



(12)发明专利申请

(10)申请公布号 CN 108256064 A

(43)申请公布日 2018.07.06

(21)申请号 201810040291.5

(22)申请日 2018.01.16

(71)申请人 北京三快在线科技有限公司
地址 100000 北京市海淀区北四环西路9号
2106-030

(72)发明人 朱敏 刘志成 张俊浩

(74)专利代理机构 北京超凡志成知识产权代理
事务所(普通合伙) 11371
代理人 宋南

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

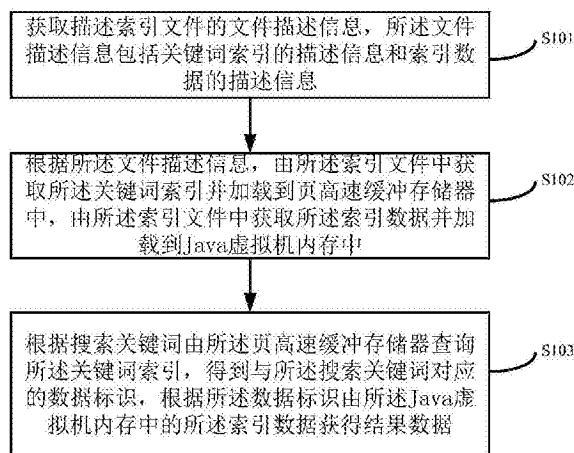
权利要求书2页 说明书8页 附图2页

(54)发明名称

一种数据搜索方法和装置

(57)摘要

本申请提供了一种数据搜索方法和装置,包括:获取描述索引文件的文件描述信息,所述文件描述信息包括关键词索引的描述信息和索引数据的描述信息;所述关键词索引表征关键词与数据标识之间的映射关系;根据所述文件描述信息,由所述索引文件中获取所述关键词索引并加载到页高速缓冲存储器中,由所述索引文件中获取所述索引数据并加载到Java虚拟机内存中;根据搜索关键词由所述页高速缓冲存储器查询所述关键词索引,得到与所述搜索关键词对应的数据标识,根据所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。



1. 一种数据搜索方法,其特征在于,该方法包括:

获取描述索引文件的文件描述信息,所述文件描述信息包括关键词索引的描述信息和索引数据的描述信息;所述关键词索引包括关键词与数据标识之间的映射关系;

根据所述文件描述信息,由所述索引文件中获取所述关键词索引并加载到页高速缓冲存储器中,由所述索引文件中获取所述索引数据并加载到Java虚拟机内存中;

根据搜索关键词由所述页高速缓冲存储器查询所述关键词索引,得到与所述搜索关键词对应的数据标识,根据所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

2. 如权利要求1所述的方法,其特征在于,所述文件描述信息包括所述关键词索引和所述索引数据分别在所述索引文件中的起始位置和数据长度;

所述根据所述文件描述信息,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中,由所述索引文件中获取所述索引数据并加载到所述Java虚拟机内存中,包括:

根据所述关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中;

根据所述索引数据在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述索引数据并加载到所述Java虚拟机内存中。

3. 如权利要求2所述的方法,其特征在于,所述关键词索引包括第一关键词索引;

根据所述关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中,包括:

根据所述第一关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述第一关键词索引并加载到所述页高速缓冲存储器中。

4. 如权利要求3所述的方法,其特征在于,所述关键词索引还包括第二关键词索引;

根据所述关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中,还包括:

根据所述索引数据在所述索引文件中的所述起始位置和所述数据长度,跳过所述索引数据;

根据所述第二关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述第二关键词索引并加载到所述页高速缓冲存储器中。

5. 如权利要求3或4所述的方法,其特征在于,所述根据搜索关键词由所述页高速缓冲存储器查询所述关键词索引,得到与所述搜索关键词对应的数据标识,根据所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据,包括:

根据所述搜索关键词由所述页高速缓冲存储器查询所述关键词索引;

若在第一关键词索引中查询到所述搜索关键词,则确定与所述搜索关键词对应的所述数据标识;

基于所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

6. 一种数据搜索装置,其特征在于,该装置包括:

