



(12) 发明专利

(10) 授权公告号 CN 111737607 B

(45) 授权公告日 2023. 11. 10

(21) 申请号 202010572653.2

(22) 申请日 2020.06.22

(65) 同一申请的已公布的文献号
申请公布号 CN 111737607 A

(43) 申请公布日 2020.10.02

(73) 专利权人 中国银行股份有限公司
地址 100818 北京市西城区复兴门内大街1号

(72) 发明人 申亚坤

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227
专利代理师 王娇娇

(51) Int. Cl.
G06F 16/9538 (2019.01)
G06F 16/953 (2019.01)

(56) 对比文件

CN 107480135 A, 2017.12.15
CN 110532451 A, 2019.12.03
WO 2019174132 A1, 2019.09.19

审查员 田小娟

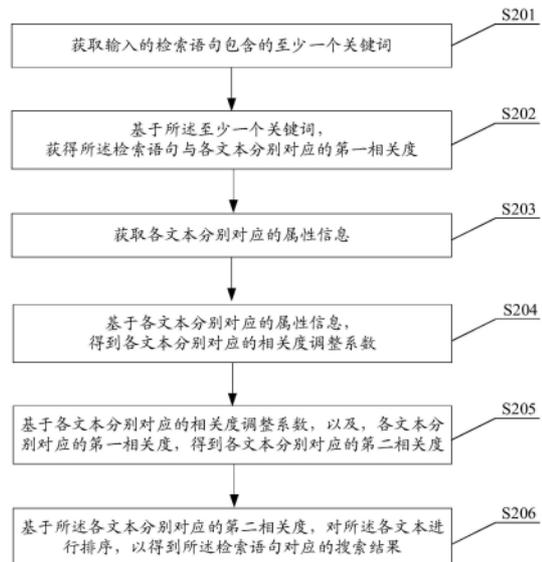
权利要求书3页 说明书14页 附图5页

(54) 发明名称

数据处理方法、装置、电子设备以及存储介质

(57) 摘要

本发明公开了一种数据处理方法、装置、电子设备以及存储介质,获取输入的检索语句包含的至少一个关键词,基于至少一个关键词,获得检索语句与各文本分别对应的第一相关度;获取各文本分别对应的属性信息,基于各属性信息得到各文本分别对应的相关度调整系数,和各文本分别对应的第一相关度,得到各文本分别对应的第二相关度;基于第二相关度对所述各文本进行排序,以得到所述检索语句对应的搜索结果。使得搜索结果中靠前的文本与检索语句的相关程度更大,用户检索一次即可得到自己所需的文本,降低了用户的检索次数,从而降低了搜索引擎的数据处理量。



1. 一种数据处理方法,其特征在于,包括:

获取输入的检索语句包含的至少一个关键词;

基于所述至少一个关键词,获得所述检索语句与各文本分别对应的第一相关度;

其中,一个文本与所述检索语句的第一相关度是指所述至少一个关键词分别位于所述一个文本的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和;一个所述关键词位于所述一个文本的词频是指所述一个文本包含所述关键词的总数目与所述一个文本包含的词语的总数目的比值;一个关键词对应的重要性调整系数与所述各文本中包含所述关键词的文本的总数目与所述各文本的总数目的比值成负相关;

获取文本库中的各文本分别对应的属性信息,所述文本库是指存储所述各文本的数据库;

其中,一个文本对应的属性信息至少包括各文本中与所述一个文本具有上下文关系的至少一个第一文本分别与所述一个文本之间的文本距离;具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段;具有同一上下文关系的两个文本的文本距离是指所述两个文本位于同一源文本中的自然段间隔;

基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数;

基于各文本分别对应的相关度调整系数,以及,各文本分别对应的第一相关度,得到各文本分别对应的第二相关度;

基于所述各文本分别对应的第二相关度,对所述各文本进行排序,以得到所述检索语句对应的搜索结果。

2. 根据权利要求1所述数据处理方法,其特征在于,所述一个文本对应的属性信息还包括:各文本中与所述一个文本具有预设关联关系的第二文本的总数目;

所述基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数,包括:

针对每一个文本,基于所述总数目以及所述至少一个第一文本分别与所述一个文本之间的文本距离,获得所述一个文本对应的相关度调整系数,以得到所述各文本分别对应的相关度调整系数。

3. 根据权利要求1所述数据处理方法,其特征在于,所述一个文本对应的属性信息还包括所述一个文本的热度属性信息;一个文本对应的热度属性信息包括展现量、点击量和点击率中的一种或多种;

所述基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数,包括:

针对每一个文本,基于所述一个文本对应的热度属性信息以及所述至少一个第一文本分别与所述一个文本之间的文本距离,获得所述一个文本对应的相关度调整系数,以得到各文本分别对应的相关度调整系数。

4. 根据权利要求1所述数据处理方法,其特征在于,所述一个文本对应的属性信息还包括所述一个文本的热度属性信息以及所述各文本中与所述一个文本具有预设关联关系的第二文本的总数目;一个文本对应的热度属性信息包括展现量、点击量和点击率中的一种或多种;

所述基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数,包括:

针对每一个文本,基于所述一个文本对应的热度属性信息、所述总数目以及所述至少一个第一文本分别与所述一个文本之间的文本距离,获得所述一个文本对应的相关度调整

系数,以得到各文本分别对应的相关度调整系数。

5. 根据权利要求1所述数据处理方法,其特征在于,所述获取各文本分别对应的属性信息,包括:

将各文本分别对应的第一相关度进行降序排序,以得到第一排序结果;

获得位于所述第一排序结果中最前面的第一预设数目个文本;

获取所述第一预设数目个文本分别对应的属性信息。

6. 根据权利要求1所述数据处理方法,其特征在于,还包括:

获取各文本分别对应的第二相关度的降序排序结果中,位于最前面的第二预设数目个文本;

基于所述第二预设数目个文本分别对应的类型标签,将所述第二预设数目个文本进行分类,以得到分类结果;

展示所述分类结果。

7. 一种数据处理装置,包括:

第一获取模块,用于获取输入的检索语句包含的至少一个关键词;

第二获取模块,用于基于所述至少一个关键词,获得所述检索语句与各文本分别对应的第一相关度;

其中,一个文本与所述检索语句的第一相关度是指所述至少一个关键词分别位于所述一个文本的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和;一个所述关键词位于所述一个文本的词频是指所述一个文本包含所述关键词的总目与所述一个文本包含的词语的总数目的比值;一个关键词对应的重要性调整系数与所述各文本中包含所述关键词的文本的总目与所述各文本的总数目的比值成负相关;

第三获取模块,用于获取文本库中的各文本分别对应的属性信息,所述文本库是指存储所述各文本的数据库;

其中,一个文本对应的属性信息至少包括各文本中与所述一个文本具有上下文关系的至少一个第一文本分别与所述一个文本之间的文本距离;具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段;具有同一上下文关系的两个文本的文本距离是指所述两个文本位于同一源文本中的自然段间隔;

第一确定模块,用于基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数;

第二确定模块,用于基于各文本分别对应的相关度调整系数,以及,各文本分别对应的第一相关度,得到各文本分别对应的第二相关度;

排序模块,用于基于所述各文本分别对应的第二相关度,对所述各文本进行排序,以得到所述检索语句对应的搜索结果。

8. 根据权利要求7所述数据处理装置,其特征在于,还包括:

第四获取模块,用于获取各文本分别对应的第二相关度的降序排序结果中,位于最前面的第二预设数目个文本;

分类模块,用于基于所述第二预设数目个文本分别对应的类型标签,将所述第二预设数目个文本进行分类,以得到分类结果;

展示模块,用于展示所述分类结果。

9. 一种电子设备,包括:

存储器,用于存储程序;

处理器,用于执行所述程序,所述程序具体用于:

获取输入的检索语句包含的至少一个关键词;

基于所述至少一个关键词,获得所述检索语句与各文本分别对应的第一相关度;

其中,一个文本与所述检索语句的第一相关度是指所述至少一个关键词分别位于所述一个文本的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和;一个所述关键词位于所述一个文本的词频是指所述一个文本包含所述关键词的总数目与所述一个文本包含的词语的总数目的比值;一个关键词对应的重要性调整系数与所述各文本中包含所述关键词的文本的总数目与所述各文本的总数目的比值成负相关;

获取文本库中的各文本分别对应的属性信息,所述文本库是指存储所述各文本的数据库;

其中,一个文本对应的属性信息至少包括各文本中与所述一个文本具有上下文关系的至少一个第一文本分别与所述一个文本之间的文本距离;具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段;具有同一上下文关系的两个文本的文本距离是指所述两个文本位于同一源文本中的自然段间隔;

基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数;

基于各文本分别对应的相关度调整系数,以及,各文本分别对应的第一相关度,得到各文本分别对应的第二相关度;

基于所述各文本分别对应的第二相关度,对所述各文本进行排序,以得到所述检索语句对应的搜索结果。

10. 一种存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时,实现如权利要求1至6任一项所述的数据处理方法中的各个步骤。

数据处理方法、装置、电子设备以及存储介质

技术领域

[0001] 本发明涉及信息检索领域,具体涉及一种数据处理方法、装置、电子设备以及存储介质。

背景技术

[0002] 随着互联网技术的发展,用户可以在搜索框中输入检索语句,搜索引擎会检索得到包含检索语句中关键词的文本。

[0003] 目前搜索引擎检索得到的文本并不是用户想要得到的文本,即搜索引擎检索得到的文本与用户输入的检索语句的相关度较低,导致用户需要反复输入检索语句,即反复进行检索,增加了搜索引擎的数据处理量。

发明内容

[0004] 有鉴于此,本发明提供了一种数据处理方法、装置、服务器以及存储介质,以克服现有技术中检索结果相关度不高的问题。

[0005] 为实现上述目的,本发明提供如下技术方案:

[0006] 一种数据处理方法,包括:

[0007] 获取输入的检索语句包含的至少一个关键词;

[0008] 基于所述至少一个关键词,获得所述检索语句与各文本分别对应的第一相关度;

[0009] 其中,一个文本与所述检索语句的第一相关度是指所述至少一个关键词分别位于所述文本的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和;一个所述关键词位于所述文本的词频是指所述文本包含所述关键词的总目与所述文本包含的词语的总数目的比值;一个关键词对应的重要性调整系数与所述各文本中包含所述关键词的文本的总目与所述各文本的总数目的比值成负相关;

[0010] 获取各文本分别对应的属性信息;

[0011] 其中,一个文本对应的属性信息至少包括各文本中与所述文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离;具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段;具有同一上下文关系的两个文本的文本距离是指所述两个文本位于同一源文本中的自然段间隔;

[0012] 基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数;

[0013] 基于各文本分别对应的相关度调整系数,以及,各文本分别对应的第一相关度,得到各文本分别对应的第二相关度;

[0014] 基于所述各文本分别对应的第二相关度,对所述各文本进行排序,以得到所述检索语句对应的搜索结果。

[0015] 一种数据处理装置,包括:

[0016] 第一获取模块,用于获取输入的检索语句包含的至少一个关键词;

[0017] 第二获取模块,用于基于所述至少一个关键词,获得所述检索语句与各文本分别

对应的第一相关度；

[0018] 其中，一个文本与所述检索语句的第一相关度是指所述至少一个关键词分别位于所述文本的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和；一个所述关键词位于所述文本的词频是指所述文本包含所述关键词的总数目与所述文本包含的词语的总数目的比值；一个关键词对应的重要性调整系数与所述各文本中包含所述关键词的文本的总数目与所述各文本的总数目的比值成负相关；

[0019] 第三获取模块，用于获取各文本分别对应的属性信息；

[0020] 其中，一个文本对应的属性信息至少包括各文本中与所述文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离；具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段；具有同一上下文关系的两个文本的文本距离是指所述两个文本位于同一源文本中的自然段间隔；

[0021] 第一确定模块，用于基于各文本分别对应的属性信息，得到各文本分别对应的相关度调整系数；

[0022] 第二确定模块，用于基于各文本分别对应的相关度调整系数，以及，各文本分别对应的第一相关度，得到各文本分别对应的第二相关度；

[0023] 排序模块，用于基于所述各文本分别对应的第二相关度，对所述各文本进行排序，以得到所述检索语句对应的搜索结果。

[0024] 一种电子设备，包括：

[0025] 存储器，用于存储程序；

[0026] 处理器，用于执行所述程序，所述程序具体用于：

[0027] 获取输入的检索语句包含的至少一个关键词；

[0028] 基于所述至少一个关键词，获得所述检索语句与各文本分别对应的第一相关度；

[0029] 其中，一个文本与所述检索语句的第一相关度是指所述至少一个关键词分别位于所述文本的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和；一个所述关键词位于所述文本的词频是指所述文本包含所述关键词的总数目与所述文本包含的词语的总数目的比值；一个关键词对应的重要性调整系数与所述各文本中包含所述关键词的文本的总数目与所述各文本的总数目的比值成负相关；

[0030] 获取各文本分别对应的属性信息；

[0031] 其中，一个文本对应的属性信息至少包括各文本中与所述文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离；具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段；具有同一上下文关系的两个文本的文本距离是指所述两个文本位于同一源文本中的自然段间隔；

[0032] 基于各文本分别对应的属性信息，得到各文本分别对应的相关度调整系数；

[0033] 基于各文本分别对应的相关度调整系数，以及，各文本分别对应的第一相关度，得到各文本分别对应的第二相关度；

[0034] 基于所述各文本分别对应的第二相关度，对所述各文本进行排序，以得到所述检索语句对应的搜索结果。

[0035] 一种存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时，实现如上述任一项所述的数据处理方法中的各个步骤。

[0036] 经由上述的技术方案可知,本发明提供了一种数据处理方法,首先得到检索语句包含的至少一个关键词,基于至少一个关键词,获得检索语句与各文本分别对应的第一相关度,可以理解的是,针对一个关键词,若该关键词在文本中出现的词频越大,说明这个关键词对于该文本越重要;各文本中包含该关键词的文本的总目数与各文本的总目数的比值越小,说明这个关键词越不常见,这个关键词越能反映这个文本的特性,针对每一文本,将所述至少一个关键词分别对应的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和,作为文本与检索语句的第一相关度,那么,文本对应的第一相关度越大,说明文本与检索语句的相关程度越大。

[0037] 进一步的,获得各文本分别对应的属性信息,一个文本对应的属性信息至少包括各文本中与所述文本具有上下文关系的文本与所述文本之间的文本距离;具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段;可以理解的是,若属于同一源文本的多个文本中一个文本与检索语句相关程度较大,那么,另一文本与检索语句的相关程度也应该较大;基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数,基于各文本分别对应的相关度调整系数,以及,各文本分别对应的第一相关度,得到各文本分别对应的第二相关度,基于所述各文本分别对应的第二相关度,对所述各文本进行排序,以得到所述检索语句对应的搜索结果。使得搜索结果中靠前的文本与检索语句的相关程度更大,用户检索一次即可得到自己所需的文本,降低了用户的检索次数,从而降低了搜索引擎的数据处理量。

附图说明

[0038] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的实施例,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,还可以根据提供的附图获得其他的附图。

[0039] 图1为本发明实施例提供的一种实施环境的架构图;

[0040] 图2为本发明实施例提供的数据处理方法的一种实现方式的流程图;

[0041] 图3为本发明实施例提供的数据处理方法的另一种实现方式的流程图;

