

(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
2020 年 8 月 6 日 (06.08.2020)



(10) 国际公布号
WO 2020/155747 A1

(51) 国际专利分类号:

G06F 16/332 (2019.01)

街道福安社区益田路 5033 号平安金融中心
23 楼, Guangdong 518000 (CN)。

(21) 国际申请号:

PCT/CN2019/117166

(22) 国际申请日:

2019 年 11 月 11 日 (11.11.2019)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201910078024.1 2019 年 1 月 28 日 (28.01.2019) CN

(71) 申请人: 平安科技 (深圳) 有限公司 (PING AN

TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];

中国广东省深圳市福田区福田街道福
安社区益田路 5033 号平安金融中心 23
楼, Guangdong 518000 (CN)。

(72) 发明人: 金戈 (JIN, Ge); 中国广东省深圳市福

田区福田街道福安社区益田路 5033 号平安
金融中心 23 楼, Guangdong 518000 (CN)。 徐亮

(XU, Liang); 中国广东省深圳市福田区福田

(74) 代理人: 广州三环专利商标代理有限公司

(SCIHEAD IP LAW FIRM); 中国广东省广州市

越秀区先烈中路 80 号汇华商贸大厦 1508

室, Guangdong 510070 (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,

GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,

JO, JP, KE, KG, KH, KN, KP, KR, KW, 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,

(54) Title: METHOD AND APPARATUS FOR RECOMMENDING ANSWER TO QUESTION, AND STORAGE MEDIUM AND SERVER

(54) 发明名称: 问题答案推荐方法、装置、存储介质及服务器

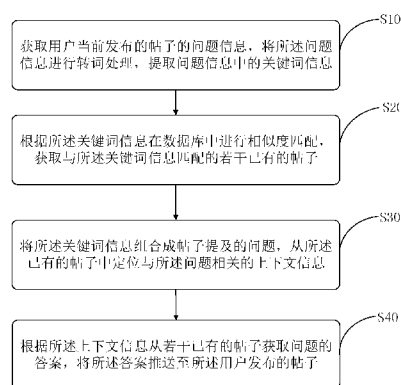


图 1

S10 Acquire question information of a post currently published by a user, carry out word conversion processing on the question information, and extract keyword information from the question information

S20 According to the keyword information, carry out similarity matching in a database to acquire several existing posts matching the keyword information

S30 Combine the keyword information into a question mentioned in the post, and locate context information, related to a question, from the existing posts

S40 Acquire, according to the context information, answers to the question from the several existing posts, and push the answers to the post published by the user

(57) Abstract: A method and apparatus for recommending an answer to a question, and a storage medium and a server. The method comprises: acquiring question information of a post currently published by a user, carrying out word conversion processing on the question information, and extracting keyword information from the question information (S10); according to the keyword information, carrying out similarity matching in a database to acquire several existing posts matching the keyword information (S20); combining the keyword information into a question mentioned in the post, and locating context information, related to a question, from the existing posts (S30); and acquiring, according to the context information, answers to the question from the several existing posts, and pushing the answers to the post published by the user (S40). The accuracy of question matching can be improved, and the answers to a question related to a post can be quickly matched.

(57) 摘要: 一种问题答案推荐方法、装置、存储介质及服务器, 所述方法包括: 获取用户当前发布的帖子的问题信息, 将所述问题信息进行转词处理, 提取问题信息中的关键词信息 (S10); 根据所述关键词信息在数据库中进行相似度匹配, 获取与所述关键词信息匹配的若干已有的帖子 (S20); 将所述关键词信息组合成帖子提及的问题, 从所述已有的帖子中定位与所述问题相关的上下文信息 (S30); 根据所述上下文信息从若干已有的帖子获取问题的答案, 将所述答案推送至所述用户发布的帖子 (S40)。能够提高问题匹配的精确度, 快速地匹配帖子相关的问题答案。

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)。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

问题答案推荐方法、装置、存储介质及服务器

本申请要求于 2019 年 01 月 28 日提交中国专利局、申请号为 201910078024.1、申请名称为“问题答案推荐方法、装置、存储介质及服务器”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本申请涉及数据处理领域，具体涉及一种问题答案推荐方法、装置、存储介质及服务器。

10 背景技术

随着互联网技术的发展，社区平台越来越受到人们的欢迎，用户可以通过社区平台爆料新闻、发表观点或寻求帮助，例如，现阶段很多论坛，用户在论坛上会发布帖子寻求帮助，而实际上论坛上很多已拥有的帖子的相关内容便可以回答该帖子的问题，论坛无法有效、精准地向当前发布帖子的用户推荐相关的答案，造成论坛上存在大量重复的帖子，造成资源的浪费。

15

发明内容

为克服以上技术问题，特别是现有技术无法高效匹配用户帖子相关问题的答案，特提出以下技术方案：

20 第一方面，本申请提供了一种问题答案推荐方法，包括：

获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息；

根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子；

25 将所述关键词信息组合成帖子提及的问题，从所述已有的帖子中定位与所述问题相关的上下文信息；

根据所述上下文信息从若干已有的帖子获取问题的答案，将所述答案推送至所述用户发布的帖子。

第二方面，本申请提供一种问题答案推荐装置，包括：

转词模块：用于获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息；

匹配模块：用于根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子；

5 定位模块：用于将所述关键词信息组合成帖子提及的问题，从所述已有的帖子中定位与所述问题相关的上下文信息；

推送模块：用于根据所述上下文信息从若干已有的帖子获取问题的答案，将所述答案推送至所述用户发布的帖子。

进一步的，所述所述定位模块还包括执行；

10 将所述关键词信息组合成帖子提及的问题，根据所述已有的帖子及所述问题合成答案表达式；从所述已有的帖子中定位与所述问题相关的上下文信息；将所述上下文信息与所述答案表达式相比较，确定所述上下文信息的效用，获取所述效用满足预设要求的上下文信息。

15 第三方面，本申请还提供了一种计算机非易失性可读存储介质，所述计算机可读存储介质上存储有计算机程序，该计算机程序被处理器执行时实现上述的问题答案推荐方法。

20 第四方面，本申请还提供了一种服务器，所述服务器包括一个或多个处理器、存储器、一个或多个计算机程序，其中所述一个或多个计算机程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行，所述一个或多个程序配置用于执行上述的问题答案推荐方法。