获取模块,用于获取描述索引文件的文件描述信息,所述文件描述信息包括关键词索引的描述信息和索引数据的描述信息;所述关键词索引包括关键词与数据标识之间的映射

关系；

加载模块,用于根据所述文件描述信息,由所述索引文件中获取所述关键词索引并加载到页高速缓冲存储器中,由所述索引文件中获取所述索引数据并加载到Java虚拟机内存中；

搜索模块,用于根据搜索关键词由所述页高速缓冲存储器查询所述关键词索引,得到与所述搜索关键词对应的数据标识,根据所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

7.如权利要求6所述的装置,其特征在于,所述文件描述信息包括所述关键词索引和所述索引数据分别在所述索引文件中的起始位置和数据长度；

所述加载模块具体用于：

根据所述关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中；

根据所述索引数据在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述索引数据并加载到所述Java虚拟机内存中。

8.如权利要求7所述的装置,其特征在于,所述关键词索引包括第一关键词索引；

所述加载模块还用于：

根据所述第一关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述第一关键词索引并加载到所述页高速缓冲存储器中。

9.如权利要求8所述的装置,其特征在于,所述搜索模块具体用于：

根据所述搜索关键词由所述页高速缓冲存储器查询所述关键词索引；

若在第一关键词索引中查询到所述搜索关键词,则确定与所述搜索关键词对应的所述数据标识；

基于所述数据标识由所述Java虚拟机内存中的所述索引数据搜索数据。

10.如权利要求8或9所述的装置,其特征在于,所述搜索模块具体用于：

根据所述搜索关键词由所述页高速缓冲存储器查询所述关键词索引；

若在第一关键词索引中查询到所述搜索关键词,则确定与所述搜索关键词对应的所述数据标识；

基于所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

一种数据搜索方法和装置

技术领域

[0001] 本申请涉及搜索技术领域,具体而言,涉及一种数据搜索方法和装置。

背景技术

[0002] 随着互联网技术的兴起,信息呈爆炸式增长,用户希望能够从海量的信息中获取自己关注的内容,由此而诞生了多种搜索引擎。搜索引擎发展伊始主要针对网页搜索,而网页也是传统互联网的主要信息载体。然而近年来,随着互联网技术的发展,用于可以方便的通过移动设备访问互联网,获取服务,由此而兴起了一批线上-线下(Online-to-Offline, O2O)本地生活化服务,极大的方便了人们的生活。与网页不同,O2O行业的信息描述载体往往都具有多个文本域,从不同的角度尝试对服务进行具体的描述,例如,餐饮服务商家可能会包括商家名称、商家注册公司名称、品牌名称、商家所处商圈、商家地址、商家主营产品、商家营业时间等等字段,这类描述性文本字段有时多达五十个以上。在搜索过程中,不仅要求若干字段命中,也会要求按照一些字段进行排序。随着索引数据量的增多,单个文档的字段数量增多,搜索引擎会面临搜索性能瓶颈,也会面临资源瓶颈。造成搜索引擎性能瓶颈的原因,是在没有增加存储设备的基础上由于数据量增加导致的,数据量增加之后会占用更多的内存,在进行数据检索时,检索效率降低。

发明内容

[0003] 有鉴于此,本申请的目的在于提供一种数据搜索方法和装置,用于解决现有技术中由于数据量增加导致的检索效率低的问题。

[0004] 第一方面,本申请实施例提供了一种数据搜索方法,该方法包括:

[0005] 获取描述索引文件的文件描述信息,所述文件描述信息包括关键词索引的描述信息和索引数据的描述信息;所述关键词索引包括关键词与数据标识之间的映射关系;

[0006] 根据所述文件描述信息,由所述索引文件中获取所述关键词索引并加载到页高速缓冲存储器中,由所述索引文件中获取所述索引数据并加载到Java虚拟机内存中;

[0007] 根据搜索关键词由所述页高速缓冲存储器查询所述关键词索引,得到与所述搜索关键词对应的数据标识,根据所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

[0008] 可选地,所述文件描述信息包括所述关键词索引和所述索引数据分别在所述索引文件中的起始位置和数据长度;

[0009] 所述根据所述文件描述信息,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中,由所述索引文件中获取所述索引数据并加载到所述Java虚拟机内存中,包括:

[0010] 根据所述关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中;

[0011] 根据所述索引数据在所述索引文件中的所述起始位置和所述数据长度,由所述索

引文件中获取所述索引数据并加载到所述Java虚拟机内存中。

[0012] 可选地,所述关键词索引包括第一关键词索引;

[0013] 根据所述关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中,包括:

[0014] 根据所述第一关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述第一关键词索引并加载到所述页高速缓冲存储器中。

