理器 340，它被配置在本地或服务器端，可以对版式文档预先进行标记，并在得到流式标记数据后进行存储。通常地，流式标记预处理器 340 获得的预处理标记可在服务器端用算法对文档进行处理，也可通过人工方式或算法与人工结合的方式对文档进行标记。通常的，预处理标记情况下，可提供一定的软件工具给厂商进行预装。

在本实施例中，上述流式标记数据可以通过标记预处理而得，也可以通过实时标记处理而得。但无论哪种方式，所得到的流式标记数据都应版式文档分开进行存储。

本实施例中，对版式文档进行流式标记进行预处理的基本过程为：首先对版式文档进行版面解析，其中的版面分析不限于算法分析、人工分析等等；

然后，对版面信息流式标记后的结果外置存储，其中的存储方式不限于云存储、数据库或本地外置文件存储。由此，通过这种预处理标记方式，就可以使得流式标记与原始的版式文档相分离。

本实施例中，对版式文档进行流式标记实时处理的过程与此类似，主要在于标记处理的时间及主体存在差别，在此不再展开说明。顺便提及的是，在实时标记版式文档来获得流式标记数据时，如涉及到某种文件格式的描述标准以及实时解析该标准的相应算法问题，请详细参考习知技术中的相关文献资料，在此不再赘述。

参考图3，同时结合图1和图2，上述版式文档的重排系统300的基本工作过程如下：

(1)对版式文档进行版面解析，对版面信息流式标记后的结果外置存储；其中版面分析的算法不限于算法分析、人工分析等，存储的方式不限于以文件或数据库记录形式进行云存储、本地外置文件存储。

(2)系统中的电子阅读器在显示的时候可以选定自己认为最优的流式逻辑信息结构，这些选定的流式逻辑信息结构可以是本地算法实现的、本地预处理的文档、用户指定的、最新标记技术标记的流式逻辑信息结构。

(3)版式文档重排的过程为：先通过某种对应关系获取到和原始文档对应的流式标记数据/文档，如采用原始文档摘要方式（不限于 md5、sha 等各种摘要方式）及其他数据库键值对来指定对应关系；再从流式标记数据结构获取到的一个个流式标记，它们记录着流式标记及其和原始文档相关内容的对应关系，这种对应不限于文档位置偏移、对象号等等；最后通过流式标记数据找到原始文档中的对应内容，直接交给排版引擎排版显示。

(4)如果原始文档找不到相应外置的流式逻辑标记数据，则通过实时版面解析系统进行分析、标记，之后交给排版引擎排版并把标记结果进行外置

存储。

这样，本申请的版式文档的重排系统通过利用流式标记数据外置存储的方式，在不修改不破坏原始文档的情况下，可以实现有效提高版式文档的重排效果和重排效率。

针对上述版式文档的重排系统，此处需要进一步补充说明以下问题：

其一，系统中可能需要识别多种不同流式逻辑信息结构的版式文档，在不识别某种流式逻辑信息结构的情况下，则认为该流式逻辑信息结构不是本地所要的流式逻辑信息结构。如果该流式逻辑信息结构是新的版本，可以在该流式逻辑信息结构里面描述版本号、是否经预处理等信息。此外，系统也可以相应通知升级阅读器版本，以便最终来识别理解该流式逻辑信息结构。

其二，系统在选定流式逻辑信息结构时，可以选用较好的对应关系是md5，它通过标记数据来唯一对应一个原始文档。具体到内容的对应关系上，其可以使用文档位置偏移、对象号等来进行描述，其详细情况可参照前述标记实例所述。

其三、系统对排版引擎没有特定要求，现在市面上的成熟排版引擎（如webkit）均可视情况选用，当然也可以自行开发排版引擎。总之，排版引擎的问题不是本申请的重点，实施本申请的技术方案时可以认为排版引擎为约定俗成的即可。

其四，系统对实时标记引擎也没有特殊要求，只要实时标记引擎处理速度较快，效果可以接受即可。具体实施时，实时标记引擎一般通过算法来实现，好处在于算法可以不停地升级，不断地改进速度和效果。考虑到服务器端有更强大的集群计算和历史数据统计、机器学习、人工智能等能力，因而实时标记引擎也可以考虑做在服务器端，这样计算速度不成问题，还

