



(12) 发明专利申请

(10) 申请公布号 CN 112559717 A

(43) 申请公布日 2021.03.26

(21) 申请号 202011555351.0

(22) 申请日 2020.12.24

(71) 申请人 北京百度网讯科技有限公司
地址 100085 北京市海淀区上地十街10号
百度大厦2层

(72) 发明人 苑浩 谢子哲 彭程

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021
代理人 鄢功军

(51) Int. Cl.

G06F 16/332 (2019.01)

G06F 16/33 (2019.01)

G06F 16/338 (2019.01)

G06F 40/216 (2020.01)

G06F 40/30 (2020.01)

权利要求书3页 说明书10页 附图5页

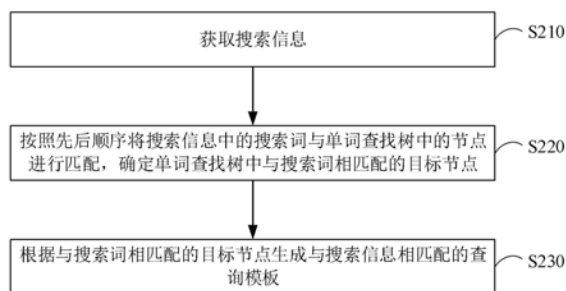
(54) 发明名称

搜索匹配方法、装置、电子设备以及存储介质

(57) 摘要

本公开公开了搜索匹配方法、装置、电子设备和存储介质,涉及计算机技术领域,尤其涉及智能搜索领域、云计算领域。具体实现方案为:获取搜索信息,其中,搜索信息中包括具有先后顺序的多个搜索词;按照先后顺序将搜索信息中的搜索词与单词查找树中的节点进行匹配,确定单词查找树中与搜索词相匹配的目标节点,其中,单词查找树中包括一种或多种类型的节点;根据与搜索词相匹配的目标节点生成与搜索信息相匹配的查询模板。

200



1. 一种搜索匹配方法,包括:
 - 获取搜索信息,其中,所述搜索信息中包括具有先后顺序的多个搜索词;
 - 按照所述先后顺序将所述搜索信息中的搜索词与单词查找树中的节点进行匹配,确定所述单词查找树中与所述搜索词相匹配的目标节点,其中,所述单词查找树中包括一种或多种类型的节点;
 - 根据与所述搜索词相匹配的目标节点生成与所述搜索信息相匹配的查询模板。
2. 根据权利要求1所述的方法,还包括:
 - 确定与所述搜索信息相匹配的查询模板对应的目标答复策略;
 - 根据所述搜索信息执行所述目标答复策略。
3. 根据权利要求2所述的方法,其中,所述根据所述搜索信息执行所述目标答复策略包括:
 - 利用所述查询模板确定所述搜索信息中的关键搜索词;
 - 基于所述关键搜索词执行所述目标答复策略。
4. 根据权利要求1所述的方法,还包括:
 - 获取模板配置操作,其中,所述模板配置操作中包括模板内容;
 - 响应于所述模板配置操作,根据所述模板内容生成新的单词查找树;以及
 - 将所述新的单词查找树与所述用户标识关联的存储在缓存服务器中。
5. 根据权利要求4所述的方法,还包括:
 - 对所述缓存服务器中与所述用户标识相关联的单词查找树的更新时间进行监控;
 - 在所述缓存服务器中与用户标识相关联的单词查找树的更新时间与上一次更新时间不同的情况下,将发生更新的与所述用户标识相关联的单词查找树存储在搜索引擎数据库中。
6. 根据权利要求4所述的方法,其中,所述获取模板配置操作包括:
 - 获取用户根据业务数据自定义的模板内容;
 - 根据所述用户根据业务数据自定义的模板内容,生成所述模板配置操作。
7. 根据权利要求1所述的方法,其中,所述根据与所述搜索词相匹配的目标节点生成与所述搜索信息相匹配的查询模板包括:
 - 将与所述搜索词相匹配的目标节点对应的字符内容,按照所述目标节点在所述单词查找树中的顺序,组合为所述查询模板。
8. 根据权利要求1所述的方法,其中,所述单词查找树中节点的类型包括以下至少一种:
 - 模糊字符节点,用于存储模糊匹配字符;
 - 实体节点,用于存储实体名称;
 - 通配符节点,用于存储通配符类型;
 - 抽取开始节点和抽取结束节点,用于标识抽取内容的起始位置和结束位置。
9. 一种搜索匹配装置,包括:
 - 第一获取模块,用于获取搜索信息,其中,所述搜索信息中包括具有先后顺序的多个搜索词;
 - 匹配模块,用于按照所述先后顺序将所述搜索信息中的搜索词与单词查找树中的节点

进行匹配,确定所述单词查找树中与所述搜索词相匹配的目标节点,其中,所述单词查找树中包括一种或多种类型的节点;

生成模块,用于根据与所述搜索词相匹配的目标节点生成与所述搜索信息相匹配的查询模板。

10. 根据权利要求9所述的装置,还包括:

确定模块,用于确定与所述搜索信息相匹配的查询模板对应的目标答复策略;

执行模块,用于根据所述搜索信息执行所述目标答复策略。

11. 根据权利要求10所述的装置,其中,所述执行模块包括:

确定单元,用于利用所述查询模板确定所述搜索信息中的关键搜索词;

执行单元,用于基于所述关键搜索词执行所述目标答复策略。

12. 根据权利要求9所述的装置,还包括:

第二获取模块,用于获取模板配置操作,其中,所述模板配置操作中包括模板内容;

响应模块,用于响应于所述模板配置操作,根据所述模板内容生成新的单词查找树;以及

第一存储模块,用于将所述新的单词查找树与用户标识关联的存储在缓存服务器中。

13. 根据权利要求12所述的装置,还包括:

监控模块,用于对所述缓存服务器中与所述用户标识相关联的单词查找树的更新时间进行监控;

第二存储模块,用于在所述缓存服务器中与所述用户标识相关联的单词查找树的更新时间与上一次更新时间不同的情况下,将发生更新的与所述用户标识相关联的单词查找树存储在搜索引擎数据库中。

14. 根据权利要求12所述的装置,其中,所述第二获取模块包括:

获取单元,用于获取用户根据业务数据自定义的模板内容;

生成单元,用于根据所述用户根据业务数据自定义的模板内容,生成所述模板配置操作。

15. 根据权利要求9所述的装置,其中,所述生成模块用于:将与所述搜索词相匹配的目标节点对应的字符内容,按照所述目标节点在所述单词查找树中的顺序,组合为所述查询模板。

16. 根据权利要求9所述的装置,其中,所述单词查找树中节点的类型包括以下至少一种:

