



(12) 发明专利申请

(10) 申请公布号 CN 103473228 A

(43) 申请公布日 2013. 12. 25

(21) 申请号 201210184615. 5

(22) 申请日 2012. 06. 06

(71) 申请人 深圳市世纪光速信息技术有限公司
地址 518057 广东省深圳市南山区粤海街道
科技中一路腾讯大厦 16 层

(72) 发明人 高飞荣 李锐 杨志峰

(74) 专利代理机构 北京中博世达专利商标代理
有限公司 11274

代理人 申健

(51) Int. Cl.

G06F 17/30 (2006. 01)

G06F 3/0481 (2013. 01)

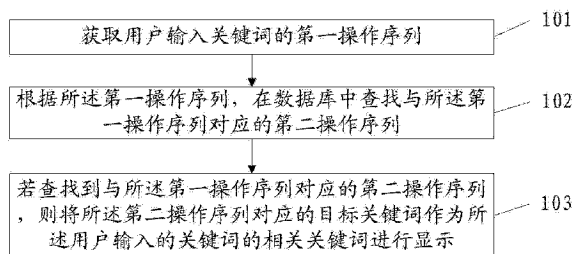
权利要求书2页 说明书5页 附图4页

(54) 发明名称

相关关键词的显示方法及装置

(57) 摘要

本发明公开了一种相关关键词的显示方法及装置,涉及信息处理领域,提高了搜索引擎为用户提供的相关关键词的准确性。本发明包括:获取用户输入关键词的第一操作序列,所述第一操作序列包含所述用户输入关键词的所有操作行为;根据所述第一操作序列,在数据库中查找与所述第一操作序列对应的第二操作序列,所述第二操作序列为与所述第一操作序列之间满足预设的序列相似度的操作序列;若查找到与所述第一操作序列对应的第二操作序列,则将所述第二操作序列对应的目标关键词作为所述用户输入的关键词的相关关键词进行显示。本发明实施例主要应用于相关关键词的显示过程中。



1. 一种相关关键词的显示方法,其特征在于,包括:

获取用户输入关键词的第一操作序列,所述第一操作序列包含所述用户输入关键词的所有操作行为;

根据所述第一操作序列,在数据库中查找与所述第一操作序列对应的第二操作序列,所述第二操作序列为与所述第一操作序列之间满足预设的序列相似度的操作序列;

若查找到与所述第一操作序列对应的第二操作序列,则将所述第二操作序列对应的目标关键词作为所述用户输入的关键词的相关关键词进行显示。

2. 根据权利要求1所述的方法,其特征在于,所述获取用户输入关键词的第一操作序列包括:

采集所述用户输入关键词的操作行为,并根据所述用户的标识信息,从已存储的操作序列中查找所述用户对应的原始操作序列,所述原始操作序列包含所述用户在输入所述关键词之前进行的所有操作行为;

若查找到所述用户对应的原始操作序列,则根据采集到的操作行为和所述原始操作序列内的操作行为以及所述采集到的操作行为和所述原始操作序列内的操作行为之间的先后顺序,生成第一操作序列,并将所述原始操作序列替换为所述第一操作序列进行存储。

3. 根据权利要求2所述的方法,其特征在于,所述采集所述用户输入关键词的操作行为包括:

当采集到所述用户输入关键词的第一个操作行为时,为所述用户设置采集起始时间戳;

按照预定时长以及所述采集起始时间戳,设置与所述采集起始时间戳对应的采集终止时间戳;

在所述采集起始时间戳以及所述采集终止时间戳之间的时间段内,采集所述用户输入关键词的操作行为。

4. 根据权利要求3所述的方法,其特征在于,还包括:

当在所述采集起始时间戳以及所述采集终止时间戳之间的时间段内采集到所述用户结束搜索的操作行为时,结束对所述用户输入关键词的操作行为的采集。

5. 根据权利要求2-4所述的方法,其特征在于,还包括:

若未查找到所述用户对应的原始操作序列,则根据采集到的操作行为生成第一操作序列,并将所述第一操作序列作为所述用户对应的原始操作序列进行存储。

6. 根据权利要求5所述的方法,其特征在于,所述序列相似度为所述第一操作序列和所述第二操作序列的最长公共子序列的长度与所述第一操作序列的长度和所述第二操作序列的长度的平均值之间的比例;

或者,所述序列相似度为将所述第一操作序列转化为所述第二操作序列所需要插入、删除或者替换的序列元素的数目。

7. 一种相关关键词的显示的装置,其特征在于,包括:

获取单元,用于获取用户输入关键词的第一操作序列,所述第一操作序列包含所述用户输入关键词的所有操作行为;

查找单元,用于在数据库中查找与所述获取单元获取的所述第一操作序列对应的第二操作序列,所述第二操作序列为与所述第一操作序列之间满足预设的序列相似度的操作序

列；

显示单元,用于在所述查找单元查找到与所述第一操作序列对应的第二操作序列时,将所述第二操作序列对应的目标关键词作为所述用户输入的关键词的相关关键词进行显示。

8. 根据权利要求 7 所述的装置,其特征在于,所述获取单元包括:

采集模块,用于采集所述用户输入关键词的操作行为;

查找模块,用于根据所述用户的标识信息,从已存储的操作序列中查找所述用户对应的原始操作序列,所述原始操作序列包含所述用户在输入所述关键词之前进行的所有操作行为;

生成模块,用于在所述查找模块查找到所述用户对应的原始操作序列时,根据所述采集模块采集到的操作行为和所述原始操作序列内的操作行为以及所述采集到的操作行为和所述原始操作序列内的操作行为之间的先后顺序,生成第一操作序列;

存储模块,用于将所述原始操作序列替换为所述生成模块生成的第一操作序列进行存储。

9. 根据权利要求 8 所述的装置,其特征在于,所述采集模块具体包括:

设置子模块,用于在采集到所述用户输入关键词的第一个操作行为时,为所述用户设置采集起始时间戳;

所述设置子模块,还用于按照预定时长以及所述采集起始时间戳,设置与所述采集起始时间戳对应的采集终止时间戳;

采集子模块,用于在所述设置子模块设置好的采集起始时间戳以及采集终止时间戳之间的时间段内,采集所述用户输入关键词的操作行为。

10. 根据权利要求 9 所述的装置,其特征在于,所述采集子模块还用于在在所述采集起始时间戳以及所述采集终止时间戳之间的时间段内采集到所述用户结束搜索的操作行为时,结束对所述用户输入关键词的操作行为的采集。

11. 根据权利要求 8 至 10 任一项所述的装置,其特征在于,所述生成模块,还用于在所述查找模块未查找到所述用户对应的原始操作序列时,根据所述采集模块采集到的操作行为生成第一操作序列;

所述存储模块,还用于将所述生成模块生成的第一操作序列作为所述用户对应的原始操作序列进行存储。

12. 根据权利要求 11 所述的装置,其特征在于,所述序列相似度为所述第一操作序列和所述第二操作序列的最长公共子序列的长度与所述第一操作序列的长度和所述第二操作序列的长度的平均值之间的比例;

或者,所述序列相似度为将所述第一操作序列转化为所述第二操作序列所需要插入、删除或者替换的序列元素的数目。

相关关键词的显示方法及装置

技术领域

[0001] 本发明涉及信息处理领域,尤其涉及一种相关关键词的显示方法及装置。

背景技术

[0002] 搜索引擎是指根据一定的策略、运用特定的计算机程序搜集互联网上的信息,在对信息进行组织和处理后,为用户提供检索服务的系统。搜索引擎作为用户获取信息的主要渠道之一,确定用户的搜索目标是首要的任务,目前的主要搜索引擎都是利用用户输入关键词确定用户的搜索目标对应的目标关键词,并将确定的目标关键词作为与用户输入的关键词相关的关键词显示给用户,以使用户进行选择。发明人在实现现有技术提供的方案过程中,发现现有技术提供的方案限制了搜索引擎可以为用户提供的相关关键词的范围,降低了搜索引擎为用户提供的相关关键词的准确性。