[0042] 图4为本发明实施例提供的数据处理方法中文本类型标签的一种实现方式的示意图;

[0043] 图5为本发明实施例提供的数据处理方法中展示分类结果的一种实现方式的示意图;

[0044] 图6为本发明实施例提供的数据处理装置的一种实现方式的结构图;

[0045] 图7为本发明实施例提供的电子设备的一种实现方式的结构图。

具体实施方式

[0046] 下面将结合本发明实施例中的附图,对本发明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发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0047] 在对本公开实施例提供的数据处理方法进行详细介绍之前,这里先对本公开实施例涉及的实施环境进行简单介绍。

[0048] 如图1所示,为本发明实施例提供的一种实施环境的架构图。如图1所示,下述数据处理方法可以应用于该实施环境中,该实施环境包括:一个或多个终端设备11以及电子设备12。

[0049] 其中,终端设备11可以为任何一种可与用户通过键盘、触摸板、触摸屏、遥控器、语音交互或手写设备等一种或多种方式进行人机交互的电子产品,例如,手机、平板电脑、掌上电脑、个人计算机、可穿戴设备、智能电视等。

[0050] 图1仅仅是一种示例,实际应用中终端设备11的数量可以按照实际需求设定,图1中示出了一个终端设备11。

[0051] 可选的,电子设备12可以为任何一种可与用户通过键盘、触摸板、触摸屏、遥控器、语音交互或手写设备等一种或多种方式进行人机交互的电子产品,例如,手机、平板电脑、掌上电脑、个人计算机、可穿戴设备、智能电视等。

[0052] 可选的,电子设备12可以为服务器,该服务器可以是一台服务器,也可以是由若干台服务器组成的服务器集群,或者是一个云计算服务中心。

[0053] 可选的,终端设备11与电子设备12为同一设备;可选的,终端设备11与电子设备12为不同的设备。

[0054] 终端设备11,用于接收用户输入的检索语句,将检索语句发送至电子设备12。

[0055] 例如,用户可以在终端设备11展示的用户界面中输入“我要办卡”这一检索语句。

[0056] 可选的,终端设备11上安装有应用程序客户端,可以基于应用程序客户端展示的用户界面接收检索语句。

[0057] 可选的,终端设备11上安装有浏览器客户端,可以通过浏览器展示的网页版客户端的用户界面接收检索语句。

[0058] 上述客户端可以为搜索引擎客户端。

[0059] 电子设备12,用于基于本发明实施例提供的数据处理方法得到所述检索语句对应的搜索结果,将搜索结果反馈至终端设备11。

[0060] 终端设备11,还用于接收到搜索结果后,展示所述搜索结果。

[0061] 下面结合上述实施环境对本发明提供的数据处理方法、装置、电子设备及存储介质进行说明。

[0062] 如图2所示,为本发明实施例提供的数据处理方法的一种实现方式的流程图,该方法包括:

[0063] 步骤S201:获取输入的检索语句包含的至少一个关键词。

[0064] 在一可选实施例中,检索语句可以为语音或文本。若检索语句为语音,需要将语音转换成文本。

[0065] 可选的,本发明实施例提供但不限于以下获得检索语句包含的至少一个关键词的方法。

[0066] 第一种获得检索语句包含的至少一个关键词的方法包括:

[0067] 步骤A1:将检索语句进行划分,以得到多个词汇。

[0068] 可选的,若检索语句为“客户购房的贷款合同”,那么,检索语句包含的词汇为:客

户、购房的、贷款合同。

[0069] 步骤A2:根据预设规则,从多个词汇中获得关键词。

[0070] 可选的,预设规则可以包括:去除步骤A1得到的多个词汇中属于停用词的词汇。假设停用词包括:的、得、了、吗、吧、在、中等等。那么,通过步骤A2得到的关键词包括:客户、购房、贷款合同。

[0071] 第二种获得检索语句包含的至少一个关键词的方法包括:基于统计特征的关键词提取方法。

[0072] 基于统计特征的关键词提取算法是利用检索语句中词语的统计信息抽取检索语句的关键词。

[0073] 第三种获得检索语句包含的至少一个关键词的方法包括:基于词图模型的关键词提取算法,如,TextRank算法。

[0074] 基于词图模型的关键词提取算法,首先要构建检索语句的语言网络图,然后对语言网络图进行分析,在语言网络图上寻找具有重要作用的词或者短语,这些短语就是检索语句的关键词。

[0075] 第四种获得检索语句包含的至少一个关键词的方法包括:基于主题模型的关键词提取算法,如,LDA算法。

[0076] 基于主题模型的关键词提取算法主要利用的是主题模型中关于主题分布的性质进行关键词提取。

[0077] 步骤S202:基于所述至少一个关键词,获得所述检索语句与各文本分别对应的第一相关度。

[0078] 其中,一个文本与所述检索语句的第一相关度是指所述至少一个关键词分别位于所述文本的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和;一个所述关键词位于所述文本的词频是指所述文本包含所述关键词的总数目与所述文本包含的词语的总数目的比值;一个关键词对应的重要性调整系数与所述各文本中包含所述关键词的文本的总数目与所述各文本的总数目的比值成负相关。

[0079] 一个关键词对应的重要性调整系数与比值成负相关是指比值越大重要性调整系数越小,比值越小重要性调整系数越大。

[0080] 一个关键词位于文本的词频=该文本包含该关键词的总数目/该文本包含的词语的总词数。

[0081] 为了突出关键词在文本中的重要性,采用重要性调整系数来表示,可选的,其重要性调整系数的表达式为:

[0082] 一个关键词对应的重要性调整系数= $\log[\text{文本库包含的文本的总数目}/(\text{文本库中包含该关键词的文本的总数目}+1)]$ 。

[0083] 本发明实施例提及的文本库是指存储所有文本的存储空间,例如数据库。

[0084] 其中,分母加1是为了防止文本库中不包括包含该关键词的文本,即文本库中包含该关键词的文本的总数目为0的情况出现。

[0085] 下面举一具体例子对一个文本与所述检索语句的第一相关度进行说明。

[0086] 假设检索语句为“银行贷款需要哪些条件”,检索语句包含:“银行”、“贷款”和“条件”三个关键词。假设文本为文本1。

[0087] 文本1与检索语句的第一相关度= $\log[\text{文本库包含的文本的总数目}/(\text{文本库中包含“银行”的文本的总数目}+1)] * \text{文本1包含“银行”的总数目}/\text{文本1包含的词语的总词数} + \log[\text{文本库包含的文本的总数目}/(\text{文本库中包含“贷款”的文本的总数目}+1)] * \text{文本1包含“条件”的总数目}/\text{文本1包含的词语的总词数} + \log[\text{文本库包含的文本的总数目}/(\text{文本库中包含“条件”的文本的总数目}+1)] * \text{文本1包含“条件”的总数目}/\text{文本1包含的词语的总词数}$ 。

[0088] 步骤S203:获取各文本分别对应的属性信息。

[0089] 其中,一个文本对应的属性信息至少包括各文本中与所述文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离;具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段;具有同一上下文关系的两个文本的文本距离是指所述两个文本位于同一源文本中的自然段间隔。

[0090] 本发明实施例中,对于任一文本而言,将与该文本具有同一上下文关系的文本称为第一文本。

[0091] 下面对“源文本”进行说明。将一个文本A拆分成多个文本B,一个文本B包括文本A中一个或多个自然段,本发明实施例中将文本A称为源文本。将多个文本B称为具有同一上下文关系的多个文本。

[0092] 将一个文本C拆分成多个文本D,一个文本D包括文本C中一个或多个自然段,本发明实施例将文本C称为源文本,将多个文本D称为具有同一上下文关系的多个文本。