模糊字符节点,用于存储模糊匹配字符;

实体节点,用于存储实体名称;

通配符节点,用于存储通配符类型;

抽取开始节点和抽取结束节点,用于标识抽取内容的起始位置和结束位置。

17. 一种电子设备,包括:

至少一个处理器;以及

与所述至少一个处理器通信连接的存储器;其中,

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行权利要求1-8中任一项所述的方法。

18. 一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行根据权利要求1-8中任一项所述的方法。

19. 一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现根据权利要求1-8中任一项所述的方法。

搜索匹配方法、装置、电子设备以及存储介质

技术领域

[0001] 本公开涉及计算机技术领域,尤其涉及智能搜索领域。

背景技术

[0002] 在人机对话系统中,例如智能客服系统、语音助手等等,对于用户输入的查询语句,需要实时将查询内容映射到某一个答案上。

[0003] 相关技术中一般使用关键词检索、语义相似度以及正则匹配方式响应用户输入,但在相关技术中对用户的查询内容进行匹配的方法存在解析耗时,效率低的技术问题。

发明内容

[0004] 本公开提供了一种搜索匹配方法、装置、电子设备以及存储介质。

[0005] 根据本公开的一方面,提供了一种搜索匹配方法,包括:获取搜索信息,其中,上述搜索信息中包括具有先后顺序的多个搜索词;按照上述先后顺序将上述搜索信息中的搜索词与单词查找树中的节点进行匹配,确定上述单词查找树中与上述搜索词相匹配的目标节点,其中,上述单词查找树中包括一种或多种类型的节点;根据与上述搜索词相匹配的目标节点生成与上述搜索信息相匹配的查询模板。

[0006] 根据本公开的另一方面,提供了一种搜索匹配装置,包括:第一获取模块、匹配模块和生成模块。

[0007] 第一获取模块,用于获取搜索信息,其中,上述搜索信息中包括具有先后顺序的多个搜索词。

[0008] 匹配模块,用于按照上述先后顺序将上述搜索信息中的搜索词与单词查找树中的节点进行匹配,确定上述单词查找树中与上述搜索词相匹配的目标节点,其中,上述单词查找树中包括一种或多种类型的节点。

[0009] 生成模块,用于根据与上述搜索词相匹配的目标节点生成与上述搜索信息相匹配的查询模板。

[0010] 根据本公开的另一方面,提供了一种电子设备,包括:至少一个处理器;以及与所述至少一个处理器通信连接的存储器;其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行如上所述的方法。

[0011] 根据本公开的另一方面,提供了一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行如上所述的方法。

[0012] 根据本公开的另一方面,提供了一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现如上所述的方法。

[0013] 应当理解,本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征,也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

[0014] 附图用于更好地理解本方案,不构成对本公开的限定。其中:

[0015] 图1示意性示出了根据本公开实施例的可以应用搜索匹配方法及装置的示例性系统架构;

[0016] 图2示意性示出了根据本公开实施例的搜索匹配方法的流程图;

[0017] 图3示意性示出了根据本公开实施例的单词查找树的示意图;

[0018] 图4示意性示出了根据本公开实施例的生成新的单词查找树的流程图;

[0019] 图5示意性示出了根据本公开实施例的在配置界面上进行模板配置的示意图;

[0020] 图6示意性示出了根据本公开实施例的更新单词查找树的示意图;

[0021] 图7示意性示出了根据本公开实施例的搜索匹配装置的框图;以及

[0022] 图8示意性示出了可以用来实施本公开的实施例的示例电子设备的示意性框图。

具体实施方式

[0023] 以下结合附图对本公开的示范性实施例做出说明,其中包括本公开实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本公开的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0024] 在实现本公开的过程中发现,基于关键词检索和语义相似度的匹配以及正则匹配可以响应用户输入,实现搜索功能。基于关键词检索的方法主要为从检索词中使用TF-IDF或者其他的关键词提取算法提取出关键词,利用检索词和候选检索词的关键词交集来检索出对应的答案。基于语义相似度的匹配方法主要为使用深度模型如Bert模型对句子的语义相似度进行度量,具体做法为将候选句子和检索词输入Bert模型中,取出向量计算语义相似度,语义相似度高的句子视为匹配成功。正则匹配主要是针对固定句式编写正则表达式,解析用户的检索词是否满足这个正则表达式,满足则匹配成功。

[0025] 但基于关键词检索的方法,由于关键词本身不能代表固定句式的局限性,所以在召回上存在问题。而基于语义相似度的匹配方法存在的问题是相似度的计算取决于相似度计算模块,准确性较低,并且,对于一些未录入系统的词,词向量可能会存在缺失和不准的问题;基于正则表达式的匹配存在性能问题,正则表达式越多,性能越差,上百条的正则表达式耗时较长。

[0026] 本公开的实施例提供了搜索匹配方法、装置、电子设备和存储介质,其中,搜索匹配方法具体实现方案为:获取搜索信息,其中,搜索信息中包括具有先后顺序的多个搜索词;按照先后顺序将搜索信息中的搜索词与单词查找树中的节点进行匹配,确定单词查找树中与搜索词相匹配的目标节点,其中,单词查找树中包括一种或多种类型的节点;根据与搜索词相匹配的目标节点生成与搜索信息相匹配的查询模板。

[0027] 图1示意性示出了根据本公开实施例的可以应用搜索匹配方法及装置的示例性系统架构。

[0028] 需要注意的是,图1所示仅为可以应用本公开实施例的系统架构的示例,以帮助本领域技术人员理解本公开的技术内容,但并不意味着本公开实施例不可以用于其他设备、系统、环境或场景。例如,在另一实施例中,可以应用搜索匹配方法及装置的示例性系统架

构可以包括终端设备,但终端设备可以无需与服务器进行交互,即可实现本公开实施例提供的搜索匹配方法及装置。

[0029] 如图1所示,根据该实施例的系统架构100可以包括终端设备101、102、103,网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型,例如有线和/或无线通信链路等等。

[0030] 用户可以使用终端设备101、102、103通过网络104与服务器105交互,以接收或发送消息等。终端设备101、102、103上可以安装有各种通讯客户端应用,例如知识阅读类应用、网页浏览器应用、搜索类应用、即时通信工具、邮箱客户端和/或社交平台软件等(仅为示例)。

[0031] 终端设备101、102、103可以是具有显示屏并且支持网页浏览的各种电子设备,包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。

[0032] 服务器105可以是提供各种服务的服务器,例如对用户利用终端设备101、102、103所浏览的内容提供支持的后台管理服务器(仅为示例)。后台管理服务器可以对接收到的用户请求等数据进行分析等处理,并将处理结果(例如根据用户请求获取或生成的网页、信息、或数据等)反馈给终端设备。