发明内容

[0003] 本发明的实施例提供一种相关关键词的显示方法及装置,提高了搜索引擎为用户提供的相关关键词的准确性。

[0004] 为达到上述目的,本发明的实施例采用如下技术方案:

[0005] 一种相关关键词的显示方法,包括:

[0006] 获取用户输入关键词的第一操作序列,所述第一操作序列包含所述用户输入关键词的所有操作行为;

[0007] 根据所述第一操作序列,在数据库中查找与所述第一操作序列对应的第二操作序列,所述第二操作序列为与所述第一操作序列之间满足预设的序列相似度的操作序列;

[0008] 若查找到与所述第一操作序列对应的第二操作序列,则将所述第二操作序列对应的目标关键词作为所述用户输入的关键词的相关关键词进行显示。

[0009] 一种相关关键词的显示的装置,包括:

[0010] 获取单元,用于获取用户输入关键词的第一操作序列,所述第一操作序列包含所述用户输入关键词的所有操作行为;

[0011] 查找单元,用于在数据库中查找与所述第一操作序列对应的第二操作序列,所述第二操作序列为与所述第一操作序列之间满足预设的序列相似度的操作序列;

[0012] 显示单元,用于在所述查找单元查找到与所述第一操作序列对应的第二操作序列时,将所述第二操作序列对应的目标关键词作为所述用户输入的关键词的相关关键词进行显示。

[0013] 本发明实施例提供的相关关键词的显示的方法及装置,通过使用用户输入关键词的操作序列来查找所述用户输入的关键词的相关关键词,与现有技术中基于关键字来查找相关关键词相比,为所述用户提供的相关关键词可以不局限于必须包含所述用户输入的关键词,使得能够为用户提供更大范围的相关关键词,缩短了用户搜索结果的路径,从而提高了搜索引擎为用户提供的相关关键词的准确性。

附图说明

[0014] 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一些实施例,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,还可以根据这些附图获得其他的附图。

[0015] 图 1 为本发明实施例 1 中的一种相关关键词的显示方法的流程图;

[0016] 图 2 为本发明实施例 1 中的另一种相关关键词的显示方法的流程图;

[0017] 图 3 为本发明实施例 2 中的一种相关关键词的显示装置的组成框图;

[0018] 图 4 为本发明实施例 2 中的另一种相关关键词的显示装置的组成框图;

[0019] 图 5 为本发明实施例 2 中的另一种相关关键词的显示装置的组成框图。

具体实施方式

[0020] 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作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。

[0021] 实施例 1

[0022] 本发明实施例提供一种相关关键词的显示方法,如图 1 所示,该方法包括:

[0023] 101、获取用户输入关键词的第一操作序列,所述第一操作序列包含所述用户输入关键词的所有操作行为。

[0024] 其中,所述操作行为可以是更换输入法、输入字符、输入空格或删除字符等操作。所述第一操作序列就是将用户输入关键词的操作行为按照用户对操作行为的执行顺序组成的序列,例如,用户输入的关键词为“饭店”,则第一操作序列为{输入“饭”,输入“店”},若用户将“饭店”修改为“饭庄”,则第一操作序列为{输入“饭”,输入“店”,删除“店”、输入“庄”}。

[0025] 其中,所述获取用户输入关键词的第一操作序列的实现方法,如图 2 所示,具体包括:

[0026] 1011、采集所述用户输入关键词的操作行为。

[0027] 其中,所述采集所述用户输入关键词的操作行为可以通过以下方法实现,具体包括:

[0028] 当采集到所述用户输入关键词的第一个操作行为时,为所述用户设置采集起始时间戳。

[0029] 按照预定时长以及所述采集起始时间戳,设置与所述采集起始时间戳对应的采集终止时间戳。

[0030] 在所述采集起始时间戳以及所述采集终止时间戳之间的时间段内,采集所述用户输入关键词的操作行为。

[0031] 其中,所述用户输入关键词的第一个操作行为可以定义为用户在使用搜索引擎时在搜索栏内输入关键字的过程中的动作,例如,用户在搜索栏内输入“我们”两个字,则输入