可以借助大数据、机器学习等方式来获取更好的标记结果，只是需要用网络来传输标记结果罢了。在服务器端同时设置实时标记引擎的情况下：若网络较好，阅读器终端可以使用服务器端的标记数据；若网络不好，阅读器终端则可以使用自身的轻量级标记系统的标记数据。

可以理解的是，这种版式文档的重排系统可以有不同的应用实例，它可以是某一网络系统，也可以是某一种单机设备（例如手机、平板电脑等移动智能终端），以下提供一种电子阅读终端作为产品实例来进行具体说明。

为方便起见，本申请中对于版式文档的重排系统及电子阅读终端等结构而言，分别用模块、装置或单元等词汇来表示类似的功能结构，以下简要进行描述。

请参见图 4，表示本申请电子阅读终端实施例的组成框图。该电子阅读终端 400 可对版式文档进行重排，其具有流式标记数据获取单元 410 及版式文档重排单元 420，其中：流式标记数据获取单元 410 可获取与版式文档分开存储的流式标记数据，该流式标记数据根据预设的逻辑信息结构与版式文档之间建立对应关系。换言之，该流式标记数据按照预设的流式逻辑信息结构标记对版式文档进行版面解析的结果；版式文档重排单元 420 根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。

上述的流式标记数据获取单元 410 可以预先查找是否存在与版式文档对应的预处理的流式标记数据：存在与版式文档对应的流式标记数据时，获取该流式标记数据；不存在与版式文档对应的流式标记数据时，按照预设的流式逻辑信息结构对版式文档进行标记，以获得流式标记数据并进行存储。这样，无论版式文档是否已经预先标记，该电子阅读终端 400 都能够有效地进行重排，之后进行显示。

以上对本申请的有关实施例进行了详细的描述,其中的版式文档重排技术方案与现有技术方案相比具有明显的优势,简要总结如下。

如前所述,现有的版式文档重排技术主要采用两种方案:一种是直接获取原始版式文档,实时版本分析、理解、标记、重排;另一种是对原始文档进行流式标记存入原始文档,显示的时候从原始文档获取流式标记进行重排。这两种现有技术方案都存在一定的缺陷,具体理由请参见前文所述。

与此不同,本申请提出的版式文档的重排方法、系统及电子阅读终端具有明显的优势,其克服决了上述两种现有技术方案在重排效果、效率方面的缺陷,解决了文档覆盖不全、文档同步难等问题。这种版式文档的重排方法、系统及电子阅读终端具有但不仅限于以下特点:

(1) 不限于以云存储、数据库或本地外置文件存储等方式外置存储流式标记数据,因而不会破坏原始文档,有助于进行专业的预处理、技术升级及数据更新等。

(2) 通过某种对应关系来获取流式标记数据,并对版式文档进行解析。这种对应关系包括但不限于各种摘要或者其他方式,它们可以在外置存储器中存储和原始文档的指定关系,因而无需通过修改原始文档的方式来把原始文档和流式逻辑信息结构强关联。

(3) 流式逻辑信息结构只描述逻辑信息,不在逻辑信息结构中存储实质文档信息。通过流式逻辑信息结构中的某种对应关系,如指定文档偏移、对象号等等来和原始文档内容产生对应关系,具有数据量小的特点。

(4) 流式逻辑结构信息是对版式文档进行版面分析的结果,这些流式逻辑结构信息不限于用算法分析或人工分析的方式来标记这些版式文档,具体的标记形式及手段较为多样。

本申请虽然以较佳实施例公开如上,但其并不是用来限定本申请,任何本领域技术人员在不脱离本申请的精神和范围内,都可以做出可能的变动和修改,因

此本申请的保护范围应当以本申请权利要求所界定的范围为准。

在一个典型的配置中，计算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。

1、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系统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暂存电脑可读媒体 (transitory media)，如调制的数据信号和载波。

2、本领域技术人员应明白，本申请的实施例可提供为系统、系统或计算机程序产品。因此，本申请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。

