

(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
2021 年 2 月 4 日 (04.02.2021)



(10) 国际公布号
WO 2021/017288 A1

(51) 国际专利分类号:
G06F 11/07 (2006.01)

(21) 国际申请号: PCT/CN2019/117684

(22) 国际申请日: 2019 年 11 月 12 日 (12.11.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910708623.7 2019 年 8 月 1 日 (01.08.2019) 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: 熊星(XIONG, Xing); 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)。

(74) 代理人: 深圳市隆天联鼎知识产权代理有限公司(SHENZHEN LUNGTIN LIANDING INTELLECTUAL PROPERTY AGENT LTD.); 中国广东省深圳市福田区南园路上田大厦 4A 刘抗美, Guangdong 518000 (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,

(54) Title: METHOD AND APPARATUS FOR REPEATEDLY IDENTIFYING SYSTEM ERROR, ELECTRONIC DEVICE, AND COMPUTER-READABLE STORAGE MEDIUM

(54) 发明名称: 对系统错误进行重复识别的方法、装置、电子设备和计算机可读存储介质

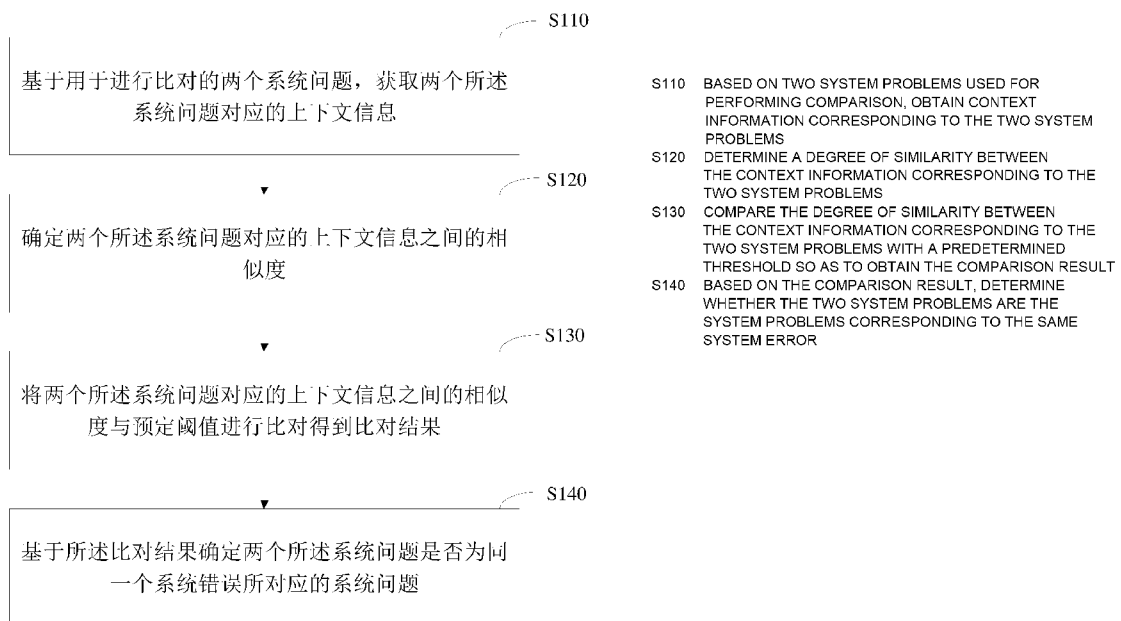


图 1

(57) Abstract: Disclosed in the present application are a method and apparatus for repeatedly identifying a system error, an electronic device, and a computer-readable storage medium, relating to the technical field of data processing. The method comprises: based on two system problems used for performing comparison, obtaining context information corresponding to the two system problems, wherein the context information comprises the texts of the causes of the system problems, the screenshots of the causes of the system problems,

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,
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))。

functional module information causing the system problems, and device information causing the system problems; determining a degree of similarity between the context information corresponding to the two system problems; and comparing the degree of similarity between the context information corresponding to the two system problems with a predetermined threshold so as to obtain the comparison result. The method provided by the present application does not depend on the subjective experience of a developer to identify whether the two system problems are the system problems caused by one system error, thereby improving the accuracy of system problem identification.

(57) 摘要: 本申请揭示了一种对系统错误进行重复识别的方法、装置、电子设备及计算机可读存储介质, 属于数据处理的技术领域, 其中所述方法包括: 基于用于进行比对的两个系统问题, 获取两个所述系统问题对应的上下文信息, 其中, 所述上下文信息包括系统问题所产生原因的文本、系统问题所产生原因的截图、产生系统问题的功能模块信息以及产生系统问题的设备信息; 确定两个所述系统问题对应的上下文信息之间的相似度; 将两个所述系统问题对应的上下文信息之间的相似度与预定阈值进行比对得到比对结果。本申请提出的方法不再依赖开发人员的主观经验去识别两个系统问题是否为一个系统错误所产生系统, 由此提高了相同系统问题识别的准确性。

对系统错误进行重复识别的方法、装置、电子设备和计算机可读存储介质

本申请要求引用 2019 年 08 月 01 日递交、发明名称为“对系统错误进行重复识别的方法和装置”的中国专利申请 201910708623.7 的优先权，在此通过引用将其全部内容合并于此。

5

技术领域

本申请涉及数据处理的技术领域，特别是涉及对系统错误进行重复识别的方法、装置、电子设备和计算机可读存储介质。

10

背景技术

在系统研发测试过程中，测试人员会根据所发现的 BUG（系统错误）提交说明包含该系统错误的相关信息的系统问题至管理系统中，系统的维护人员根据测试人员所提交的系统问题去解决该系统错误。由于不同的测试人员可能会就同一个系统错误提交相似的系统问题至管理系统中，而这些相似的系统问题，则需要进行重复识别，从而避免对同一个系统错误进行重复处理。

15

本申请的发明人意识到，通过维护人员去对所提交的相似的系统问题进行重复识别，并基于维护人员主观经验将其归类为同一个系统错误所对应的系统问题，该过程全靠维护人员的主观经验，其准确度不高。

20

技术问题

基于此，为解决现有技术中避免依靠维护人员的主观经验对所提交的相似的系统问题进行重复识别，且识别的准确度不高的技术问题，本申请提供了一种对系统错误进行重复识别的方法、装置、电子设备及计算机可读存储介质。

25

发明解决方案

第一方面，提供了一种对系统错误进行重复识别的方法，包括：

基于用于进行比对的两个系统问题，获取两个所述系统问题对应的上下文信息，其中，所述上下文信息包括系统问题所产生原因的文本、系统问题所产生原因的截图、产生系统问题的功能模块信息以及产生系统问题的设备信息；

30

确定两个所述系统问题对应的上下文信息之间的相似度；

将两个所述系统问题对应的上下文信息之间的相似度与预定阈值进行比对得到比对

结果；

基于所述比对结果确定两个所述系统问题是否为同一个系统错误所对应的系统问题。

第二方面，提供了一种对系统错误进行重复识别的装置，包括：

获取单元，用于基于用于进行比对的两个系统问题，获取两个所述系统问题对应的上下
5 下文信息，其中，所述上下文信息包括系统问题所产生原因的文本、系统问题所产生原因的
截图、产生系统问题的功能模块信息以及产生系统问题的设备信息；

第一执行单元，用于确定两个所述系统问题对应的上下文信息之间的相似度；

比对单元，用于将两个所述系统问题对应的上下文信息之间的相似度与预定阈值进行
比对得到比对结果；

10 第二执行单元，用于基于所述比对结果确定两个所述系统问题是否为同一个系统错误
所对应的系统问题。

第三方面，提供了一种电子设备，包括处理器；以及存储器，用于存储所述处理器的
对系统错误进行重复识别的程序；其中，所述处理器配置为执行如本申请第一方面或第一
方面任意可能实现方式中的方法。

