



(12)发明专利申请

(10)申请公布号 CN 106980634 A

(43)申请公布日 2017.07.25

(21)申请号 201610149047.3

(22)申请日 2016.03.16

(30)优先权数据

201641001810 2016.01.18 IN

(71)申请人 维布络有限公司

地址 印度卡纳塔克邦班加罗尔

(72)发明人 P·瑞丽 S·H·古鲁斯哇玛帕

R·巴拉苏布拉玛尼安

(74)专利代理机构 上海思微知识产权代理事务

所(普通合伙) 31237

代理人 智云

(51)Int.Cl.

G06F 17/30(2006.01)

G06Q 10/06(2012.01)

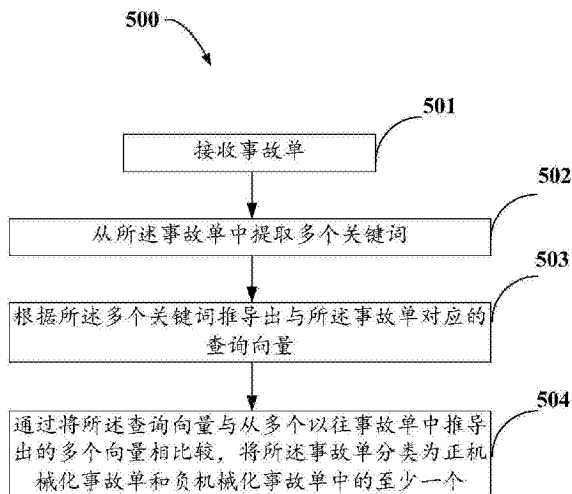
权利要求书2页 说明书15页 附图7页

(54)发明名称

用于分类和解决软件生产事故单的系统和方法

(57)摘要

本发明涉及一种用于分类和解决软件生产事故单的系统和方法。所述软件生产事故单分类方法包括:接收一事故单;从所述事故单中提取出多个关键词;以及根据所述多个关键词推导出与所述事故单对应的查询向量;以及通过将所述查询向量与从多个以往事故单中推导出的多个向量相比较,将所述事故单分类为正机械化事故单和负机械化事故单中的至少一个,其中,根据所述多个以往事故单的多个关键词及其在所述多个以往事故单中的相应出现次数,推导出所述多个向量。



1. 一种软件生产事故单分类方法,其特征在于,该方法包括:
 - 由一处理器接收一事故单;
 - 由所述处理器从所述事故单中提取出多个关键词;
 - 由所述处理器根据所述多个关键词推导出与所述事故单对应的查询向量;以及
 - 由所述处理器通过将所述查询向量与从多个以往事故单中推导出的多个向量相比较,将所述事故单分类为正机械化事故单和负机械化事故单中的至少一个,其中,根据所述多个以往事故单的多个关键词及其在所述多个以往事故单中的相应出现次数,推导出所述多个向量。
2. 如权利要求1所述的方法,其特征在于,提取出所述多个关键词包括,预处理所述事故单,以提取词干或去除至少一个停止词。
3. 如权利要求1所述的方法,其特征在于,还包括针对所述多个关键词中的每个关键词确定唯一数字表示符。
4. 如权利要求1所述的方法,其特征在于,从所述多个以往事故单推导出所述多个向量包括:
 - 根据机械化状态,将一事故库中的所述多个以往事故单分类为至少两个类别;
 - 针对每个类别,从所述多个以往事故单中提取多个关键词;以及
 - 针对每个类别,根据所述多个以往事故单的多个关键词及其在所述多个以往事故单中的相应出现次数,推导出所述多个向量。
5. 如权利要求4所述的方法,其特征在于,推导出所述多个向量包括,推导出所有关键词、名词关键词和动词关键词中的至少一个的多个向量。
6. 如权利要求4所述的方法,其特征在于,推导出所述多个向量包括,推导出所述多个关键词的众数、中数和值域中的至少一个的多个向量。
7. 如权利要求4所述的方法,其特征在于,推导出所述多个向量包括,将所述多个向量中的每个向量归一化。
8. 如权利要求1所述的方法,其特征在于,将所述事故单分类包括,根据所述查询向量和所述多个向量之间的相似度和相异度中的至少一个,实施最邻近分类。
9. 如权利要求8所述的方法,其特征在于,所述相似度包括余弦相似度,所述相异度包括欧几里得距离,实施所述最邻近分类包括根据该余弦相似度和欧几里得距离确定决定参数。
10. 如权利要求8所述的方法,其特征在于,将所述事故单分类还包括,根据所述多个向量推导出一个或多个特性矩阵,以及将所述查询向量与该一个或多个特性矩阵相比较。
11. 如权利要求1所述的方法,其特征在于,还包括利用现有解决方案解决所述事故单中的正机械化事故单,包括:
 - 从按所述事故单索引编制于知识库内的多个现有解决方案中识别出所述正机械化事故单的现有解决方案;以及
 - 调用与所述现有解决方案相关联的一个或多个脚本。
12. 如权利要求11所述的方法,其特征在于,还包括通过所述事故单以及所述正机械化事故单的现有解决方案,更新所述事故库。
13. 一种软件生产事故单分类系统,其特征在于,该系统包括:

至少一个处理器;以及

存有指令的计算机可读介质,所述指令在由所述至少一个处理器执行时使得该至少一个处理器实施操作,该操作包括:

接收一事故单;

从所述事故单中提取出多个关键词;

根据所述多个关键词推导出与所述事故单对应的查询向量;以及

通过将所述查询向量与从多个以往事故单中推导出的多个向量相比较,将所述事故单分类为正机械化事故单和负机械化事故单中的至少一个,其中,根据所述多个以往事故单的多个关键词及其在所述多个以往事故单中的相应出现次数,推导出所述多个向量。

14.如权利要求13所述的系统,其特征在于,所述操作还包括,针对所述多个关键词中的每个关键词确定唯一数字表示符。

15.如权利要求13所述的系统,其特征在于,从所述多个以往事故单推导出所述多个向量包括:

根据机械化状态,将一事故库中的所述多个以往事故单分类为至少两个类别;

针对每个类别,从所述多个以往事故单中提取多个关键词;以及

针对每个类别,根据所述多个以往事故单的多个关键词及其在所述多个以往事故单中的相应出现次数,推导出所述多个向量。

16.如权利要求15所述的系统,其特征在于,推导出所述多个向量包括,推导出所有关键词、名词关键词和动词关键词中的至少一个的多个向量。

17.如权利要求15所述的系统,其特征在于,推导出所述多个向量包括,推导出所述多个关键词的众数、中数和值域中的至少一个的多个向量。

18.如权利要求13所述的系统,其特征在于,将所述事故单分类包括,根据所述查询向量和所述多个向量之间的相似度和相异度中的至少一个,实施最邻近分类;所述相似度包括余弦相似度,所述相异度包括欧几里得距离,实施所述最邻近分类包括根据该余弦相似度和欧几里得距离确定决定参数。

19.如权利要求13所述的系统,其特征在于,所述操作还包括利用现有解决方案解决所述事故单中的正机械化事故单,包括:

从按所述事故单索引编制于知识库内的多个现有解决方案中识别出所述正机械化事故单的现有解决方案;以及

调用与所述现有解决方案相关联的一个或多个脚本。

20.一种存有计算机可执行指令的非暂时性计算机可读介质,其特征在于,所述指令用于:

接收一事故单;

从所述事故单中提取出多个关键词;

根据所述多个关键词推导出与所述事故单对应的查询向量;以及

通过将所述查询向量与从多个以往事故单中推导出的多个向量相比较,将所述事故单分类为正机械化事故单和负机械化事故单中的至少一个,其中,根据所述多个以往事故单的多个关键词及其在所述多个以往事故单中的相应出现次数,推导出所述多个向量。

用于分类和解决软件生产事故单的系统和方法

技术领域

[0001] 本发明总体涉及信息技术基础架构管理,尤其涉及用于分类和解决软件生产事故单的系统和方法。

背景技术

[0002] 在当前商业环境中,各行业和各企业均在不断寻求可高效运营且能满足不断发展的客户需求的方法和模型。上述方法中的一种为由信息技术(IT)领域中的各种进展实现的行业或企业内各种流程和活动的数字化。上述数字化通过IT基础架构实现,而该IT基础架构涉及设备和软件解决方案的复杂组合。然而,为了实现数字化的各种益处,上述IT基础架构需要平稳运行。

[0003] 现今已设计和开发出各种工具,用于对此类IT基础架构中的任何异常或故障进行监测和/或预测,以期实现此类异常的快速主动解决。然而,虽然获得了很大进展,但是支持团队所提供的解决方案仍然经常出现延误和/或准确性较差的问题。而且,由于各种原因,上述延误逐渐累积。在这些原因当中,某些纯粹是由与记录于IT基础架构内以提醒支持团队对其进行快速解决的各种异常或缺陷对应的无用和/或冗余警报或事故单引起的。各种用于解决此问题的现有优化工具和技术仍然未能有效解决此问题。因此,技术支持团队在待解决的无用警报或事故单的攻势下疲于应付,从而使得用于解决真正严重警报或事故单的时间受到限制。这些限制进一步对组织或企业的整体运作造成影响。采用自动化是解决上述无用警报或事故单问题的一种方法,该方法无需人工干预,从而提高了问题解决的效率并降低了成本。然而,从事故单中识别出待由自动化解解决的候选事故单以及对相应事故解决脚本的调用仍然需要手动完成。

发明内容

[0004] 在一种实施方式中,本发明公开了一种软件生产事故单分类方法。在一个实施例中,所述方法包括接收一事故单;从所述事故单中提取出多个关键词;根据所述多个关键词,推导出与所述事故单对应的查询向量;以及通过将所述查询向量与从多个以往事故单中推导出的多个向量相比较,将所述事故单分类为正机械化事故单和负机械化事故单中的至少一个,其中,所述多个向量根据所述多个以往事故单的多个关键词及其在所述多个以往事故单中的相应出现次数推导得出。