[0033] 需要说明的是,本公开实施例所提供的搜索匹配方法一般可以由终端设备101、102、或103执行。相应地,本公开实施例所提供的搜索匹配装置也可以设置于终端设备101、102、或103中。

[0034] 或者,本公开实施例所提供的搜索匹配方法一般也可以由服务器105执行。相应地,本公开实施例所提供的搜索匹配装置一般可以设置于服务器105中。本公开实施例所提供的搜索匹配方法也可以由不同于服务器105且能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群执行。相应地,本公开实施例所提供的搜索匹配装置也可以设置于不同于服务器105且能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群中。

[0035] 应该理解,图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器的。

[0036] 图2示意性示出了根据本公开实施例的搜索匹配方法的流程图。

[0037] 如图2所示,该方法200包括操作S210~S230。

[0038] 在操作S210,获取搜索信息,其中,搜索信息中包括具有先后顺序的多个搜索词。

[0039] 在操作S220,按照先后顺序将搜索信息中的搜索词与单词查找树中的节点进行匹配,确定单词查找树中与搜索词相匹配的目标节点,其中,单词查找树中包括一种或多种类型的节点。

[0040] 在操作S230,根据与搜索词相匹配的目标节点生成与搜索信息相匹配的查询模板。

[0041] 根据本公开的实施例,搜索信息的类型不作限定。例如,可以是语音信息、文本信息、图像信息。其中,图像信息中可以包括文字信息。

[0042] 根据本公开的实施例,搜索信息可以是一句话或者一段话,搜索信息中包括具有先后顺序的多个搜索词。例如,“帮我查一下广州的HU2341航班”这一句搜索信息中,“帮我”、“查一下”、“广州”、“的”、“HU2341航班”具有先后顺序。根据本公开的实施例,该先后顺

序可以是搜索信息被输入的顺序。

[0043] 根据本公开的实施例,在获取到搜索信息之后,可以将搜索信息进行分词处理,得到一个或多个搜索词。当仅存在一个搜索词的时候,也可以将该搜索词与单词查找树中的节点进行匹配,确定单词查找树中与搜索词相匹配的目标节点。

[0044] 根据本公开的实施例,搜索词的顺序可以根据搜索词在搜索信息的句子中的位置进行确定。根据本公开的实施例,可以按照搜索词的先后顺序,依次与单词查找树中的节点进行匹配,每个搜索词可以具有与之匹配的目标节点。

[0045] 根据本公开的实施例,单词查找树中可以包括类型相同或不同的多个节点。

[0046] 根据本公开的实施例,在确定出与搜索词相匹配的目标节点之后,可以根据所有的目标节点生成与搜索信息相匹配的查询模板。例如,根据所有的目标节点所表征的内容,按照目标节点在单词查找树中的先后顺序自动生成查询模板。

[0047] 通过本公开的实施例,按照搜索词的先后顺序将搜索词与单词查找树中的节点进行匹配,根据与搜索词相匹配的目标节点生成与搜索信息相匹配的查询模板。本公开的实施例基于单词查找树进行搜索匹配,可以利用字符串的公共前缀来降低查询时间的开销来达到提升效率的目的。解决了相关技术中对用户的查询内容进行匹配的方法存在解析耗时,效率低的技术问题。

[0048] 根据本公开的实施例,单词查找树中可以包括一种或多种类型的节点。根据本公开的实施例,单词查找树中节点的类型包括以下至少一种:模糊字符节点,用于存储模糊匹配字符;实体节点,用于存储实体名称;通配符节点,用于存储通配符类型;抽取开始节点和抽取结束节点,用于标识抽取内容的起始位置和结束位置。

[0049] 根据本公开的实施例,单词查找树中除了上述类型的节点之外,还可以包含一些普通节点,例如,表示“的”这类普通词。根据本公开的实施例,单词查找树中可以不包括模糊字符节点,为了实现模糊匹配,可以作为一个模糊属性添加在其他节点中。

[0050] 根据本公开的实施例,利用单词查找树中的字符节点,实体节点,通配符节点,抽取开始节点和抽取结束节点,可以较为方便且灵活的设计查询模板。

[0051] 例如,通过上述类型的节点,单词查找树具备如下丰富的功能:1) 支持引用实体和词库:例如,“[@sys_number][\$my_vocab]”,其中@sys_number为数字实体,\$my_vocab为用户自定义词)。2) 支持通配符:例如,[W:0-20^{abc,edf}]不包含abc和edf的0-20个任意字符。3) 支持抽取:例如“[W:0-10]{[@sys_number]name='number'}”,输入“我要买123”,返回结果{'number':123}。4) 支持模糊匹配:例如,“~aabb”,输入任意aabb的字符串就可以匹配。

[0052] 根据本公开的实施例,单词查找树还可以称为键树,思想是空间换时间,利用字符串的公共前缀来降低查询时间的开销来达到提升效率的目的,主要用于统计和排序大量的字符串。

[0053] 根据本公开的实施例,在单词查找树中可以包括一个根节点,根节点不包含字符,除根节点之外的每个节点可以包含字符,可选地,每个节点可以仅包含一个字符。

[0054] 根据本公开的实施例,从根节点到某一个节点,路径上经过的字符链接起来,为该节点对应的字符串,该字符串组合起来可以形成查询模板;每个节点的所有子节点包含的字符可以不相同,当然,子节点包含的字符也可以相同。

[0055] 图3示意性示出了根据本公开实施例的单词查找树的示意图。

[0056] 如图3所示,该单词查找树中包括一个根节点,其中,w:0-10表示通配符,支持0-10个任意字符。w:0-5表示通配符,支持0-5个任意字符。~表示模糊匹配字符。C表示特征character,start表示抽取开始节点,end表示抽取结束节点,E表示实体,num表示数字。

[0057] 根据本公开的实施例,从根节点到某一个节点,路径上经过的字符链接起来,可以得到多个模板,例如,如图3所示,从根节点到最后一个节点,可以得到查询模板一~查询模板四。

[0058] 根据本公开的实施例,根据每个节点所表征的字符含义不同,可以设计不同的查询模板。例如,查询模板一~查询模板四表征的查询语句可以是如下。

[0059] 查询模板一,[W:0-10]查一下[@sys_location]的[@sys_filgth_num]航班。

[0060] 查询模板二,[W:0-10]查一下[@sys_location]的[@sys_airport]。

[0061] 查询模板三,[W:0-5][@sys_location]的[@sys_refer]。

[0062] 查询模板四,~[@sys_location]的[@weather]。

[0063] 根据本公开的实施例,查询模板一~查询模板四支持的查询句子包括但不限于如下。

[0064] 帮我查一下广州的HU2341航班;您好,我想查一下北京的FH403航班;哈喽哈喽,我很着急,查一下南京的HY3242航班;我想今天起飞,帮我查一下成都的机场;您好,成都的天气怎么样;冷不冷啊,石家庄的天气;求北京的最早的航班;我想坐武汉最晚一趟的航班。