15 第四方面，提供了一种计算机可读存储介质，其上存储有对系统错误进行重复识别的
程序，所述对系统错误进行重复识别的程序被处理器执行时实现如本申请第一方面或第一
方面任意可能实现方式中的方法。

第五方面，提供了一种计算机程序产品，所述计算机程序产品包括存储了计算机程序
的非瞬时性计算机可读存储介质，所述计算机程序可操作来使计算机执行如本申请第一方
20 面所述的方法或第一方面任意可能实现方式中的方法。

有益效果

本申请的实施例提供的技术方案可以包括以下有益效果：

25 可以实现自动判定两个系统问题是否为同一个系统错误对应的系统问题，不再依赖开
发人员的主观经验去识别两个系统问题是否为一个系统错误所产生系统，提高了相同系统
问题识别的准确性。

应当理解的是，以上的一般描述和后文的细节描述仅是示例性的，并不能限制本申请。

附图说明

30 图 1 为本申请一个实施例示出的一种对系统错误进行重复识别的方法的实现流程图。

图 2 为本申请一个实施例示出的一种对系统错误进行重复识别的方法中的步骤 S120

的具体实现流程图。

图 3 为本申请一个实施例示出的一种对系统错误进行重复识别的装置的框图。

图 4 示意性示出一种用于实现上述对系统错误进行重复识别的方法的电子设备示例框图。

5 图 5 示意性示出一种用于实现上述对系统错误进行重复识别的方法的计算机可读存储介质。

具体实施方式

为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本
10 申请进行进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

本申请实施例中的对系统错误进行重复识别的方法的执行主体为电子设备，该电子设备具体可以为对系统问题进行管理的管理系统，以下以管理系统为例来进行说明。

参考图 1 所示，根据本申请的一个实施例的对系统错误进行重复识别的方法，包括
15 以下步骤 S110 至步骤 S140，详细说明如下：

步骤 S110，基于用于进行比对的两个系统问题，获取两个所述系统问题对应的上下文信息，其中，所述上下文信息包括所述系统问题所产生原因的文本信息、所述系统问题所产生原因的截图、产生所述系统问题的功能模块信息以及产生所述系统问题的设备信息。

步骤 S120，确定两个所述系统问题对应的上下文信息之间的相似度。

20 步骤 S130，将两个所述系统问题对应的上下文信息之间的相似度与预定阈值进行比对得到比对结果。

步骤 S140，基于所述比对结果确定两个所述系统问题是否为同一个系统错误所对应的系统问题。

在步骤 S110 中，上述系统问题指的是某个业务系统在研发测试终端上进行研发测试
25 的过程中，测试人员基于在业务系统中所发现的 BUG（系统错误）所提交的说明该系统错误相关的上下文信息的系统问题。其中，该系统问题对应的上下文信息包括系统问题所产生原因的文本信息、系统问题所产生原因的截图、产生系统问题的功能模块信息以及产生系统问题的设备信息。

需要说明的是，管理系统获取系统问题所产生原因的文本时，可以从系统问题产生时
30 的系统日志中获取系统问题所产生原因的文本。

管理系统获取系统问题所产生原因的截图时，可以通过在研发测试终端出现系统问题的情况下，对研发测试终端显示的原因提示进行截图来实现。

管理系统获取产生系统问题的功能模块信息时，可以从系统问题产生时的系统日志中获取。

管理系统获取产生系统问题的设备信息时,可以通过从产生系统问题的研发测试终端的内存中的出厂配置模块中调取。

在步骤 S120 中,管理系统基于两个系统问题对应的上下文信息中所包含的系统问题所产生原因的文本、系统问题所产生原因的截图、产生系统问题的功能模块信息以及产生系统问题的设备信息这四个维度的信息确定两个系统问题对应的上下文信息之间的相似度。需要说明的是,两个系统问题对应的上下文信息的相似度越高,则说明两个系统问题为同一个系统错误所对应的系统问题的可能性越大。

参考图 2,图 2 为本申请一个实施例示出的一种对系统错误进行重复识别的方法中的步骤 S120 的具体实现流程图,其中,所述确定两个所述系统问题对应的上下文信息之间的相似度的步骤 S120,包括:

步骤 S1201,比对两个所述系统问题对应的上下文信息中系统问题所产生原因的文本,确定第一相似分数。

步骤 S1202,比对两个所述系统问题对应的上下文信息中系统问题所产生原因的截图,确定第二相似分数。

步骤 S1203,比对两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息,确定第三相似分数。

步骤 S1204,比对两个所述系统问题对应的上下文信息中产生系统问题的设备信息,确定第四相似分数。

步骤 S1205,基于所述第一相似分数、所述第二相似分数、所述第三相似分数以及所述第四相似分数,确定两个所述系统问题对应的上下文信息之间的相似度。

在步骤 S1201 中,管理系统将两个系统问题对应的上下文信息中系统问题所产生原因的文本进行比对,来确定上下文信息中系统问题所产生原因的文本之间的第一相似分数。

在一个实施例中,第一相似分数基于以下方法确定:

将两个所述系统问题对应的上下文信息中系统问题所产生原因的文本中的标题和内容分别进行分词;

针对每个所述系统问题对应的上下文信息中系统问题所产生原因的文本中的标题和内容,确定分出的每个词在分出的词中的出现次数,将出现次数高于预定出现次数阈值的词作为所述系统问题对应的上下文信息中系统问题所产生原因中的关键词;

基于两个所述系统问题对应的关键词的交集的关键词数除以两个所述系统问题对应的关键词的并集的关键词数,得到第一相似分数。

在本实施例中,在确定第一相似分数时,管理系统获取得到两个系统问题对应的上下文信息中系统问题所产生原因的文本,对两个文本中的标题和内容进行分词得到多个词,其中分词的方法可以采用结巴分词。

管理系统确定分出的每个词在所有分出的词中的出现次数,将出现次数高于预定出现次数阈值的词作为系统问题对应的上下文信息中系统问题所产生原因中的关键词,进而获

取得到反映系统问题所产生原因的关键信息。

管理系统确定两个系统问题对应的关键词的交集集中的关键词数,以及确定两个系统问题对应的关键词的并集中的关键词数,将两个系统问题对应的关键词的交集集中的关键词数除以两个系统问题对应的关键词的并集中的关键词数得到第一相似分数,进而根据第一相似分数确定两个系统问题对应的上下文信息中系统问题所产生原因的文本是否一致。需要说明的是,当两个系统问题为同一个系统问题时,系统问题所产生原因的文本记录的信息将基本一致,由此对应的关键词相同的概率较大,则第一相似分数越高;反之,当两个系统问题不是同一个系统问题时,系统问题所产生原因的文本记录的信息将不一致,由此对应的关键词相同的概率较小,则第一相似分数越低。

在步骤 S1202 中,管理系统将两个系统问题对应的上下文信息中系统问题所产生原因的截图进行比对,来确定上下文信息中系统问题所产生原因的文本之间的第二相似分数。

在一个实施例中,所述第二相似分数基于以下方法确定:

对两个所述系统问题对应的上下文信息中系统问题所产生原因的截图进行光学字符识别,获得从所述截图中识别出的字符;