[0005] 在一种实施方式中,本发明还公开了一种软件生产事故单分类系统。在一个实施例中,所述系统包括,至少一个处理器,以及以可通信方式连接至所述至少一个处理器的存储器。所述存储器存有处理器可执行指令,该指令在执行时使得所述至少一个处理器实施操作,该操作包括:接收一事故单;从所述事故单中提取出多个关键词;根据所述多个关键词推导出与所述事故单对应的查询向量;以及通过将所述查询向量与从多个以往事故单中推导出的多个向量相比较,将所述事故单分类为正机械化事故单和负机械化事故单中的至少一个,其中,所述多个向量根据所述多个以往事故单的多个关键词及其在所述多个以往

事故单中的相应出现次数推导得出。

[0006] 在一种实施方式中,本发明还公开了一种非暂时性计算机可读介质,该介质存有用于分类软件生产事故单的计算机可执行指令。在一个实施例中,所述被存指令用于:接收一事故单;从所述事故单中提取出多个关键词;根据所述多个关键词推导出与所述事故单对应的查询向量;以及通过将所述查询向量与从多个以往事故单中推导出的多个向量相比较,将所述事故单分类为正机械化事故单和负机械化事故单中的至少一个,其中,所述多个向量根据所述多个以往事故单的多个关键词及其在所述多个以往事故单中的相应出现次数推导得出。

[0007] 应该理解的是,以上概略描述与以下详细描述均仅在于例示和说明,而不在于限制所要求保护的发明。

附图说明

[0008] 所附各图并入本发明之内并构成本发明的一部分,用于对例示实施方式进行描述,并与说明书一道阐明所公开的原理。

[0009] 图1为根据本发明一些实施方式的用于分类和解决软件生产事故单的例示系统的框图。

[0010] 图2为根据本发明一些实施方式的分类及解决引擎的功能框图。

[0011] 图3所示为保存于解决方案知识库内的例示知识表示符。

[0012] 图4所示为用于为图3中知识表示符编制索引的例示二维KD树。

[0013] 图5为根据本发明一些实施方式用于分类和解决软件生产事故单的例示方法流程图。

[0014] 图6为根据本发明一些实施方式用于分类和解决软件生产事故单的详细例示方法流程图。

[0015] 图7为用于实施符合本发明实施方式的例示计算机系统框图。

具体实施方式

[0016] 以下,参考附图,对例示实施方式进行描述。在任何方便之处,各图中均采用相同附图标记指代相同或类似部件。虽然本文中描述了所公开原理的实施例和特征,但是在不脱离所公开实施方式的精神和范围的前提下,还可进行修改、调整以及做出其他实施方式。以下具体描述意在仅视作例示,而真正的范围及精神如下附权利要求书所述。

[0017] 现在参考图1,该图所示为根据本发明一些实施方式的用于分类和解决软件生产事故单(incident tickets)的例示系统100。所述事故单对应于信息技术(IT)基础架构中的异常、缺陷或任何其他障碍。具体而言,系统100实施一种技术,该技术用于根据事故单的机械化或自动化可能性对其进行分类,并用于为可机械化或自动化的事故单提供自动化解决方案。

[0018] 系统100包括支持组织或企业的IT基础架构101。IT基础架构101可包括由通信网络相互连接的IT系统或设备,该系统或设备运行有软件应用程序主机,所述软件应用程序用于管理所述组织或企业、实现归档和检索等信息处理、以及实现工作流程等。由此可见,IT基础架构101为硬件资源、软件资源和网络资源的集合。所述硬件资源可以为服务器、计

计算机、交换机、路由器、接受器、集线器等。所述软件资源可以为操作系统、企业资源规划(ERP)解决方案、客户关系管理(CRM)解决方案、生产力应用程序等。所述网络资源可以为以太网、互联网、防火墙、通信协议等。网络管理员、开发人员、最终用户等多个用户102可在完成所述组织或企业的日常工作中使用IT基础架构101。

[0019] 系统100还包括IT服务管理(ITSM)工具103。ITSM工具103为业务层事故管理工具,在该工具中所有事故均记录为事故库中的事故单。ITSM工具103接收和处理事故单,以提供快速准确的解决方案。在一些实施方式中,ITSM工具103可以为基于IT基础架构库(ITIL)指导方针的软件版本。ITSM工具103例如可包括,但不限于,Wipro®eHelpline、BMC™ Remedy 9、ServiceNOW™等。所述事故单既可在监测和处理后自动记录于ITSM工具103中,也可手工录入。由此可见,由任何监测方案产生的符合条件的事件将被转化为事故单。同样地,用户102等人类用户也可将事故单录入ITSM工具103中。一般而言,系统所生成的事件或事故单涉及IT基础架构101的一个或多个部件的可用性、能力和性能(例如,存储空间使用率、CPU使用率等)。此外,所述事件或事故单也可涉及软件或应用程序基础架构故障(例如,自动系统调度程序问题、数据库刷新失败、长时间运行查询问题等),或者可涉及硬件故障(例如,网络端口问题、LAN问题等)。

[0020] 系统100还包括用于监测IT基础架构101的监测工具104。监测工具104以预定义阈值为相对参考,和/或以预定义时间为间隔,对特定基础架构部件(服务器、应用程序等)的行为进行检测。应该注意的是,在许多情况下,人们可根据常规模式或行为中的变化容易地发现IT基础架构101中的问题。例如,每当发生达到阈值的事件时,监测工具104均可生成针对该事件的警报。由此可见,监测工具104对与IT基础架构101中潜在问题相对应的问题或事件进行拾取。监测工具104例如可包括,但不限于,Wipro®Viking、HP®OpenView、IBM®Tivoli等。

[0021] 系统100还包括事件关联引擎105,该引擎用于对监测工具104发现的多个事件进行关联,并根据一组规则识别出具有相关性和重要性的一些事件。例如,当路由器宕机时,其底层设备或应用程序(如计算机、ERP)也均会发生宕机。之后,监测工具104可生成针对所述路由器以及其他底层设备的警报。而且,对这些事件的关联结果将表明问题在于所述路由器。除了对所述多个事件进行处理之外,关联引擎105还分析事件之间的关系,并根据该分析生成上述一组规则,以对无用事件进行抑制,以防其被录入为事故单。举例而言,所述规则可包括,但不限于,该警报已存在、多对一警报、通过断电抑制等。

[0022] 系统100还包括用于分类和解决软件生产事故单的事故分类及解决引擎106。如以下将结合图2更加详细描述,事故分类及解决引擎106根据事故单的机械化可能性对其进行分类,然后为可机械化的事故单提供自动化解决方案。事故分类及解决引擎106包括初步分析器107、运行时间分析器108、以及解决方案知识库(Solution Knowledge, SK库)及配置管理数据库(Configuration Management DataBase, CMDB)109。

[0023] 初步分析器107根据以前记录于ITSM 103中所有事故单的机械化状态对所有事故单进行分析,并通过自然语言处理推导出向量或特性矩阵,该向量或特性矩阵用于后续识别和确定出可能可机械化的候选事故单。其后,初步分析器107利用所述向量和特性矩阵对SK库及CMDB 109中的SK库进行更新。此外,初步分析器107还将所述机械化解决方案表示为

所述SK库中的知识,搜索该SK库中不同机械化解方案的可用性,以及相应更新该SK库。

[0024] 运行时间分析器108对来自所述事件关联引擎105及ITSM工具103的警报进行分析,并将这些警报分类为可机械化的候选事故单,或分类为不使用自然语言处理且非基于初步分析器107构建的所述向量或特性矩阵的事故单。此外,运行时间分析器108还从SK库中获取关于正机械化候选事故单的可用机械化解方案的知识,并将这些正机械化候选事故单映射至所述可用机械化解方案。此外,运行时间分析器108还对ITSM工具103中的解决方案信息进行更新,然后将该解决方案实施或应用于IT基础架构101。此外,本领域技术人员可以理解的是,运行时间分析器108还可使用负机械化候选事故单对ITSM工具103进行更新。在此之后,由用户110从ITSM工具103中捡取并手动解决这些事故单,并且将相应解决方案在ITSM工具103中更新,以便IT基础架构101中的执行应用程序对其进行后续使用。

[0025] SK库为向量和/或特性矩阵以及自动化解决方案(例如脚本)的知识库。所述解决方案索引至警报/事故单。SK库及CMDB 109中的CMDB为配置库,该配置库包括与所述IT基础架构的具体位置(例如IP地址等)、性能(例如服务器性能、存储空间容量等)以及其他此类信息的配置项(CI)相关的信息。ITSM工具103利用该信息执行IT基础架构101内的解决方案。

[0026] 现在参考图2,该图为根据本发明一些实施方式由图1中系统100实现的事事故分类及解决引擎200(类似于事故分类及解决引擎106)的示意功能框图。如上所述,事故分类及解决引擎200包括初步分析器201(与初步分析器107类似)、运行时间分析器202(与运行时间分析器108类似)以及SK库及CMDB 203(与SK库及CMDB 109类似)。初步分析器201根据以往事故单推导出向量和/或特性矩阵,而运行时间分析器202判断当前事故单为机械化候选事故单,或者为非基于所述向量和/或特性矩阵的事故单。此外,SK库及CMDB 203存有向量、特性矩阵、现有解决方案或解决方法、正机械化事故单与所述现有解决方案之间的映射、CI相关信息以及其他此类信息。

[0027] 在步骤204中,初步分析器201将所有以往(即以前记录的)事故单作为训练数据,从所述ITSM工具对其进行获取。在一些实施方式中,所述训练数据可根据该信息经验知识手动选择。所述手动选择可通过咨询自动化领域专家、操作工程师等领域专家而实现。在此之后,在步骤206中,根据以往事故单205的已知机械化状态对其进行分类。如此,所有所述以往事故单均分类至至少两个类别或等级中:类别1为可机械化的事故单(即有自动化解决方案的正样本);类别2为不可机械化的事故单(即无自动化解决方案的负样本)。举例而言,指出“设备被盗”的事故单因所述设备需要以物理方式更换,因此不是机械化候选事故单。然而,指出“存储空间使用率已超阈值”的事故单因可通过运行脚本清理存储空间而得以解决,因此可以为机械化候选事故单。在一些实施方式中,还可根据所述以往事故单的机械化程度将其分类至多于两个的类别中:可完全机械化;不可机械化;可部分机械化(例如,40%机械化、60%机械化等)。