[0015] 可选地,所述关键词索引还包括第二关键词索引;

[0016] 根据所述关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中,还包括:

[0017] 根据所述索引数据在所述索引文件中的所述起始位置和所述数据长度,跳过所述索引数据;

[0018] 根据所述第二关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述第二关键词索引并加载到所述页高速缓冲存储器中。

[0019] 可选地,所述根据搜索关键词由所述页高速缓冲存储器查询所述关键词索引,得到与所述搜索关键词对应的数据标识,根据所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据,包括:

[0020] 根据所述搜索关键词由所述页高速缓冲存储器查询所述关键词索引;

[0021] 若在所述第一关键词索引中查询到所述搜索关键词,则确定与所述搜索关键词对应的所述数据标识;

[0022] 基于所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

[0023] 第二方面,本申请实施例提供了一种数据搜索装置,该装置包括:

[0024] 获取模块,用于获取描述索引文件的文件描述信息,所述文件描述信息包括关键词索引的描述信息和索引数据的描述信息;所述关键词索引包括关键词与数据标识之间的映射关系;

[0025] 加载模块,用于根据所述文件描述信息,由所述索引文件中获取所述关键词索引并加载到页高速缓冲存储器中,由所述索引文件中获取所述索引数据并加载到Java虚拟机内存中;

[0026] 搜索模块,用于根据搜索关键词由所述页高速缓冲存储器查询所述关键词索引,得到与所述搜索关键词对应的数据标识,根据所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

[0027] 可选地,所述文件描述信息具体包括所述关键词索引和所述索引数据分别在所述索引文件中的起始位置和数据长度;

[0028] 所述加载模块具体用于:

[0029] 根据所述关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中;

[0030] 根据所述索引数据在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述索引数据并加载到所述Java虚拟机内存中。

[0031] 可选地,所述关键词索引包括第一关键词索引;

[0032] 所述加载模块还用于:

[0033] 根据所述第一关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述第一关键词索引并加载到所述页高速缓冲存储器中。

[0034] 可选地,所述搜索模块具体用于:

[0035] 根据所述搜索关键词由所述页高速缓冲存储器查询所述关键词索引;

[0036] 若在第一关键词索引中查询到所述搜索关键词,则确定与所述搜索关键词对应的所述数据标识;

[0037] 基于所述数据标识由所述Java虚拟机内存中的所述索引数据搜索数据。

[0038] 可选地,所述搜索模块具体用于:

[0039] 根据所述搜索关键词由所述页高速缓冲存储器查询所述关键词索引;

[0040] 若在第一关键词索引中查询到所述搜索关键词,则确定与所述搜索关键词对应的所述数据标识;

[0041] 基于所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

[0042] 本申请提供的数据搜索方法和装置,通过将索引数据加载在Java虚拟机内存中、将关键词索引加载在页高速缓冲存储器中,减少了内存的消耗,进而在用户检索数据时,可以根据用户的搜索关键词从页高速缓冲存储器中快速获取数据标识,根据确定的数据标识从索引数据中搜索数据,提高了检索效率,同时也减轻了检索系统对内存的要求,使得排序、分类、函数查询等搜索中非常常见的功能的处理时间更短。

[0043] 为使本申请的上述目的、特征和优点能更明显易懂,下文特举较佳实施例,并配合所附附图,作详细说明如下。

附图说明

[0044] 为了更清楚地说明本申请实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本申请的某些实施例,因此不应被看作是对范围的限定,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他相关的附图。

[0045] 图1为本申请一实施例提供的一种数据搜索方法的第一种流程示意图;

[0046] 图2为本申请一实施例提供的一种数据搜索方法的第二种流程示意图;

[0047] 图3为本申请一实施例提供的一种数据搜索装置的结构示意图;

[0048] 图4为本申请实施例提供的一种计算设备的结构示意图。

具体实施方式

[0049] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合本申请实施例中附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。通常在此处附图中描述和示出的本申请实施例的组件可以以各种不同的配置来布置和设计。因此,以下对在附图中提供的本申请的实施例的详细描述并非旨在限制要求保护的本申请的范围,而是仅仅表示本申请的选定实施例。基于本申请的实施例,本领域技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0050] 本申请实施例提供了一种数据搜索方法,如图1所示,该方法包括以下步骤:

[0051] S101,获取描述索引文件的文件描述信息,所述文件描述信息包括关键词索引的描述信息和索引数据的描述信息;所述关键词索引包括关键词与数据标识之间的映射关系;

[0052] 具体地,索引文件可以为全文搜索引擎(Lucene)提供的索引文件,该索引文件一般为将等多个文件整合在一起形成的,例如,将关键词索引文件、索引数据文件等整合在一起,关键词索引可以为正排序索引或倒排序索引,索引数据可以为正排索引数据;关键词索引的描述信息为关键词索引在索引文件中的起始位置和数据长度,索引数据的描述信息为索引数据的在索引文件中的起始位置和数据长度;数据标识可以为存储数据的文档的标识,如,文档的名称等。

[0053] 在具体实施中,索引文件的格式可以是cfs格式,索引文件的文件描述信息可以是cfe格式的,在构建索引数据时,一般会建立document=>field value的映射表,本申请对此不予限制。

[0054] S102,根据所述文件描述信息,由所述索引文件中获取所述关键词索引并加载到页高速缓冲存储器中,由所述索引文件中获取所述索引数据并加载到Java虚拟机内存中;

[0055] 可选地,在执行步骤S102中,所述根据所述文件描述信息,由所述索引文件中获取所述关键词索引并加载到页高速缓冲存储器中,由所述索引文件中获取所述索引数据并加载到Java虚拟机内存中,包括:

[0056] 根据所述文件描述信息中的关键词索引在所述索引文件中的起始位置和数据长度,由所述索引文件中获取所述关键词索引并加载到页高速缓冲存储器中;

[0057] 根据所述文件描述信息中的索引数据在所述索引文件中的起始位置和数据长度,由所述索引文件中获取所述索引数据并加载到Java虚拟机内存中。

[0058] 在具体实施中,在获取到索引文件和索引文件的文件描述信息后,创建流对象,基于流对象将索引文件中的内容加载到页高速缓冲存储器(PageCache)中或者加载到Java虚拟机内存中。以下进行详细介绍。

[0059] 关键词索引包括第一关键词索引和第二关键词索引,第一关键词索引一般为倒排序索引,以方便进行检索,第二关键词索引可以为存储字段索引,一般为数据标识与数据片段之间的映射关系;在加载第一关键词索引时,根据所述第一关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述第一关键词索引并加载到所述页高速缓冲存储器中。

[0060] 在加载索引数据时,根据所述索引数据在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述索引数据并加载到所述Java虚拟机内存中。

[0061] 在加载第二关键词索引时,根据所述索引数据在所述索引文件中的所述起始位置和所述数据长度,跳过所述索引数据;

[0062] 根据所述第二关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述第二关键词索引并加载到所述页高速缓冲存储器中。

[0063] 具体地,在加载索引数据时,根据索引数据的文件描述信息中的索引子数据在索引数据中的起始位置和数据长度,由索引数据中获取索引子数据并加载到Java虚拟机内存中。

[0064] 在具体实施中,流对象在根据文件描述信息中的第一关键词索引在所述索引文件

中的起始位置和数据长度,由所述索引文件中获取所述第一关键词索引并加载到页高速缓冲存储器中后,可以通过文件描述信息确定索引文件中位于第一关键词索引之后的内容是否为索引数据,在确定索引文件中位于第一关键词索引之后的内容为索引数据后,从索引文件的描述信息中获取索引数据的数据长度,将流对象中的数据前进参数值设置为索引数据的数据长度,这样,流对象可以直接跳过索引数据,根据文件描述信息中的第二关键词索引在索引文件中的起始位置和数据长度,由索引文件中获取第二关键词索引并加载到页高速缓冲存储器中。

[0065] 在将第一关键词索引和第二关键词索引加载到页高速缓冲存储器中后,上述流对象可以指定下一个要读取的字节在索引文件中的位置,例如,将下一个要读取的字节设置为索引文件开始的第一个字节,这样等同与重新打开了索引文件,也可以重新创建流对象,本申请对比不予限制。

[0066] 流对象基于索引文件的文件描述信息判断索引文件当前内容是否为索引数据,若当前内容不是索引数据,则获取当前加载内容的数据长度,将流对象的数据前进参数值设置为等于上述数据长度,这样可以跳过当前的加载内容,直接加载当前加载内容之后的索引数据,将索引数据加载到Java虚拟机内存中。在将索引数据加载到Java虚拟机内存中时,基于索引数据的文件描述信息中的索引子数据的起始位置和数据长度,由索引数据中获取索引子文件并依次将索引子数据加载到Java虚拟机内存中。