本申请与现有技术相比，具有以下有益效果：

25 本申请提供了一种基于帖子、问题、答案三种元素组合的答案推荐方法，适用于论坛发布帖子寻求解答的应用场景，用户在论坛上发布帖子时，获取用户当前发布的帖子的问题信息，并将所述问题信息进行转词处理，，确定问题信息中的关键词，提取关键词信息从而形式对所述问题信息的描述，每一个发布的帖子均有对应的关键词信息存储在数据库中，根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子，由于所述关键词信息能够组合成帖子提及的问题，从匹配到的已有的帖子中定位与所述问题相关的上下文信息，根据所述上下文信息便可清晰地了解已有帖子的内

容是否与用户当前发布的帖子的问题相关,在从若干个已有的帖子中获取所述上下文信息后,便能得到不同帖子中与用户当前发布的帖子的问题的相关答案,然后将所述答案推送至所述用户发布的帖子中,用户从推送的答案中便能直观地了解到已有的帖子发布的相关内容,查看已有的帖子的内容信息是否与自己的所需解答的问题相关,从而快速地查找到所需的答案,提高问题匹配的精确度,同时,避免论坛中相同问题或类似问题出现多次,造成论坛重复的帖子过多,从而节省了服务器有限的资源。

本申请附加的方面和优点将在下面的描述中部分给出,这些将从下面的描述中变得明显,或通过本申请的实践了解到。

10

附图说明

本申请上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

图 1 为本申请问题答案推荐方法的一实施例流程示意图;

15 图 2 为本申请问题答案推荐方法的另一实施例流程示意图;

图 3 为本申请问题答案推荐装置的一实施例示意图;

图 4 为本申请服务器的一实施例结构示意图。

具体实施方式

20 下面详细描述本申请的实施例,所述实施例的示例在附图中示出。

本技术领域技术人员可以理解,除非特意声明,这里使用的单数形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是,本申请的说明书中使用的措辞“包括”是指存在所述特征、整数、步骤、操作,但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作的 25 情况。

本申请实施例提供一种问题答案推荐方法,如图 1 所示,所述方法包括以下步骤:

S10: 获取用户当前发布的帖子的问题信息,将所述问题信息进行转词处理,提取问题信息中的关键词信息。

在论坛中，用户可以通过发布帖子爆料新闻、发表观点或寻求帮助，当用户在论坛上发布帖子时，尤其是发布寻求问题解答的帖子，为了避免论坛中相同问题或相似问题出现多次，造成论坛重复的帖子过多，本实施例中，在用户发布帖子后，例如用户点击“提交”按钮后，获取用户当前发布的帖子的内容信息，本实施例将帖子的内容信息定义为问题信息，然后将用户发布的帖子的问题信息进行转词处理，在一种实施方式中，通过爬取帖子的内容信息，然后将所述内容信息保存为 txt 文件，再对所述 txt 文件进行分词处理，提取帖子的内容信息中的关键词信息，并将所述关键词信息转化为相应的词向量，将每一个关键词转化为词向量方便后续对关键词进行计算处理。例如帖子发布的内容为“从广州过去香港机场，怎么最方便？”，将帖子的内容信息进行转词处理，提取的关键词信息包括广州、去、香港、机场、方便。

S20: 根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子。

本实施例中，每一个发布的帖子都会进行将帖子的内容信息转词处理，转化为词向量并提取内容信息中的关键词信息的操作，然后将每一个帖子的关键词信息保存在数据库，当有新的帖子发布时，根据所述关键词信息在数据库中进行相似度匹配，通过计算关键词信息对应的词向量进行匹配，例如关键词信息“北京”的词向量为 [5.0, 3.0, 9.0]，关键词信息“中国首都”的词向量为 [5.1, 2.9, 9.0]，计算该两个词向量的差值，通过所述差值便可得到两个词向量的相似度，优选的，所述相似度匹配包括中英文的匹配、拼音匹配、近义词的匹配、相似词的匹配，然后从数据库中获取与所述关键词信息匹配的若干已有的帖子，通过关键词信息进行匹配后，确定与当前帖子具有匹配的关键词信息，然后通过匹配到的关键词信息定位到已发布的帖子，例如将广州与羊城、花城、GZ、020 进行匹配，将香港与 HK、HKG 等进行匹配，获取已发布的帖子中存在该些关键词信息的帖子。

S30: 将所述关键词信息组合成帖子提及的问题，从所述已有的帖子中定位与所述问题相关的上下文信息。

在得到与所述关键词信息匹配的若干个已发布的帖子后，将所述关键词信息组合成用户当前发布的帖子提及的问题，从而更加确定问题的内容，然后从所述已有的帖子中定位与所述问题相关的上下文信息，由于一个帖子中的回复不一定全部与该帖子相关，因此需要对已有帖子的回复进行筛选，从已有的帖子中确定与所述问题相关的上下文信息，例如在已有的帖子中，将点赞数或者推荐数最高的回复作为与所述问题相关的答案之一，同时获取该回复的前后文回复，即从已有的帖子中定位与所述问题相关的上下文信息，根据所述上下文信息便可清晰地了解已有帖子的内容是否与用户当前发布的帖子的问题相关。

S40：根据所述上下文信息从若干已有的帖子获取问题的答案，将所述答案推送至所述用户发布的帖子。

在从若干个已有的帖子中获取所述上下文信息后，便能得到不同帖子中与用户当前发布的帖子的问题的相关答案，本实施例中，根据所述上下文信息从若干已有的帖子获取问题的答案，然后将所述答案推送至所述用户发布的帖子中，在一种实施方式中，可以从若干个已有的帖子中根据上下文信息获取到问题的若干个答案，然后将该若干个答案进行整合，将整合后的答案推送至用户当前发布的帖子；在另一种实施方式中，可以从若干个已有的帖子中根据上下文信息获取到问题的若干个答案，然后将该若干个答案一一推送至用户当前发布的帖子，并且从相关的答案便可定位到所述已发布的帖子中，用户从推送的答案中便能直观地了解到已有的帖子发布的相关内容，查看已有的帖子的内容信息是否与自己的所需解答的问题相关，从而快速地查找到所需的答案。

本实施例提供了一种基于帖子、问题、答案三种元素组合的答案推荐方法，适用于论坛发布帖子寻求解答的应用场景，用户在论坛上发布帖子时，获取用户当前发布的帖子的问题信息，并将所述问题信息进行转词处理，确定问题信息中的关键词，提取关键词信息从而形式对所述问题信息的描述，每一个发布的帖子均有对应的关键词信息存储在数据库中，根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干个已有的帖子，由于所述关键词信息能够组合成帖子提及的问题，从