[0028] 在步骤207中,针对每个类别,可对每个事故单的描述进行初始预处理,从而在保留所述事故单的内容、背景及含义的同时,确定一简洁或代表性描述。在一些实施方式中,所述预处理可包括去除停止词(例如,标点符号、数字等),这是因为所述停止词对事故单的内容、背景及含义的贡献极小或没有贡献,从而导致后续分类处理的准确性降低。例如,对于具有“a device hasbeen stolen”(“设备被盗”)这一描述的事故单,可将“has”、“been”

以及所有标点符号等停止词从该描述中去除。此外,在一些实施方式中,所述预处理可包括在所述描述中同一单词的不同形式可由同一单词表示时(例如,device/devices、utilization/utilizes/utilizing等)提取词干。应该注意的是,所述预处理还可以包括其他文本处理技术。如此,针对具有“memory utilization has exceeded the threshold value on window server”这一描述的事故单,所述预处理可提供“memory utilization exceed threshold window server”或“memory utilization”这一简洁或代表性描述。由上可知,对于每个类别,我们均具有与属于该给定类别的多个事故单相对应的多个简洁或代表性描述。

[0029] 此外,在步骤208中,对于每个类别,还从属于该给定类别的多个简洁描述中抽取一些关键字。例如,可从简洁描述“device stolen”中抽取关键词“device”和“stolen”,以及可从简洁描述“memory utilization exceed threshold window server”中抽取关键字“memory”、“utilization”、“exceed”、“threshold”、“window”和“server”。在一些实施方式中,利用上述关键词,可至少建立三个不同的训练语料库,即包括所有关键词(例如,“device”、“stolen”、“memory”、“utilization”等)的训练语料库;包括所有名词关键词(例如,“device”、“CPU”、“memory”等)的训练语料库;包括所有动词关键词(例如,“stolen”、“utilization”等)的训练语料库。应该注意的是,在各种其他实施方式中,还可采用取决于各种其他关键词分类方法的各种其他训练语料库。此外,在一些实施方式中,为了后续处理的方便性,可使用唯一数字表示符代表每个所述关键词。例如,在一些实施例中,可以为每个字母分配唯一数字,然后可根据每个字母在关键词内的位置对该字母实施指数和减法运算,从而求得所述唯一数字表示符。应该注意的是,之所以选择指数和减法运算的原因在于这些运算的非结合性。如此,可按以下描述确定关键词“device”的唯一数字表示符:

[0030]

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

[0031] 可知, $D \rightarrow 4(p-1)$, $e \rightarrow 5(p-2)$, $v \rightarrow 22(p-3)$, $i \rightarrow 9(p-4)$, $c \rightarrow 3(p-5)$, $e \rightarrow 5(p-6)$, 其中 p 表示相应字母在所述关键词内的位置。因此,关键词“device”的唯一数字表示符为 $4^1 - 5^2 - 22^3 - 9^4 - 3^5 - 5^6$, 即-33098。同样地,关键词“CPU”的唯一数字表示符为 $3^1 - 16^2 - 21^3$, 即-9514; 关键词“memory”的唯一数字表示符为 $13^1 - 5^2 - 13^3 - 15^4 - 18^5 - 25^6$, 即-246083027; 关键词“utilization”的唯一数字表示符为 $21^1 - 20^2 - 9^3 - 12^4 - 9^5 - 26^6 - 1^7 - 20^8 - 9^9 - 15^{10} - 14^{11}$, 即-4652511977448。

[0032] 此外,在步骤209中,对于每个类别,根据多个关键词及其在多个以往事故单内的相应出现次数,推导出多个向量。在一些实施方式中,可针对所述多个关键词的众数(mode)、中数(median)和值域(range)中的至少一个导出所述向量。因此,可通过反复提取最常见的关键词(即众数)、在按出现频率的升序或降序排列的关键词中反复提取出现频率处于中间的关键词(即中数)、在按出现频率的升序或降序排列的关键词中反复提取最常见和最少见的关键词(即值域),推导出所述向量。每个所述向量的大小(即该向量中的关键词数)可根据预设数目(例如,50个关键字)确定,或当满足一些预设条件(例如,至少出现50次

的关键词)时确定。应该注意的是,在各种其他实施方式中,可推导出取决于以往事故单中关键词出现次数的各种其他向量。在一些实施方式中,可针对每个所述训练语料库(即所有关键词训练语料库、名词关键词训练语料库和动词关键词训练语料库),推导出所述多个向量。如此,在一些实施方式中,如下所示,每个类别有9个向量,共有18个向量:

[0033] 向量1=类别1(可机械化)的{所有关键词的众数},大小为m1;

[0034] 向量2=类别1的{所有名词关键词的众数},大小为m2;

[0035] 向量3=类别1的{所有动词关键词的众数},大小为m3;

[0036] 向量4=类别1的{所有关键词的中数},大小为m4;

[0037] 向量5=类别1的{所有名词关键词的中数},大小为m5;

[0038] 向量6=类别1的{所有动词关键词的中数},大小为m6;

[0039] 向量7=类别1的{所有关键词的值域},大小为m7;

[0040] 向量8=类别1的{所有名词关键词的值域},大小为m8;

[0041] 向量9=类别1的{所有动词关键词的值域},大小为m9;

[0042] 向量10=类别2(不可机械化)的{所有关键词的众数},大小为m10;

[0043] 向量11=类别2的{所有名词关键词的众数},大小为m11;

[0044] 向量12=类别2的{所有动词关键词的众数},大小为m12;

[0045] 向量13=类别2的{所有关键词的中数},大小为m13;

[0046] 向量14=类别2的{所有名词关键词的中数},大小为m14;

[0047] 向量15=类别2的{所有动词关键词的中数},大小为m15;

[0048] 向量16=类别2的{所有关键词的值域},大小为m16;

[0049] 向量17=类别2的{所有名词关键词的值域},大小为m17;

[0050] 向量18=类别2的{所有动词关键词的值域},大小为m18。

[0051] 此外,在一些实施方式中,对于每个类别,可根据所述多个向量,推导出一个或多个特性矩阵。每个特性矩阵均为所述多个向量中的一组向量。例如,在一些实施方式中,共有6个矩阵(每个类别3个矩阵),其中,矩阵1可以为分别表示类别1的所有关键词、名词关键词和动词关键词的众数的向量1、向量2和向量3的集合,矩阵2可以为分别表示类别1的所有关键词、名词关键词和动词关键词的中数的向量4、向量5和向量6的集合,矩阵3可以为分别表示类别1的所有关键词、名词关键词和动词关键词的值域的向量7、向量8和向量9的集合,矩阵4可以为分别表示类别2的所有关键词、名词关键词和动词关键词的众数的向量10、向量11和向量12的集合,矩阵5可以为分别表示类别2的所有关键词、名词关键词和动词关键词的中数的向量13、向量14和向量15的集合,矩阵6可以为分别表示类别2的所有关键词、名词关键词和动词关键词的值域的向量16、向量17和向量18的集合。上述导出的向量和/或特性矩阵保存于SK库及CMDB203中,以便运行时间分析器202在后续将其用于将所述当前事故单分类为机械化候选事故单或非机械化候选事故单。此外,如以下将结合图3和图4更加详细描述,初步分析器201可在SK库及CMDB 203中的SK库(即知识库)中对已有机械化解决方案进行索引编制、保存及更新,以便可机械化事故单的后续解决。

[0052] 在步骤210中,运行时间分析器202将所述事故单作为测试数据或运行时间数据,从所述ITSM工具或事件关联引擎对其进行接收。如上所述,所述事件关联引擎向运行时间分析器202提供自动捕获事故单,而所述ITSM工具向运行时间分析器202提供手动录入事故

单。其后,在步骤212中,对事故单211进行预处理,以从最初提供的描述中提取该事故单的简洁或代表性描述。预处理步骤212与上述预处理步骤207类似,而且可也涉及去除停止词(如标点符号、数字等)、提取词干及其他此类文本处理技术。此外,在步骤213中,从所述简洁或代表性描述中提取若干关键词。此外,在一些实施方式中,为了后续处理的方便性,每个所述关键词均可由唯一数字表示符代表。步骤213类似于上述用于提取关键词并求出被提关键词的唯一数字表示符的步骤208。此外,在步骤214中,根据所述多个关键词,推导出查询向量。

[0053] 在此之后,在步骤215中,通过将所述查询向量与所述多个向量或特性矩阵相比较,将所述事故单分类至正机械化事故单与负机械化事故单当中的至少一个。在一些实施方式中,可将所述事故单分类至多于两个的类别中,即可完全机械化;不可机械化;可部分机械化(例如,40%机械化、60%机械化等)。在分类时,首先从SK库及CMDB 203对所导出的多个向量和/或特性矩阵进行访问。在一些实施方式中,可使用下式将每个所导出的向量和/或特性矩阵归一化:

$$[0054] \quad \text{归一化值} = \frac{\text{向量}(i,j)}{\sqrt{\sum_{i=1}^n \sum_{j=1}^m (\text{向量}(i,j))^2}}$$

[0055] 其中,i和j表示各单词表示符在所述向量和/或特性矩阵中的位置,而n和m表示所述向量和/或特性矩阵的大小。应该注意的是,所述向量的i的值为1,因此n的值也为1,而所述特性矩阵因包括3个向量,因此其n值等于3。此外,所述查询向量的大小(即该查询向量内的关键词数)调整至与待与其比较的所述向量或特性矩阵的大小相同。这可通过利用伪关键词(如“00000”)填充所述查询向量,直至调整后的查询向量的大小与待与其比较的所述向量或特性矩阵的大小相同而实现。此外,还对上述调整后的查询向量内的各原始关键词的位置进行进一步调整,以使其与此类各关键词在待与其比较的所述向量或特性矩阵内的位置相对应(即该各关键词在待与所述查询向量比较的所述向量内处于相同位置)。举例而言,假如所述查询向量包含关键字“device”和“stolen”,而待与其比较的向量具有100个关键词,其中“device”处于第4号位置,“stolen”处于第99号位置,那么,调整后的查询向量也将包含100个关键词,其中“device”处于第4号位置,“stolen”处于第99号位置,其余位置为伪关键词。