基于对两个所述系统问题识别出的字符中重合的字符数除以对两个所述系统问题识别出的字符数之和,得到第二相似分数。

在本实施例中,在确定第二相似分数时,管理系统获取两个系统问题对应的上下文信息中系统问题所产生原因的截图。

对于获取的系统问题所产生原因的截图,分别通过光学字符识别,获得从截图中识别出的字符。

管理系统确定两个系统问题对应的上下文信息中系统问题所产生原因的截图中识别出的字符重合的字符数,以及两个系统问题对应的上下文信息中系统问题所产生原因的截图中识别出的字符数之和,将两个系统问题识别出的字符重合的字符数除以两个系统问题识别出的字符数之和得到第二相似分数,进而根据第二相似分数确定两个系统问题对应的上下文信息中系统问题所产生原因的截图是否一致。需要说明的是,当两个系统问题为同一个系统问题时,系统问题所产生原因的截图将基本一致,由此对应的字符相同的概率较大,则第二相似分数越高;反之,当两个系统问题不是同一个系统问题时,系统问题所产生原因的截图将不一致,由此对应的字符相同的概率较小,则第二相似分数越低。

在步骤 S1203 中,比对两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息,确定第三相似分数。

在步骤 S1204,比对两个所述系统问题对应的上下文信息中产生系统问题的设备信息,确定第四相似分数。

在一个实施例中,所述第三相似分数基于以下方法确定:

将两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息进行比对;

若两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息一致,则获取

预设的第一分数作为第三相似分数；

若两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息不一致，则确定所述第三相似分数为零；

所述第四相似分数基于以下方法确定：

5 将两个所述系统问题对应的上下文信息中产生系统问题的设备信息进行比对；

若两个所述系统问题对应的上下文信息中产生系统问题的设备信息一致，则获取预设的第二分数作为第四相似分数；

若两个所述系统问题对应的上下文信息中产生系统问题的设备信息不一致，则确定所述第四相似分数为零。

10 在本实施例中，在确定第三相似分数时，管理系统在将两个系统问题对应的上下文信息中产生系统问题的功能模块信息进行比对，若两个系统问题对应的上下文信息中产生系统问题的功能模块信息一致，则获取预设的第一分数作为第三相似分数。在本实施例中，当产生系统问题的功能模块信息一致时，则两个系统问题才可能是同一系统错误所产生，当产生系统问题的功能模块信息不一致时，则两个系统问题不可能是同一系统错误所产生；
15 由此，当产生系统问题的功能模块信息一致时，可以将第一分数设置为 1，则第三相似分数为 1；当两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息不一致，则可以将第三相似分数设置为零。

在确定第四相似分数时，管理系统将两个系统问题对应的上下文信息中产生系统问题的设备信息进行比对，若两个系统问题对应的上下文信息中产生系统问题的设备信息一致，
20 则获取预设的第二分数作为第四相似分数。在本实施例中，当产生系统问题的设备信息一致时，则两个系统问题才可能是同一系统错误所产生，当产生系统问题的设备信息不一致时，则两个系统问题不可能是同一系统错误所产生；由此，当产生系统问题的设备信息一致时，可以将第二分数设置为 1，则第四相似分数为 1；当两个所述系统问题对应的上下文信息中产生系统问题的设备信息不一致，则可以将第四相似分数设置为零。

25 在步骤 S1205 中，在管理系统比对两个系统问题对应的上下文信息之间的相似度时，具体将通过上下文信息中的系统问题所产生原因的文本、系统问题所产生原因的截图、产生系统问题的设备信息以及产生系统问题的设备信息四个维度的信息分别确定相应的第一相似分数、第二相似分数、第三相似分数以及第四相似分数，由所确定的第一相似分数、第二相似分数、第三相似分数以及第四相似分数确定两个系统问题对应的上下文信息之间的相似度。
30

在一个实施例中，所述基于所述第一相似分数、所述第二相似分数、所述第三相似分数以及所述第四相似分数，确定两个所述系统问题对应的上下文信息之间的相似度步骤 S1205，包括：

基于以下公式确定两个系统问题之间的相似度：

$$S = \frac{\ln(A + B + e)}{\alpha} + \beta (C + D)$$

，其中， α 和 β 为大于1的正常数，A为所述第一相似分数，B为所述第二相似分数，C为所述第三相似分数，D为所述第四相似分数，S为两个系统问题之间的相似度，e为自然常数。

在本实施例中，根据两个系统问题的上下文信息之间的相似度来确定两个系统问题是否为同一个系统错误对应的系统问题的过程中，当两者上下文信息中产生系统问题的功能模块信息或系统问题的设备信息不一致时，说明两个系统问题可以为同一个系统错误对应的系统问题的可能性较小，因此系统问题的功能模块信息或系统问题的设备信息这两个维度可以作为两个系统问题可以为同一个系统错误对应的系统问题的前提，因此第三相似分数、第四相似分数在上下文信息之间的相似度的计算中影响较大，所以对应的权重较大。需要说明的是，由于产生系统问题的功能模块信息或系统问题的设备信息一致时，由此也不能准确地判断两个系统问题一定为同一个系统错误对应的系统问题，因此基于所产生原因的文本以及所产生原因的截图确定两个系统问题可以为同一个系统错误对应的系统问题作为可选方案，由此第一相似分数、第二相似分数在上下文信息之间的相似度的计算中影响较小，所以对应的权重较小。

还请继续参考图1，在步骤S130中，管理系统将系统问题对应的上下文信息之间的相似度与预定阈值进行比对得到比对结果，其中比对结果为系统问题对应的上下文信息之间的相似度小于预定阈值、系统问题对应的上下文信息之间的相似度等于预定阈值和系统问题对应的上下文信息之间的相似度大于预定阈值的一种。

在步骤S140中，管理系统基于上述比对结果确定两个上述系统问题是否为同一个系统错误所对应的系统问题。

可选地，在本申请的一个实施例中，所述基于所述比对结果确定两个所述系统问题是否为同一个系统错误所对应的系统问题的步骤S140，包括如下步骤：

若所述比对结果为两个所述系统问题对应的上下文信息之间的相似度大于或者等于预定阈值，则确定两个所述系统问题为同一个系统错误所对应的系统问题。

若所述比对结果为两个所述系统问题对应的上下文信息之间的相似度小于预定阈值，则确定两个所述系统问题不是同一个系统错误所对应的系统问题。

管理系统在基于比对结果确定两个系统问题是否为同一个系统错误所对应的系统问题时，当两个系统问题对应的上下文信息之间的相似度大于或者等于预定阈值，则说明两个系统问题对应的上下文信息中所包括信息基本一致，那么为同一个系统错误对应的系统问题的可能性较大，进而确定两个系统问题为同一个系统错误所对应的系统问题；当两个系统问题对应的上下文信息之间的相似度小于预定阈值，则说明两个系统问题对应的上下文信息中所包括信息存在较大差异，那么为同一个系统错误对应的系统问题的可能性较小，则确定两个系统问题不是同一个系统错误所对应的系统问题。

以上可以看出，在两个系统问题是否为同一个系统错误所对应的系统问题时，根据两个系统问题对应的上下文信息中用于反映系统问题所产生原因的文本、系统问题所产生原

因的截图、产生系统问题的功能模块信息以及产生系统问题的设备信息等多个维度的信息确定两个系统问题对应的上下文之间的相似度,由此可以准确地确定两个系统问题是否为同一个系统错误所对应的系统问题,进而可以实现自动判定两个系统问题是否为同一个系统错误对应的系统问题,不再依赖开发人员的主观经验去识别两个系统问题是否为一个系统错误所产生系统,由此提高了相同系统问题识别的准确性。

参考图 3,图 3 是本申请一个实施例提供的一种对系统错误进行重复识别的装置,所述对系统错误进行重复识别的装置可以集成于上述的电子设备中,具体可以包括获取单元 110、第一执行单元 120、比对单元 130 以及第二执行单元 140。

获取单元 110,用于基于用于进行比对的两个系统问题,获取两个所述系统问题对应的上下文信息,其中,所述上下文信息包括系统问题所产生原因的文本、系统问题所产生原因的截图、产生系统问题的功能模块信息以及产生系统问题的设备信息;