匹配到的已有的帖子中定位与所述问题相关的上下文信息，根据所述上下文信息便可清晰地了解已有帖子的内容是否与用户当前发布的帖子的问题相关，在从若干个已有的帖子中获取所述上下文信息后，便能得到不同帖子中与用户当前发布的帖子的问题的相关答案，然后将所述答案推送至所述用户发布的帖子中，用户从推送的答案中便能直观地了解到已有的帖子发布的相关内容，查看已有的帖子的内容信息是否与自己的所需解答的问题相关，从而快速地查找到所需的答案，同时，避免论坛中相同问题或类似问题出现多次，造成论坛重复的帖子过多，从而节省了服务器有限的资源。

10 本申请的一种实施例，所述根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子，包括：

获取匹配时间段，根据所述关键词信息在数据库中进行相似度匹配，并基于所述匹配时间段获取与所述关键词信息匹配的若干已有的帖子。

随着时间的推移，相同的问题在不同时间段具有不同的答案，久远的
15 答案已不适用相同的问题，为了提供更准确的答案，本实施例中，在根据所述关键词信息在数据库中进行相似度匹配时，获取匹配时间段，限定所匹配的帖子的发布时间，在根据所述关键词信息在数据库中进行相似度匹配之后，确定了若干个已有的帖子，同时对帖子基于时间进行排序，然后基于所述匹配时间段获取与所述关键词信息匹配的若干已有的帖子，在一
20 种实施方式中，将帖子发布时间距离当前时间越近的帖子作为相似度更高的帖子，优先检索最近时间段的帖子。

进一步的，所述基于所述匹配时间段获取与所述关键词信息匹配的若干已有的帖子之后，还包括：

接收用户对所述帖子关联的答案的反馈信息，根据所述反馈信息调整
25 所述匹配时间段。

本实施例中，在匹配到已有的帖子，并根据匹配到的帖子将答案推送给用户后，接收用户的反馈信息，若用户反馈答案不准确，则修改所述匹配时间段，匹配更早期发布的帖子及相关的答案。进一步的，在接收到用户的反馈信息而修改匹配时间段后，在匹配的结果中隐藏用户已浏览过或

已反馈过答案对应的帖子，避免两次推送的答案中具有多个相同的答案，造成资源的浪费。

本申请的一种实施例，所述根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子，包括：

- 5 所述根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子，从所述已有的帖子中筛选回复数满足要求的帖子。

10 本实施例中，在根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子后，从匹配的帖子中筛选回复数达到一定数量的帖子，优选的，回复数越多的帖子的相似度更高，对于没有回复的帖子，该帖子亦不具备问题所需的答案，需要将该些帖子剔除，避免信息搜索的冗余度。

进一步的，所述从所述已有的帖子中筛选回复数满足要求的帖子之后，还包括：

- 15 接收用户对所述帖子关联的答案的反馈信息，根据所述反馈信息调整回复数要求。

20 本实施例中，在筛选出回复数满足要求的帖子，根据匹配到的帖子将答案推送给用户后，接收用户的反馈信息，若用户反馈答案不准确，则修改所述回复数要求，匹配回复数较低的帖子及相关的答案。例如，在第一次匹配时，筛选出回复数超过 n （比如 n 为 30）的若干个已有的帖子，若用户反馈答案不准确，则在第二次筛选时，筛选回复数超过 m （比如 m 为 10）的帖子。更佳的，在接收到用户的反馈信息而修改回复数要求后，在匹配的结果中隐藏用户已浏览过或已反馈过答案对应的帖子，避免两次推送的答案中具有多个相同的答案，造成资源的浪费。

- 25 本实施例的又一种实施方式，若接收到用户对所述帖子关联的答案的正反馈信息，例如接收到用户确定所推送的答案有效的反馈信息，该反馈信息视为正反馈信息，此时，增加所述答案与所述上下文信息、所述上下文信息与上述问题，所述问题与上述关键词信息的关联度，后续若有相似的帖子询问相似的问题，优先向用户推送关联度更高的已有的帖子所确定

的答案。

本申请的一种实施例，如图 2 所示，所述将所述关键词信息组合成帖子提及的问题，从所述已有的帖子中定位与所述问题相关的上下文信息，包括：

- 5 S31：将所述关键词信息组合成帖子提及的问题，根据所述已有的帖子及所述问题合成答案表达式；从所述已有的帖子中定位与所述问题相关的上下文信息；将所述上下文信息通过所述答案表达式进行计算，确定所述上下文信息的效用，获取所述效用满足预设要求的上下文信息。

- 10 本实施例中，通过所述关键词组合成用户发布的帖子所需要解答的问题，然后再根据已有的帖子及所述问题合成答案表达式，答案表达式确定了答案应该包括哪些词向量，词向量的计算规则，以及所要得到的结果的期望值；同时，在已有的帖子中定位与所述问题相关的上下文信息，所述上下文信息包含有与所述答案表达式具有相近的词向量；利用所述答案表达式计算定位到的上下文信息，确定所述上下文信息的效用，即得到该上
15 下文信息输入答案表达式中的期望值，然后获取期望值满足预设要求的上下文信息，该些上下文信息便能够提供更加准确的与问题相关的答案，从而更准确地向用户推送所需的答案。

本申请的一种实施例，所述将所述问题信息进行转词处理，提取问题信息中的关键词信息，包括：

- 20 将所述问题信息进行转词处理，提取问题信息中的关键词信息，将所述关键词信息转为关键词向量，将所述关键词向量组成帖子问题的向量矩阵；

所述根据所述关键词信息在数据库中进行相似度匹配，包括：

根据所述向量矩阵在数据库中进行相似度匹配。

- 25 通过爬取所述帖子的问题信息，然后提取帖子的问题信息中的关键词信息时，所提取的关键词信息为文本 string 型的数据，对于文本 string 型的数据，为了方便计算以进行相似度的匹配，本实施例中，将所述关键词信息转为关键词向量，每一个关键词都转化为一个关键词向量，将多个关键词组合便能得到帖子的问题信息的具体内容，为了方便计算，组合时将所

述关键词向量组成向量矩阵，每句话便可以通过词向量组装成向量矩阵，组合后的向量矩阵便能被识别为帖子的问题信息，在一种实施方式中，向量矩阵的第一行是第一个词向量，第二行是第二个，那么若干个帖子的问题的相似度便可以通过的矩阵加减法计算，本实施例中，在根据所述关键词信息在数据库中进行相似度匹配，则根据所述向量矩阵在数据库中进行相似度匹配，匹配出两个向量矩阵的差值小于预设值的结果，从而得到与
5 所述关键词信息匹配的若干已有的帖子。

