

(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 12 月 10 日 (10.12.2020)



(10) 国际公布号
WO 2020/244336 A1

(51) 国际专利分类号:
G06F 16/35 (2019.01)

(21) 国际申请号: PCT/CN2020/086878

(22) 国际申请日: 2020 年 4 月 24 日 (24.04.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910482036.0 2019年6月4日 (04.06.2019) CN

(71) 申请人: 深圳前海微众银行股份有限公司
(**WEBANK CO., LTD.**) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路1号A栋201室, Guangdong 518027 (CN)。

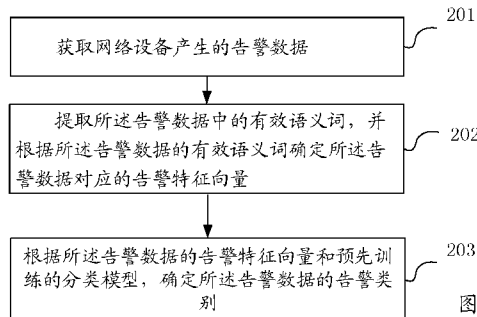
(72) 发明人: 黄楷(**HUANG, Kai**); 中国广东省深圳市前海深港合作区前湾一路1号A栋201室, Guangdong 518027 (CN)。 杨俊杰(**YANG, Junjie**); 中国广东省深圳市前海深港合作区前湾一路1号A栋201室, Guangdong 518027 (CN)。 卢道和(**LU, Daohe**); 中国广东省深圳市前海深港合作区前湾一路1号A栋201室, Guangdong 518027 (CN)。

(74) 代理人: 北京同达信恒知识产权代理有限公司
(**TDIP & PARTNERS**); 中国北京市西城区裕民路18号北环中心A座2002, Beijing 100029 (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,

(54) **Title:** ALARM CLASSIFICATION METHOD AND DEVICE, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 一种告警分类方法、装置、电子设备及存储介质



201 OBTAINING ALARM DATA GENERATED BY A NETWORK DEVICE
202 EXTRACTING EFFECTIVE SEMANTIC WORDS IN THE ALARM DATA, AND DETERMINING AN ALARM FEATURE VECTOR CORRESPONDING TO THE ALARM DATA ACCORDING TO THE VALID SEMANTIC WORDS OF THE ALARM DATA
203 DETERMINING AN ALARM CATEGORY OF THE ALARM DATA ACCORDING TO THE ALARM FEATURE VECTOR OF THE ALARM DATA AND A PRE-TRAINED CLASSIFICATION MODEL

图 2

(57) **Abstract:** An alarm classification method and device, an electronic device, and a storage medium, wherein the method comprises: obtaining alarm data generated by a network device (201); extracting effective semantic words in the alarm data, and determining an alarm feature vector corresponding to the alarm data according to the valid semantic words of the alarm data (202); wherein, each valid semantic word corresponds to one feature element in the alarm feature vector of the alarm data; and determining an alarm category of the alarm data according to the alarm feature vector of the alarm data and a pre-trained classification model (203); wherein the classification model is determined according to an estimated probability of historical alarm data under each alarm category and the alarm category corresponding to the historical alarm data, and the estimated probability of the historical alarm data under each alarm category is determined according to the conditional probability of each feature element of the historical alarm data under each alarm category.

(57) **摘要:** 一种告警分类方法、装置、电子设备及存储介质, 该方法包括: 获取网络设备产生的告警数据 (201); 提取所述告警数据中的有效语义词, 并根据所述告警数据的有效语义词确定所述告警数据对应的告警特征向量 (202); 其中, 每个有效语义词对应所述告警数据的告警特征向量中的一个特征元素; 根据所述告警数据的告警特征向量和预先训练的分类模型, 确定所述告警数据的告警类别 (203); 所述分类模型为根据历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别确定的, 所述历史告警数据在各告警类别下的预估概率是根据所述历史告警数据的每个特征元素在每个告警类别的条件概率确定的。