[0093] 可以理解的是,由于文本B和文本D对应的源文本不同,所以文本B和文本D不属于具有同一上下文关系的多个文本。

[0094] 可选的,本发明实施例提及的各文本可以包括源文本,也可以不包括源文本。

[0095] 在一可选实施例中,具有上下文关系的每一文本对应一属性标签,一个文本对应的属性标签包括:该文本包含的内容位于源文本的位置以及源文本的标识,例如,该文本包含的内容属于源文本哪个自然段。

[0096] 可选的,源文本的标识可以包括:源文本的名称、源文本的存储地址、源文本的作者中的一种或多种。

[0097] 至少两个文本同属于同一个源文本时,称至少两个文本属于具有同一上下文关系的文本。具有同一上下文关系的两个文本之间的文本距离是指两个文本分别位于源文本的位置的差值。

[0098] 可选的,具有同一上下文关系的两个文本之间的文本距离是指两个文本分别包含的在源文本中距离最近的两个自然段的差值。

[0099] 例如,文本1和文本2属于具有同一上下文关系的文本,文本1包括源文本中第一个自然段至第三个自然段,文本2包括源文本中第五个自然段至第六个自然段。

[0100] 文本1和文本2的文本距离=第五个自然段-第三个自然段=2。

[0101] 可选的,具有同一上下文关系的两个文本之间的文本距离是指两个文本分别包含的在源文本中距离最近的两个自然段的差值,与,两个文本分别包含的在源文本中距离最远的两个自然段的差值,的平均值。

[0102] 仍以上述文本1和文本2为例,文本1和文本2的文本距离= $\{(\text{第五个自然段}-\text{第三个自然段})+(\text{第六个自然段}-\text{第一个自然段})\}/2=7/2$ 。

[0103] 上述两种得到文本距离的方法仅为示例,本发明并不对具有同一上下文关系的两

个文本之间的文本距离的具体实现方式进行限定。

[0104] 步骤S204:基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数。

[0105] 下面以一个文本对应的属性信息至少包括各文本中与所述文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离为例,对得到相关度调整系数的过程进行说明。

[0106] 假设所有文本中文本1、文本2以及文本3属于具有同一上下文关系的多个文本;文本4和文本5属于具有同一上下文关系的多个文本。

[0107] 假设文本1和文本2的文本距离为5,文本1和文本3的文本距离为6,文本2和文本3的文本距离为4;文本4和文本5的文本距离为7。

[0108] 可以理解的是,具有同一上下文关系的两个文本在源文本中的位置越近,两个文本可能越相关,即两个文本的文本距离越近,说明两个文本的相关度可能越大。基于此可以对文本距离进行预设运算,使得相关度调整系数与文本距离呈负相关,即文本距离越大相关度调整系数越小,文本距离越小相关度调整系数越大。

[0109] 可选的,预设运算可以为取倒数运算。

[0110] 可以理解的是,若各文本中包含的多个文本对应同一源文本,若属于同一源文本的多个文本都与检索语句具有一定的相关度,说明属于同一源文本的多个文本与检索语句的相关度可能更高。

[0111] 对于任一文本而言,可以基于与该文本具有同一上下文关系的至少一个第一文本分别对应的文本距离得到第一子相关度调整系数。基于第一子相关度调整系数以得到相关度调整系数。

[0112] 可选的,基于文本1、文本2和文本3分别对应的文本距离,得到文本1、文本2和文本3分别对应的相关度调整系数的过程可以如下:

[0113] 文本1的第一子相关度调整系数 = $(1/5+1/6)$ *第一权重;文本2的第一子相关度调整系数 = $(1/5+1/4)$ *第一权重;文本3第一子相关度调整系数 = $(1/4+1/6)$ *第一权重。

[0114] 第一权重是小于或等于1的任意正数。在实际应用中可以调节第一权重,以使得各文本分别对应的第二相关度之和等于1。

[0115] 可以理解的是,各文本可能包括不具有同一上下文关系的文本,例如,一个文本不是通过源文本拆分得到的,那么,这个文本为不具有同一上下文关系的文本,对于这个文本而言,若属性信息仅仅包括各文本中与所述文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离,那么,这个文本对应的相关度调整系数为0。

[0116] 步骤S205:基于各文本分别对应的相关度调整系数,以及,各文本分别对应的第一相关度,得到各文本分别对应的第二相关度。

[0117] 对于任一文本而言,基于该文本的相关度调整系数以及第一相关度,得到第二相关度的方法有多种,本发明实施例提供但不限于以下几种。

[0118] 第一种:将该文本的相关度调整系数与第一相关度相加,得到第二相关度。

[0119] 第二种:将该文本的相关度调整系数与第一相关度的平均值,作为第二相关度。

[0120] 步骤S206:基于所述各文本分别对应的第二相关度,对所述各文本进行排序,以得到所述检索语句对应的搜索结果。

[0121] 可以理解的是,基于各文本分别对应的第二相关度,对各文本进行降序排序,并将排序结果作为检索语句的搜索结果。

[0122] 可以理解的是,搜索结果中排序越靠前的文本与检索语句的相关度越大。

[0123] 本发明实施例提供了一种数据处理方法,首先得到检索语句包含的至少一个关键词,基于至少一个关键词,获得检索语句与各文本分别对应的第一相关度,可以理解的是,针对一个关键词,若该关键词在文本中出现的词频越大,说明这个关键词对于该文本越重要;各文本中包含该关键词的文本的总数目与各文本的总数目的比值越小,说明这个关键词越不常见,这个关键词越能反映这个文本的特性,针对每一文本,将所述至少一个关键词分别对应的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和,作为文本与检索语句的第一相关度,那么,文本对应的第一相关度越大,说明文本与检索语句的相关程度越大。

[0124] 进一步的,获得各文本分别对应的属性信息,一个文本对应的属性信息至少包括各文本中与所述文本具有上下文关系的文本与所述文本之间的文本距离;具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段;可以理解的是,若属于同一源文本的多个文本中一个文本与检索语句相关程度较大,那么,另一文本与检索语句的相关程度也应该较大;基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数,基于各文本分别对应的相关度调整系数,以及,各文本分别对应的第一相关度,得到各文本分别对应的第二相关度,基于所述各文本分别对应的第二相关度,对所述各文本进行排序,以得到所述检索语句对应的搜索结果。使得搜索结果中靠前的文本与检索语句的相关程度更大,用户检索一次即可得到自己所需的文本,降低了用户的检索次数,从而降低了搜索引擎的数据处理量。

[0125] 在一可选实施例中,本发明实施例中提及的一个文本对应的属性信息还可以包括多种信息,本发明实施例提供但不限于以下几种。

[0126] 第一种属性信息:针对每一文本,该文本对应的属性信息包括各文本中与该文本具有预设关联关系的第二文本的总数目,以及,各文本中与该文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离。

[0127] 可选的,预设关联关系可以是指:属于同一作者、属于同一类型中的一种或多种。

[0128] 步骤S204包括:针对每一文本,基于所述总数目以及所述至少一个第一文本分别与所述文本之间的文本距离,获得所述文本对应的相关度调整系数,以得到所述各文本分别对应的相关度调整系数。

[0129] 本发明实施例中,对于任一文本而言,将与该文本具有预设关联关系的文本称为第二文本。

[0130] 可以理解的是,若各文本中不包括与该文本具有预设关联关系的第二文本,那么总数目为0;若各文本中不包括中与该文本具有上下文关系的第一文本,那么,文本距离可以设置为空值,或,文本距离为任意表征文本距离不存在的标识。

[0131] 可以理解的是,若各文本中包含的具有预设关联关系的多个文本均与检索语句具有一定的相关性,说明具有预设关联关系的多个文本与检索语句的相关度可能更高。可选的,一个文本对应的相关度调整系数与和该文本具有预设关联关系的第二文本的总数目呈正相关,即和该文本具有预设关联关系的第二文本的总数目越大,该文本的相关度调整系