本申请的一种实施例，所述获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息，包括：

10 获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息，将所述关键词信息转为关键词向量；

所述将所述关键词信息组合成帖子提及的问题，根据所述已有的帖子及所述问题合成答案表达式，包括：

15 根据所述关键词向量组成帖子提及的问题的向量矩阵，根据所述已有的帖子及所述向量矩阵合成答案表达式。

本实施例中，在爬取到用户当前发布的帖子的问题信息后，将所述问题信息进行转词处理，将完整的问题信息切分成一个一个的词语组合，然后提取问题信息中的关键词信息，并将所述关键词信息转为关键词向量，在根据所述关键词信息组合成帖子提及的问题时，将所述关键词向量进行
20 组合，组合成帖子提及的问题的向量矩阵，然后再根据已有的帖子，已有的帖子中也包含有关键词信息转化得到的关键词向量，以及已有的帖子所得到的问题的向量矩阵，计算当前帖子提及的问题的向量矩阵与已有的帖子提及的问题的向量矩阵，合成相应的答案表达式，后续根据该答案表达式对已有的帖子中的关键词向量进行筛选，便可得到已有的帖子中满足要
25 求的上下文信息，即当前帖子提及的问题所需的答案。

如图 3 所示，在另一种实施例中，本申请提供了一种问题答案推荐装置，包括：

转词模块 10：用于获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息；

匹配模块 20：用于根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子；

定位模块 30：用于将所述关键词信息组合成帖子提及的问题，从所述已有的帖子中定位与所述问题相关的上下文信息；

5 推送模块 40：用于根据所述上下文信息从若干已有的帖子获取问题的答案，将所述答案推送至所述用户发布的帖子。

本申请的一种实施例，所述匹配模块 20 还包括执行；

获取匹配时间段，根据所述关键词信息在数据库中进行相似度匹配，并基于所述匹配时间段获取与所述关键词信息匹配的若干已有的帖子。

10 本申请的一种实施例，所述匹配模块 20 还包括执行；

根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子，从所述已有的帖子中筛选回复数满足要求的帖子。

本申请的一种实施例，所述定位模块 30 还包括执行；

15 将所述关键词信息组合成帖子提及的问题，根据所述已有的帖子及所述问题合成答案表达式；从所述已有的帖子中定位与所述问题相关的上下文信息；将所述上下文信息与所述答案表达式相比较，确定所述上下文信息的效用，获取所述效用满足预设要求的上下文信息。

本申请的一种实施例，所述转词模块 10 还包括执行：

20 将所述问题信息进行转词处理，提取问题信息中的关键词信息，将所述关键词信息转为关键词向量，将所述关键词向量组成帖子问题的向量矩阵；

所述匹配模块 20 还包括执行：

根据所述向量矩阵在数据库中进行相似度匹配。

25 在另一种实施例中，本申请提供了一种计算机可读存储介质，计算机可读存储介质上存储有计算机程序，该计算机程序被处理器执行时实现上述实施例所述的问题答案推荐方法。其中，所述计算机可读存储介质包括但不限于任何类型的盘（包括软盘、硬盘、光盘、CD-ROM、和磁光盘）、ROM (Read-Only Memory, 只读存储器)、RAM (Random Access Memory,

随即存储器)、EPROM (EraSable Programmable Read-Only Memory, 可擦写可编程只读存储器)、EEPROM (Electrically EraSable Programmable Read-Only Memory, 电可擦可编程只读存储器)、闪存、磁性卡片或光线卡片。也就是, 存储设备包括由设备(例如, 计算机、手机)以能够读的形式存储或传输信息的任何介质, 可以是只读存储器, 磁盘或光盘等。

本申请实施例提供的一种计算机可读存储介质, 可实现获取用户当前发布的帖子的问题信息, 将所述问题信息进行转词处理, 提取问题信息中的关键词信息; 根据所述关键词信息在数据库中进行相似度匹配, 获取与所述关键词信息匹配的若干已有的帖子; 将所述关键词信息组合成帖子提及的问题, 从所述已有的帖子中定位与所述问题相关的上下文信息; 根据所述上下文信息从若干已有的帖子获取问题的答案, 将所述答案推送至所述用户发布的帖子。通过提供一种基于帖子、问题、答案三种元素组合的答案推荐方法, 适用于论坛发布帖子寻求解答的应用场景, 用户在论坛上发布帖子时, 获取用户当前发布的帖子的问题信息, 并将所述问题信息进行转词处理, , 确定问题信息中的关键词, 提取关键词信息从而形式对所述问题信息的描述, 每一个发布的帖子均有对应的关键词信息存储在数据库中, 根据所述关键词信息在数据库中进行相似度匹配, 获取与所述关键词信息匹配的若干已有的帖子, 由于所述关键词信息能够组合成帖子提及的问题, 从匹配到的已有的帖子中定位与所述问题相关的上下文信息, 根据所述上下文信息便可清晰地了解已有帖子的内容是否与用户当前发布的帖子的问题相关, 在从若干个已有的帖子中获取所述上下文信息后, 便能得到不同帖子中与用户当前发布的帖子的问题的相关答案, 然后将所述答案推送至所述用户发布的帖子中, 用户从推送的答案中便能直观地了解到已有的帖子发布的相关内容, 查看已有的帖子的内容信息是否与自己的所需解答的问题相关, 从而快速地查找到所需的答案, 同时, 避免论坛中相同问题或类似问题出现多次, 造成论坛重复的帖子过多, 从而节省了服务器有限的资源。

本申请实施例提供的计算机可读存储介质可以实现上述问题答案推荐方法的实施例, 具体功能实现请参见方法实施例中的说明, 在此不再赘述。

此外，在又一种实施例中，本申请还提供一种服务器，如图 4 所示，所述服务器包括处理器 403、存储器 405、输入单元 407 以及显示单元 409 等器件。本领域技术人员可以理解，图 4 示出的结构器件并不构成对所有服务器的限定，可以包括比图示更多或更少的部件，或者组合某些部件。