第一执行单元 120,用于确定两个所述系统问题对应的上下文信息之间的相似度。

比对单元 130,用于将两个所述系统问题对应的上下文信息之间的相似度与预定阈值进行比对得到比对结果。

第二执行单元 140,用于基于所述比对结果确定两个所述系统问题是否为同一个系统错误所对应的系统问题。

可选地,所述第一执行单元 120 被配置为:

比对两个所述系统问题对应的上下文信息中系统问题所产生原因的文本,确定第一相似分数;

比对两个所述系统问题对应的上下文信息中系统问题所产生原因的截图,确定第二相似分数;

比对两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息,确定第三相似分数;

比对两个所述系统问题对应的上下文信息中产生系统问题的设备信息,确定第四相似分数;

基于所述第一相似分数、所述第二相似分数、所述第三相似分数以及所述第四相似分数,确定两个所述系统问题对应的上下文信息之间的相似度。

可选地,所述第一执行单元 120 被配置为:

基于以下公式确定两个系统问题之间的相似度:

$$S = \frac{\ln(A + B + e)}{\alpha} + \beta (C + D)$$

,其中, α 和 β 为大于 1 的正常数,A 为所述第一相似分数,B 为所述第二相似分数,C 为所述第三相似分数,D 为所述第四相似分数,S 为两个系统问题对应的上下文信息之间的相似度,e 为自然常数。

可选地,所述第一执行单元 120 被配置为:

将两个所述系统问题对应的上下文信息中系统问题所产生原因的文本中的标题和內容分别进行分词;

针对每个所述系统问题对应的上下文信息中系统问题所产生原因的文本中的标题和内容，确定分出的每个词在分出的词中的出现次数，将出现次数高于预定出现次数阈值的词作为所述系统问题对应的上下文信息中系统问题所产生原因中的关键词；

5 基于两个所述系统问题对应的关键词的交集集中的关键词数除以两个所述系统问题对应的关键词的并集中的关键词数，得到第一相似分数。

可选地，所述第一执行单元 120 被配置为：

对两个所述系统问题对应的上下文信息中系统问题所产生原因的截图进行光学字符识别，获得从所述截图中识别出的字符；

10 基于对两个所述系统问题识别出的字符中重合的字符数除以对两个所述系统问题识别出的字符数之和，得到第二相似分数。

可选地，所述第一执行单元 120 被配置为：

将两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息进行比对；

若两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息一致，则获取预设的第一分数作为第三相似分数；

15 若两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息不一致，则确定所述第三相似分数为零；

将两个所述系统问题对应的上下文信息中产生系统问题的设备信息进行比对；

若两个所述系统问题对应的上下文信息中产生系统问题的设备信息一致，则获取预设的第二分数作为第四相似分数；

20 若两个所述系统问题对应的上下文信息中产生系统问题的设备信息不一致，则确定所述第四相似分数为零。

可选地，所述第二执行单元 140 被配置为：

若所述比对结果为两个所述系统问题对应的上下文信息之间的相似度大于或者等于预定阈值，则确定两个所述系统问题为同一个系统错误所对应的系统问题；

25 若所述比对结果为两个所述系统问题对应的上下文信息之间的相似度小于预设阈值，则确定两个所述系统问题不是同一个系统错误所对应的系统问题。

上述装置中各个模块的功能和作用的实现过程具体详见上述对系统错误进行重复识别的方法中对应步骤的实现过程，在此不再赘述。

30 应当注意，尽管在上文详细描述中提及了用于动作执行的设备的若干模块或者单元，但是这种划分并非强制性的。实际上，根据本申请的实施方式，上文描述的两个或更多模块或者单元的特征和功能可以在一个模块或者单元中具体化。反之，上文描述的一个模块或者单元的特征和功能可以进一步划分为由多个模块或者单元来具体化。

35 此外，尽管在附图中以特定顺序描述了本申请中方法的各个步骤，但是，这并非要求或者暗示必须按照该特定顺序来执行这些步骤，或是必须执行全部所示的步骤才能实现期望的结果。附加的或备选的，可以省略某些步骤，将多个步骤合并为一个步骤执行，以及

/或者将一个步骤分解为多个步骤执行等。

通过以上的实施方式的描述，本领域的技术人员易于理解，这里描述的示例实施方式可以通过软件实现，也可以通过软件结合必要的硬件的方式来实现。因此，根据本申请实施方式的技术方案可以以软件产品的形式体现出来，该软件产品可以存储在一个非易失性存储介质（可以是 CD-ROM，U 盘，移动硬盘等）中或网络上，包括若干指令以使得一台计算设备（可以是个人计算机、服务器、移动终端、或者网络设备等）执行根据本申请实施方式的方法。

在本申请的示例性实施例中，还提供了一种能够实现上述方法的计算机设备。

所属技术领域的技术人员能够理解，本发明的各个方面可以实现为系统、方法或程序产品。因此，本申请的各个方面可以具体实现为以下形式，即：完全的硬件实施方式、完全的软件实施方式（包括固件、微代码等），或硬件和软件方面结合的实施方式，这里可以统称为“电路”、“模块”或“系统”。

参考图 4，图 4 是根据本申请的这种实施方式的电子设备 400。图 4 显示的设备 400 仅仅是一个示例，不应对本申请实施例的功能和使用范围带来任何限制。

如图 4 所示，电子设备 400 以通用计算设备的形式表现。设备 400 的组件可以包括但不限于：上述至少一个处理器 410、上述至少一个存储器 420、连接不同系统组件（包括存储器 420 和处理器 410）的总线 430。

其中，所述存储器存储有程序代码，所述程序代码可以被所述处理器 410 执行，使得所述处理器 410 执行本说明书上述“示例性方法”部分中描述的根据本申请各种示例性实施方式的步骤。例如，所述处理器 410 可以执行如图 1 中所示的步骤 S110：基于用于进行比对的两个系统问题，获取两个所述系统问题对应的上下文信息，其中，所述上下文信息包括系统问题所产生原因的文本、系统问题所产生原因的截图、产生系统问题的功能模块信息以及产生系统问题的设备信息；步骤 S120：确定两个所述系统问题对应的上下文信息之间的相似度；步骤 S130：将两个所述系统问题对应的上下文信息之间的相似度与预定阈值进行比对得到比对结果；步骤 S140：基于所述比对结果确定两个所述系统问题是否为同一个系统错误所对应的系统问题。

存储器 420 可以包括易失性存储单元形式的可读介质，例如随机存取存储单元（RAM）4201 和/或高速缓存存储单元 4202，还可以包括只读存储单元（ROM）4203。

存储器 420 还可以包括具有一组（至少一个）程序模块 4205 的程序/实用工具 4204，这样的程序模块 4205 包括但不限于：操作系统、一个或者多个应用程序、其它程序模块以及程序数据，这些示例中的每一个或某种组合中可能包括网络环境的实现。

总线 430 可以为表示几类总线结构中的一种或多种，包括存储单元总线或者存储单元控制器、外围总线、图形加速端口、处理单元或者使用多种总线结构中的任意总线结构的局域总线。

电子设备 400 也可以与一个或多个外部设备 600（例如键盘、指向设备、蓝牙设备等）

通信,还可与一个或者多个使得用户能与该电子设备 400 交互的设备通信,和/或与使得该电子设备 400 能与一个或多个其它计算设备进行通信的任何设备(例如路由器、调制解调器等等)通信。这种通信可以通过输入/输出(I/O)接口 440 进行。并且,电子设备 400 还可以通过网络适配器 460 与一个或者多个网络(例如局域网(LAN),广域网(WAN)和/或公共网络,例如因特网)通信。如图所示,网络适配器 460 通过总线 430 与电子设备 400 的其它模块通信。应当明白,尽管图中未示出,可以结合电子设备 400 使用其它硬件和/或软件模块,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID 系统、磁带驱动器以及数据备份存储系统等。