数越大,和该文本具有预设关联关系的第二文本的总数目越小,该文本的相关度调整系数越小。

[0132] 可选的,步骤S204具体可以包括:对于任一文本,可以基于与该文本具有预设关联关系的第二文本的总数目得到第二子相关度调整系数;基于与该文本具有同一上下文关系的至少一个第一文本分别对应的文本距离得到第一子相关度调整系数;基于该文本对应的第一子相关度调整系数以及第二子相关度调整系数,得到该文本对应的的相关度调整系数。

[0133] 本发明实施例中对于任一文本,基于与该文本具有预设关联关系的第二文本的总数目得到第二子相关度调整系数的方式可以有多种,本发明实施例提供但不限于:一个文本的第二子相关度调整系数=该文本对应的总数目*第二权重。

[0134] 例如,对于文本A而言,假设文本库包含的所有文本中与文本A有预设关联关系的第二文本的总数目为25,那么,文本的第二子相关度调整系数=25*第二权重。

[0135] 可选的,第二权重为小于1的任意正数。在实际应用中需要调整第二权重,使得各文本分别对应的第二相关度之和等于1。

[0136] 可选的,基于第一子相关度调整系数和第二子相关度调整系数得到相关度调整系数的方法有多种,本发明实施例提供但不限于以下几种。

[0137] 第一种:相关度调整系数=第一子相关度调整系数+第二子相关度调整系数。

[0138] 第二种:相关度调整系数=(第一子相关度调整系数+第二子相关度调整系数)/2。

[0139] 第三种:相关度调整系数=第一子相关度调整系数*第三权重+第二子相关度调整系数*第四权重。

[0140] 其中,第三权重和第四权重可以相同,也可以不同,且第三权重和第四权重均为小于1的任意正数。

[0141] 第二种属性信息:针对每一文本,该文本对应的属性信息包括该文本的热度属性信息,以及,各文本中与该文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离。

[0142] 可选的,一个文本对应的热度属性信息包括展现量、点击量和点击率中的一种或多种。

[0143] 步骤S204包括:针对每一文本,基于所述文本对应的热度属性信息以及所述至少一个第一文本分别与所述文本之间的文本距离,获得所述文本对应的的相关度调整系数,以得到各文本分别对应的的相关度调整系数。

[0144] 下面对展现量、点击量和点击率的关系进行说明。

[0145] 可选的,点击率=点击量/展现量。

[0146] 本发明实施例中,一个文本的展现量是指该文本对应的链接的展现量,一个文本的点击量是指该文本对应的链接的点击量。

[0147] 在一可选实施例中,步骤S204可以具体包括:针对每一文本,基于该文本对应的热度属性信息得到第三子相关度调整系数;基于与该文本具有同一上下文关系的至少一个第一文本分别对应的文本距离得到第一子相关度调整系数;基于该文本对应的第一子相关度调整系数以及第三子相关度调整系数,得到该文本对应的的相关度调整系数。

[0148] 本发明实施例中基于针对每一文本,基于该文本对应的热度属性信息得到第三子相关度调整系数的方法可以有多种,本发明实施例提供但不限于:一个文本对应的第三子

相关度调整系数=该文本对应的热度属性信息*第五权重。

[0149] 可选的,第五权重为小于1的任意正数。在实际应用中需要调整第五权重,使得各文本分别对应的第二相关度之和等于1。

[0150] 可选的,基于第一子相关度调整系数和第三子相关度调整系数得到相关度调整系数的方法有多种,本发明实施例提供但不限于以下几种。

[0151] 第一种:相关度调整系数=第一子相关度调整系数+第三子相关度调整系数。

[0152] 第二种:相关度调整系数=(第一子相关度调整系数+第三子相关度调整系数)/2。

[0153] 第三种:相关度调整系数=第一子相关度调整系数*第六权重+第三子相关度调整系数*第七权重。

[0154] 其中,第六权重和第七权重可以相同,也可以不同,且第六权重和第七权重均为小于1的任意正数。

[0155] 第三种属性信息:针对每一文本,该文本对应的属性信息包括各文本中与该文本具有预设关联关系的第二文本的总数目,各文本中与该文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离,以及,该文本对应的热度属性信息。

[0156] 可选的,步骤S204具体可以包括:对于任一文本,可以基于与该文本具有预设关联关系的第二文本的总数目得到第二子相关度调整系数;基于与该文本具有同一上下文关系的至少一个第一文本分别对应的文本距离得到第一子相关度调整系数;基于该文本对应的热度属性信息得到第三子相关度调整系数;基于该文本对应的第一子相关度调整系数、第二子相关度调整系数以及第三子相关度调整系数,得到该文本对应的相关度调整系数。

[0157] 对于任一文本而言,得到该文本的第一子相关度调整系数、第二子相关度调整系数以及第三子相关度调整系数的方法可以参见上述方式,这里不再赘述。

[0158] 可选的,基于第一子相关度调整系数、第二子相关度调整系数以及第三子相关度调整系数得到相关度调整系数的方法有多种,本发明实施例提供但不限于以下几种。

[0159] 第一种:相关度调整系数=第一子相关度调整系数+第二子相关度调整系数+第三子相关度调整系数。

[0160] 第二种:相关度调整系数=(第一子相关度调整系数+第二子相关度调整系数+第三子相关度调整系数)/3。

[0161] 第三种:相关度调整系数=第一子相关度调整系数*第八权重+第二子相关度调整系数*第九权重+第三子相关度调整系数*第十权重。

[0162] 其中,第八权重、第九权重和第十权重可以相同,也可以不同,且第八权重、第九权重和第十权重均为小于1的任意正数。

[0163] 如图3所示,为本发明实施例提供的数据处理方法的另一种实现方式的流程图,该方法包括:

[0164] 步骤S301:获取输入的检索语句包含的至少一个关键词。

[0165] 步骤S302:基于所述至少一个关键词,获得所述检索语句与各文本分别对应的第一相关度。

[0166] 其中,一个文本与所述检索语句的第一相关度是指所述至少一个关键词分别位于所述文本的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和;一个所述关键词位于所述文本的词频是指所述文本包含所述关键词的总数目与所述文本包含的词

语的总数目的比值；一个关键词对应的重要性调整系数与所述各文本中包含所述关键词的文本的总数目与所述各文本的总数目的比值成负相关。

[0167] 步骤S301和步骤S302可以参见步骤S201和步骤S202,这里不再赘述。

[0168] 步骤S303:将各文本分别对应的第一相关度进行降序排序,以得到第一排序结果。

[0169] 步骤S304:获得位于所述第一排序结果中最前面的第一预设数目个文本,以及,所述第一预设数目个文本分别对应的属性信息。

[0170] 其中,一个文本对应的属性信息至少包括各文本中与所述文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离;具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段;具有同一上下文关系的两个文本的文本距离是指所述两个文本位于同一源文本中的自然段间隔。

[0171] 步骤S304和步骤S203过程一致,只是对应的文本数不同。

[0172] 例如,若文本库中包括1000个文本,假设第一预设数目为500,那么,可以得到第一排序结果中前500个文本。

[0173] 本发明实施例中,通过步骤S303至步骤S304可以去除部分与检索语句相关度较低的文本。

[0174] 步骤S305:基于第一预设数目个文本分别对应的属性信息,得到第一预设数目个文本分别对应的相关度调整系数。

[0175] 可选的,采用本发明实施例提供的方法,对于图2所示的流程而言,步骤S204中的各文本是指第一预设数目个文本中的各文本。

[0176] 步骤S306:基于第一预设数目个文本分别对应的相关度调整系数,以及,分别对应的第一相关度,得到第一预设数目个文本分别对应的第二相关度。

[0177] 步骤S307:基于第一预设数目个文本分别对应的第二相关度,对第一预设数目个文本进行排序,以得到所述检索语句对应的搜索结果。

[0178] 可选的,该搜索结果包括第一预设数目个文本。

[0179] 在一可选实施例中,可以理解的是,用户一般仅看排序靠前的文本,基于此,可以对搜索结果进行以下处理:获取各文本分别对应的第二相关度的降序排序结果中,位于最前面的第二预设数目个文本;展示所述第二预设数目个文本。