5 所述存储器 405 可用于存储计算机程序 401 以及各功能模块，所述处理器 403 运行存储在存储器 405 的计算机程序 401，从而执行设备的各种功能应用以及数据处理。所述存储器 405 可以是内存储器或外存储器，或者包括内存储器 and 外存储器两者。内存储器可以包括只读存储器(ROM)、可编程 ROM(PROM)、电可编程 ROM(EPROM)、电可擦写可编程 ROM(EEPROM)、快闪存储器、或者随机存储器。外存储器可以包括硬盘、软盘、ZIP 盘、U 盘、磁带等。本申请所公开的存储器包括但不限于这些类型的存储器。本申请所公开的存储器 405 只作为例子而非作为限定。

输入单元 407 用于接收信号的输入及接收用户的输入，输入单元 407 可包括触控面板以及其它输入设备，触控面板可收集用户在其上或附近的触摸操作（比如用户使用手指、触笔等任何适合的物体或附件在触控面板上或在触控面板附近的操作），并根据预先设定的程序驱动相应的连接装置；其它输入设备可以包括但不限于物理键盘、功能键（比如播放控制按键、开关按键等）、轨迹球、鼠标、操作杆等中的一种或多种。显示单元 409 可用于显示用户输入的信息或提供给用户的信息以及计算机设备的各种菜单。显示单元 409 可采用液晶显示器、有机发光二极管等形式。处理器 403 是计算机设备的控制中心，利用各种接口和线路连接整个电脑的各个部分，通过运行或执行存储在存储器 403 内的软件程序和/或模块，以及调用存储在存储器内的数据，执行各种功能和处理数据。

25 在一种实施方式中，所述服务器包括一个或多个处理器 403，以及一个或多个存储器 405，一个或多个计算机程序 401，其中所述一个或多个计算机程序 401 被存储在存储器 405 中并被配置为由所述一个或多个处理器 403 执行，所述一个或多个计算机程序 401 配置用于执行以上实施例所述的问题答案推荐方法。图 4 中所示的一个或多个处理器 403 能够执行、实现图 3 中所示的转词模块 10、匹配模块 20、定位模块 30、推送模块 40 的

功能。

本申请实施例提供的一种服务器，可实现获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息；根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子；将所述关键词信息组合成帖子提及的问题，从所述已有的帖子中定位与所述问题相关的上下文信息；根据所述上下文信息从若干已有的帖子获取问题的答案，将所述答案推送至所述用户发布的帖子。通过提供一种基于帖子、问题、答案三种元素组合的答案推荐方法，适用于论坛发布帖子寻求解答的应用场景，用户在论坛上发布帖子时，获取用户当前发布的帖子的问题信息，并将所述问题信息进行转词处理，，确定问题信息中的关键词，提取关键词信息从而形式对所述问题信息的描述，每一个发布的帖子均有对应的关键词信息存储在数据库中，根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子，由于所述关键词信息能够组合成帖子提及的问题，从匹配到的已有的帖子中定位与所述问题相关的上下文信息，根据所述上下文信息便可清晰地了解已有帖子的内容是否与用户当前发布的帖子的问题相关，在从若干个已有的帖子中获取所述上下文信息后，便能得到不同帖子中与用户当前发布的帖子的问题的相关答案，然后将所述答案推送至所述用户发布的帖子中，用户从推送的答案中便能直观地了解到已有的帖子发布的相关内容，查看已有的帖子的内容信息是否与自己的所需解答的问题相关，从而快速地查找到所需的答案，同时，避免论坛中相同问题或类似问题出现多次，造成论坛重复的帖子过多，从而节省了服务器有限的资源。

本申请实施例提供的服务器可以实现上述提供的问题答案推荐方法的实施例，具体功能实现请参见方法实施例中的说明，在此不再赘述。

以上所述仅是本申请的部分实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本申请原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本申请的保护范围。

权利要求

1.一种问题答案推荐方法，其特征在于，包括：

获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息；

5 根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子；

将所述关键词信息组合成帖子提及的问题，从所述已有的帖子中定位与所述问题相关的上下文信息；

10 根据所述上下文信息从若干已有的帖子获取问题的答案，将所述答案推送至所述用户发布的帖子。

2.根据权利要求1所述的方法，其特征在于，所述根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子，包括：

15 获取匹配时间段，根据所述关键词信息在数据库中进行相似度匹配，并基于所述匹配时间段获取与所述关键词信息匹配的若干已有的帖子；

所述基于所述匹配时间段获取与所述关键词信息匹配的若干已有的帖子之后，还包括：

接收用户对所述帖子关联的答案的反馈信息，根据所述反馈信息调整所述匹配时间段。

20 3.根据权利要求1所述的方法，其特征在于，所述根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子，包括：

根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子，从所述已有的帖子中筛选回复数满足要求的帖子；

25 所述从所述已有的帖子中筛选回复数满足要求的帖子之后，还包括：

接收用户对所述帖子关联的答案的反馈信息，根据所述反馈信息调整回复数要求。

4.根据权利要求1所述的方法，其特征在于，所述将所述关键词信息组合成帖子提及的问题，从所述已有的帖子中定位与所述问题相关的上下文信息，

包括：

将所述关键词信息组合成帖子提及的问题，根据所述已有的帖子及所述问题合成答案表达式；从所述已有的帖子中定位与所述问题相关的上下文信息；将所述上下文信息与所述答案表达式相比较，确定所述上下文信息的效用，获取所述效用满足预设要求的上下文信息。

5.根据权利要求 1 所述的方法，其特征在于，所述将所述问题信息进行转词处理，提取问题信息中的关键词信息，包括：

将所述问题信息进行转词处理，提取问题信息中的关键词信息，将所述关键词信息转为关键词向量，将所述关键词向量组成帖子问题的向量矩阵；

10 所述根据所述关键词信息在数据库中进行相似度匹配，包括：

根据所述向量矩阵在数据库中进行相似度匹配。

6.根据权利要求 4 所述的方法，其特征在于，所述获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息，包括：

15 获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息，将所述关键词信息转为关键词向量；

所述将所述关键词信息组合成帖子提及的问题，根据所述已有的帖子及所述问题合成答案表达式，包括：

根据所述关键词向量组成帖子提及的问题的向量矩阵，根据所述已有的帖子及所述向量矩阵合成答案表达式。

20 7、根据权利要求 1-6 任一项所述的方法，其特征在于，所述提取帖子的问题信息中的关键词信息时，所提取的关键词信息为文本 string 型的数据。

8.一种问题答案推荐装置，其特征在于，包括：

25 转词模块：用于获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息；

匹配模块：用于根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子；

定位模块：用于将所述关键词信息组合成帖子提及的问题，从所述已有的帖子中定位与所述问题相关的上下文信息；

推送模块：用于根据所述上下文信息从若干已有的帖子获取问题的答案，将所述答案推送至所述用户发布的帖子。

9.根据权利要求 8 所述的装置，其特征在于，所述定位模块还包括执行；
将所述关键词信息组合成帖子提及的问题，根据所述已有的帖子及所述问题合成答案表达式；从所述已有的帖子中定位与所述问题相关的上下文信息；
5 将所述上下文信息与所述答案表达式相比较，确定所述上下文信息的效用，获取所述效用满足预设要求的上下文信息。