[0067] S103,在搜索数据时,根据搜索关键词由所述页高速缓冲存储器查询所述关键词索引,得到与所述搜索关键词对应的数据标识,根据所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据;

[0068] 可选地,在执行步骤S103时,在搜索数据时,根据搜索关键词由所述页高速缓冲存储器查询所述关键词索引,得到与所述搜索关键词对应的数据标识,根据所述数据标识由所述Java虚拟机内存中的所述索引数据搜索数据,通过步骤S201-S204实现,具体包括以下步骤:

[0069] S201,根据所述搜索关键词由所述页高速缓冲存储器查询所述关键词索引;

[0070] S202,若在所述第一关键词索引中查询到所述搜索关键词,则确定与所述搜索关键词对应的所述数据标识;

[0071] 可选地,若在所述第一关键词索引中未查询到所述搜索关键词,则查询所述第二关键词索引;若在所述第二关键词索引中查询到所述搜索关键词,则确定与所述搜索关键词对应的数据片段。

[0072] S203,基于所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

[0073] 在具体实施时,若索引数据为学生的成绩,若用户希望搜索2017级前100名学生的各科成绩及住址,搜索关键词可以设置为年级字段取值为2017、正排打字段为成绩的前100名、返回字段为住址,根据年级字段取值为2017搜索关键词索引,在第一关键词索引中查询到2017后,获取与2017对应的数据标识,也就是存储数据的文档的文档标识,基于获取的文档标识,查询Java虚拟机内存中的索引数据,获取所有符合年级为2017的学生的各科成绩,排序并找出前100名,基于确定的文档标识和住址字段查询页高速缓冲存储器中的第二关键词索引,从第二关键词索引中获取对应文档标识中的住址字段对应的数据片段。

[0074] 本申请提供的数据搜索方法,通过将索引数据加载在Java虚拟机内存中、将关键词索引加载在页高速缓冲存储器中,减少了内存的消耗,进而在用户检索数据时,可以根据用户的搜索关键词从页高速缓冲存储器中快速获取数据标识,根据确定的数据标识从索引数据中搜索数据,提高了检索效率,同时也减轻了检索系统对内存的要求,使得排序、分类、函数查询等搜索中非常常见的功能的处理时间更短。

[0075] 本申请实施例提供一种数据搜索装置,如图3所示,该装置包括:

[0076] 获取模块31,用于获取描述索引文件的文件描述信息,所述文件描述信息包括关键词索引的描述信息和索引数据的描述信息;所述关键词索引包括关键词与数据标识之间的映射关系;

[0077] 加载模块32,用于根据所述文件描述信息,由所述索引文件中获取所述关键词索引并加载到页高速缓冲存储器中,由所述索引文件中获取所述索引数据并加载到Java虚拟机内存中;

[0078] 搜索模块33,用于在搜索数据时,根据搜索关键词由所述页高速缓冲存储器查询所述关键词索引,得到与所述搜索关键词对应的数据标识,根据所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

[0079] 可选地,所述文件描述信息具体包括所述关键词索引和所述索引数据分别在所述索引文件中的起始位置和数据长度;

[0080] 所述加载模块32具体用于:

[0081] 根据所述关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述关键词索引并加载到所述页高速缓冲存储器中;

[0082] 根据所述索引数据在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述索引数据并加载到所述Java虚拟机内存中。

[0083] 可选地,所述关键词索引包括第一关键词索引;

[0084] 所述加载模块32还用于:

[0085] 根据所述第一关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述第一关键词索引并加载到所述页高速缓冲存储器中。

[0086] 可选地,所述关键词索引还包括第二关键词索引;

[0087] 所述加载模块32还用于:

[0088] 根据所述索引数据在所述索引文件中的所述起始位置和所述数据长度,跳过所述索引数据;

[0089] 根据所述第二关键词索引在所述索引文件中的所述起始位置和所述数据长度,由所述索引文件中获取所述第二关键词索引并加载到所述页高速缓冲存储器中。

[0090] 可选地,所述搜索模块33具体用于:

[0091] 根据所述搜索关键词由所述页高速缓冲存储器查询所述关键词索引;

[0092] 若在所述第一关键词索引中查询到所述搜索关键词,则确定与所述搜索关键词对应的所述数据标识;