“我”这一动作为该用户的第一个操作行为。此时,服务器要在该用户输入“我”这一操作动作的同时为该用户设定采集起始时间戳。

[0032] 其中,采集起始时间戳可以设置为采集到用户第一个操作行为时的时间戳,也可以独立设置新的时间戳,本发明实施例对此不进行限制。

[0033] 值得说明的是,采集终止时间戳的设置可以设定为采集起始时间戳之后的 5 分钟或 10 分钟。当然,对于采集终止时间戳的设置,本发明实施例对此不进行限制。

[0034] 另外,需要说明的是,当在所述采集起始时间戳以及所述采集终止时间戳之间的时间段内采集到所述用户结束搜索的操作行为时,结束对所述用户输入关键词的操作行为的采集。

[0035] 1012、根据所述用户的标识信息,从已存储的操作序列中查找所述用户对应的原始操作序列,所述原始操作序列包含所述用户在输入所述关键词之前进行的所有操作行为。若查找到所述用户对应的原始操作序列,则执行步骤 1013 ;否则执行步骤 1014。

[0036] 其中,所述用户的标识信息可以为用户的 IP 地址等具有标识功能的信息,本发明实施例在此不一一列举。

[0037] 其中,所述原始操作序列与所述第一操作序列的结构相同,并且其包含的所述用户在输入所述关键词之前进行的所有操作行为与所述用户输入所述关键词同属于同一次搜索。

[0038] 1013、根据采集到的操作行为和所述原始操作序列内的操作行为以及所述采集到的操作行为和所述原始操作序列内的操作行为之间的先后顺序,生成第一操作序列,并将所述原始操作序列替换为所述第一操作序列进行存储,并执行步骤 102。

[0039] 1014、根据采集到的操作行为生成第一操作序列,并将所述第一操作序列作为所述用户对应的原始操作序列进行存储,并执行步骤 102。

[0040] 102、根据所述第一操作序列,在数据库中查找与所述第一操作序列对应的第二操作序列。

[0041] 其中,所述第二操作序列为与所述第一操作序列之间满足预设的序列相似度的操作序列,所述序列相似度可以设置为所述第一操作序列和所述第二操作序列的最长公共子序列的长度与所述第一操作序列的长度和所述第二操作序列的长度的平均值之间的比例。例如,第一操作序列为 { 输入“春”、输入“天”、输入“在”、输入“哪”、输入“里”},第二操作序列为输入“夏”、输入“天”、输入“在”、输入“哪”、输入“里”},则这两个操作序列的公共子序列为 { 输入“天”、输入“在”、输入“哪”、输入“里”},长度为 4。该第一操作序列的长度和该第二操作序列的长度的平均值为 5,则此时,第一操作序列和第二操作序列的序列相似度为 80%。

[0042] 所述序列相似度也可以设置为将所述第一操作序列转化为所述第二操作序列所需要插入、删除或者替换的序列元素的数目。例如,第一操作序列为 { 输入“春”、输入“天”、输入“在”、输入“哪”、输入“里”},第二操作序列为输入“夏”、输入“天”、输入“在”、输入“哪”、输入“里”},则将第一操作序列转换为第二操作序列需要将操作行为“输入‘春’”替换为“输入‘夏’”,此时,进行替换的序列元素为 1。

[0043] 基于上述两种序列相似度的设置方法,在使用上述两种设置方法时,都需要设置两种设置方法各自对应的阈值,例如,可以将序列相似度的阈值设置为 60%,大于或者等于

60% 认为满足预设的序列相似度, 小于 60% 则认为不满足预设的序列相似度, 或者也可以将序列相似度的阈值设置为 4, 等于或者少于 4 个, 则认为满足预设的序列相似度, 大于 4 个则不满足预设的序列相似度。本发明实施例对此不进行限制。