权利要求书

1、一种版式文档的重排方法，其特征在于，包括：

获取与版式文档分开存储的流式标记数据，该流式标记数据根据预设的逻辑信息结构与版式文档之间建立对应关系；

根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。

2、如权利要求1所述的版式文档的重排方法，其特征在于，流式标记数据包括与版式文档的文档内容相对应的逻辑信息，未包括版式文档的实质内容。

3、如权利要求2所述的版式文档的重排方法，其特征在于，流式标记数据包括版式文档的摘要内容。

4、如权利要求1所述的版式文档的重排方法，其特征在于，预先查找是否存在与版式文档对应的预处理的流式标记数据；

若是，获取该流式标记数据；

若否，按照预设的流式逻辑信息结构对版式文档进行标记，以获取流式标记数据并进行存储。

5、如权利要求4所述的版式文档的重排方法，其特征在于，通过算法分析或人工分析或算法分析与人工分析相结合的方式对版式文档进行版面解析，在按照预设的流式逻辑信息结构进行标记后获得相应的流式标记数据。

6、如权利要求4所述的版式文档的重排方法，其特征在于，流式标记数据以文件或数据库记录的形式外置存储于服务器端或本地。

7、如权利要求1所述的版式文档的重排方法，其特征在于，通过本地选定的流式逻辑信息结构，根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。

8、如权利要求7所述的版式文档的重排方法，其特征在于，本地选定的流式逻辑信息结构对应于本地算法实现的、本地预处理的、用户指定的、或最新标记技术

标记的流式逻辑信息结构。

9、如权利要求7所述的版式文档的重排方法，其特征在于，通过本地选定的流式逻辑信息结构确定的流式标记数据与版式文档的对应关系，从流式标记数据中获取全部或部分的流式标记，针对每一流式标记查找到版式文档中的对应文档内容，交由排版引擎重新排版及显示。

10、一种版式文档的重排系统，其特征在于，包括：

流式标记析取器，被配置为获取与流式标记数据，该流式标记数据根据预设的逻辑信息结构与版式文档之间建立对应关系；

存储器，被配置为存储流式标记数据，该流式标记数据与版式文档分开存储；

排版引擎，被配置为根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。

11、如权利要求10所述的版式文档的重排系统，其特征在于，所述流式标记析取器包括流式标记查找模块、流式标记读取模块及实时标记引擎模块，其中：

所述流式标记查找模块，被配置为预先查找是否存在与版式文档对应的流式标记数据；

所述流式标记读取模块，被配置为存在与版式文档对应的流式标记数据时，获取该流式标记数据；

所述实时标记引擎模块，被配置为不存在与版式文档对应的流式标记数据时，按照预设的流式逻辑信息结构对版式文档进行标记，以获得流式标记数据并进行存储。

12、如权利要求11所述的版式文档的重排系统，其特征在于，所述实时标记引擎模块，被配置为通过算法分析或人工分析或算法分析与人工分析相结合的方式对版式文档进行版面解析，在按照预设的流式逻辑信息结构进行标记后获得相应的流式标记数据。

13、如权利要求11所述的版式文档的重排系统，其特征在于，所述实时标记引擎模块被配置在本地或服务器端。

14、如权利要求10所述的版式文档的重排系统，其特征在于，所述存储器为云存储器或本地存储器，以文件或数据库记录的形式外置存储流式标记数据。

15、如权利要求10所述的版式文档的重排系统，其特征在于，所述排版引擎被配置为，通过本地选定的流式逻辑信息结构，根据流式标记数据查找版式文档中对应的文档内容，以便来对版式文档进行重排。

16、如权利要求15所述的版式文档的重排系统，其特征在于，所述排版引擎被配置为，通过本地选定的流式逻辑信息结构确定的流式标记数据与版式文档的对应关系，从流式标记数据中获取全部或部分的流式标记，针对每一流式标记查找到版式文档中的对应文档内容，交由排版引擎重新排版及显示。