[0093] 基于所述数据标识由所述Java虚拟机内存中的所述索引数据获得结果数据。

[0094] 可选地,所述搜索模块33还用于:

[0095] 根据所述搜索关键词由所述页高速缓冲存储器查询所述关键词索引;

[0096] 若在所述第一关键词索引中未查询到所述搜索关键词,则查询所述第二关键词索引;

[0097] 若在所述第二关键词索引中查询到所述搜索关键词,则确定与所述搜索关键词对应的数据片段。

[0098] 对应于图1中的数据搜索方法,本发明实施例还提供了一种计算机设备,如图4所示,该设备包括存储器1000、处理器2000及存储在该存储器1000上并可在该处理器2000上运行的计算机程序,其中,上述处理器2000执行上述计算机程序时实现上述数据搜索方法的步骤。

[0099] 具体地,上述存储器1000和处理器2000能够为通用的存储器和处理器,这里不做具体限定,当处理器2000运行存储器1000存储的计算机程序时,能够执行上述数据搜索方法,从而解决现有技术中由于数据量增加导致的检索效率低的问题,通过将索引数据加载在Java虚拟机内存中、将关键词索引加载在页高速缓冲存储器中,减少了内存的消耗,进而在用户检索数据时,可以根据用户的搜索关键词从页高速缓冲存储器中快速获取数据标识,根据确定的数据标识从索引数据中搜索数据,提高了检索效率,同时也减轻了检索系统对内存的要求,使得排序、分类、函数查询等搜索中非常常见的功能的处理时间更短。

[0100] 在本申请所提供的实施例中,应该理解到,所揭露装置和方法,可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,又例如,多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些通信接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0101] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0102] 另外,在本申请提供的实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。

[0103] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备)执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0104] 应注意到:相似的标号和字母在下面的附图中表示类似项,因此,一旦某一项在一个附图中被定义,则在随后的附图中不需要对其进行进一步定义和解释,此外,术语“第一”、“第二”、“第三”等仅用于区分描述,而不能理解为指示或暗示相对重要性。

[0105] 最后应说明的是:以上所述实施例,仅为本申请的具体实施方式,用以说明本申请的技术方案,而非对其限制,本申请的保护范围并不局限于此,尽管参照前述实施例对本申

请进行了详细的说明,本领域的普通技术人员应当理解:任何熟悉本技术领域的技术人员在本申请揭露的技术范围内,其依然可以对前述实施例所记载的技术方案进行修改或可轻易想到变化,或者对其中部分技术特征进行等同替换;而这些修改、变化或者替换,并不使相应技术方案的本质脱离本申请实施例技术方案的精神和范围。都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应所述以权利要求的保护范围为准。