[0044] 其中, 所述数据库中存储有操作序列以及与所述操作序列对应的目标关键词。这些存储的目标关键词都是用户在使用搜索引擎进行搜索时在搜索栏内输入的内容。例如, 用户在搜索引擎的搜索栏内输入“我们”并进行了一次搜索, 则确定“我们”为本次搜索行为的目标关键词。而在用户输入这些目标关键词的过程中采集到的操作行为则组成与这些目标关键词存在对应关系的操作序列。通常情况下, 我们会选择存储这些目标关键词中出现次数比较多的目标关键词以及对应的操作序列。当然也可以将所有的内容都保存下来, 本发明实施例对此不进行限制。

[0045] 103、若查找到与所述第一操作序列对应的第二操作序列, 则将所述第二操作序列对应的目标关键词作为所述用户输入的关键词的相关关键词进行显示。

[0046] 其中, 所述显示所述第二操作序列对应的目标关键词的实现方法可以为将这些目标关键词以列表的形式进行显示。

[0047] 需要说明的是, 若在数据库中未查找到与所述第一操作序列对应的第二操作序列, 则可以不作任何显示, 或者使用现有技术中的根据基于关键字的相关关键词显示方法将相关关键词进行显示, 本发明实施例对此不进行限制。

[0048] 本发明实施例提供的相关关键词的显示的方法, 通过使用用户输入关键词的操作序列来查找所述用户输入的关键词的相关关键词, 与现有技术中基于关键字来查找相关关键词相比, 为所述用户提供的相关关键词可以不局限于必须包含所述用户输入的关键词, 使得能够为用户提供更大范围的相关关键词, 缩短了用户搜索结果的路径, 从而提高了搜索引擎为用户提供的相关关键词的工作效率和准确性。

[0049] 实施例 2

[0050] 本发明实施例提供了一种相关关键词的显示装置, 如图 3 所示, 该装置包括: 获取单元 21、查找单元 22、显示单元 23。

[0051] 获取单元 21, 用于获取用户输入关键词的第一操作序列, 所述第一操作序列包含所述用户输入关键词的所有操作行为。

[0052] 查找单元 22, 用于在数据库中查找与所述获取单元 21 获取的所述第一操作序列对应的第二操作序列, 所述第二操作序列为与所述第一操作序列之间满足预设的序列相似度的操作序列。

[0053] 显示单元 23, 用于在所述查找单元 22 查找到与所述第一操作序列对应的第二操作序列时, 将所述第二操作序列对应的目标关键词作为所述用户输入的关键词的相关关键词进行显示。

[0054] 可选的是, 如图 4 所示, 所述获取单元 21 包括: 采集模块 211、查找模块 212、生成模块 213、存储模块 214。

[0055] 采集模块 211, 用于采集所述用户输入关键词的操作行为。

[0056] 查找模块 212, 用于根据所述用户的标识信息, 从已存储的操作序列中查找所述用户对应的原始操作序列, 所述原始操作序列包含所述用户在输入所述关键词之前进行的所有操作行为。

[0057] 生成模块 213,用于在所述查找模块 212 查找到所述用户对应的原始操作序列时,根据所述采集模块 211 采集到的操作行为和所述原始操作序列内的操作行为以及所述采集到的操作行为和所述原始操作序列内的操作行为之间的先后顺序,生成第一操作序列。

[0058] 存储模块 214,用于将所述原始操作序列替换为所述生成模块 213 生成的第一操作序列进行存储。

[0059] 可选的是,如图 5 所示,所述采集模块 211 还包括:设置子模块 2111、采集子模块 2112。

[0060] 设置子模块 2111,用于在采集到所述用户输入关键词的第一个操作行为时,为所述用户设置采集起始时间戳。

[0061] 所述设置子模块 2111,还用于按照预定时长以及所述采集起始时间戳,设置与所述采集起始时间戳对应的采集终止时间戳。

[0062] 采集子模块 2112,用于在所述设置子模块 2111 设置好的采集起始时间戳以及采集终止时间戳之间的时间段内,采集所述用户输入关键词的操作行为。

[0063] 可选的是,所述集子模块 2112,还用于在在所述采集起始时间戳以及所述采集终止时间戳之间的时间段内采集到所述用户结束搜索的操作行为时,结束对所述用户输入关键词的操作行为的采集。