[0065] 根据本公开的实施例,根据与搜索词相匹配的目标节点生成与搜索信息相匹配的查询模板包括:将与搜索词相匹配的目标节点对应的字符内容,按照目标节点在单词查找树中的顺序,组合为查询模板。

[0066] 根据本公开的实施例,以“我要钱低于123的车”的搜索信息为例,设计模板为:“[W:0-10]钱低于{[@sys_number]name=‘number’}的车”。

[0067] 根据本公开的实施例,从单词查找树的root节点开始匹配,第一个节点为通配符,则会跳过“我”节点继续匹配查询信息中剩余部分,同理也会跳过“要”节点,这时刚好匹配后面的节点“钱”、“低”、“于”,持续沿当前的分支匹配,接着进入抽取开始节点,此时的查询信息中剩余部分会进入实体识别模块,实体识别模块会返回123作为一个数字实体,则匹配到单词查找树中的实体节点,下一步进入抽取结束节点,后续的字符节点挨个匹配至结束。搜索信息和模板匹配成功,同时返回结果{‘number’:123}。

[0068] 根据本公开的实施例,模板匹配的任务是将一系列用户的查询词对应到一个模板上,是建立一种多对一的映射关系,这样面对不同的用户查询词,使用有限的模板就可以进行覆盖。

[0069] 根据本公开的实施例,建立多对一的映射关系的示例如下:查一下广州停电信息;帮我查一下里水有没有停电;帮我查一下东莞有没有停电;查一下看看汕头最近停水了嘛。

[0070] 上述四句话实际上都精确对应于一个模板,查一下|<城市>|服务(停电/停水)。

[0071] 需要说明的是,该单词查找树仅是示例性实施例,但不限于此,还可以包括本领域已知的其他形式的单词查找树,只要能够支持并实现上述功能即可。

[0072] 根据本公开的实施例,在得到与搜索信息相匹配的查询模板之后,确定与搜索信息相匹配的查询模板对应的目标答复策略;根据搜索信息执行目标答复策略。

[0073] 根据本公开的实施例,目标答复策略可以是预先设计好的,通过将查询模板与查询接口进行关联,基于查询接口对搜索信息进行响应。

[0074] 根据本公开的实施例,根据搜索信息执行目标答复策略包括:利用查询模板确定搜索信息中的关键搜索词;基于关键搜索词执行目标答复策略。

[0075] 根据本公开的实施例,在得到与搜索信息相匹配的查询模板之后,还可以返回搜索信息中的关键搜索词。关键搜索词包括但不限于数字、实体名称、位置信息等等。

[0076] 下面参考图4~图6,结合具体实施例对图2所示的方法做进一步说明。

[0077] 图4示意性示出了根据本公开实施例的生成新的单词查找树的流程图。

[0078] 如图4所示,该方法400包括操作S410~S430。

[0079] 在操作S410,获取模板配置操作,其中,模板配置操作中包括模板内容。

[0080] 在操作S420,响应于模板配置操作,根据模板内容生成新的单词查找树。

[0081] 在操作S430,将新的单词查找树与用户标识关联的存储在缓存服务器中。

[0082] 根据本公开的实施例,可以获取用户输入的模板配置操作,用户可以根据实际业务场景配置模板内容。

[0083] 根据本公开的实施例,可以将模板内容切分为多个词,定义每个词对应一种类型的节点,将词与对应类型的节点关联,根据每个词生成新的单词查找树。

[0084] 根据本公开的实施例,获取模板配置操作包括:获取用户根据业务数据自定义的模板内容;根据用户根据业务数据自定义的模板内容,生成模板配置操作。

[0085] 根据本公开的实施例,业务数据例如可以是银行相关业务,银行相关业务例如可以包括办卡业务。根据业务数据自定义的模板内容例如可以包括办借记卡、申请信用卡、开卡、开卡手续费等等。

[0086] 根据本公开的实施例,业务数据例如还可以是租车相关业务,根据业务数据自定义的模板内容例如可以包括租车去北京,租车去目标城市费用等等。

[0087] 图5示意性示出了根据本公开实施例的在配置界面上进行模板配置的示意图。

[0088] 如图5所示,在需要配置模板时,可以先展示配置界面,配置界面中可以包括规则实体、复合实体、意图/问答模板。用户可以在配置界面中输入内容,例如,用户表述为:我想租车去北京,啦啦啦.....

[0089] 根据本公开的实施例,电子设备可以基于用户输入内容生成模板内容。例如,模板内容:租车去【@sys_location】。

[0090] 根据本公开的实施例,电子设备可以基于规则实体、复合实体、意图/问答模板等进行验证,得到验证结果:符合规则,可以命中。

[0091] 根据本公开的实施例,将单词查找树作为基本数据结构,利用空间换取时间的思想,使得搜索匹配的解析速度快,请求响应速率是基于模型匹配的方式的几十到几百倍。

[0092] 根据本公开的实施例,用户可以实现实时干预模板设计且人工成本低,用户无需进行标注,只需添加新的模板,系统监控到模板改动,会自动更新模板树(即单词查找树),模板实体生效。

[0093] 根据本公开的实施例,提供自定义的模板,用户可以根据自己的需求来定义模板完成实体和同义词的匹配,以及模糊匹配,并可以抽取相应的实体。基于模板的匹配还可以按需生成搜索词。

[0094] 根据本公开的实施例,用户标识可以是产生模板配置操作的用户对应的用户标识,对于每个用户,由于用户自己配置的模板不同,每个用户会有一个用户标识,这样保证了用户在调用模板匹配服务时,响应的模板不会产生冲突。

[0095] 根据本公开的实施例,可以将新的单词查找树与用户标识关联起来,然后将关联后的新的单词查找树与用户标识存储在缓存服务器中。

[0096] 根据本公开的实施例,可以对缓存服务器中与用户标识相关联的单词查找树的更新时间进行监控;在缓存服务器中与用户标识相关联的单词查找树的更新时间与上一次更新时间不同的情况下,将发生更新的与用户标识相关联的单词查找树存储在搜索引擎数据库中。