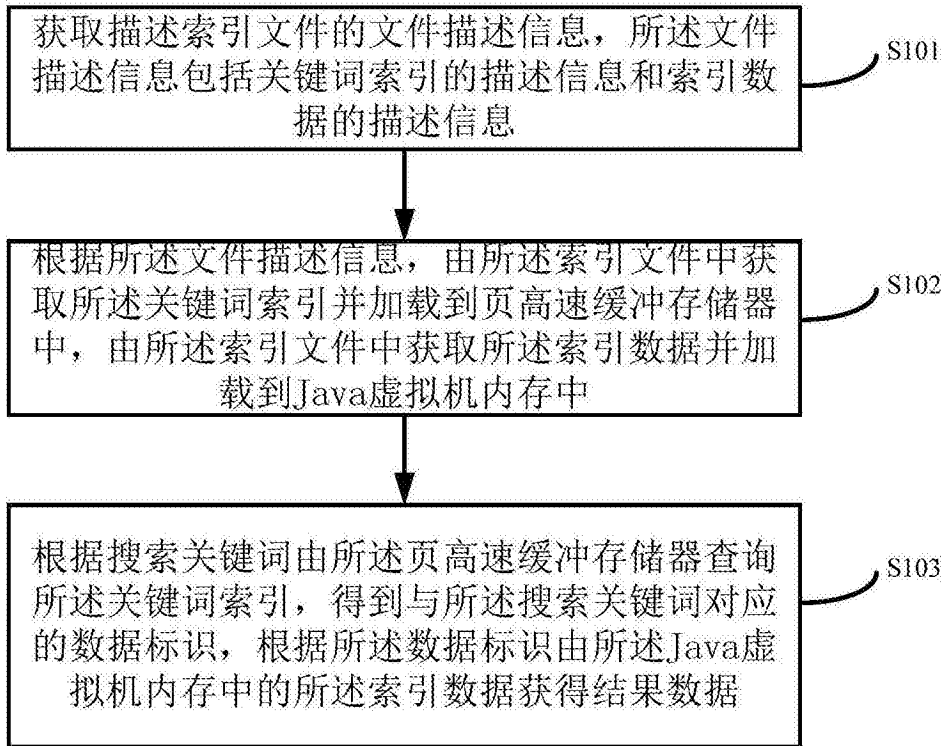


图1

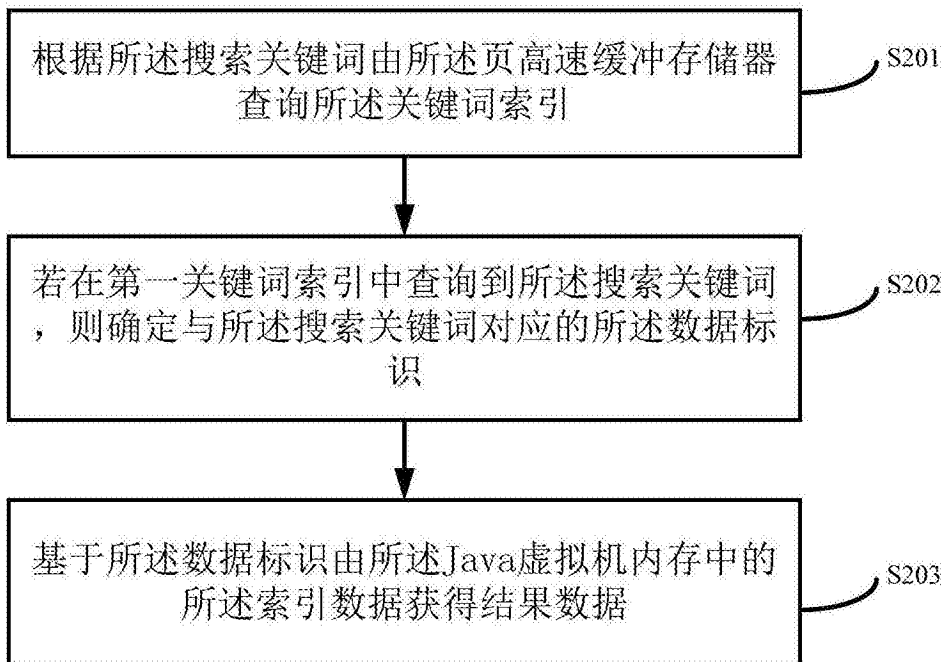


图2

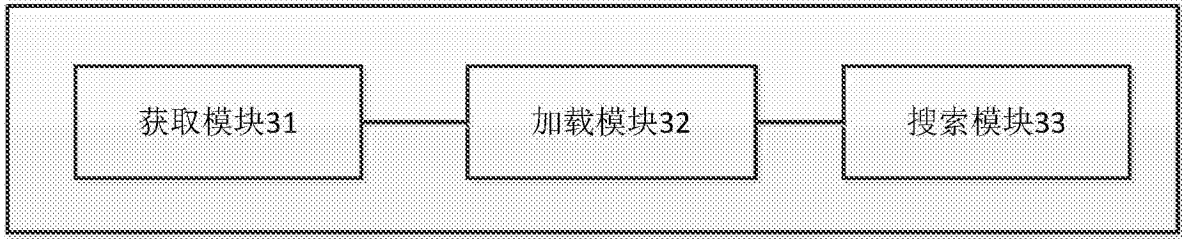


图3

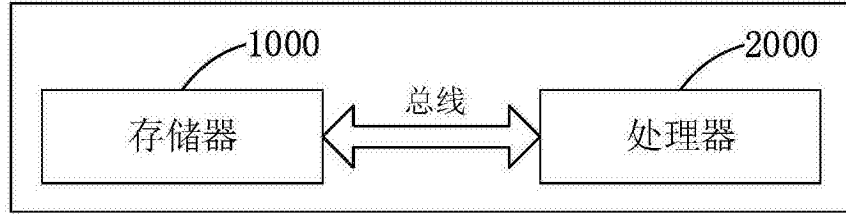


图4