17、如权利要求10所述的版式文档的重排系统，其特征在于，包括流式标记预处理器，被配置为对版式文档预先进行标记，在得到流式标记数据后进行存储。

18、如权利要求17所述的版式文档的重排系统，其特征在于，所述流式标记预处理器被配置在本地或服务器端。

19、一种电子阅读终端，可对版式文档进行重排，其特征在于，所述电子阅读终端被配置为：

获取与版式文档分开存储的流式标记数据，该流式标记数据根据预设的逻辑信息结构与版式文档之间建立对应关系；以及

根据流式标记数据查找版式文档中对应的文档内容，来对版式文档进行重排。

20、如权利要求19所述的电子阅读终端，其特征在于，所述电子阅读终端被配置为：

预先查找是否存在与版式文档对应的预处理的流式标记数据；

存在与版式文档对应的流式标记数据时，获取该流式标记数据；以及

不存在与版式文档对应的流式标记数据时,按照预设的流式逻辑信息结构对版式文档进行标记,以获得流式标记数据并进行存储。

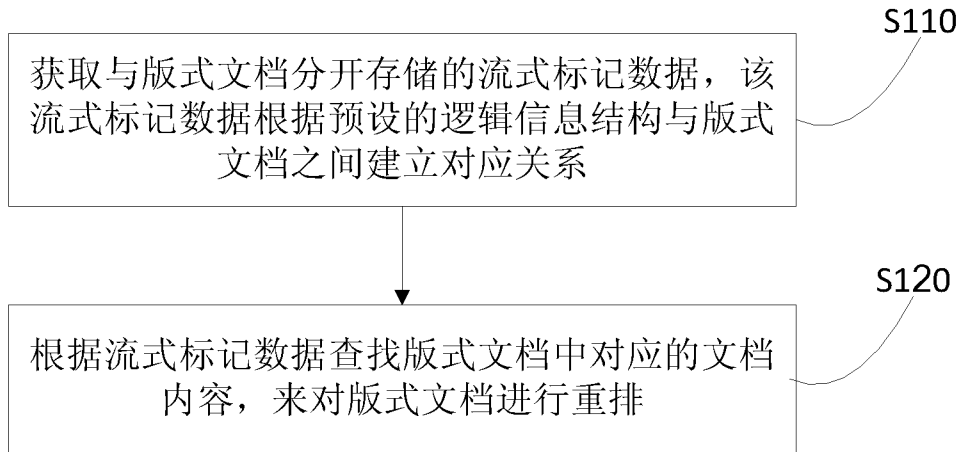


图 1

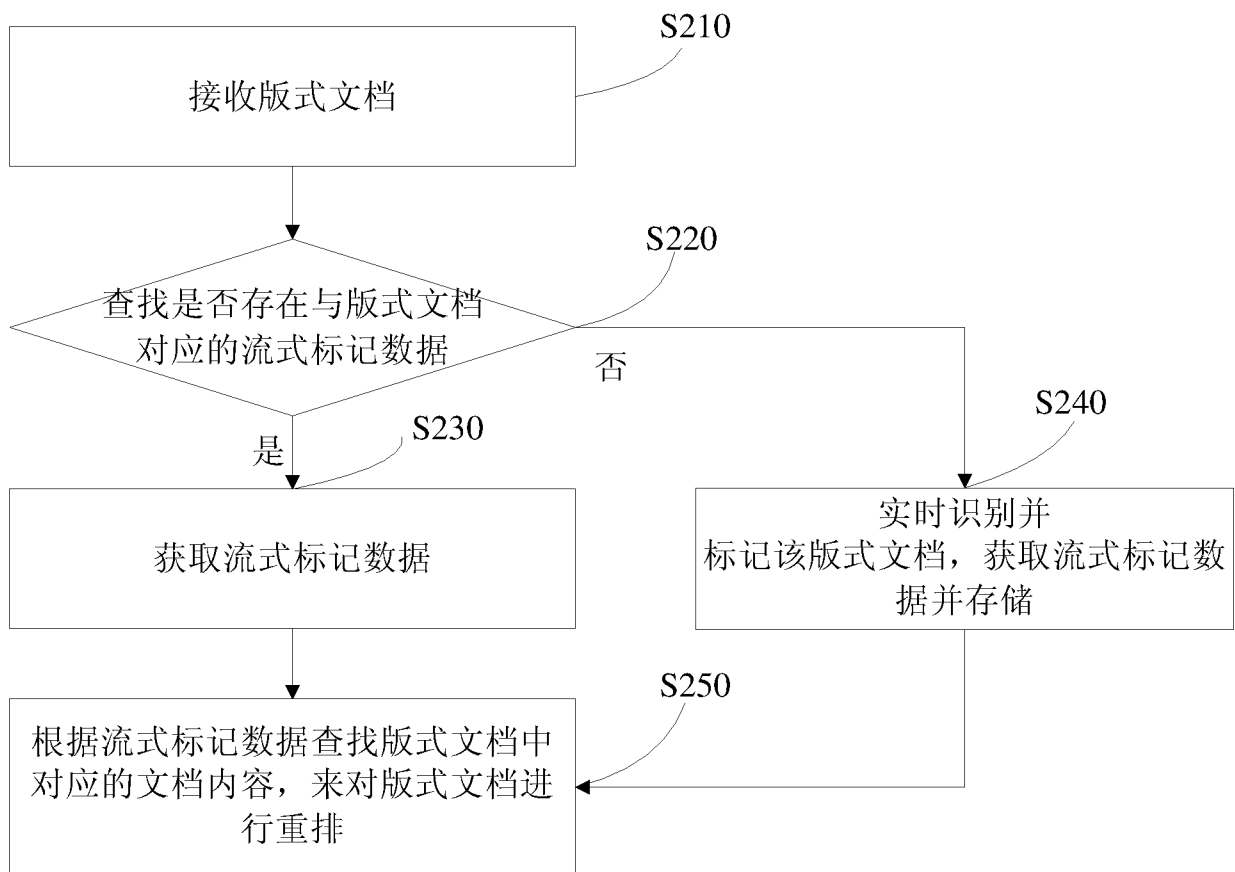


图 2

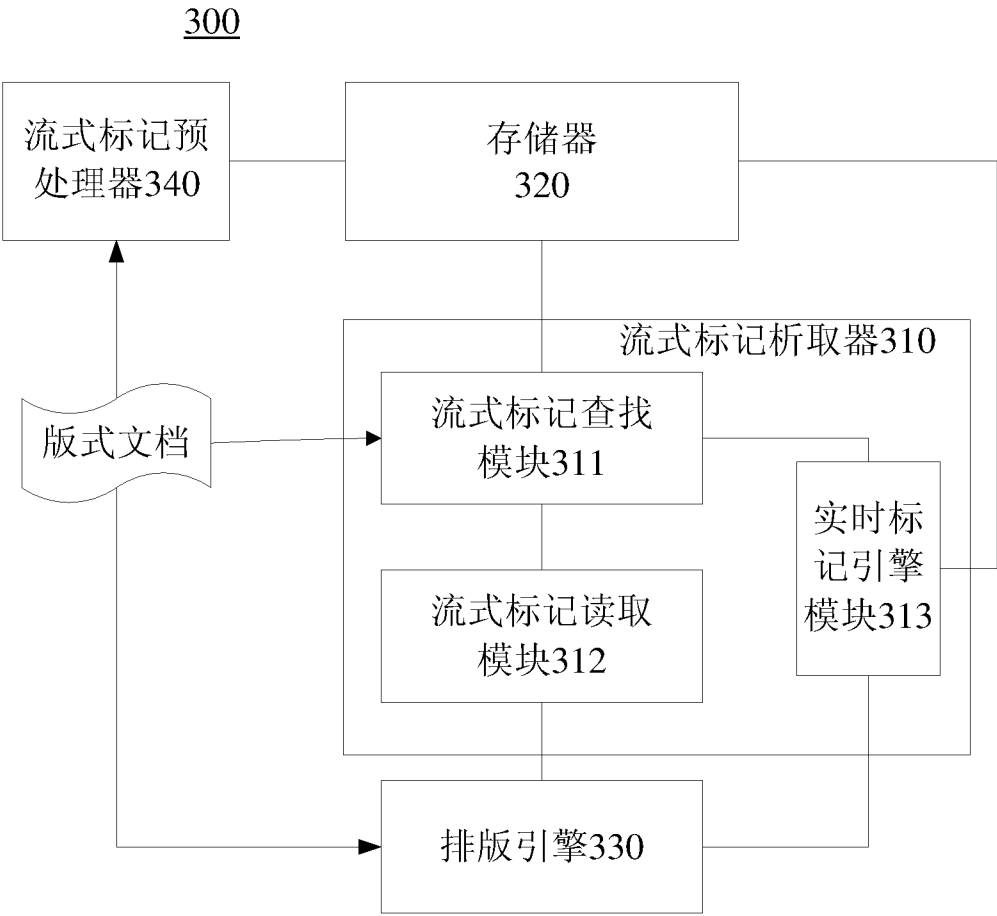


图 3

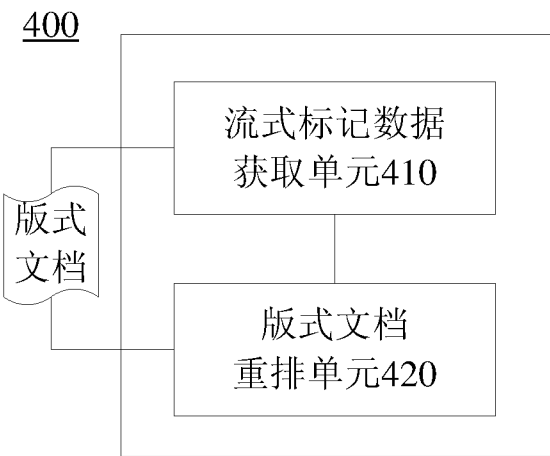


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/081626

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/24 (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)

WPI, EPODOC, CNKI, CNPAT, IEEE: reformat, logic, size, XML, layout, structure, formed, streaming, position, relation, correspond+, screen

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 101887413 A (PEKING UNIVERSITY FOUNDER GROUP CORP. et al.), 17 November 2010 (17.11.2010), abstract, and description, paragraphs 28, 32-38 and 40-42	1-20