通过以上的实施方式的描述,本领域的技术人员易于理解,这里描述的示例实施方式可以通过软件实现,也可以通过软件结合必要的硬件的方式来实现。因此,根据本申请实施方式的技术方案可以以软件产品的形式体现出来,该软件产品可以存储在一个非易失性存储介质(可以是 CD-ROM, U 盘,移动硬盘等)中或网络上,包括若干指令以使得一台计算设备(可以是个人计算机、服务器、终端装置、或者网络设备等)执行根据本申请实施方式的方法。

在本申请的示例性实施例中,还提供了一种计算机可读存储介质,其上存储有能够实现本说明书上述方法的程序产品。在一些可能的实施方式中,本申请的各个方面还可以实现为一种程序产品的形式,其包括程序代码,当所述程序产品在终端设备上运行时,所述程序代码用于使所述终端设备执行本说明书上述“示例性方法”部分中描述的根据本申请各种示例性实施方式的步骤。

参考图 5 所示,描述了根据本申请的实施方式的用于实现上述方法的程序产品 500,其可以采用便携式紧凑盘只读存储器(CD-ROM)并包括程序代码,并可以在电子设备,例如个人电脑上运行。然而,本申请的程序产品不限于此,在本文件中,可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

所述程序产品可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以为但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。上述计算机可读存储介质可以为非易失性可读存储介质,例如存储于 CD-ROM、U 盘或移动硬盘等设备中的非易失性可读存储介质,可以包括若干指令以使得一台计算设备(可以是个人计算机、服务器、终端装置、或者网络设备等)执行根据本申请实施方式的方法。可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM 或闪存)、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了可读程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光

信号或上述的任意合适的组合。可读信号介质还可以是可读存储介质以外的任何可读介质，该可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于无线、有线、
5 光缆、RF 等等，或者上述的任意合适的组合。

可以以一种或多种程序设计语言的任意组合来编写用于执行本申请操作的程序代码，所述程序设计语言包括面向对象的设计语言—诸如 Java、C++ 等，还包括常规的过程式程序设计语言—诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算
10 设备上部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中，远程计算设备可以通过任意种类的网络，包括局域网（LAN）或广域网（WAN），连接到用户计算设备，或者，可以连接到外部计算设备（例如利用因特网服务提供商来通过因特网连接）。

本申请一实施例提供了一种计算机程序产品，上述计算机程序产品包括存储了计算机
15 程序的非瞬时性计算机可读存储介质，计算机程序可操作来使计算机执行如本申请前述各个实施例中的对系统错误进行重复识别的方法。

此外，上述附图仅是根据本申请示例性实施例的方法所包括的处理的示意性说明，而不是限制目的。易于理解，上述附图所示的处理并不表明或限制这些处理的时间顺序。另外，也易于理解，这些处理可以是例如在多个模块中同步或异步执行的。

20 本领域技术人员在考虑说明书及实践这里公开的申请后，将容易想到本申请的其他实施例。本申请旨在涵盖本公开的任何变型、用途或者适应性变化，这些变型、用途或者适应性变化遵循本申请的一般性原理并包括本申请未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的，本申请的真正范围和精神由权利要求指出。

权利要求

1、一种对系统错误进行重复识别的方法，所述方法包括：

基于用于进行比对的两个系统问题，获取两个所述系统问题对应的上下文信息，其中，
5 所述上下文信息包括系统问题所产生原因的文本、系统问题所产生原因的截图、产生系统问题的功能模块信息以及产生系统问题的设备信息；

确定两个所述系统问题对应的上下文信息之间的相似度；

将两个所述系统问题对应的上下文信息之间的相似度与预定阈值进行比对得到比对结果；

10 基于所述比对结果确定两个所述系统问题是否为同一个系统错误所对应的系统问题。

2、根据权利要求1所述的方法，其中，所述确定两个所述系统问题对应的上下文信息之间的相似度的步骤，包括：

比对两个所述系统问题对应的上下文信息中系统问题所产生原因的文本，确定第一相似分数；

15 比对两个所述系统问题对应的上下文信息中系统问题所产生原因的截图，确定第二相似分数；

比对两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息，确定第三相似分数；

20 比对两个所述系统问题对应的上下文信息中产生系统问题的设备信息，确定第四相似分数；

基于所述第一相似分数、所述第二相似分数、所述第三相似分数以及所述第四相似分数，确定两个所述系统问题对应的上下文信息之间的相似度。

25 3、根据权利要求2所述的方法，其中，所述基于所述第一相似分数、所述第二相似分数、所述第三相似分数以及所述第四相似分数，确定两个所述系统问题对应的上下文信息之间的相似度，包括：

基于以下公式确定两个系统问题之间的相似度：

$$S = \frac{\ln(A + B + e)}{\alpha} + \beta (C + D)$$
，其中， α 和 β 为大于1的正常数，A为所述第一相似分数，B为所述第二相似分数，C为所述第三相似分数，D为所述第四相似分数，S为两个系统问题对应的上下文信息之间的相似度，e为自然常数。

30 4、根据权利要求2所述的方法，其中，所述第一相似分数基于以下方法确定：

将两个所述系统问题对应的上下文信息中系统问题所产生原因的文本中的标题和内容分别进行分词；

针对每个所述系统问题对应的上下文信息中系统问题所产生原因的文本中的标题和内容，确定分出的每个词在分出的词中的出现次数，将出现次数高于预定出现次数阈值的
35 词作为所述系统问题对应的上下文信息中系统问题所产生原因中的关键词；

基于两个所述系统问题对应的关键词的交集的关键词数除以两个所述系统问题对应的关键词的并集的关键词数，得到第一相似分数。

5、根据权利要求2所述的方法，其特征在于，所述第二相似分数基于以下方法确定：

对两个所述系统问题对应的上下文信息中系统问题所产生原因的截图进行光学字符识别，获得从所述截图中识别出的字符；

基于对两个所述系统问题识别出的字符中重合的字符数除以对两个所述系统问题识别出的字符数之和，得到第二相似分数。

6、根据权利要求2所述的方法，其特征在于，所述第三相似分数基于以下方法确定：

将两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息进行比对；

10 若两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息一致，则获取预设的第一分数作为第三相似分数；

若两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息不一致，则确定所述第三相似分数为零；

所述第四相似分数基于以下方法确定：

15 将两个所述系统问题对应的上下文信息中产生系统问题的设备信息进行比对；

若两个所述系统问题对应的上下文信息中产生系统问题的设备信息一致，则获取预设的第二分数作为第四相似分数；

若两个所述系统问题对应的上下文信息中产生系统问题的设备信息不一致，则确定所述第四相似分数为零。

20 7、根据权利要求1所述的方法，其特征在于，所述基于所述比对结果确定两个所述系统问题是否为同一个系统错误所对应的系统问题的步骤，包括：

若所述比对结果为两个所述系统问题对应的上下文信息之间的相似度大于或者等于预定阈值，则确定两个所述系统问题为同一个系统错误所对应的系统问题；

25 若所述比对结果为两个所述系统问题对应的上下文信息之间的相似度小于预定阈值，则确定两个所述系统问题不是同一个系统错误所对应的系统问题。

8、一种对系统错误进行重复识别的装置，所述装置包括：

获取单元，用于基于用于进行比对的两个系统问题，获取两个所述系统问题对应的上下文信息，其中，所述上下文信息包括系统问题所产生原因的文本、系统问题所产生原因的截图、产生系统问题的功能模块信息以及产生系统问题的设备信息；