[0097] 图6示意性示出了根据本公开实施例的更新单词查找树的示意图。

[0098] 如图6所示,用户修改模板可以采用触发式Redis(缓存服务器601)更新,可以将用户标识与当前更新时间进行关联。用户可以通过管理页面增加、删除或修改模板,使用定时器602监控Reids的改动,定时器602可以采用生产者和消费者模式。

[0099] 根据本公开的实施例,数据库603本地可以维护一个Redis数据的副本,监视Redis中的单词查找树的更新时间。当Redis的更新时间比数据库本地更新时,代表数据库603本地模版数据已经过时,可以将发生修改的用户标识加入到阻塞队列中;进一步地,从阻塞队列中一次性获取所有用户标识,批量更新这些用户标识对应的模版,即从数据库读取对应用户标识下的所有模版,在内存中创建单词查找树,并将新创建的单词查找树提供于用户。

[0100] 根据本公开的实施例,实体识别模块604可以用于识别搜索信息中的实体,模板识别模块605可以用于识别与搜索信息对应的模板。

[0101] 根据本公开的实施例,用户在进行搜索时,可以基于查询页面输入搜索信息,搜索系统可以基于数据库603中的单词查找树,利用实体识别模块604和模板识别模块605对用户输入信息进行响应。

[0102] 图7示意性示出了根据本公开实施例的搜索匹配装置的框图。

[0103] 如图7所示,搜索匹配装置700包括:第一获取模块710、匹配模块720和生成模块730。

[0104] 第一获取模块710,用于获取搜索信息,其中,搜索信息中包括具有先后顺序的多个搜索词。

[0105] 匹配模块720,用于按照先后顺序将搜索信息中的搜索词与单词查找树中的节点进行匹配,确定单词查找树中与搜索词相匹配的目标节点,其中,单词查找树中包括一种或多种类型的节点。

[0106] 生成模块730,用于根据与搜索词相匹配的目标节点生成与搜索信息相匹配的查询模板。

[0107] 通过本公开的实施例,按照搜索词的先后顺序将搜索词与单词查找树中的节点进行匹配,根据与搜索词相匹配的目标节点生成与搜索信息相匹配的查询模板。本公开的实施例基于单词查找树进行搜索匹配,可以利用字符串的公共前缀来降低查询时间的开销来达到提升效率的目的。解决了相关技术中对用户的查询内容进行匹配的方法存在解析耗时,效率低的技术问题。

[0108] 根据本公开的实施例,搜索匹配装置700还包括:确定模块和执行模块。

- [0109] 确定模块,用于确定与搜索信息相匹配的查询模板对应的目标答复策略。
- [0110] 执行模块,用于根据搜索信息执行目标答复策略。
- [0111] 根据本公开的实施例,执行模块包括:确定单元和执行单元。
- [0112] 确定单元,用于利用查询模板确定搜索信息中的关键搜索词;
- [0113] 执行单元,用于基于关键搜索词执行目标答复策略。
- [0114] 根据本公开的实施例,搜索匹配装置700还包括:第二获取模块、响应模块和第一存储模块。
- [0115] 第二获取模块,用于获取模板配置操作,其中,模板配置操作中包括模板内容。
- [0116] 响应模块,用于响应于模板配置操作,根据模板内容生成新的单词查找树。
- [0117] 第一存储模块,用于将新的单词查找树与用户标识关联的存储在缓存服务器中。
- [0118] 根据本公开的实施例,搜索匹配装置700还包括:监控模块和第二存储模块。
- [0119] 监控模块,用于对缓存服务器中与用户标识相关联的单词查找树的更新时间进行监控。
- [0120] 第二存储模块,用于在缓存服务器中与用户标识相关联的单词查找树的更新时间与上一次更新时间不同的情况下,将发生更新的与用户标识相关联的单词查找树存储在搜索引擎数据库中。
- [0121] 根据本公开的实施例,第二获取模块包括:获取单元和生成单元。
- [0122] 获取单元,用于获取用户根据业务数据自定义的模板内容;
- [0123] 生成单元,用于根据用户根据业务数据自定义的模板内容,生成模板配置操作。
- [0124] 根据本公开的实施例,生成模块用于:将与搜索词相匹配的目标节点对应的字符内容,按照目标节点在单词查找树中的顺序,组合为查询模板。
- [0125] 根据本公开的实施例,单词查找树中节点的类型包括以下至少一种:模糊字符节点,用于存储模糊匹配字符;实体节点,用于存储实体名称;通配符节点,用于存储通配符类型;抽取开始节点和抽取结束节点,用于标识抽取内容的起始位置和结束位置。
- [0126] 根据本公开的实施例,本公开还提供了一种电子设备、一种可读存储介质和一种计算机程序产品。
- [0127] 根据本公开的实施例,还提供了一种电子设备,包括:至少一个处理器;以及与至少一个处理器通信连接的存储器;其中,存储器存储有可被至少一个处理器执行的指令,指令被至少一个处理器执行,以使至少一个处理器能够执行如上所述的方法。
- [0128] 根据本公开的实施例,还提供了一种存储有计算机指令的非瞬时计算机可读存储介质,其中,计算机指令用于使计算机执行如上所述的方法。
- [0129] 根据本公开的实施例,还提供了一种计算机程序产品,包括计算机程序,计算机程序在被处理器执行时实现如上所述的方法。
- [0130] 图8示出了可以用来实施本公开的实施例的示例电子设备800的示意性框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例,并且不意在限制本文中描述的和/或者要求的本公开的实现。

[0131] 如图8所示,设备800包括计算单元801,其可以根据存储在只读存储器 (ROM) 802中的计算机程序或者从存储单元808加载到随机访问存储器 (RAM) 803中的计算机程序,来执行各种适当的动作和处理。在RAM 803中,还可存储设备800操作所需的各种程序和数据。计算单元801、ROM 802以及RAM 803通过总线804彼此相连。输入/输出 (I/O) 接口805也连接至总线804。

[0132] 设备800中的多个部件连接至I/O接口805,包括:输入单元806,例如键盘、鼠标等;输出单元807,例如各种类型的显示器、扬声器等;存储单元808,例如磁盘、光盘等;以及通信单元809,例如网卡、调制解调器、无线通信收发机等。通信单元809允许设备800通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

[0133] 计算单元801可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元801的一些示例包括但不限于中央处理单元 (CPU)、图形处理单元 (GPU)、各种专用的人工智能 (AI) 计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器 (DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元801执行上文所描述的各个方法和处理,例如搜索匹配方法。例如,在一些实施例中,上述方法可被实现为计算机软件程序,其被有形地包含于机器可读介质,例如存储单元808。在一些实施例中,计算机程序的部分或者全部可以经由ROM 802和/或通信单元809而被载入和/或安装到设备800上。当计算机程序加载到RAM 803并由计算单元801执行时,可以执行上文描述的方法的一个或多个步骤。备选地,在其他实施例中,计算单元801可以通过其他任何适当的方式 (例如,借助于固件) 而被配置为执行上述方法。