A	CN 101308488 A (PEKING UNIVERSITY FOUNDER GROUP CORP. et al.), 19 November 2008 (19.11.2008), the whole document	1-20
A	CN 101923723 A (HANVON TECHNOLOGY CO., LTD.), 22 December 2010 (22.12.2010), the whole document	1-20
A	CN 102591849 A (PEKING UNIVERSITY FOUNDER GROUP CORP. et al.), 18 July 2012 (18.07.2012), the whole document	1-20

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 19 August 2015 (19.08.2015)	Date of mailing of the international search report 15 September 2015 (15.09.2015)
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 WANG, Jing Telephone No.: (86-10) 62413685

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2015/081626

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101887413 A	17 November 2010	None	
CN 101308488 A	19 November 2008	EP 2291010 A1	02 March 2011
		WO 2009146657 A1	10 December 2009
		JP 2011523133 A	04 August 2011
		US 2011087959 A1	14 April 2011
CN 101923723 A	22 December 2010	None	
CN 102591849 A	18 July 2012	None	

国际检索报告

国际申请号

PCT/CN2015/081626

A. 主题的分类

G06F 17/24(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))

WPI, EPODOC, CNKI, CNPAT, IEEE: 排版, 版式, 重排, 流式, 逻辑, 结构, 屏幕, 尺寸, XML, layout, structure, formed, streaming, position, relation, correspond+, screen

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 101887413 A (北大方正集团有限公司等) 2010年 11月 17日 (2010 - 11 - 17) 摘要, 说明书第28, 32-38, 40-42段	1-20
A	CN 101308488 A (北大方正集团有限公司等) 2008年 11月 19日 (2008 - 11 - 19) 全文	1-20
A	CN 101923723 A (汉王科技股份有限公司) 2010年 12月 22日 (2010 - 12 - 22) 全文	1-20
A	CN 102591849 A (北大方正集团有限公司等) 2012年 7月 18日 (2012 - 07 - 18) 全文	1-20

☐ 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

2015年 8月 19日

国际检索报告邮寄日期

2015年 9月 15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)
北京市海淀区蓟门桥西土城路6号
100088 中国

传真号 (86-10)62019451

授权官员

王晶

电话号码 (86-10)62413685

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2015/081626

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	101887413	A	2010年 11月 17日	无			
CN	101308488	A	2008年 11月 19日	EP	2291010	A1	2011年 3月 2日
				WO	2009146657	A1	2009年 12月 10日
				JP	2011523133	A	2011年 8月 4日
				US	2011087959	A1	2011年 4月 14日
CN	101923723	A	2010年 12月 22日	无			
CN	102591849	A	2012年 7月 18日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)