[0180] 对于图2所示实施例而言,第二预设数目小于文本库包含的所有文本的总数目;对于图3所示实施例而言,第二预设数目小于第一预设数目。

[0181] 在一可选实施例中,可以对搜索结果进行分类展示,分类展示的方法包括:基于第二预设数目个文本分别对应的类型标签,将所述第二预设数目个文本进行分类,以得到分类结果;展示所述分类结果。

[0182] 在一可选实施例中,一个文本对应的类型标签包括但不限于业务类型、客户类型中的一种或多种。

[0183] 可选的,客户类型是指:个人类型、企业类型、VIP类型。

[0184] 业务类型是指:贷款业务类型、理财业务类型、还款业务类型、开户业务类型等。

[0185] 可选的,在文本库中,任一文本都有相对应的类型标签,如图4所示,为本发明实施例提供的数据处理方法中文本类型标签的一种实现方式的示意图。

[0186] 在文本的右端显示该文本所属的业务类型和客户类型,基于此,对第二预设数目

个文本进行分类。

[0187] 例如,有文本1、文本2、文本3、文本4、文本5、文本6、文本7、文本8和文本9,假设文本1对应的类型标签为“个人”和“贷款”,文本2对应的类型标签为“个人”和“还款”,文本3对应的类型标签为“个人”和“贷款”,文本4对应的类型标签为“企业”和“理财”,文本5对应的类型标签为“企业”和“理财”,文本6对应的类型标签为“企业”和“还款”,文本7、文本8及文本9分别对应的类型标签为“VIP”和“开户”。

[0188] 在一可选实施例中,将分类结果以多级菜单的形式展示,并在菜单的叶子节点处,添加文本索引链接,以便快速定位。如图5所示,为本发明实施例提供的数据处理方法中展示分类结果的一种实现方式的示意图。

[0189] 基于文本对应的类型标签对上述9个文本进行分类。按照客户类型来进行分类的话,可将这9个文本分为“个人”、“企业”和“VIP”,“个人”类型下有文本1、文本2和文本3,“企业”类型下有文本4、文本5和文本6,“VIP”类型下有文本7、文本8和文本9;按照业务类型来分类的话,可将这9个文本分为“贷款”、“理财”、“还款”和“开户”,“贷款”类型下有文本1和文本3,“理财”类型下有文本4和文本5,“还款”类型下有文本2和文本6,“开户”类型下有文本7、文本8和文本9。

[0190] 可选的,如图5所示,在展示第二预设数目个文本时,可以展示第二预设数目个文本对应的链接,图5中文本链接1是指文本1对应的链接;文本链接2是指文本对应的链接,后续依次类推,这里不再赘述。用户点击链接后,可以跳转进入相应的界面,通过该界面可以看到相应的文本,例如,用户点击文本链接1后,跳转至文本1所在界面,用户通过该界面可以看到文本1。

[0191] 上述本发明公开的实施例中详细描述了方法,对于本发明的方法可采用多种形式的装置实现,因此本发明还公开了一种装置,下面给出具体的实施例进行详细说明。

[0192] 如图6所示,为本发明实施例提供的数据处理装置的一种实现方式的结构图,该装置包括:

[0193] 第一获取模块61,用于获取输入的检索语句包含的至少一个关键词;

[0194] 第二获取模块62,用于基于所述至少一个关键词,获得所述检索语句与各文本分别对应的第一相关度;

[0195] 其中,一个文本与所述检索语句的第一相关度是指所述至少一个关键词分别位于所述文本的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和;一个所述关键词位于所述文本的词频是指所述文本包含所述关键词的总数目与所述文本包含的词语的总数目的比值;一个关键词对应的重要性调整系数与所述各文本中包含所述关键词的文本的总数目与所述各文本的总数目的比值成负相关;

[0196] 第三获取模块63,用于获取各文本分别对应的属性信息;

[0197] 其中,一个文本对应的属性信息至少包括各文本中与所述文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离;具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段;具有同一上下文关系的两个文本的文本距离是指所述两个文本位于同一源文本中的自然段间隔;

[0198] 第一确定模块64,用于基于各文本分别对应的属性信息,得到各文本分别对应的相关度调整系数;

[0199] 第二确定模块65,用于基于各文本分别对应的相关度调整系数,以及,各文本分别对应的第一相关度,得到各文本分别对应的第二相关度;

[0200] 排序模块66,用于基于所述各文本分别对应的第二相关度,对所述各文本进行排序,以得到所述检索语句对应的搜索结果。

[0201] 可选的,所述一个文本对应的属性信息还包括:各文本中与所述文本具有预设关联关系的第二文本的总数目,第一确定模块包括:

[0202] 第一确定单元,用于针对每一文本,基于所述总数目以及所述至少一个第一文本分别与所述文本之间的文本距离,获得所述文本对应的相关度调整系数,以得到所述各文本分别对应的相关度调整系数。

[0203] 可选的,所述一个文本对应的属性信息还包括所述文本的热度属性信息,一个文本对应的热度属性信息包括展现量、点击量和点击率中的一种或多种,第一确定模块包括:

[0204] 第二确定单元,用于针对每一文本,基于所述文本对应的热度属性信息以及所述至少一个第一文本分别与所述文本之间的文本距离,获得所述文本对应的相关度调整系数,以得到各文本分别对应的相关度调整系数。

[0205] 可选的,所述一个文本对应的属性信息还包括所述文本的热度属性信息以及所述各文本中与所述文本具有预设关联关系的第二文本的总数目,一个文本对应的热度属性信息包括展现量、点击量和点击率中的一种或多种,第一确定模块包括:

[0206] 第三确定单元,用于针对每一文本,基于所述文本对应的热度属性信息、所述总数目以及所述至少一个第一文本分别与所述文本之间的文本距离,获得所述文本对应的相关度调整系数,以得到各文本分别对应的相关度调整系数。

[0207] 可选的,第三获取模块包括:

[0208] 第四确定单元,用于将各文本分别对应的第一相关度进行降序排序,以得到第一排序结果;

[0209] 第一获取单元,用于获得位于所述第一排序结果中最前面的第一预设数目个文本;

[0210] 第二获取单元,用于获取所述第一预设数目个文本分别对应的属性信息。

[0211] 可选的,还包括:

[0212] 第四获取模块,用于获取各文本分别对应的第二相关度的降序排序结果中,位于最前面的第二预设数目个文本;

[0213] 分类模块,用于基于所述第二预设数目个文本分别对应的类型标签,将所述第二预设数目个文本进行分类,以得到分类结果;

[0214] 展示模块,用于展示所述分类结果。

[0215] 如图7所示,为本发明实施例提供的电子设备的一种实现方式的结构图,该电子设备包括:

[0216] 存储器71,用于存储程序;

[0217] 处理器72,用于执行所述程序,所述程序具体用于:

[0218] 获取输入的检索语句包含的至少一个关键词;

[0219] 基于所述至少一个关键词,获得所述检索语句与各文本分别对应的第一相关度;

[0220] 其中,一个文本与所述检索语句的第一相关度是指所述至少一个关键词分别位于

所述文本的词频与所述至少一个关键词分别对应的重要性调整系数的乘积的和；一个所述关键词位于所述文本的词频是指所述文本包含所述关键词的总数目与所述文本包含的词语的总数目的比值；一个关键词对应的重要性调整系数与所述各文本中包含所述关键词的文本的总数目与所述各文本的总数目的比值成负相关；