[0056] 在一些实施方式中,所述分类包括,根据所述查询向量与所述多个向量中每一向量之间的相似度和相异度中的至少一个,实施最邻近分类。在一些实施方式中,求取所述相似度包括求出以下余弦相似度:

$$[0057] \quad \text{余弦相似度} = \frac{\text{向量1} \cdot \text{向量2}}{\|\text{向量1}\| \|\text{向量2}\|}$$

[0058] 此外,在一些实施方式中,求取所述相异度包括求出以下欧几里得距离:

$$[0059] \quad \text{欧几里得距离} = \sqrt{(\text{向量1} - \text{向量2})^2}$$

[0060] 其中,向量1为所述查询向量,向量2为与所述查询向量相比较的向量。之后,通过根据所述余弦相似度和欧几里得距离求得以下决定参数,实施所述最邻近分类:

[0061]
$$\text{决定参数} = \frac{\text{余弦相似度}}{\text{欧几里得距离}}$$

[0062] 如此,举例而言,对于以下两个向量:

[0063]

向量1	7	2	5	0	9
向量2	0	2	3	2	4

[0064]
$$\text{余弦相似度} = \frac{7.0 + 2.2 + 5.3 + 0.2 + 9.4}{\sqrt{7^2 + 2^2 + 5^2 + 0^2 + 9^2} \times \sqrt{0^2 + 2^2 + 3^2 + 2^2 + 4^2}}$$

[0065]
$$\text{余弦相似度} = \frac{55}{12.60 \times 5.7}$$

[0066] 余弦相似度=0.76

[0067]

$$\text{欧几里得距离} = \sqrt{(7-0)^2 + (2-2)^2 + (5-3)^2 + (0-2)^2 + (9-4)^2}$$

[0068] 欧几里得距离=9.05

[0069]
$$\text{决定参数} = \frac{0.76}{9.05}$$

[0070] 决定参数=0.0839

[0071] 在一些实施方式中,根据所述查询向量与上述18个向量(类别1的9个向量以及类别2的9个向量)中的每个向量或者与上述6个特性矩阵(类别1的3个特性矩阵以及类别2的3个特性矩阵),计算出多个决定参数。之后,计算类别1的9个向量或类别1的3个特性矩阵的决定参数的平均值。其后,同样地,计算类别2的9个向量或类别2的3个特性矩阵的决定参数的平均值。所述事故单归类于平均决定因子值较高的类别。

[0072]

类别 1	类别 2
决定参数 1	决定参数 10

[0073]

决定参数 2	决定参数 11
决定参数 3	决定参数 12
决定参数 4	决定参数 13
决定参数 5	决定参数 14
决定参数 6	决定参数 15
决定参数 7	决定参数 16
决定参数 8	决定参数 17
决定参数 9	决定参数 18
类别 1 的平均决定参数	类别 2 的平均决定参数

[0074] 如果所述事故单归类于代表不可机械化事故单类别2,则其为负机械化事故单216,并在步骤217中通过此事故单对所述ITSM或该ITSM内的事故库进行更新。之后,可捡取此类事故单,并通过手动方法对其进行解决。然而,如果所述事故单归类于代表可机械化事故单类别1,则其为正机械化事故单218。在一些实施方式中,运行时间分析器202在步骤219中通过现有解决方案对此类事故单进行解决。如以下将结合图3和图4更加详细描述,这可通过首先从索引编制于SK库(即所述知识库)的多个现有解决方案中识别出所述现有解决方案,然后调用与该现有解决方案相关联的一个或多个脚本对该事故单进行解决的方式实现。解决之后,在步骤220中,通过所述事故单以及所述被识别及执行的现有解决方案,更新所述ITSM或该ITSM内的事故库。

[0075] 现在参考图3,事故分类及解决引擎200存有在SK库中按各种正机械化事故单的所有可用机械化解方案索引编制的知识表示符300。如图所示,每个表示正机械化事故单的用例301均可与不同描述302相关联。此外,每个用例301可具有一个或多个作为解决方案的关联脚本303。用例301的表示符(例如向量)与所述事故单的表示符相类似。举例而言,用例301可以为CPU使用率、存储空间使用率、数据库备份等。或者,用例301可以为CPU使用率、存储空间使用率、数据库备份等的唯一数字表示符。由此可见,用例301为具有不同描述但属于同一大类的不同事故单类别表示符。与同一用例301相关的所有描述302均合并为一个语料库。例如,CPU使用率用例或类别表示符可对应于“CPU UTILIZATION has exceeded the threshold value on Windows Server”(“CPU使用率已超出Windows Server的阈值”)、“CPU UTILIZATION has exceeded the threshold value on VMware VM Windows 2008Server Device”(“CPU使用率已超出Windows 2008Server Device的VMware虚拟机阈值”)、“CPU UTILIZATION has exceeded”(“CPU使用率已超出”)等描述。类似地,存储空间使用率用例或类别表示符可对应于“MEMORY UTILIZATION has exceeded the threshold value on Windows Server”(“存储空间使用率已超出Windows Server的阈值”)、“MEMORY UTILIZATION has exceeded the threshold value on Windows 2008Server Device”(“CPU使用率已超出Windows 2008Server Device的阈值”)、“CPU UTILIZATION, MEMORY UTILIZATION, WinWPX has exceeded the threshold value on VMware VM Windows

2003Server Device”(“CPU使用率、存储空间使用率、WinWPX已超出Windows 2008Server Device的VMware虚拟机阈值”)等描述。类似地,数据库备份用例或类别表示符可对应于“OVO PS Description:BACKUP failed to complete the command BACKUP DATABASE”(“OVOPS描述:备份功能未能完成数据库备份命令”)、“MS SQL SERVER Description:BACKUP failed to complete the command BACKUP DATABASE”(“MS SQL SERVER描述:备份功能未能完成数据库备份命令”)等描述。因此,与用于识别所述现有解决方案的查询向量相匹配的正是用例301。一旦识别出所述用例,则与该用例相关联的脚本即被调用,以对相应事故单进行解决。

[0076] 现在参考图4,事故分类及解决引擎200采用K维树(或称KD树)400为图3中的知识表示符300编制索引。KD树400为一种多维多层次的索引结构,因此可支持k维数据点。在一些实施方式中,初步分析器201可采用每个所述大类中事故单的所有关键词的众数值(例如,两个或三个最常见关键词)作为用于为所述用例编制索引的键值。然而,应该注意的是,也可选择关键词的任何其他值(例如,中数值或值域值)作为表示用例的键值。在所示图中,KD树400为二维树(即 $k=2$),其中,使用两个众数值作为用于为所述用例(例如,CPU使用率、存储空间使用率、数据库备份等)编制索引的键值。此外,所述唯一数字表示符可用于表示所述键值。如此,举例而言,由于“CPU”的唯一数字表示符为-9514,“memory”为-246083027,“utilization”为-4652511977448,因此用例“CPU utilization”的键值为(-9514,-4652511977448),“memory utilization”的键值为(-246083027,-4652511977448)。在此之后,便可通过插入法创建KD树400。其后,可将KD树400平衡化,以实现用例及其关联解决方案的快速检索。应该注意的是,KD树400的维数以及键值选择可由用户自定义且取决于表示用例所需的单词数。举例而言,在图示的二维KD树400中,所有用例均由两个单词表示。然而,在一些实施例中,也可用三个单词或 n 个单词表示所述用例,而相应的KD树也即为三维或 n 维KD树。应该注意的是,在一些实施方式中,所有用例应该由相同个数的关键词表示。或者,在一些实施例中,也可使用不同个数关键词表示用例,但是为了实现通过KD树400编制用例索引,还可通过使用缺省单词(例如“00000”)将用例单词个数均衡化。

[0077] 本领域技术人员可以理解的是,多种方法可用于分类和解决软件生产事故单。例如,例示系统100以及与其关联的事故分类及解决引擎200可通过本文所述方法实现软件生产事故单的分类和解决。具体而言,正如本领域技术人员可以理解的,用于实施本文所述技术和步骤的控制逻辑和/或自动化程序可由系统100以及与其关联的事故分类及解决引擎200通过硬件、软件或硬件及软件的组合实现。例如,系统100内的一个或多个处理器可对合适的代码进行访问和执行,以实现本文所述的部分或所有技术。类似地,系统100内的所述一个或多个处理器内还可纳入用于实施本文所述的部分或所有方法的专用集成电路(ASIC)。

[0078] 举例而言,现在参考图5,该图通过根据本发明的一些实施方式的流程图展示了利用系统100等系统分类和解决软件生产事故单的例示控制逻辑500。如该流程图所示,控制逻辑500包括如下步骤:在步骤501中,接收事故单;在步骤502中,从所述事故单中提取多个关键词;在步骤503中,根据所述多个关键词推导出与所述事故单对应的查询向量。控制逻辑500还包括如下步骤:在步骤504中,通过将所述查询向量与从多个以往事故单中推导出的多个向量相比较,将所述事故单分类为正机械事故单和负机械事故单中的至少一

个。所述多个向量根据多个关键词及其在所述多个以往事故单中的相应出现次数推导得出。

[0079] 在一些实施方式中,控制逻辑500还包括为所述多个关键词中的每个关键词确定求得唯一数字表示符的步骤。此外,在一些实施方式中,控制逻辑500还包括,通过根据机械化状态将事故库中的所述多个以往事故单分类至至少两个类别内、从每个类别的多个以往事故单中提取多个关键词、以及根据所述多个关键词及其在每个类别的多个以往事故单中的相应出现次数推导出所述多个向量,从所述多个以往事故单推导出所述多个向量的步骤。在一些实施方式中,推导出所述多个向量包括,推导出所有关键词、名词关键词和动词关键词中的至少一个的多个向量。此外,在一些实施方式中,推导出所述多个向量包括,推导出所述多个关键词的众数、中数和值域中的至少一个的多个向量。此外,在一些实施方式中,推导出所述多个向量包括,将所述多个向量中的每个向量归一化。