[0064] 可选的是,所述生成模块 213,还用于在所述查找模块 212 未查找到所述用户对应的原始操作序列时,根据所述采集模块 211 采集到的操作行为生成第一操作序列;

[0065] 所述存储模块 214,还用于将所述生成模块 213 生成的第一操作序列作为所述用户对应的原始操作序列进行存储。

[0066] 值得说明的是,本发明实施例中描述的序列相似度为为所述第一操作序列和所述第二操作序列的最长公共子序列的长度与所述第一操作序列的长度和所述第二操作序列的长度的平均值之间的比例;或者,所述序列相似度为将所述第一操作序列转化为所述第二操作序列所需要插入、删除或者替换的序列元素的数目。

[0067] 本发明实施例提供的相关关键词的显示的装置,通过使用用户输入关键词的操作序列来查找所述用户输入的关键词的相关关键词,与现有技术中基于关键字来查找相关关键词相比,为所述用户提供的相关关键词可以不局限于必须包含所述用户输入的关键词,使得能够为用户提供更大范围的相关关键词,缩短了用户搜索结果的路径,从而提高了搜索引擎为用户提供的相关关键词的工作效率和准确性。

[0068] 通过以上的实施方式的描述,所属领域的技术人员可以清楚地了解到本发明可借助软件加必需的通用硬件的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献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可读取的存储介质中,如计算机的软盘,硬盘或光盘等,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,或者网络设备等)执行本发明各个实施例所述的方法。