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,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, 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))。

一种告警分类方法、装置、电子设备及存储介质

相关申请的交叉引用

本申请要求在2019年06月04日提交中国专利局、申请号为201910482036.0、申请名称为“一种告警分类方法、装置、电子设备及存储介质”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。

技术领域

本发明涉及金融科技（Fintech）技术领域，尤其涉及一种告警分类方法、装置、电子设备及存储介质。

背景技术

随着计算机技术的发展，越来越多的技术应用在金融领域，传统金融业正在逐步向金融科技（Fintech）转变，告警分类技术也不例外，但由于金融行业的安全性、实时性要求，也对技术提出的更高的要求。

对金融行业的网络运维而言，现有的技术方案都是通过运维人员配置告警关键字来实现对告警信息的分类。

每一个分类都有对应的关键字。严格地利用关键字去匹配告警信息。多 IDC 的大型金融公司，为了规避风险，经常需要多厂商的网络设备，设备上架之后，由于多厂商带来的复杂性，具体的告警分类需要由以下几部组成，先进行设备的上架操作，检查新设备是否是新的品牌型号，检查品牌型号的告警关键字是否发生了变化，最后按需配置告警关键字。

因此，告警分类需要投入运维开发的人力，告警的配置系统需要运维开发的人力长期维护与更新。新的厂商新的型号新的软件版本甚至更换了网络管理系统的供应商都可能都需要重新配置，极大的浪费了资源。

发明内容

本发明实施例提供一种告警分类方法、装置、电子设备及存储介质，以解决现有技术中内容分类准确、比较复杂的问题。

本发明实施例提供的具体技术方案如下：

一方面，本发明实施例提供了一种告警分类方法，包括：

获取网络设备产生的告警数据；

提取所述告警数据中的有效语义词，并根据所述告警数据的有效语义词确定所述告警数据对应的告警特征向量；其中，每个有效语义词对应所述告警数据的告警特征向量中的一个特征元素；

根据所述告警数据的告警特征向量和预先训练的分类模型，确定所述告警数据的告警类别；所述分类模型为根据历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别确定的；所述历史告警数据在各告警类别下的预估概率是根据所述历史告警数据的每个特征元素在每个告警类别的条件概率确定的。

一种可能的实现方式，所述分类模型为根据历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别确定的，包括：

获取所述网络设备的历史告警数据集；所述历史告警数据集包括多个历史告警数据及各历史告警数据对应的告警类别；

针对每个历史告警数据，提取所述历史告警数据中的有效语义词，将所述有效语义词作为所述历史告警数据的告警特征向量的特征元素；

针对同一告警类别，确定每个特征元素在所述告警类别的条件概率；

针对每个历史告警数据，根据所述历史告警数据的告警特征向量中的每个特征元素在每个告警类别下的条件概率确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率；

根据各历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别进行所述分类模型的训练。

一种可能的实现方式，所述根据各历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别进行所述分类模型的训练，包括：

根据各历史告警数据在各告警类别下的预估概率，确定各历史告警数据的预测分类；

根据所述预测分类及所述各历史告警数据对应的告警类别，调整所述分类模型的参数，以使所述预测分类的置信度大于预设阈值。

一种可能的实现方式，所述针对每个历史告警数据，根据所述历史告警数据的告警特征向量中的每个特征元素在每个告警类别下的条件概率确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率，为根据以下公式确定：

$$P(x|y_i)P(y_i) = P(a_1|y_i) P(a_2|y_i) \dots P(a_m|y_i)P(y_i) = P(y_i) \prod_{j=1}^m P(a_j|y_i)$$