[0080] 在一些实施方式中,在步骤502中提取所述多个关键词包括,预处理所述事故单,以提取词干或去除至少一个停止词(例如,标点符号、数字等)。此外,在一些实施方式中,步骤504中的分类包括,根据所述查询向量和所述多个向量之间的相似度和相异度中的至少一个,实施最邻近分类。在一些实施方式中,所述相似度包括余弦相似度,所述相异度包括欧几里得距离,实施所述最邻近分类包括根据该余弦相似度和欧几里得距离求得决定参数。此外,在一些实施方式中,步骤504中的分类包括,根据所述多个向量推导出一个或多个特性矩阵,以及将所述查询向量与该一个或多个特性矩阵相比较。

[0081] 在一些实施方式中,控制逻辑500还包括从按所述事故单索引编制于知识库内的多个现有解决方案中识别出所述正机械化事故单的现有解决方案并利用该现有解决方案解决所述事故单,以及调用与所述现有解决方案相关联的一个或多个脚本的步骤。此外,在一些实施方式中,控制逻辑500还包括,通过所述正机械化事故单以及该事故单的现有解决方案更新所述事故库。

[0082] 现在参考图6,该图通过根据本发明的一些实施方式的流程图,更加详细地展示了用于分类和解决软件生产事故单的例示控制逻辑600。如该流程图所示,控制逻辑600包括如下步骤:在步骤601中,从所述ITSM获取以往事故单;在步骤602中,根据所述以往事故单的相应机械化状态将其分类至至少两个类别中。控制逻辑600还包括如下步骤:在步骤603中,预处理每个类别的事故单;在步骤604中,从预处理后的事故单中提取关键词;在步骤605中,为每个类别的多个关键词中的每个关键词求得唯一数字表示符。控制逻辑600还包括如下步骤:在步骤606中,根据所述多个关键词及其在每个类别的多个以往事故单中的相应出现次数,推导出多个向量。其中,可针对关键词的不同训练语料库(例如所有关键词训练语料库、名词关键词训练语料库、动词关键词训练语料库等)或者出现次数的不同类型(如众数、中数、值域等),推导所述向量。在一些实施方式中,可根据每个类别的多个向量推导出特性矩阵。控制逻辑600还包括如下步骤:在步骤607中,将所述多个向量和/或特性矩阵中的每个向量和/或特性矩阵归一化。

[0083] 此外,控制逻辑600包括如下步骤:在步骤608中,从所述事件关联引擎或所述ITSM接收事故单;在步骤609中,对该事故单进行预处理;在步骤610中,从预处理后的事故单中提取关键词;在步骤611中,为所述多个关键词中的每个关键词求得唯一数字表示符;在步骤612中,根据所述多个关键词,推导出查询向量。控制逻辑600还包括如下步骤:在步骤613

中,通过将所述查询向量与所述多个向量或特性矩阵相比较,将所述事故单分类为正机械事故单和负机械事故单当中的至少一个。所述分类包括,通过根据余弦相似度和欧几里得距离求得决定参数而实施最邻近分类。

[0084] 控制逻辑600还包括如下步骤:在步骤614中,判断所述事故单是否为正机械事故单。如果所述事故单为正机械事故单,控制逻辑600则包括如下步骤:在步骤615中,使用现有解决方案解决所述事故单。在一些实施方式中,步骤615中的解决可包括:在步骤616中,从索引编制于知识库(即SK库)内的多个现有解决方案中识别出上述现有解决方案;在步骤617中,调用与上述现有解决方案相关联的一个或多个脚本。控制逻辑600还包括如下步骤:在步骤618中,通过所述事故单及其关联解决方案(即上述现有解决方案)更新ITSM内的所述事故库。在后续中,所述ITSM可实施上述提供的所述解决方案。然而,如果所述事故单不是正机械事故单,控制逻辑600直接进至步骤618,并通过该事故单更新ITSM内的所述事故库。在此情况下,该事故单通过手动方式解决。

[0085] 还应理解的是,上述技术可采用如下形式:计算机或控制器实现的方法;以及用于实施这些方法的装置。本发明还可以以含有指令的计算机程序代码的形式实施,所述指令包含于软盘、CD-ROM、硬盘驱动器或其他任何计算机可读存储介质等有形介质中,其中,当所述计算机程序代码载入计算机或控制器内并由该计算机或控制器执行时,所述计算机即成为一种用于实施本发明的装置。本发明还可以以计算机程序代码或信号的形式实施,所述计算机程序代码或信号例如存储于存储介质中,或者载入计算机或控制器内并由该计算机或控制器执行,或者经电线或电缆、光纤或电磁辐射等传输介质传输,其中,当所述计算机程序代码载入计算机内并由该计算机执行时,所述计算机即成为一种用于实施本发明的装置。当在通用微处理器中实施时,所述计算机程序代码的代码段对所述微处理器进行配置,以创建出特定的逻辑电路。

[0086] 上述公开的方法和系统可在个人计算机(PC)或服务器计算机等常规或通用计算机系统内实施。现在参考图7,该图所示为用于实施符合本发明实施方式的例示计算机系统701的框图。计算机系统701的各种变形可用于实现用于分类和解决软件生产事故单的系统100和事故分类及解决引擎200。计算机系统701可包括中央处理单元(“CPU”或“处理器”)702。处理器702可包括至少一个用于执行程序组件的数据处理器,所述程序组件用于执行用户或系统生成的请求。用户可包括个人,使用设备(例如,本发明范围内的设备)的个人或此类设备本身。所述处理器可包括集成系统(总线)控制器、存储器管理控制单元、浮点单元、图形处理单元、数字信号处理单元等专用处理单元。所述处理器可包括AMD速龙(Athlon)、毒龙(Duron)或皓龙(Opteron)、ARM应用处理器、嵌入式或安全处理器、IBM PowerPC、Intel Core、安腾(Itanium)、至强(Xeon)、赛扬(Celeron)或其他处理器产品线等微处理器。处理器702可通过主机、分布式处理器、多核、并行、网格或其他架构实现。一些实施方式可采用专用集成电路(ASIC)、数字信号处理器(DSP)、现场可编程门阵列(FPGA)等嵌入式技术。

[0087] 处理器702可配置为通过输入/输出(I/O)接口703与一个或多个I/O设备进行通信。I/O接口703可采用通信协议/方法,例如但不限于,音频、模拟、数字、单声道、RCA、立体声、IEEE-1394、串行总线、通用串行总线(USB)、红外、PS/2、BNC、同轴、组件、复合、数字视觉接口(DVI)、高清晰度多媒体接口(HDMI)、射频天线、S-视频、VGA、IEEE 802.n/b/g/n/x、蓝

牙、蜂窝(例如码分多址(CDMA)、高速分组接入(HSPA+)、移动通信全球系统(GSM)、长期演进(LTE)、WiMax等)等。

[0088] 通过使用I/O接口703,计算机系统701可与一个或多个I/O设备进行通信。举例而言,输入设备704可以为天线、键盘、鼠标、操纵杆、(红外)遥控、摄像头、读卡器、传真机、加密狗、生物计量阅读器、麦克风、触摸屏、触摸板、轨迹球、传感器(例如加速度计、光传感器、GPS、陀螺仪、接近传感器等)、触控笔、扫描仪、存储设备、收发器、视频设备/视频源、头戴式显示器等。输出设备705可以为打印机、传真机、视频显示器(例如阴极射线管(CRT)、液晶显示器(LCD)、发光二极管(LED)、等离子等)、音频扬声器等。在一些实施方式中,收发器706可与处理器702连接。所述收发器可促进各类无线传输或接收。例如,所述收发器可包括以可操作方式连接至收发器芯片(例如德州仪器(Texas Instruments)WiLink WL1283、博通(Broadcom)BCM4750IUB8、英飞凌科技(Infineon Technologies)X-Gold 618-PMB9800等)的天线,以实现IEEE 802.11a/b/g/n、蓝牙、频率调制(FM)、全球定位系统(GPS)、2G/3G HSDPA/HSUPA通信等。

[0089] 在一些实施方式中,处理器702可配置为通过网络接口707与通信网络708通信。网络接口707可与通信网络708通信。所述网络接口可采用连接协议,包括但不限于,直接连接、以太网(例如双绞线10/100/1000BaseT)、传输控制协议/网际协议(TCP/IP)、令牌环、IEEE 802.11a/b/g/n/x等。通信网络708可包括,但不限于,直接互连、局域网(LAN)、广域网(WAN)、无线网络(例如使用无线应用协议)、因特网等。通过网络接口707和通信网络708,计算机系统701可与设备709、710和711通信。这些设备可包括,但不限于,个人计算机,服务器,传真机,打印机,扫描仪,以及蜂窝电话、智能电话(例如苹果(Apple)iPhone、黑莓(Blackberry)、基于安卓(Android)系统的电话等)、平板电脑、电子书阅读器(亚马逊(Amazon)Kindle,Nook等)、膝上型计算机、笔记本电脑、游戏机(微软(Microsoft)Xbox、任天堂(Nintendo)DS,索尼(Sony)PlayStation等)等各种移动设备。在一些实施方式中,计算机系统701可本身包含一个或多个上述设备。

[0090] 在一些实施方式中,处理器702可配置为通过存储接口712与一个或多个存储设备(例如RAM 713、ROM 714等)通信。所述存储接口可采用串行高级技术连接(SATA)、集成驱动电子设备(IDE)、IEEE 1394、通用串行总线(USB)、光纤通道、小型计算机系统接口(SCSI)等连接协议连接至存储设备,该存储设备包括,但不限于,存储驱动器、可移除磁盘驱动器等。所述存储驱动器还可包括磁鼓、磁盘驱动器、磁光驱动器、光盘驱动器、独立磁盘冗余阵列(RAID)、固态存储设备、固态驱动器等。