30 第一执行单元，用于确定两个所述系统问题对应的上下文信息之间的相似度；

比对单元，用于将两个所述系统问题对应的上下文信息之间的相似度与预定阈值进行比对得到比对结果；

第二执行单元，用于基于所述比对结果确定两个所述系统问题是否为同一个系统错误所对应的系统问题。

35 9、根据权利要求8所述的装置，其中，所述第一执行单

元被配置为：

比对两个所述系统问题对应的上下文信息中系统问题所产生原因的文本，确定第一相似分数；

5 比对两个所述系统问题对应的上下文信息中系统问题所产生原因的截图，确定第二相似分数；

比对两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息，确定第三相似分数；

比对两个所述系统问题对应的上下文信息中产生系统问题的设备信息，确定第四相似分数；

10 基于所述第一相似分数、所述第二相似分数、所述第三相似分数以及所述第四相似分数，确定两个所述系统问题对应的上下文信息之间的相似度。

10、根据权利要求9所述的系统错误进行重复识别的装置，其中，所述第一执行单元120被配置为：

基于以下公式确定两个系统问题之间的相似度：

15
$$S = \frac{\ln(A + B + e)}{\alpha} + \beta (C + D)$$
，其中， α 和 β 为大于1的正常数，A为所述第一相似分数，B为所述第二相似分数，C为所述第三相似分数，D为所述第四相似分数，S为两个系统问题对应的上下文信息之间的相似度，e为自然常数。

11、根据权利要求9所述的系统错误进行重复识别的装置，其中，所述第一执行单元被配置为：

20 将两个所述系统问题对应的上下文信息中系统问题所产生原因的文本中的标题和内容分别进行分词；

针对每个所述系统问题对应的上下文信息中系统问题所产生原因的文本中的标题和内容，确定分出的每个词在分出的词中的出现次数，将出现次数高于预定出现次数阈值的词作为所述系统问题对应的上下文信息中系统问题所产生原因中的关键词；

25 基于两个所述系统问题对应的关键词的交集的关键词数除以两个所述系统问题对应的关键词的并集中的关键词数，得到第一相似分数。

12、根据权利要求9所述的系统错误进行重复识别的装置，其中，所述第一执行单元被配置为：

30 对两个所述系统问题对应的上下文信息中系统问题所产生原因的截图进行光学字符识别，获得从所述截图中识别出的字符；

基于对两个所述系统问题识别出的字符中重合的字符数除以对两个所述系统问题识别出的字符数之和，得到第二相似分数。

13、根据权利要求9所述的系统错误进行重复识别的装置，其中，所述第一执行单元被配置为：

35 将两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息进行比对；

若两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息一致,则获取预设的第一分数作为第三相似分数;

若两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息不一致,则确定所述第三相似分数为零;

5 将两个所述系统问题对应的上下文信息中产生系统问题的设备信息进行比对;

若两个所述系统问题对应的上下文信息中产生系统问题的设备信息一致,则获取预设的第二分数作为第四相似分数;

若两个所述系统问题对应的上下文信息中产生系统问题的设备信息不一致,则确定所述第四相似分数为零。

10 14、根据权利要求 8 所述的系统错误进行重复识别的装置,其中,所述第二执行单元被配置为:

若所述比对结果为两个所述系统问题对应的上下文信息之间的相似度大于或者等于预定阈值,则确定两个所述系统问题为同一个系统错误所对应的系统问题;

15 若所述比对结果为两个所述系统问题对应的上下文信息之间的相似度小于预设阈值,则确定两个所述系统问题不是同一个系统错误所对应的系统问题。

15、一种电子设备,包括处理器;以及存储器,用于存储所述处理器的对系统错误进行重复识别的程序;其中,所述处理器配置为经由执行所述对系统错误进行重复识别的程序来执行以下处理方法:

20 基于用于进行比对的两个系统问题,获取两个所述系统问题对应的上下文信息,其中,所述上下文信息包括系统问题所产生原因的文本、系统问题所产生原因的截图、产生系统问题的功能模块信息以及产生系统问题的设备信息;

确定两个所述系统问题对应的上下文信息之间的相似度;

将两个所述系统问题对应的上下文信息之间的相似度与预定阈值进行比对得到比对结果;

25 基于所述比对结果确定两个所述系统问题是否为同一个系统错误所对应的系统问题。

16、根据权利要求 15 所述的电子设备,其中,所述确定两个所述系统问题对应的上下文信息之间的相似度,包括:

比对两个所述系统问题对应的上下文信息中系统问题所产生原因的文本,确定第一相似分数;

30 比对两个所述系统问题对应的上下文信息中系统问题所产生原因的截图,确定第二相似分数;

比对两个所述系统问题对应的上下文信息中产生系统问题的功能模块信息,确定第三相似分数;

35 比对两个所述系统问题对应的上下文信息中产生系统问题的设备信息,确定第四相似分数;

基于所述第一相似分数、所述第二相似分数、所述第三相似分数以及所述第四相似分数，确定两个所述系统问题对应的上下文信息之间的相似度。

- 17、根据权利要求 16 所述的电子设备，其中，所述基于所述第一相似分数、所述第二相似分数、所述第三相似分数以及所述第四相似分数，确定两个所述系统问题对应的上下文信息之间的相似度，包括：

基于以下公式确定两个系统问题之间的相似度：

$$S = \frac{\ln(A + B + e)}{\alpha} + \beta (C + D)$$

，其中， α 和 β 为大于 1 的正常数，A 为所述第一相似分数，B 为所述第二相似分数，C 为所述第三相似分数，D 为所述第四相似分数，S 为两个系统问题对应的上下文信息之间的相似度，e 为自然常数。

- 18、根据权利要求 16 所述的电子设备，其中，所述第一相似分数基于以下方法确定：

将两个所述系统问题对应的上下文信息中系统问题所产生原因的文本中的标题和内容分别进行分词；

- 针对每个所述系统问题对应的上下文信息中系统问题所产生原因的文本中的标题和内容，确定分出的每个词在分出的词中的出现次数，将出现次数高于预定出现次数阈值的词作为所述系统问题对应的上下文信息中系统问题所产生原因中的关键词；

基于两个所述系统问题对应的关键词的交集中的关键词数除以两个所述系统问题对应的关键词的并集中的关键词数，得到第一相似分数。

- 19、根据权利要求 16 所述的电子设备，其中，所述第二相似分数基于以下方法确定：

- 对两个所述系统问题对应的上下文信息中系统问题所产生原因的截图进行光学字符识别，获得从所述截图中识别出的字符；

基于对两个所述系统问题识别出的字符中重合的字符数除以对两个所述系统问题识别出的字符数之和，得到第二相似分数。

- 20、一种计算机可读存储介质，其上存储有对系统错误进行重复识别的程序，其特征在于，所述对系统错误进行重复识别的程序被处理器执行时实现权利要求 1 至 7 任一项所述的方法。