[0221] 获取各文本分别对应的属性信息；

[0222] 其中，一个文本对应的属性信息至少包括各文本中与所述文本具有上下文关系的至少一个第一文本分别与所述文本之间的文本距离；具有同一上下文关系的两个文本是指所述两个文本分别包含的内容为同一源文本中的不同自然段；具有同一上下文关系的两个文本的文本距离是指所述两个文本位于同一源文本中的自然段间隔；

[0223] 基于各文本分别对应的属性信息，得到各文本分别对应的相关度调整系数；

[0224] 基于各文本分别对应的相关度调整系数，以及，各文本分别对应的第一相关度，得到各文本分别对应的第二相关度；

[0225] 基于所述各文本分别对应的第二相关度，对所述各文本进行排序，以得到所述检索语句对应的搜索结果。

[0226] 处理器72可能是一个中央处理器CPU，或者是特定集成电路ASIC (Application Specific Integrated Circuit)。

[0227] 第一服务器还可以包括通信接口73以及通信总线74，其中，存储器71、处理器72以及通信接口73通过通信总线74完成相互间的通信。

[0228] 本发明实施例还提供了一种存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时，实现如上述任一所述的数据处理方法实施例中包含的各个步骤。

[0229] 需要说明的是，本说明书中的各个实施例中记载的特征可以相互替换或者组合。对于装置或系统类实施例而言，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。

[0230] 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、物品或者设备中还存在另外的相同要素。

[0231] 结合本文中所公开的实施例描述的方法或算法的步骤可以直接用硬件、处理器执行的软件模块，或者二者的结合来实施。软件模块可以置于随机存储器 (RAM)、内存、只读存储器 (ROM)、电可编程ROM、电可擦除可编程ROM、寄存器、硬盘、可移动磁盘、CD-ROM、或技术领域内所公知的任意其它形式的存储介质中。

[0232] 对所公开的实施例的上述说明，使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的，本文中所定义的一般原理可以在不脱离本发明的精神或范围的情况下，在其它实施例中实现。因此，本发明将不会被限制于本文所示的这些实施例，而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