[0091] 所述存储设备可存储一系列程序或数据库组件,包括但不限于,操作系统716、用户界面应用程序717、网页浏览器718、邮件服务器719、邮件客户端720、用户/应用程序数据721(例如本发明中所述的任何数据变量或数据记录)等。操作系统716可促进计算机系统701的资源管理和运行。操作系统例如包括,但不限于,苹果Macintosh OS X、Unix、类Unix系统套件(例如伯克利软件套件(BSD)、FreeBSD、NetBSD、OpenBSD等)、Linux套件(如红帽(Red Hat)、Ubuntu、Kubuntu等)、IBM OS/2、微软Windows(XP,Vista/7/8等)、苹果iOS、谷歌(Google)安卓、黑莓操作系统等。用户界面717可利用文本或图形工具促进程序组件的显示、执行、互动、操控或操作。例如,用户界面可在以可操作方式连接至计算机系统701的显示系统上提供光标、图标、复选框、菜单、滚动条、窗口、窗口部件等计算机交互界面元件。此

外,还可采用图形用户界面(GUI),包括但不限于,苹果Macintosh操作系统的Aqua、IBM OS/2、微软Windows(例如Aero、Metro等)、Unix X-Windows、网页界面库(例如ActiveX、Java、Javascript、AJAX、HTML、Adobe Flash等)等。

[0092] 在一些实施方式中,计算机系统701可执行网页浏览器718存储的程序组件。所述网页浏览器可以为超文本浏览应用程序,如微软网络探路者(Internet Explorer)、谷歌浏览器(Chrome)、谋智火狐(MozillaFirefox)、苹果浏览器(Safari)等。安全网页浏览可通过HTTPS(安全超文本传输协议)、安全套接字层(SSL)、安全传输层(TLS)等实现。网页浏览器可使用AJAX、DHTML、Adobe Flash、JavaScript、Java、应用程序编程接口(API)等工具。在一些实施方式中,计算机系统701可执行邮件服务器719存储的程序组件。所述邮件服务器可以为微软Exchange等因特网邮件服务器。所述邮件服务器可使用ASP、ActiveX、ANSI C++/C#、微软.NET、CGI脚本、Java、JavaScript、PERL、PHP、Python、WebObjects等工具。所述邮件服务器还可使用因特网信息访问协议(IMAP)、邮件应用程序编程接口(MAPI)、微软Exchange、邮局协议(POP)、简单邮件传输协议(SMTP)等通信协议。在一些实施方式中,计算机系统701可执行邮件客户端720存储的程序组件。所述邮件客户端可为苹果Mail、微软Entourage、微软Outlook、谋智Thunderbird等邮件查看程序。

[0093] 在一些实施方式中,计算机系统701可存储用户/应用程序数据721,例如本发明中所述数据、变量、记录等(例如,事故单,关键词,关键词的唯一数字表示符,向量,特性矩阵,机械化状态,用例,描述,脚本,配置项,KD树等)。此类数据库可以为容错、关系、可扩展、安全数据库,例如甲骨文(Oracle)或赛贝斯(Sybase)。或者,上述数据库可通过数组、散列、链表、结构、结构化文本文件(例如XML)、表格等标准化数据结构实现,或者实施为面向对象的数据库(例如通过ObjectStore、Poet、Zope等)。上述数据库可以为合并或分布数据库,有时分布于本发明所讨论的上述各种计算机系统之间。应该理解的是,上述任何计算机或数据库组件的结构及操作可以任何可工作的组合形式组合、合并或分布。

[0094] 本领域技术人员可以理解的是,上述各实施方式中描述的技术可基于智能分类实现事故单的自动、有效、快速解决。所述智能分类实现了自动化候选事故单的识别,从而降低了提供准确解决方案时的人为错误和时间延误。换句话说,上述各实施方式中描述的技术在提高自动化解决方案生产能力的同时,减少了手工工作量。

[0095] 本说明书已对用于分类和解决软件生产事故单的系统和方法进行了描述。所示步骤用于说明所述例示实施方式,并且应当预想到的是,随着技术的不断发展,特定功能的执行方式也将发生改变。本文所呈现的上述实施例用于说明而非限制目的。此外,为了描述的方便性,本文对各功能构建模块边界的定义为任意性的,只要其上述功能及其关系能够获得适当执行,也可按其他方式定义边界。根据本申请的启示内容,替代方案(包括本申请所述解的等同方案、扩展方案、变形方案、偏差方案等)对于相关领域技术人员是容易理解的。这些替代方案均落入所公开实施方式的范围和精神内。

[0096] 此外,一个或多个计算机可读存储介质可用于实施本发明的实施方式。计算机可读存储介质是指可对处理器可读取的信息或数据进行存储的任何类型的物理存储器。因此,计算机可读存储介质可对由一个或多个处理器执行的指令进行存储,包括用于使处理器执行根据本申请实施方式的步骤或阶段的指令。“计算机可读介质”一词应理解为包括有形物件且不包括载波及瞬态信号,即为非临时性介质,例如随机存取存储器(RAM)、只读存