[0069] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以所述权利要求的保护范围为准。

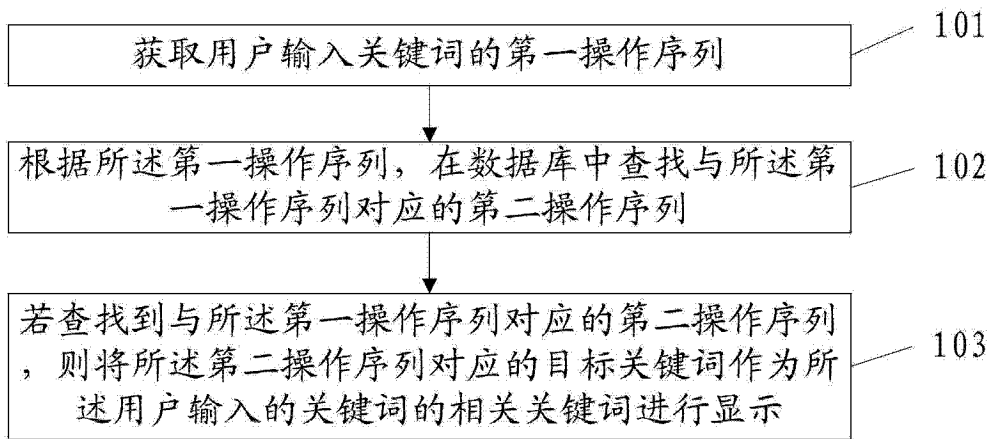


图 1

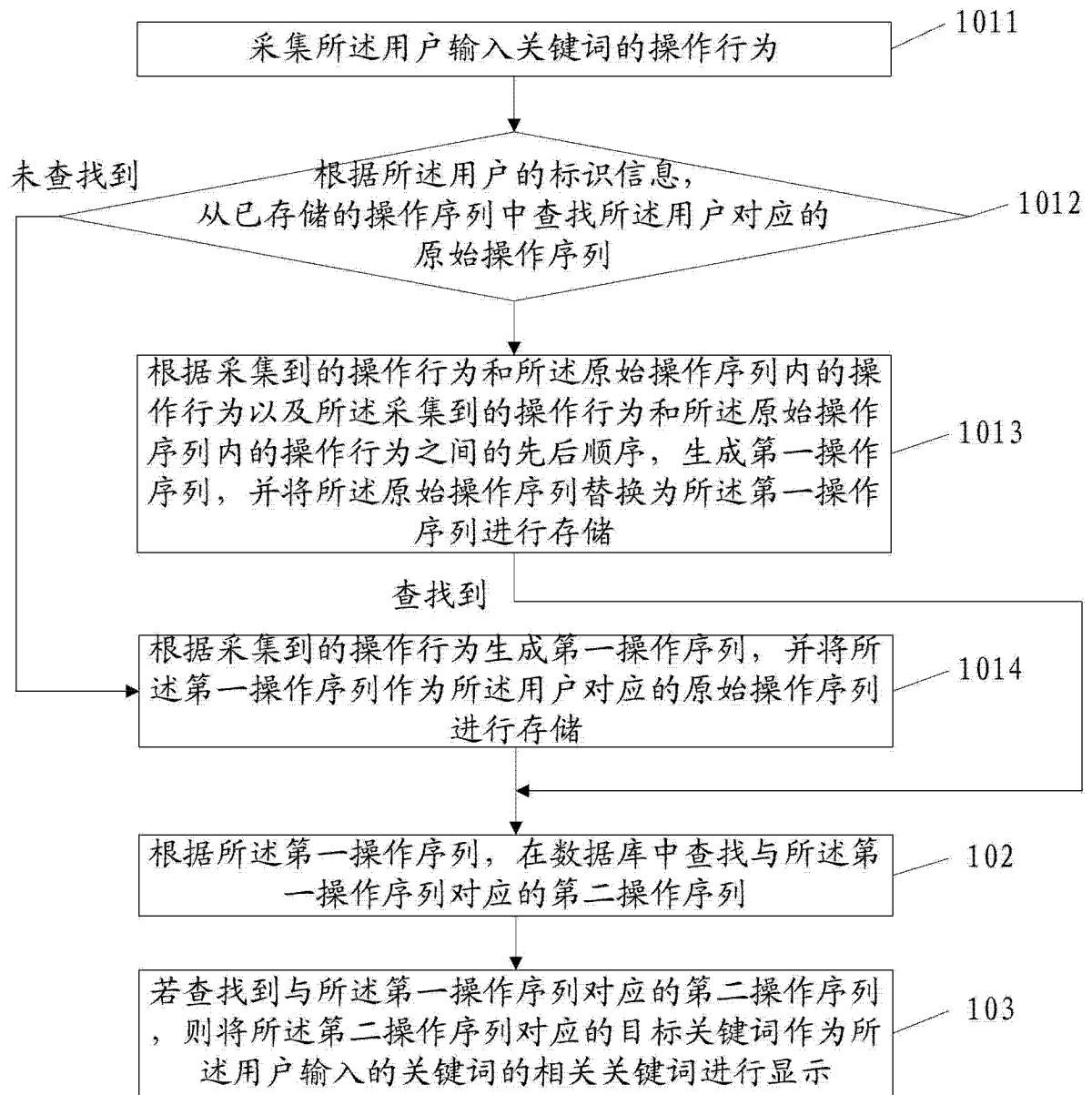