10.根据权利要求 8 所述的装置，其特征在于，所述匹配模块具体用于：
获取匹配时间段，根据所述关键词信息在数据库中进行相似度匹配，并基于所述匹配时间段获取与所述关键词信息匹配的若干已有的帖子；
10 所述匹配模块还用于：
接收用户对所述帖子关联的答案的反馈信息，根据所述反馈信息调整所述匹配时间段。

11.根据权利要求 8 所述的装置，其特征在于，所述匹配模块具体用于：
15 根据所述关键词信息在数据库中进行相似度匹配，获取与所述关键词信息匹配的若干已有的帖子，从所述已有的帖子中筛选回复数满足要求的帖子；
所述匹配模块还用于：
接收用户对所述帖子关联的答案的反馈信息，根据所述反馈信息调整回复数要求。

20 12.根据权利要求 8 所述的装置，其特征在于，所述转词模块在将所述问题信息进行转词处理，提取问题信息中的关键词信息时，具体用于：
将所述问题信息进行转词处理，提取问题信息中的关键词信息，将所述关键词信息转为关键词向量，将所述关键词向量组成帖子问题的向量矩阵；
所述匹配模块在根据所述关键词信息在数据库中进行相似度匹配时，具体
25 用于：
根据所述向量矩阵在数据库中进行相似度匹配。

13.根据权利要求 9 所述的装置，其特征在于，所述转词模块具体用于：
获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息，将所述关键词信息转为关键词向量；

所述定位模块在将所述关键词信息组合成帖子提及的问题,根据所述已有的帖子及所述问题合成答案表达式时,具体用于:

根据所述关键词向量组成帖子提及的问题的向量矩阵,根据所述已有的帖子及所述向量矩阵合成答案表达式。

5 14、根据权利要求 8-13 任一项所述的装置,其特征在于,所述提取帖子的问题信息中的关键词信息时,所提取的关键词信息为文本 string 型的数据。

15.一种计算机非易失性可读存储介质,其特征在于,所述计算机非易失性可读存储介质上存储有计算机程序,该计算机程序被处理器执行时实现权利要求 1 至 7 任一项所述的问题答案推荐方法。

10 16.一种服务器,其特征在于,包括:

一个或多个处理器;

存储器;

一个或多个计算机程序,其中所述一个或多个计算机程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行,所述一个或多个计算机程序

15 配置用于执行以下步骤:

获取用户当前发布的帖子的问题信息,将所述问题信息进行转词处理,提取问题信息中的关键词信息;

根据所述关键词信息在数据库中进行相似度匹配,获取与所述关键词信息匹配的若干已有的帖子;

20 将所述关键词信息组合成帖子提及的问题,从所述已有的帖子中定位与所述问题相关的上下文信息;

根据所述上下文信息从若干已有的帖子获取问题的答案,将所述答案推送至所述用户发布的帖子。

25 17.根据权利要求 16 所述的服务器,其特征在于,所述根据所述关键词信息在数据库中进行相似度匹配,获取与所述关键词信息匹配的若干已有的帖子时,所述一个或多个计算机程序配置用于执行以下步骤:

获取匹配时间段,根据所述关键词信息在数据库中进行相似度匹配,并基于所述匹配时间段获取与所述关键词信息匹配的若干已有的帖子;

所述基于所述匹配时间段获取与所述关键词信息匹配的若干已有的帖子

之后，所述一个或多个计算机程序还被配置用于执行以下步骤：

接收用户对所述帖子关联的答案的反馈信息，根据所述反馈信息调整所述匹配时间段。

5 18.根据权利要求 16 所述的服务器，其特征在于，所述将所述关键词信息组合成帖子提及的问题，从所述已有的帖子中定位与所述问题相关的上下文信息时，所述一个或多个计算机程序配置用于执行以下步骤：

10 将所述关键词信息组合成帖子提及的问题，根据所述已有的帖子及所述问题合成答案表达式；从所述已有的帖子中定位与所述问题相关的上下文信息；将所述上下文信息与所述答案表达式相比较，确定所述上下文信息的效用，获取所述效用满足预设要求的上下文信息。

19.根据权利要求 16 所述的服务器，其特征在于，所述将所述问题信息进行转词处理，提取问题信息中的关键词信息时，所述一个或多个计算机程序配置用于执行以下步骤：

15 将所述问题信息进行转词处理，提取问题信息中的关键词信息，将所述关键词信息转为关键词向量，将所述关键词向量组成帖子问题的向量矩阵；

所述根据所述关键词信息在数据库中进行相似度匹配时，所述一个或多个计算机程序配置用于执行以下步骤：

根据所述向量矩阵在数据库中进行相似度匹配。

20 20.根据权利要求 18 所述的服务器，其特征在于，所述获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息时，所述一个或多个计算机程序配置用于执行以下步骤：