储器(ROM)、易失性存储器、非易失性存储器、硬盘驱动器、只读光盘存储器(CD-ROM)、DVD、闪存驱动器、磁盘以及其他任何已知物理存储介质。

[0097] 以上发明及实施例旨在于仅视为示例性内容及实施例,所公开实施方式的真正范围和精神由以下权利要求指出。

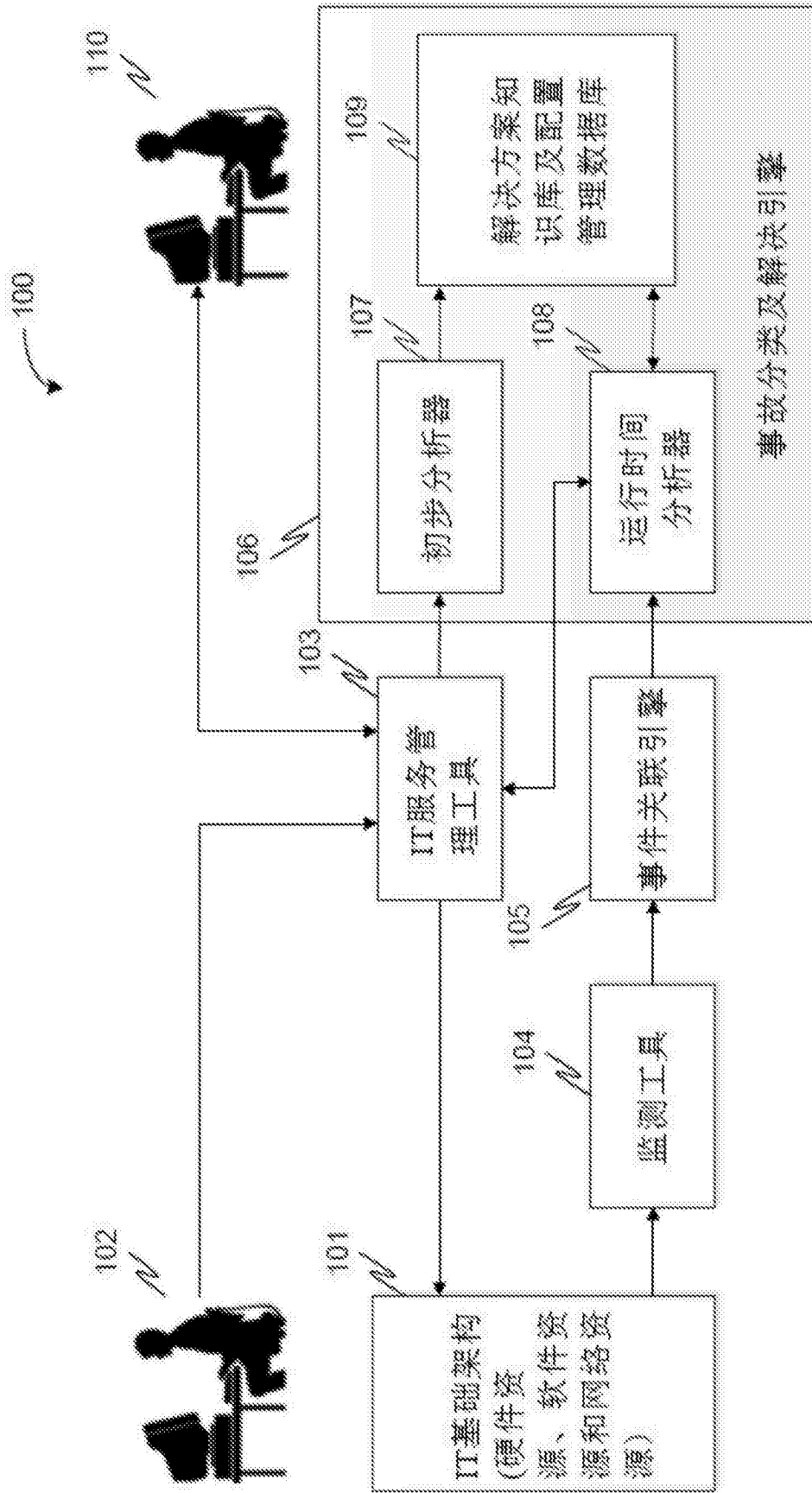


图1

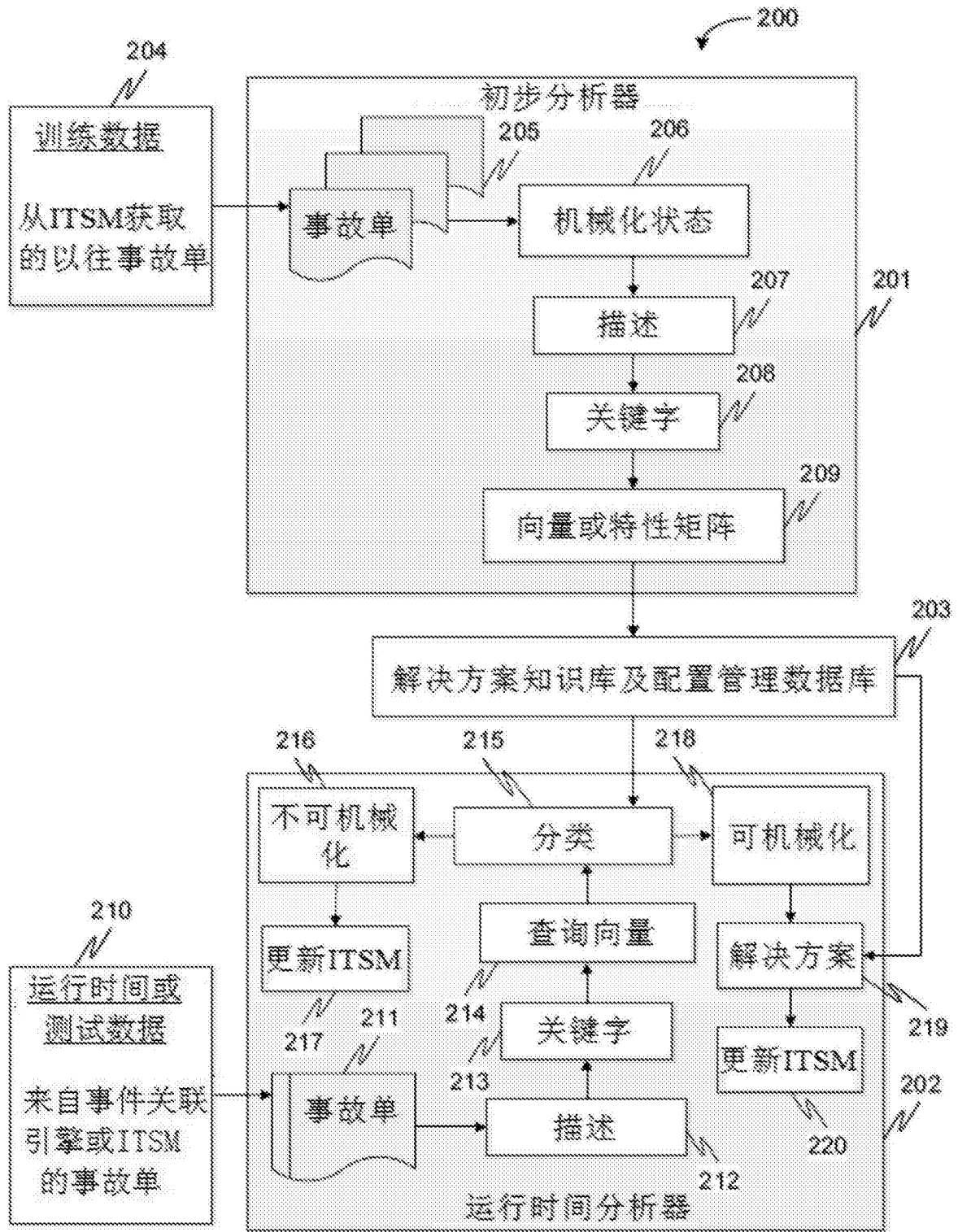


图2