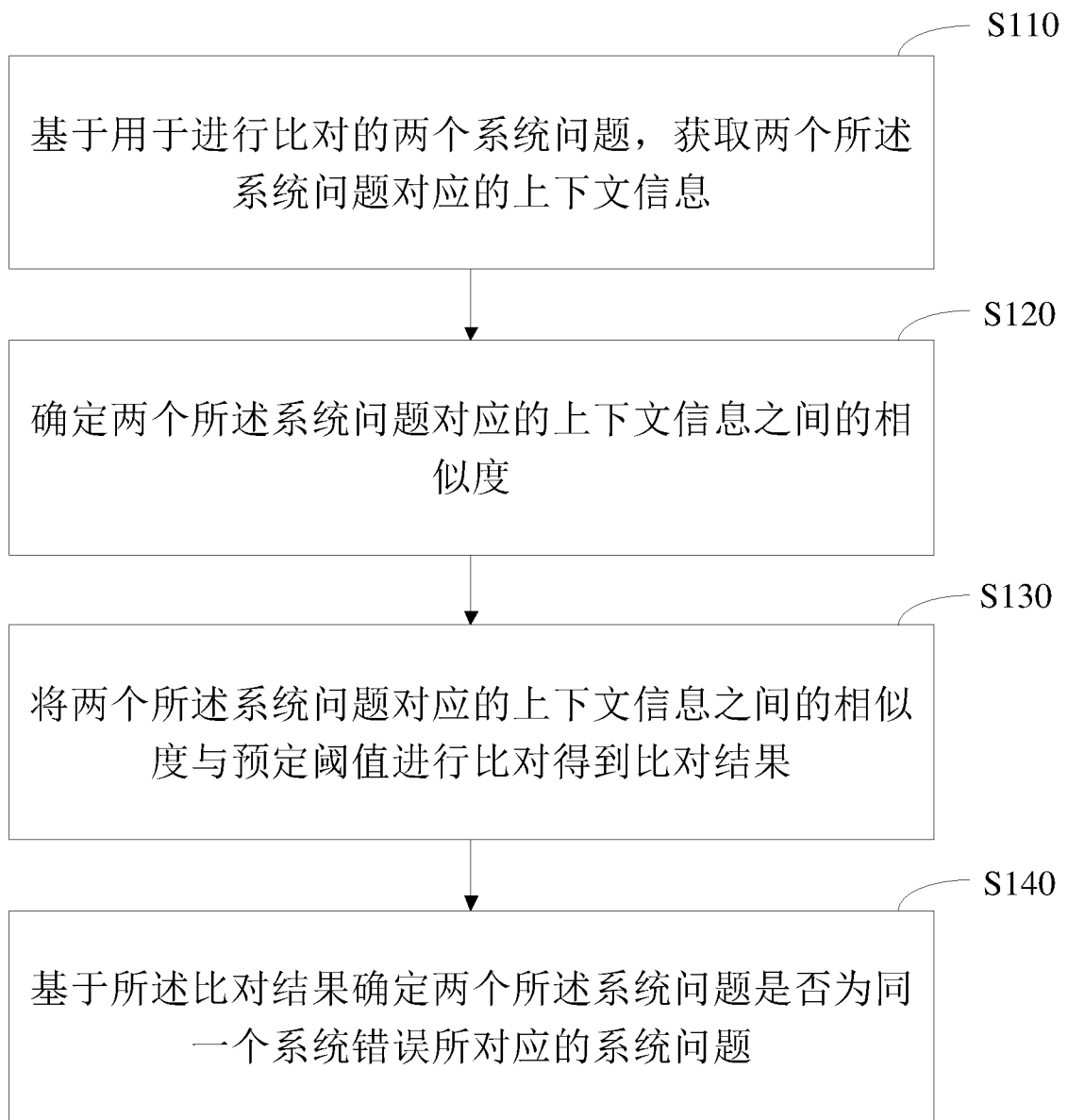


图 1

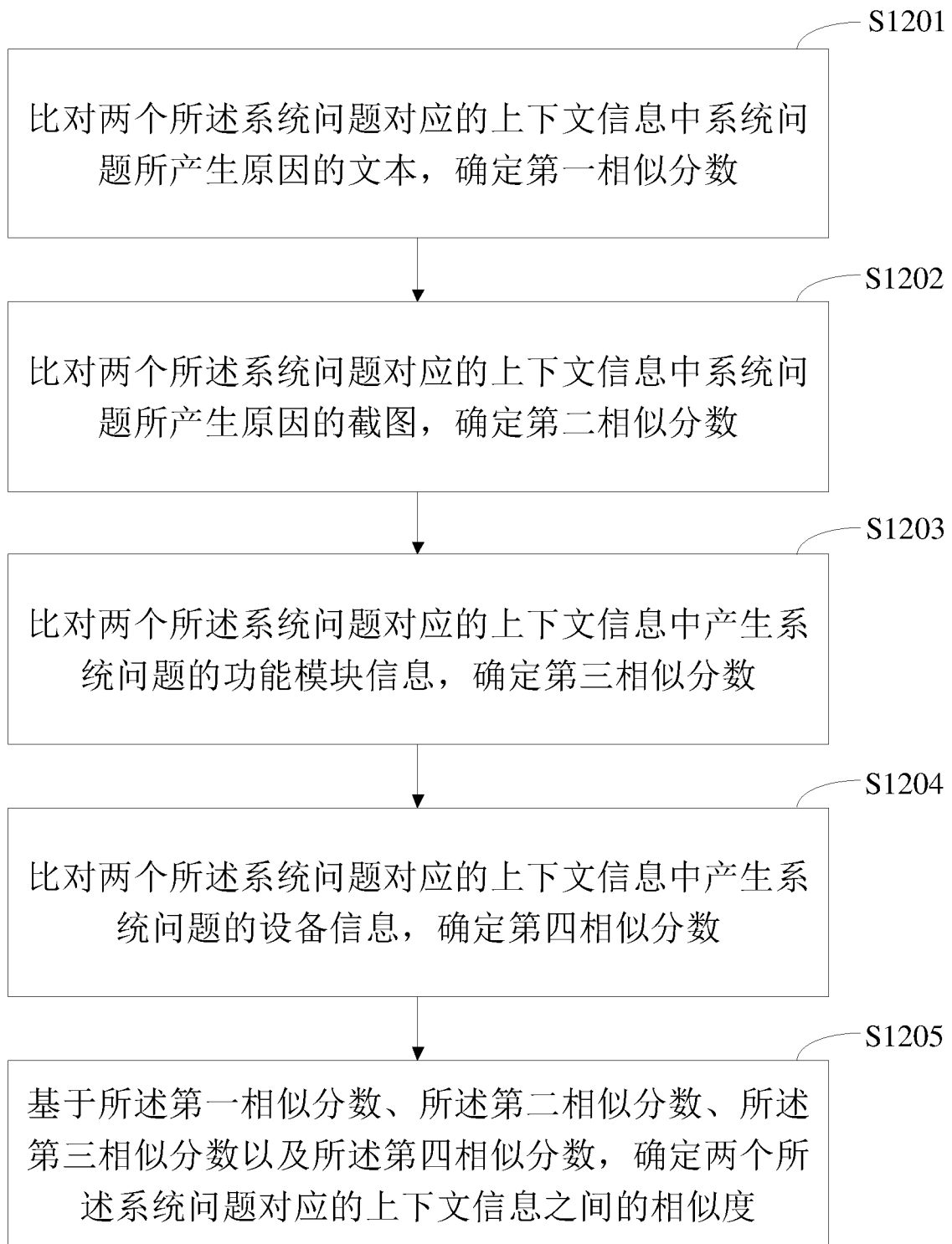


图 2



图 3

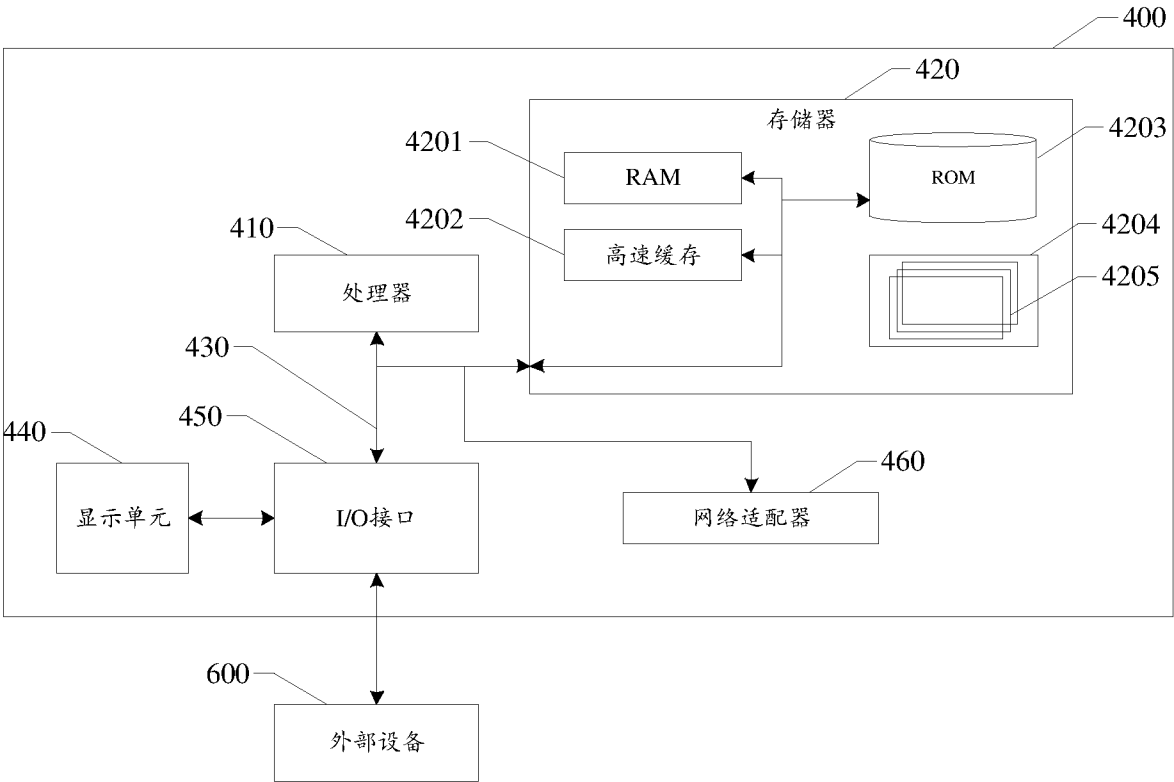


图 4

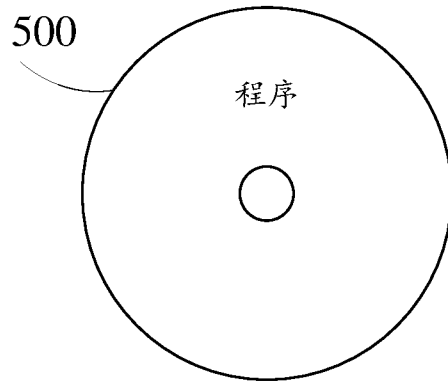


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/117684

A. CLASSIFICATION OF SUBJECT MATTER

G06F 11/07(2006.01)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; G06Q

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)

CNABS; CNTXT; CNKI; DWPI; WOTXT; USTXT; EPTXT: 问题, 上下文, 系统, 截图, 文本, 识别, 相似度, 阈值, 分数, 故障, 异常, 错误, 原因, 缺陷, 失效, 语境, 情境, 场景, 情景, 会话, 历史, 装置, 设备, 平台, 模块, 截屏, 截取, 图片, 图像, 画面, 屏幕, 图象, 抓取, 录屏, 快照, 文字, 文档, 字符, 内容, 去重, 冗余, 重新, 二次, 检测, 判断, 辨别, 认证, 标识, 辨识, 验证, 匹配, 鉴别, 确定, 相似性, 相关度, 匹配度, 相关性, 距离, 近似度, 比对, 相似程度, 比较, 门限, 阈值, 预设值, 预定, 上限, 设定值, 预定值, 范围, 评分, 得分, 打分, 分值, 评价, 成绩, 权重, 等级, context, bug, question, fault, error, system, device, screenshot, OCR, screencap, image, text, repeat, discrimin+, identif+, distinguish+, similarity, likelihood, threshold, limit, class+, divid+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 107085744 A (THE BOEING COMPANY) 22 August 2017 (2017-08-22) description, paragraphs 27-48	1-20
Y	CN 107167497 A (BEIJING BICOTEST TECH. CO., LTD.) 15 September 2017 (2017-09-15) description, paragraphs 81-84	1-20
A	US 2017278067 A1 (IBM) 28 September 2017 (2017-09-28) 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

19 April 2020

Date of mailing of the international search report

28 April 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
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/CN2019/117684

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	107085744	A	22 August 2017	EP	3206177	A1	16 August 2017
				US	2017236075	A1	17 August 2017
				US	9747564	B1	29 August 2017
				JP	2017199342	A	02 November 2017
CN	107167497	A	15 September 2017	None			
US	2017278067	A1	28 September 2017	US	2017278040	A1	28 September 2017