图 2

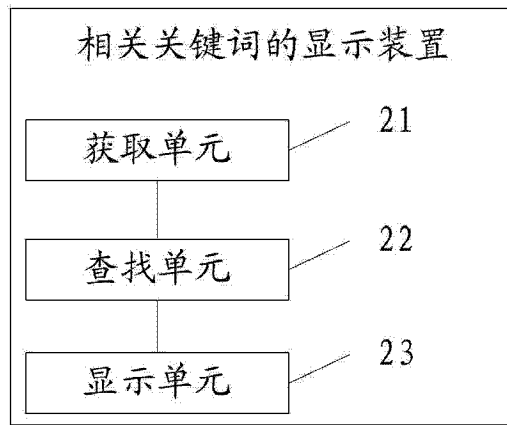


图 3

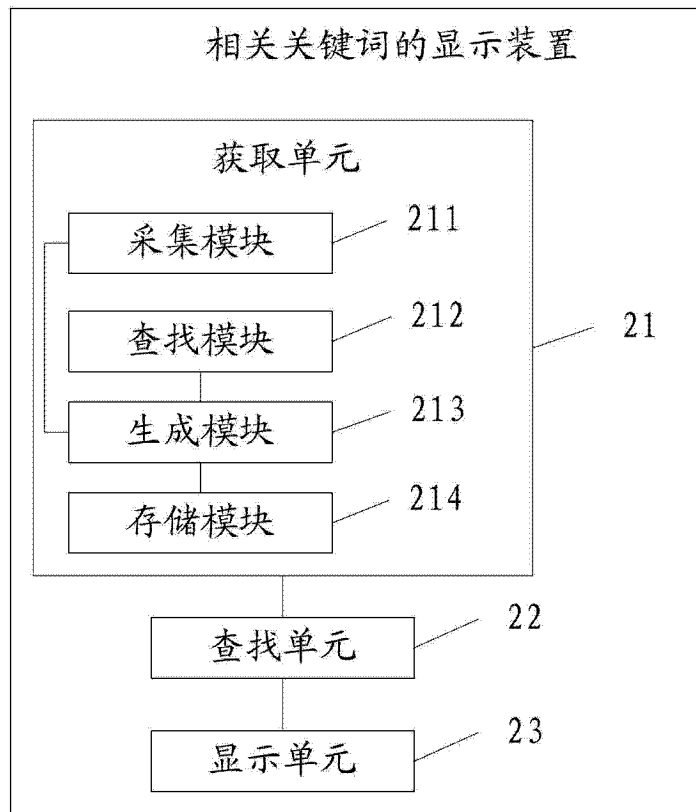


图 4

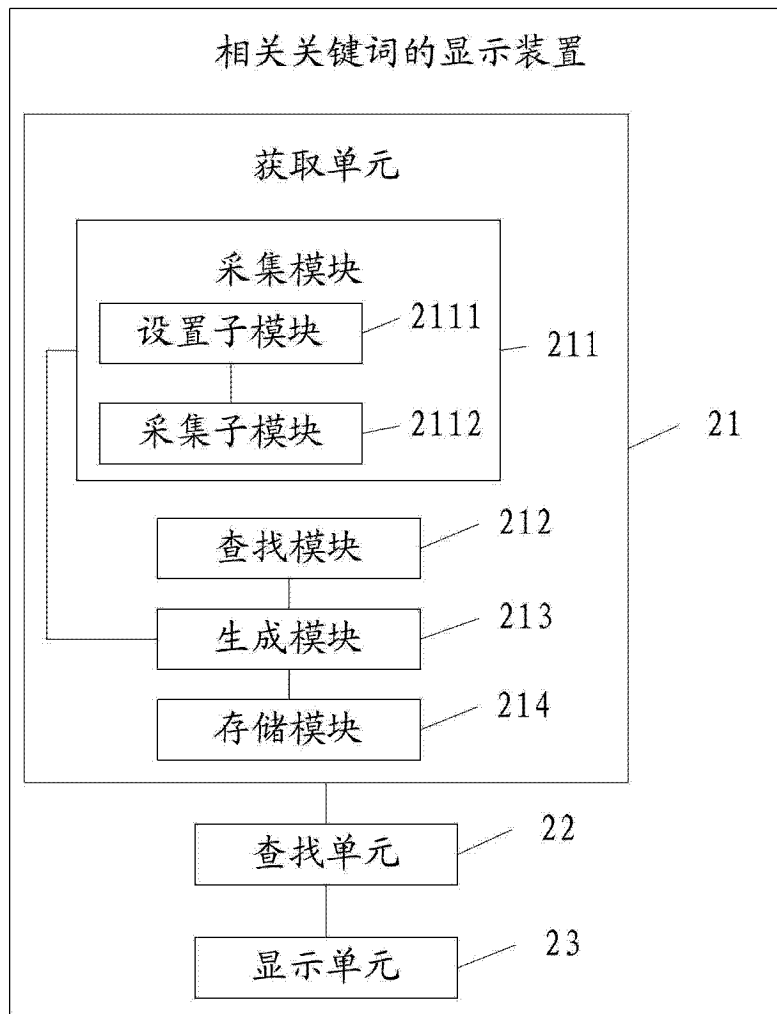


图 5