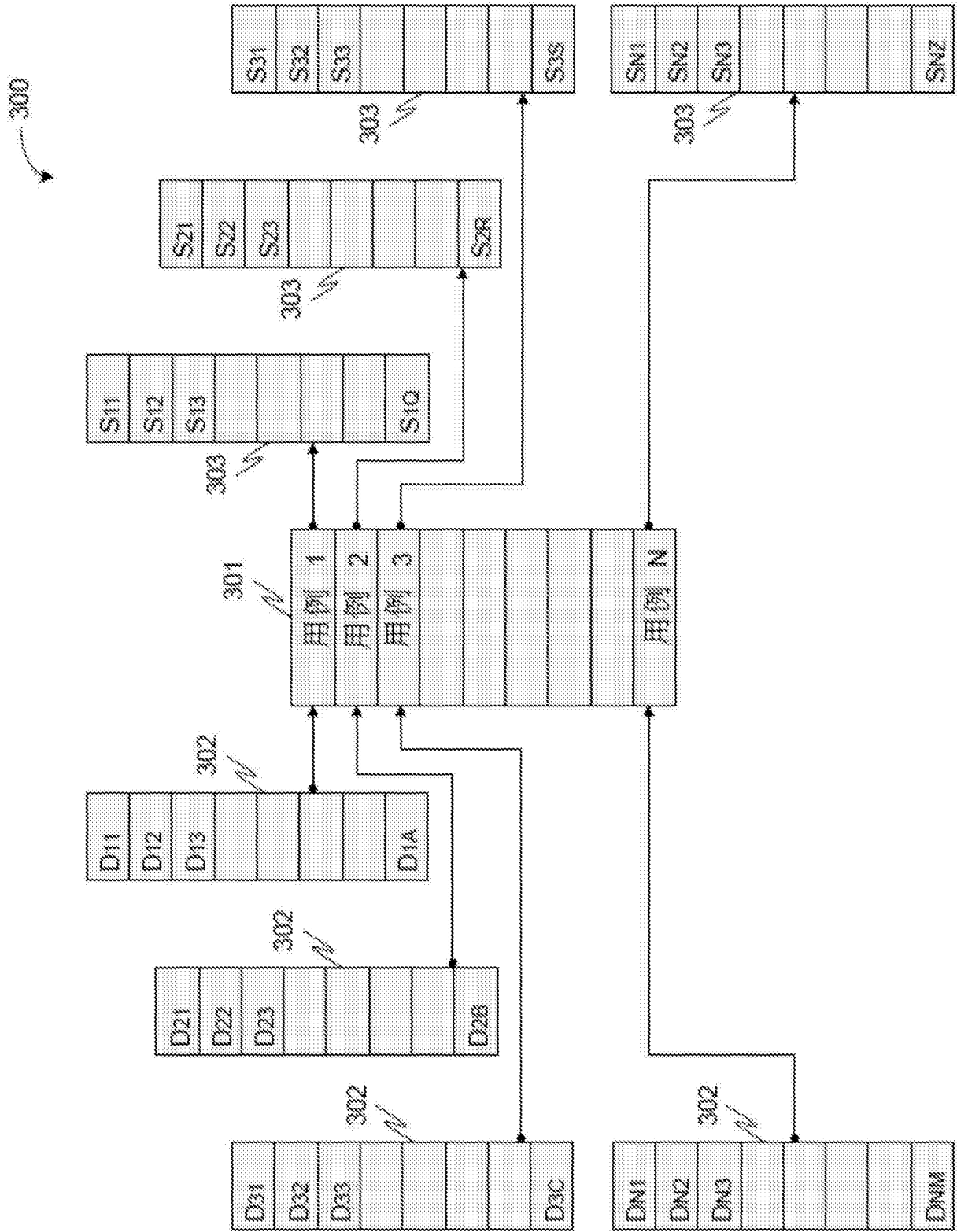


图3

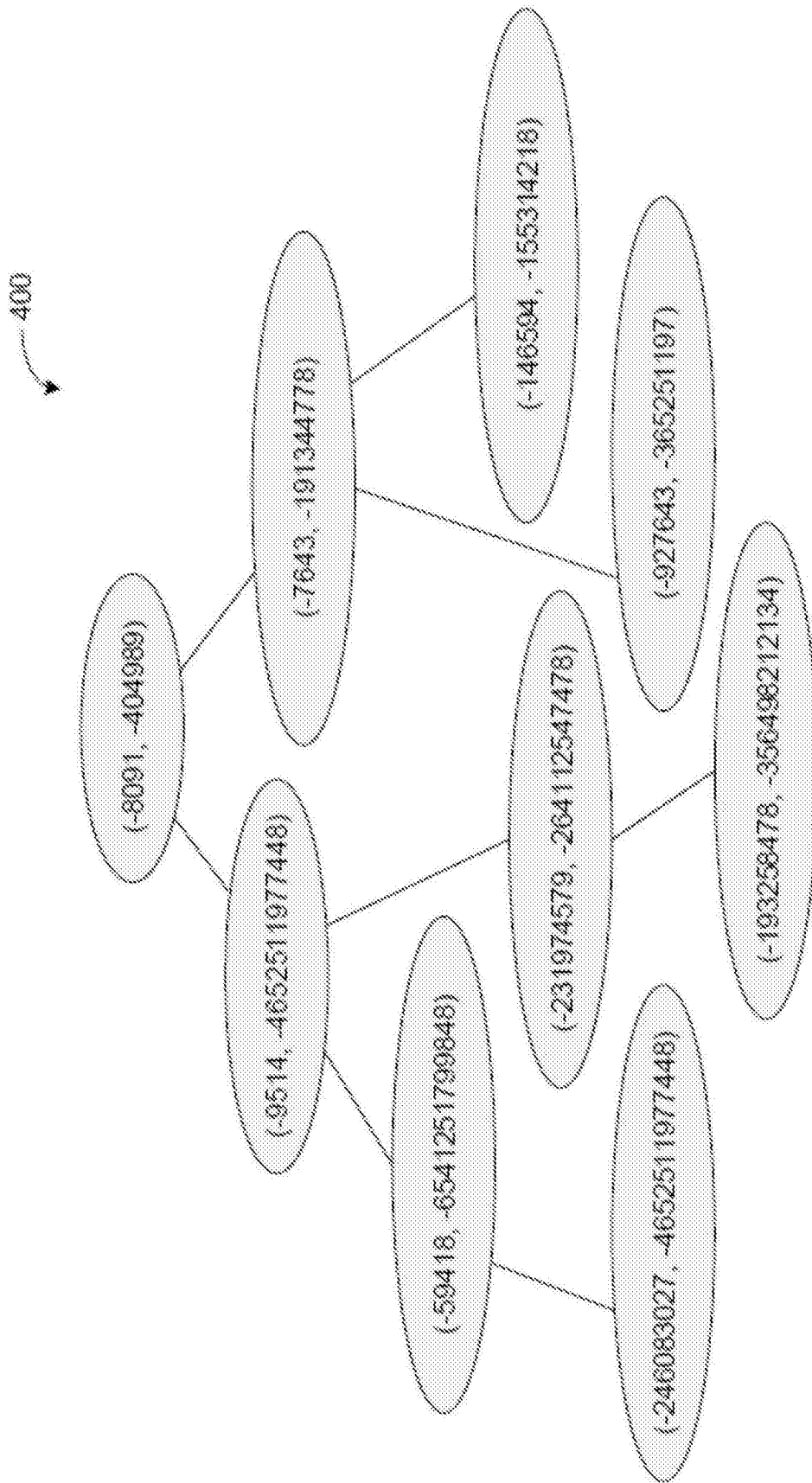


图4

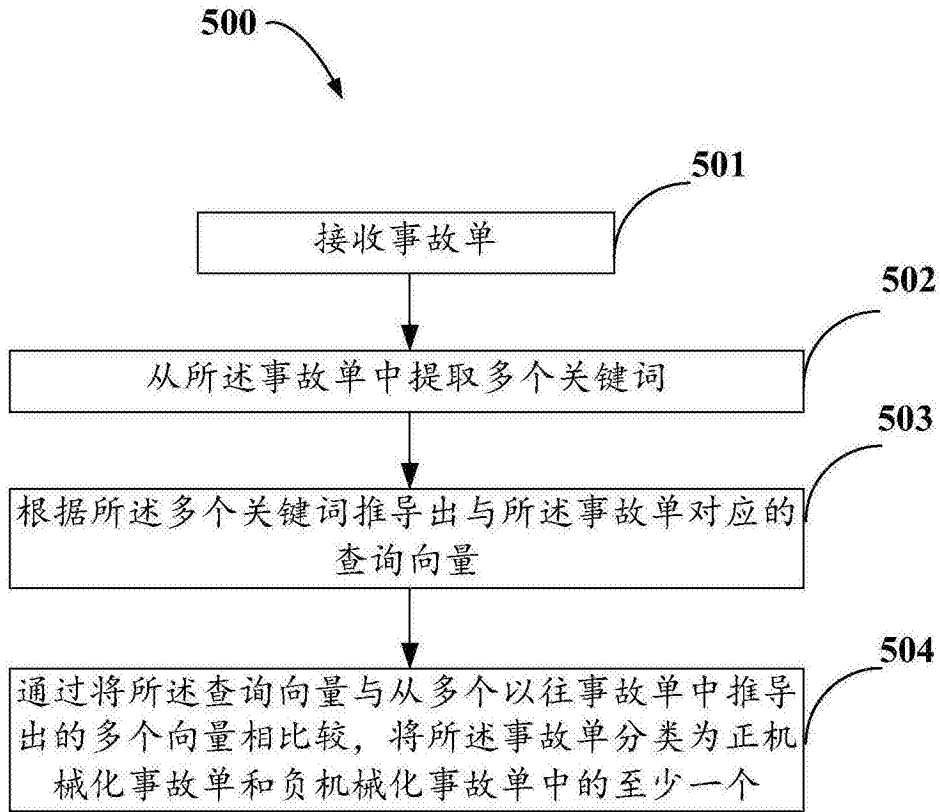


图5

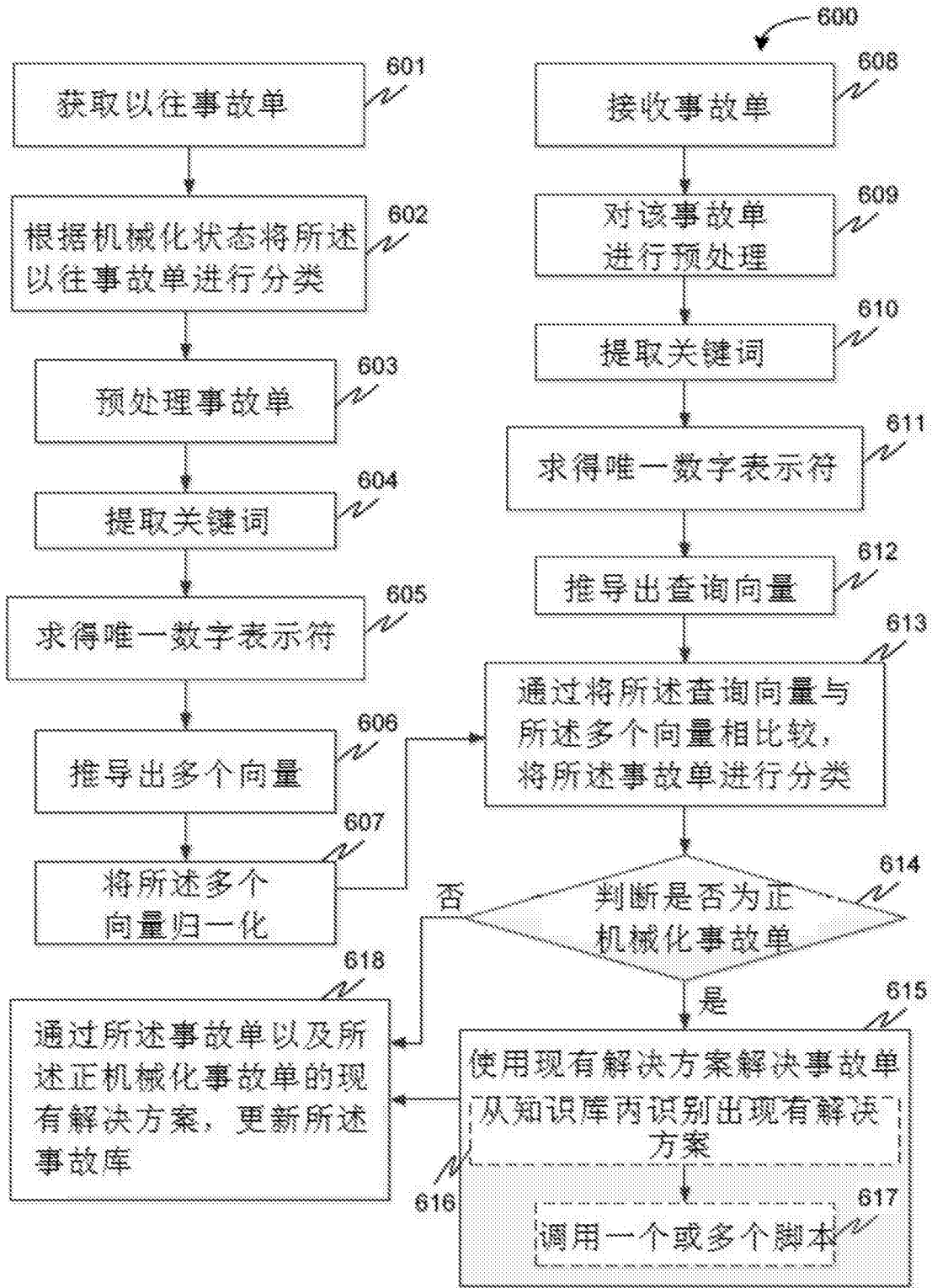


图6

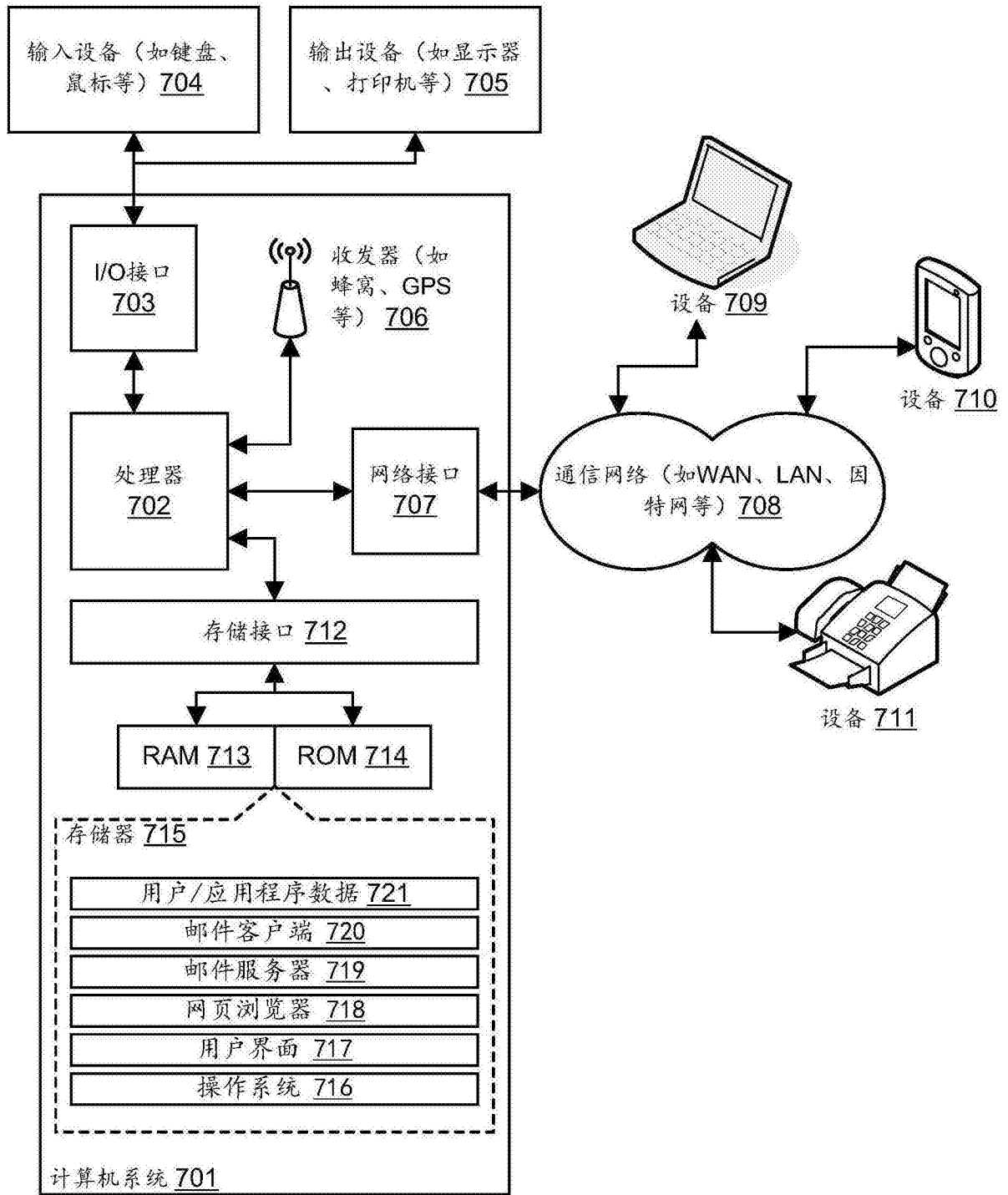


图7