获取用户当前发布的帖子的问题信息，将所述问题信息进行转词处理，提取问题信息中的关键词信息，将所述关键词信息转为关键词向量；

25 所述将所述关键词信息组合成帖子提及的问题，根据所述已有的帖子及所述问题合成答案表达式时，所述一个或多个计算机程序配置用于执行以下步骤：

根据所述关键词向量组成帖子提及的问题的向量矩阵，根据所述已有的帖子及所述向量矩阵合成答案表达式。

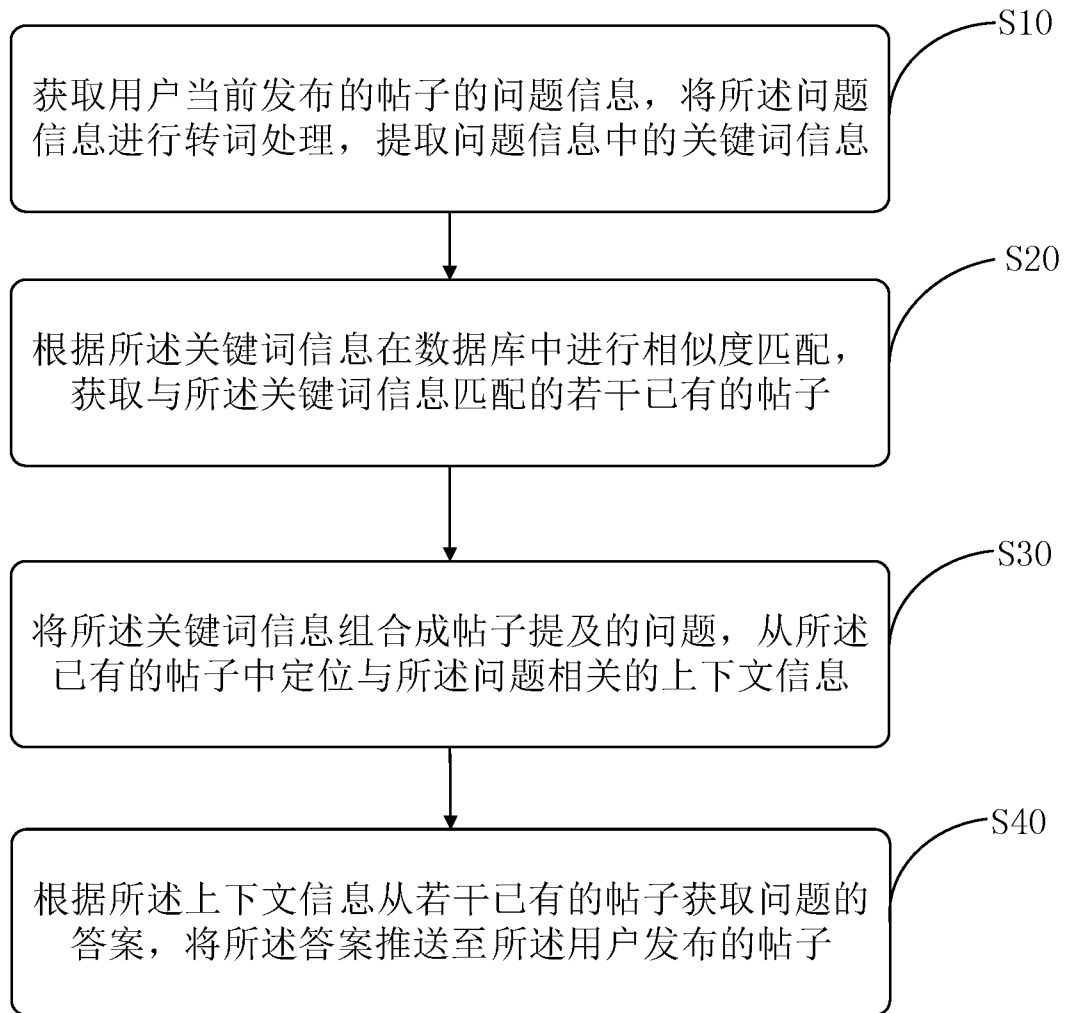


图 1

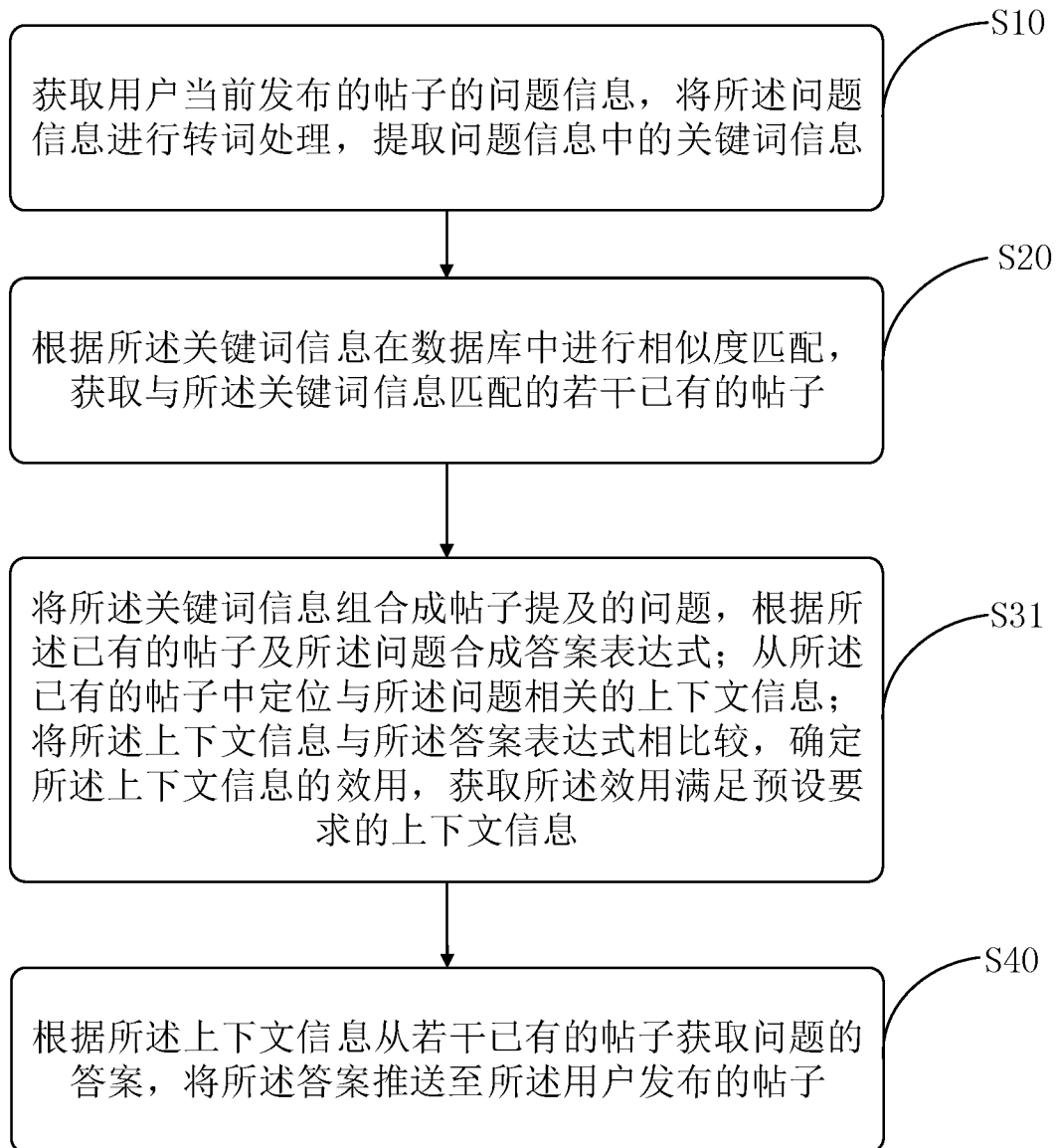


图 2

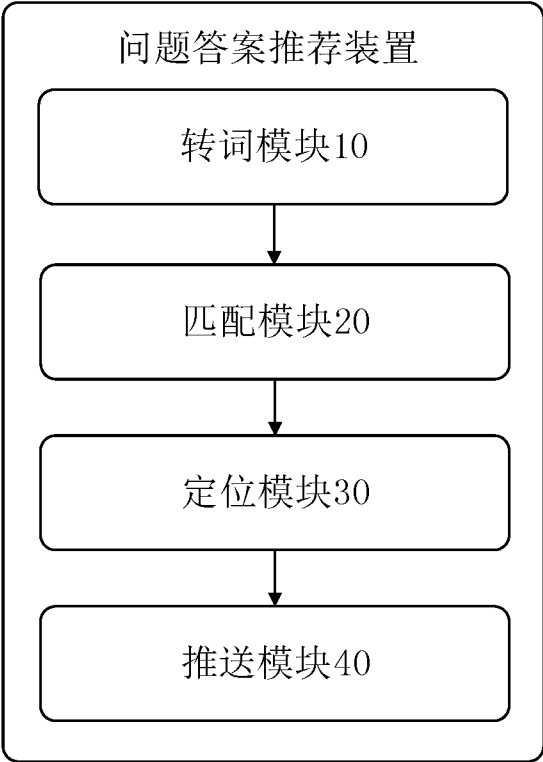


图 3

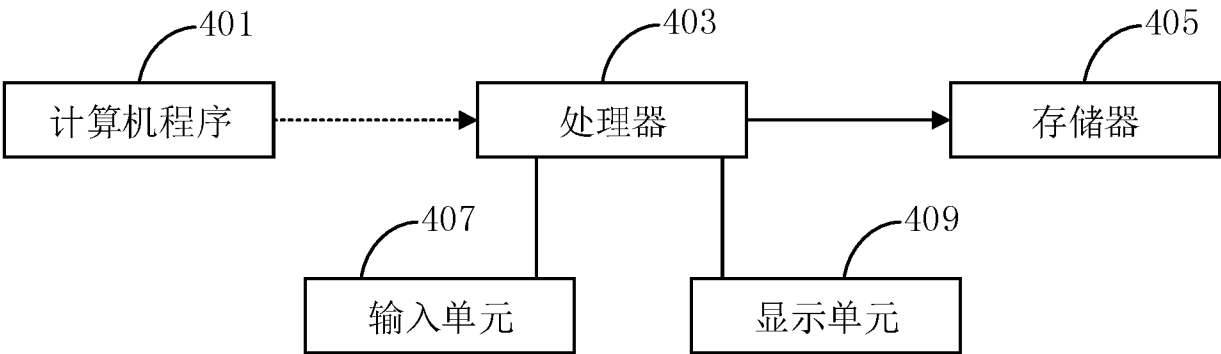


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/117166

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/332(2019.01)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)

CNPAT, WPI, EPODOC, CNKI, IEEE: 智能应答, 问答, 提问, 问题, 答案, 相似度, 匹配, 关键词, 关键字, 帖子, 文章, 论坛, 网站, 上下文, intelligent question answering system, QA, question, answer, similarity, keywords, post, article, forum, website, context

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109871438 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 11 June 2019 (2019-06-11) claims 1-10, and description, paragraph [0078]	1-20
X	CN 105893391 A (ALIBABA GROUP HOLDING LIMITED) 24 August 2016 (2016-08-24) description, paragraphs [0153]-[0170]	1-20
A	CN 108268450 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 10 July 2018 (2018-07-10) entire document	1-20
A	CN 105005564 A (BEIJING SOGOU TECHNOLOGY DEVELOPMENT CO., LTD.) 28 October 2015 (2015-10-28) entire document	1-20
A	US 2002169872 A1 (NOMIYAMA, H.) 14 November 2002 (2002-11-14) entire document	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“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)

“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

“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

“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

“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