A. 主题的分类 G06F 11/07 (2006.01) 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; G06Q 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)) CNABS; CNTXT; CNKI; DWPI; WOTXT; USTXT; EPTXT: 问题, 上下文, 系统, 截图, 文本, 识别, 相似度, 阈值, 分数, 故障, 异常, 错误, 原因, 缺陷, 失效, 语境, 情境, 场景, 情景, 会话, 历史, 装置, 设备, 平台, 模块, 截屏, 截取, 图片, 图像, 画面, 屏幕, 图象, 抓取, 录屏, 快照, 文字, 文档, 字符, 内容, 去重, 冗余, 重新, 二次, 检测, 判断, 辨别, 认证, 标识, 辨识, 验证, 匹配, 鉴别, 确定, 相似性, 相关度, 匹配度, 相关性, 距离, 近似度, 比对, 相似程度, 比较, 门限, 阈值, 预设值, 预定, 上限, 设定值, 预定值, 范围, 评分, 得分, 打分, 分值, 评价, 成绩, 权重, 等级, context, bug, question, fault, error, system, device, screenshot, OCR, screencap, image, text, repeat, discrimin+, identif+, distinguish+, similarity, likelihood, threshold, limit, class+, divid+														
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>CN 107085744 A (波音公司) 2017年 8月 22日 (2017 - 08 - 22) 说明书第27-48段</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 107167497 A (北京必可测科技股份有限公司) 2017年 9月 15日 (2017 - 09 - 15) 说明书第81-84段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 2017278067 A1 (IBM) 2017年 9月 28日 (2017 - 09 - 28) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	Y	CN 107085744 A (波音公司) 2017年 8月 22日 (2017 - 08 - 22) 说明书第27-48段	1-20	Y	CN 107167497 A (北京必可测科技股份有限公司) 2017年 9月 15日 (2017 - 09 - 15) 说明书第81-84段	1-20	A	US 2017278067 A1 (IBM) 2017年 9月 28日 (2017 - 09 - 28) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求												
Y	CN 107085744 A (波音公司) 2017年 8月 22日 (2017 - 08 - 22) 说明书第27-48段	1-20												
Y	CN 107167497 A (北京必可测科技股份有限公司) 2017年 9月 15日 (2017 - 09 - 15) 说明书第81-84段	1-20												
A	US 2017278067 A1 (IBM) 2017年 9月 28日 (2017 - 09 - 28) 全文	1-20												
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。														
<table border="0"> <tr> <td> * 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 </td> <td> “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 </td> </tr> </table>			* 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										
* 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 4月 19日		国际检索报告邮寄日期 2020年 4月 28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 郭永强 电话号码 86-010-62412084												

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/117684

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	107085744	A	2017年 8月 22日	EP	3206177	A1	2017年 8月 16日
				US	2017236075	A1	2017年 8月 17日
				US	9747564	B1	2017年 8月 29日
				JP	2017199342	A	2017年 11月 2日
CN	107167497	A	2017年 9月 15日	无			
US	2017278067	A1	2017年 9月 28日	US	2017278040	A1	2017年 9月 28日