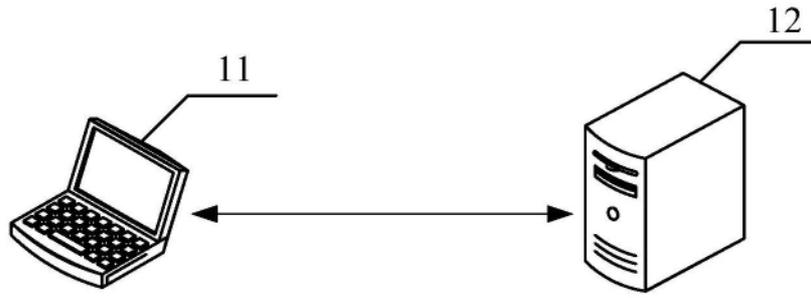


图1

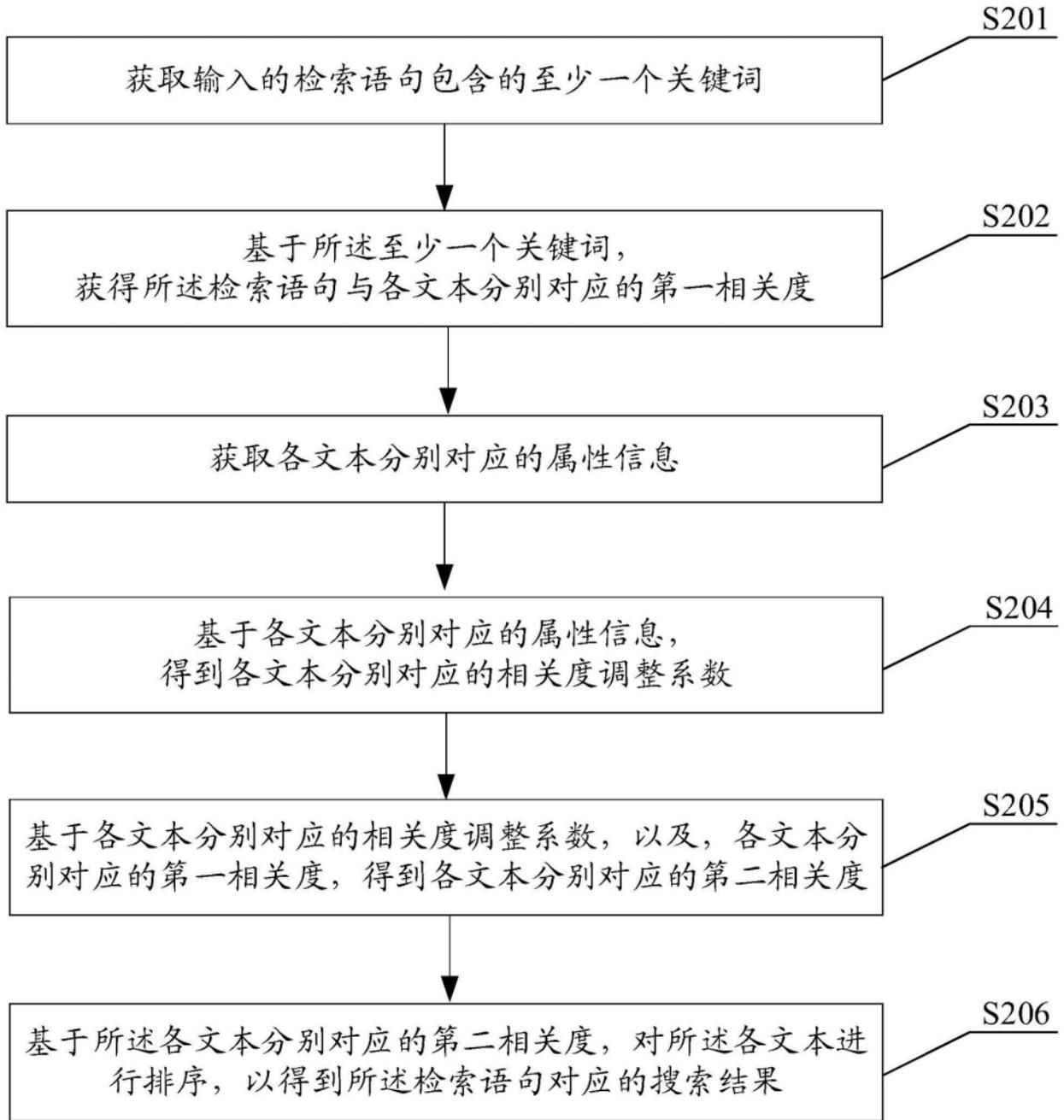


图2

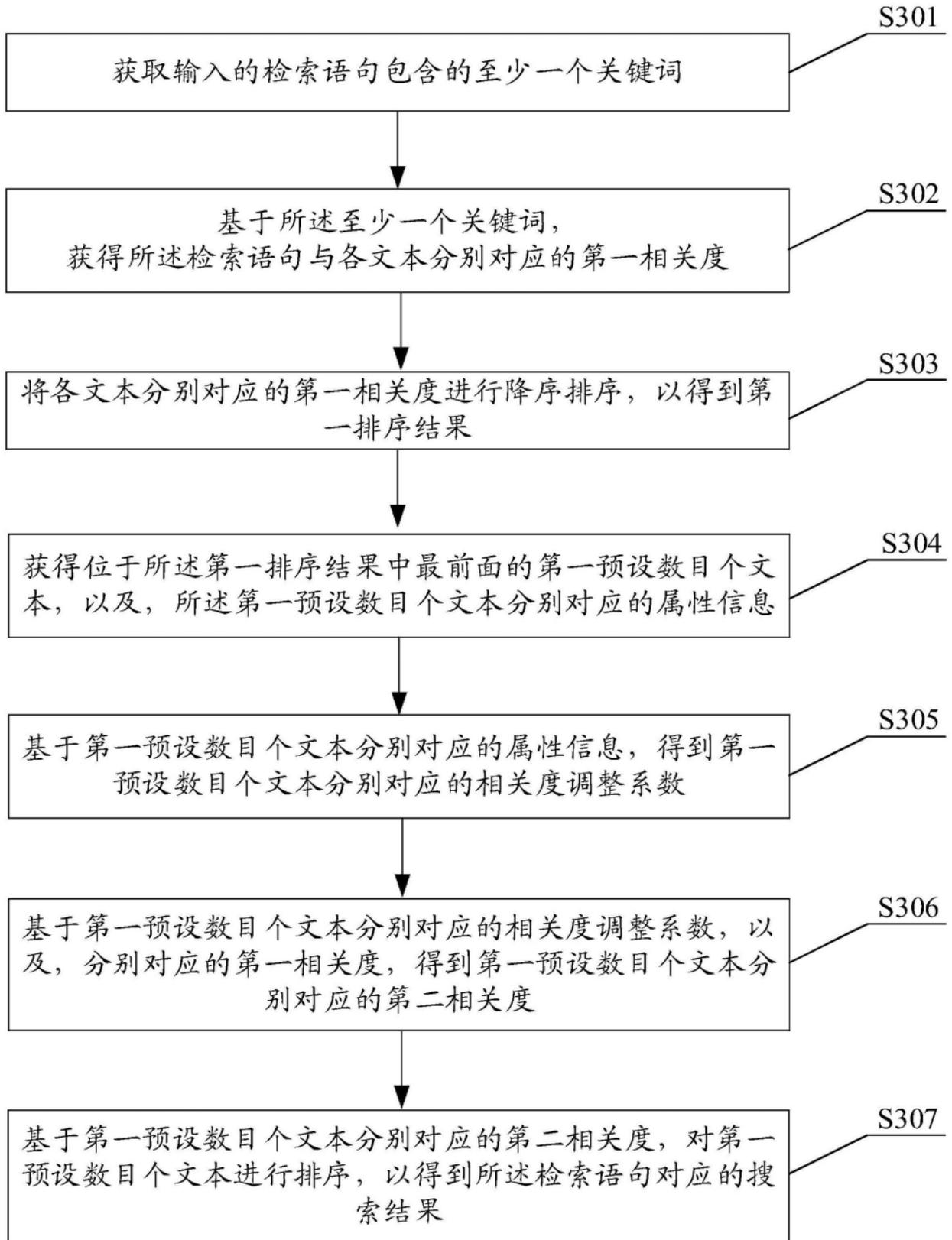


图3

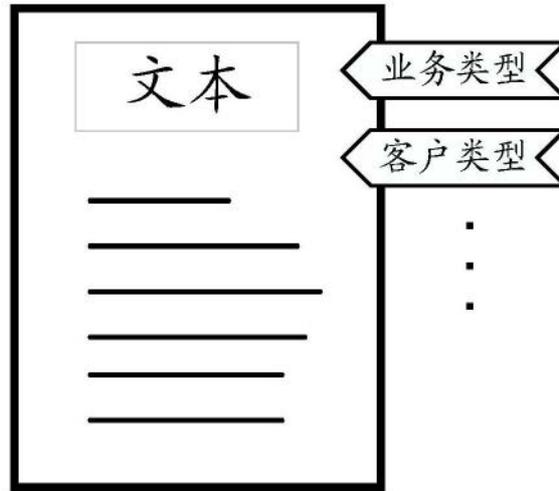


图4

个人	企业	VIP
文本链接1	文本链接4	文本链接7
文本链接2	文本链接5	文本链接8
文本链接3	文本链接6	文本链接9
.	.	.
.	.	.
.	.	.

贷款	理财	还款	开户
文本链接1	文本链接4	文本链接2	文本链接7
文本链接3	文本链接5	文本链接6	文本链接8
.	.	.	文本链接9
.	.	.	.
.	.	.	.

图5

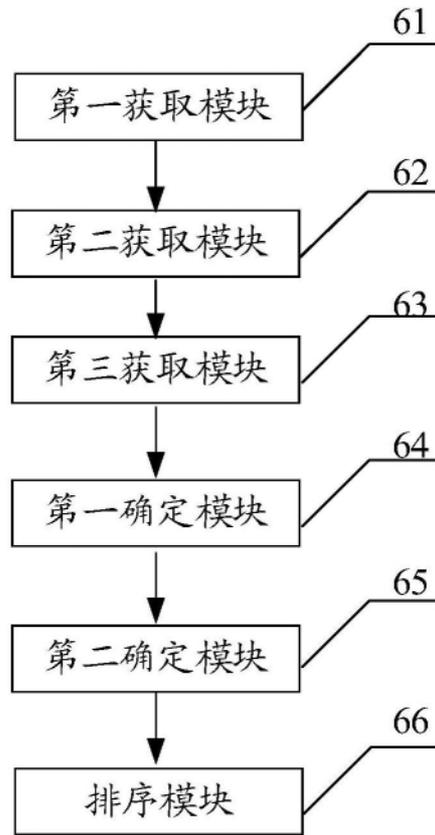


图6

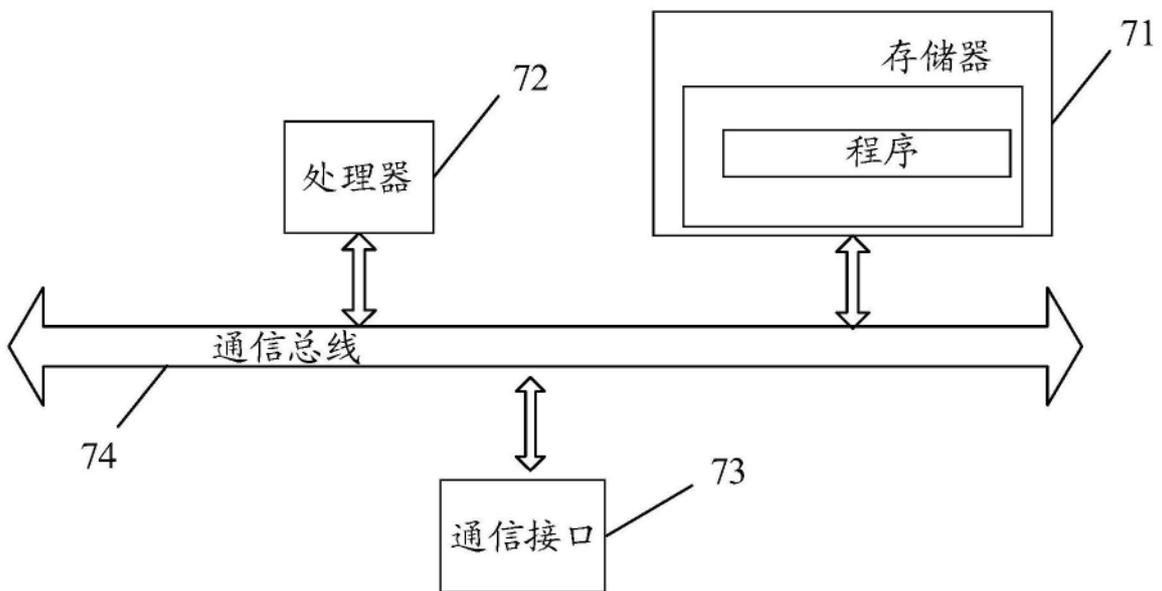


图7