[0134] 本文中以上描述的系统和技术和各种实施方式可以在数字电子电路系统、集成电路系统、场可编程门阵列 (FPGA)、专用集成电路 (ASIC)、专用标准产品 (ASSP)、芯片上系统的系统 (SOC)、负载可编程逻辑设备 (CPLD)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0135] 用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理或控制器,使得程序代码当由处理器或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行,作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

[0136] 在本公开的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器 (RAM)、只读存储器 (ROM)、可擦除可编程只读存储器 (EPROM 或快闪存储器)、光纤、便捷式紧凑盘只读存储器 (CD-ROM)、光学储存设备、磁储存设备、或

上述内容的任何合适组合。

[0137] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0138] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0139] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。

[0140] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本发公开中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本公开公开的技术方案所期望的结果,本文在此不进行限制。

[0141] 上述具体实施方式,并不构成对本公开保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等,均应包含在本公开保护范围之内。

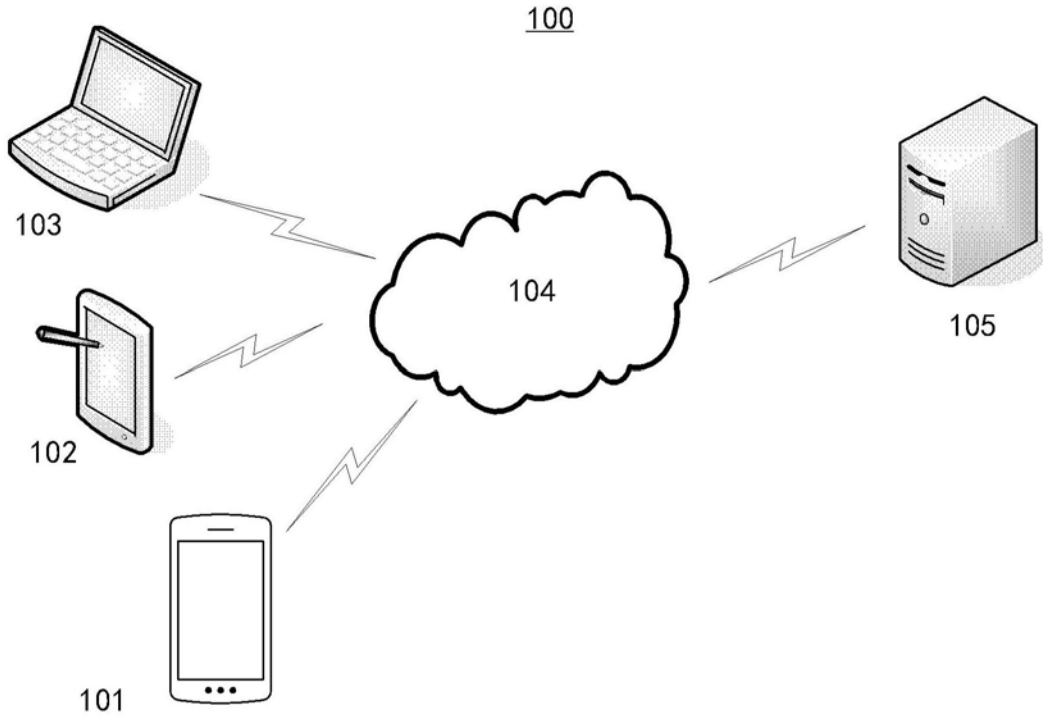


图1

200

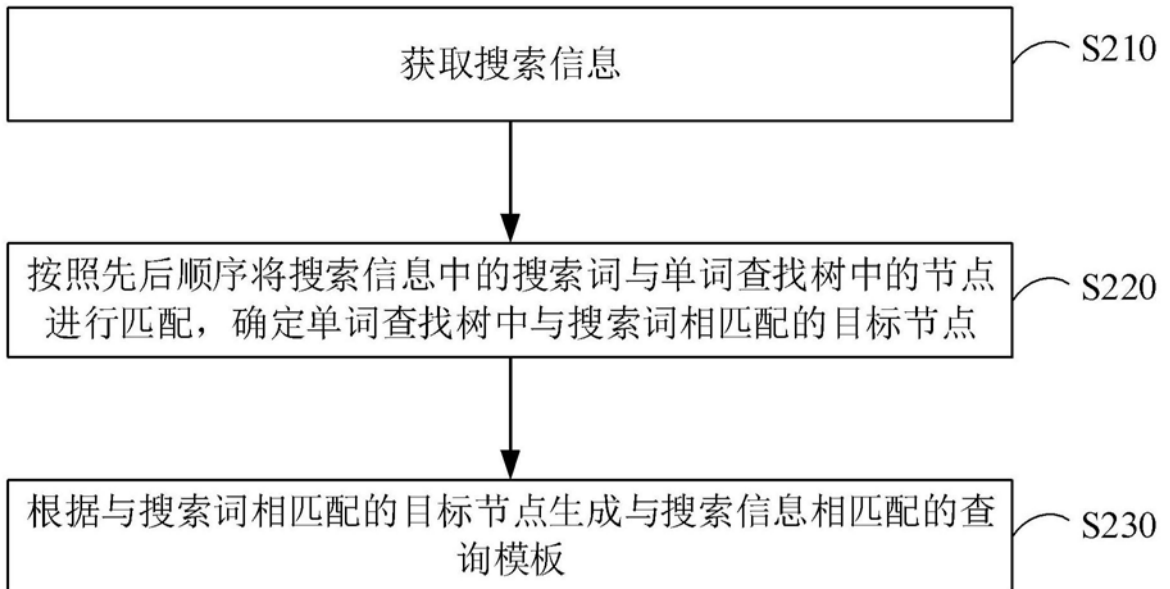


图2

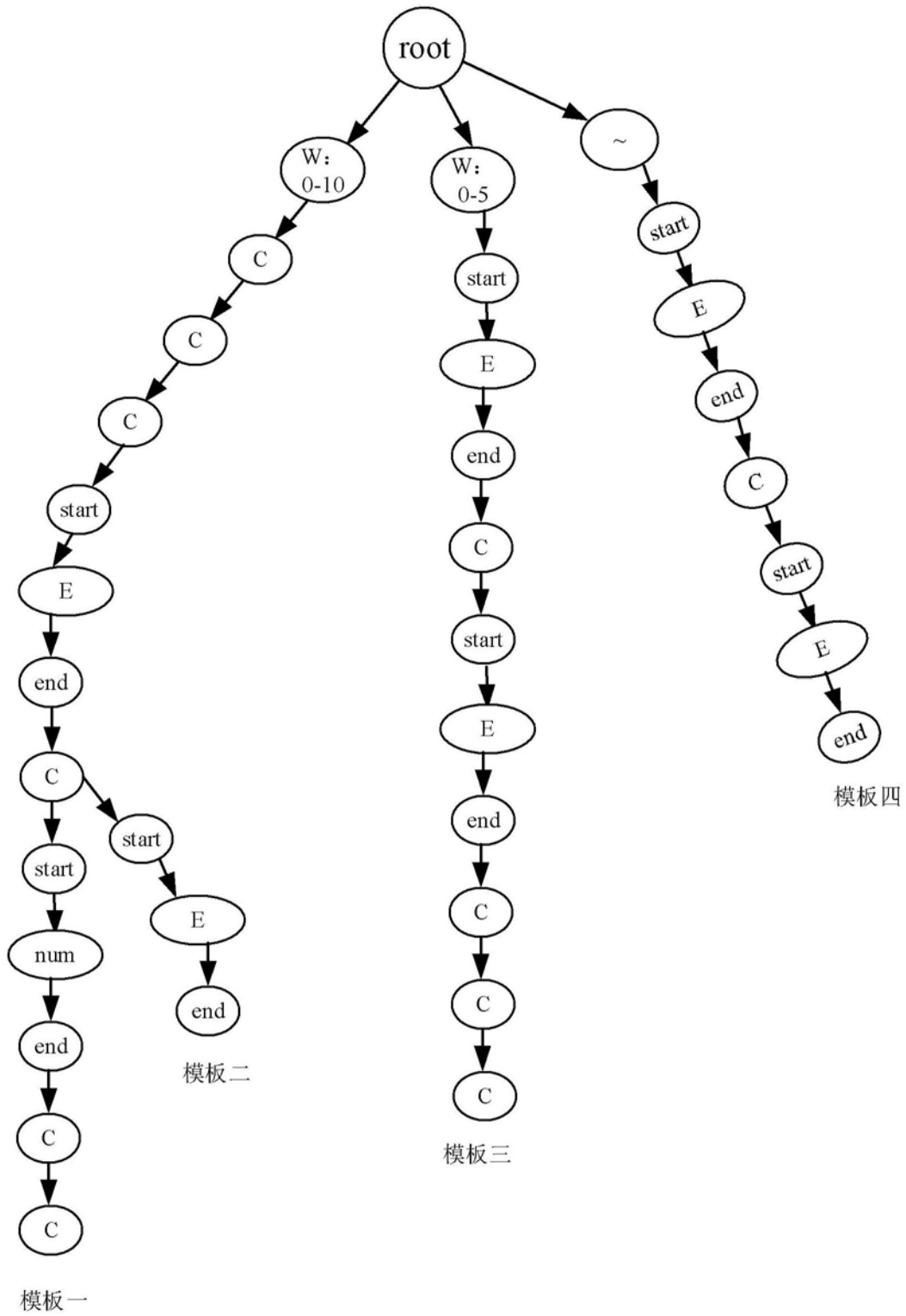


图3

400

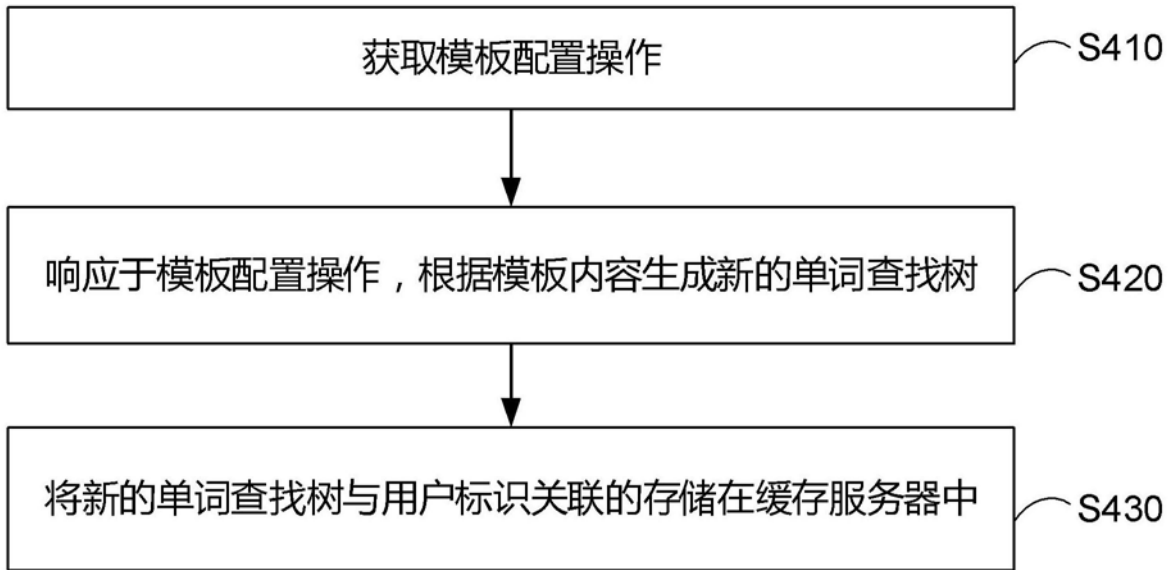


图4

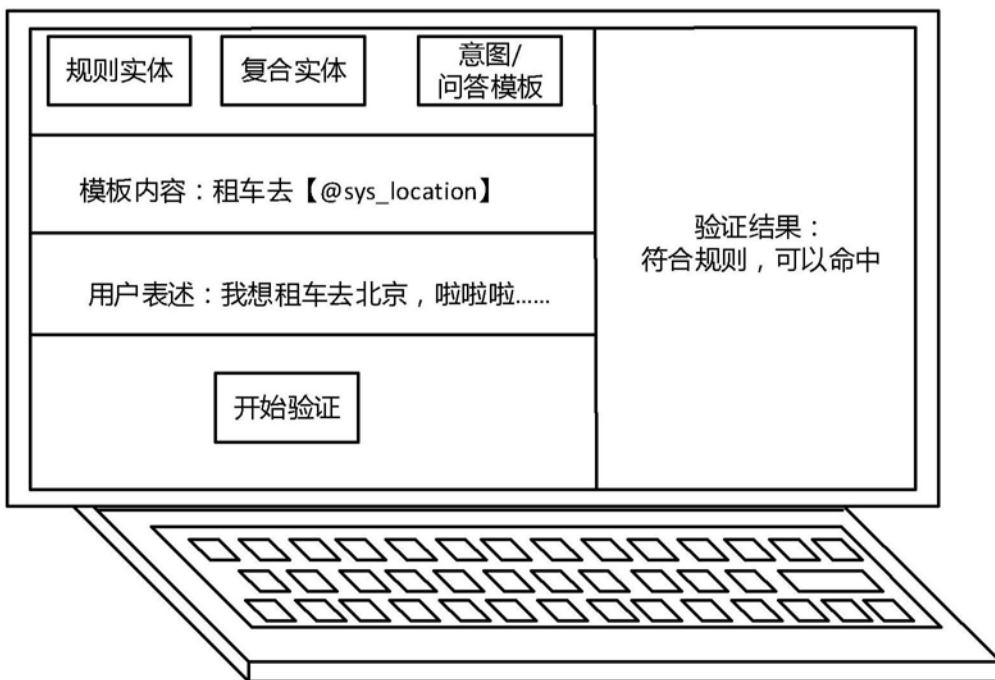


图5

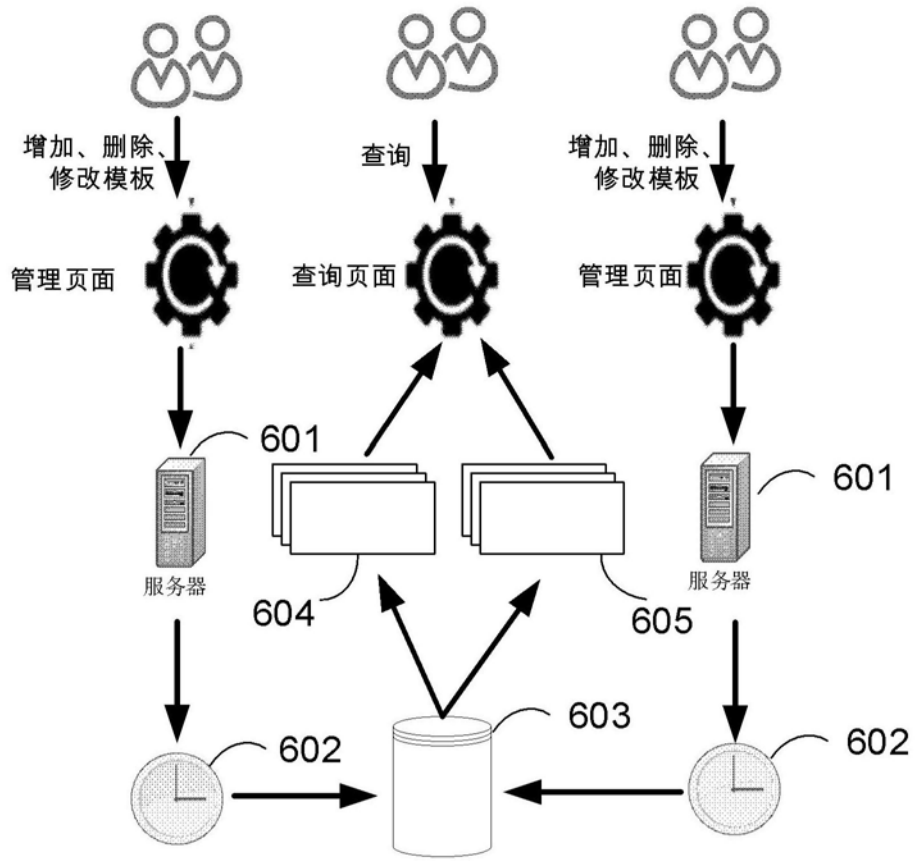


图6

700

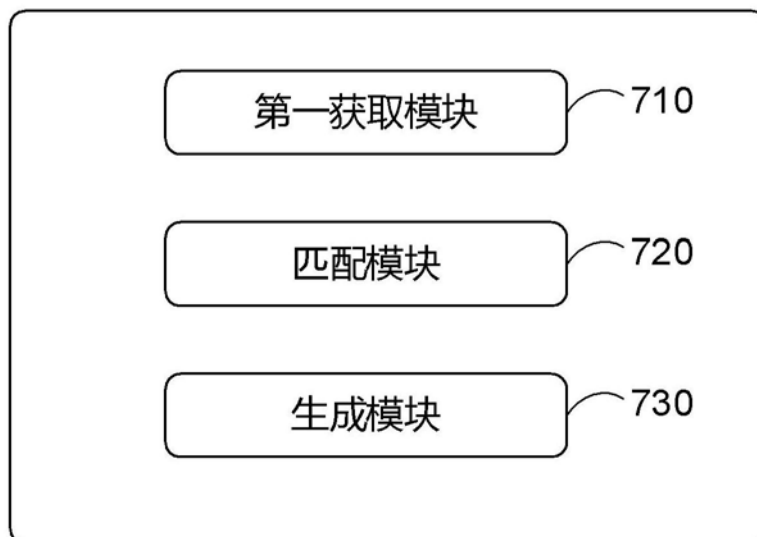


图7

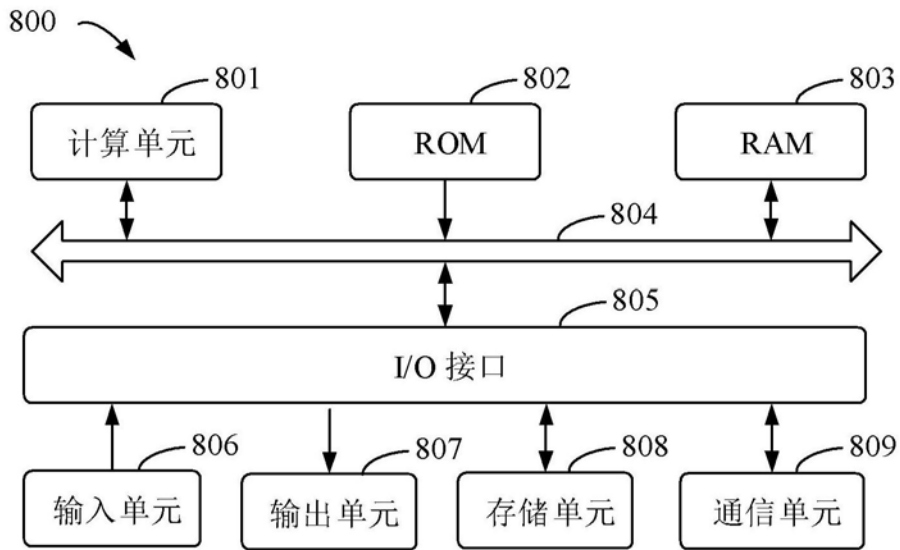


图8