“&” document member of the same patent family

Date of the actual completion of the international search

10 January 2020

Date of mailing of the international search report

31 January 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/117166

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109871438	A	11 June 2019	None			
CN	105893391	A	24 August 2016	None			
CN	108268450	A	10 July 2018	None			
CN	105005564	A	28 October 2015	None			
US	2002169872	A1	14 November 2002	JP	2001325272	A	22 November 2001

国际检索报告

国际申请号

PCT/CN2019/117166

A. 主题的分类

G06F 16/332(2019.01) 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))

CNPAT, WPI, EPDOC, CNKI, IEEE: 智能应答, 问答, 提问, 问题, 答案, 相似度, 匹配, 关键词, 关键字, 帖子, 文章, 论坛, 网站, 上下文, intelligent question answering system, QA, question, answer, similarity, keywords, post, article, forum, website, context

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
PX	CN 109871438 A (平安科技深圳有限公司) 2019年 6月 11日 (2019 - 06 - 11) 权利要求1-10, 说明书第[0078]段	1-20
X	CN 105893391 A (阿里巴巴集团控股有限公司) 2016年 8月 24日 (2016 - 08 - 24) 说明书第[0153]-[0170]段	1-20
A	CN 108268450 A (百度在线网络技术北京有限公司) 2018年 7月 10日 (2018 - 07 - 10) 全文	1-20
A	CN 105005564 A (北京搜狗科技发展有限公司) 2015年 10月 28日 (2015 - 10 - 28) 全文	1-20
A	US 2002169872 A1 (NOMIYAMA, Hiroshi) 2002年 11月 14日 (2002 - 11 - 14) 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

2020年 1月 10日

国际检索报告邮寄日期

2020年 1月 31日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

杨栋

电话号码 86-(10)-53961344

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/117166

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	109871438	A	2019年 6月 11日	无	
CN	105893391	A	2016年 8月 24日	无	
CN	108268450	A	2018年 7月 10日	无	
CN	105005564	A	2015年 10月 28日	无	
US	2002169872	A1	2002年 11月 14日	JP 2001325272 A	2001年 11月 22日