其中， $x = \{a_1, a_2, \dots, a_m\}$ 为所述历史告警数据的告警特征向量， a_m 为 x 告警特征向量中的特征元素， y_i 为第 i 个告警类别； $P(x|y_i)$ 为告警类别为 y_i 出现告警特征向量 x 的概率； $P(y_i)$ 为第 i 个告警类别在历史告警数据集中的发生概率； $P(a_j|y_i)$ 为第 a_j 个特征元素在第 i 个告警类别的条件概率； $\prod_{j=1}^m P(a_j|y_i)$ 为告警类别为 y_i 出现每个特征元素的条件概率的乘积。

一种可能的实现方式，所述方法还包括：

若确定所述历史告警数据的告警特征向量中的第一特征元素在第一告警类别下的条件概率为零，则将所述历史告警数据的告警特征向量中的所述第一特征元素在所述第一告警类别的条件概率增加预定偏置项；所述第一告警类别为所述告警类别中的任一告警类别；所述第一特征元素为所述历史告警数据的告警特征向量中的任一特征元素。

一方面，本发明实施例提供了一种告警分类装置，包括：

收发单元，用于获取网络设备产生的告警数据；

处理单元，用于提取所述告警数据中的有效语义词，并根据所述告警数据的有效语义词确定所述告警数据对应的告警特征向量；其中，每个有效语义词对应所述告警数据的告警特征向量中的一个特征元素；根据所述告警数据的告警特征向量和预先训练的分类模型，确定所述告警数据的告警类别；所述分类模型为根据历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别确定的；所述历史告警数据在各告警类别下的预估概率是根据所述历史告警数据的每个特征元素在每个告警类别的条件概率确定的。

一种可能的实现方式，所述收发单元，还用于：

获取所述网络设备的历史告警数据集；所述历史告警数据集包括多个历史告警数据及各历史告警数据对应的告警类别；

所述处理单元，还用于：针对每个历史告警数据，提取所述历史告警数据中的有效语义词，将所述有效语义词作为所述历史告警数据的告警特征向量的特征元素；针对同一告警类别，确定每个特征元素在所述告警类别的条件概率；针对每个历史告警数据，根据所述历史告警数据的告警特征向量中的每个特征元素在每个告警类别下的条件概率确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率；根据各历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别进行所述分类模型的训练。

一种可能的实现方式，所述处理单元，具体用于：

根据各历史告警数据在各告警类别下的预估概率，确定各历史告警数据的预测分类；根据所述预测分类及所述各历史告警数据对应的告警类别，调整所述分类模型的参数，以使所述预测分类的置信度大于预设阈值。

一种可能的实现方式，所述处理单元，具体用于根据以下公式确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率：

$$P(x|y_i)P(y_i) = P(a_1|y_i) P(a_2|y_i) \dots P(a_m|y_i)P(y_i) = P(y_i) \prod_{j=1}^m P(a_j|y_i)$$

其中， $x = \{a_1, a_2, \dots, a_m\}$ 为所述历史告警数据的告警特征向量， a_m 为 x 告警特征向量中的特征元素， y_i 为第 i 个告警类别； $P(x|y_i)$ 为告警类别为 y_i 出现告警特征向量 x 的概率； $P(y_i)$ 为第 i 个告警类别在历史告警数据集中的发生概率； $P(a_j|y_i)$ 为第 a_j 个特征元素在第 i 个告警类别的条件概率； $\prod_{j=1}^m P(a_j|y_i)$ 为告警类别为 y_i 出现每个特征元素的条件概率的乘积。

一种可能的实现方式，所述处理单元，还用于：

若确定所述历史告警数据的告警特征向量中的第一特征元素在第一告警类别下的条件概率为零，则将所述历史告警数据的告警特征向量中的所述第一特征元素在所述第一告警类别的条件概率增加预定偏置项；所述第一告警类别为所述告警类别中的任一告警类别；所述第一特征元素为所述历史告警数据的告警特征向量中的任一特征元素。

本发明另一个实施例提供了一种电子设备，包括：

至少一个存储器，用于存储程序指令；

至少一个处理器，用于调用所述存储器中存储的程序指令，按照获得的程序指令执行上述任一种告警分类方法。

一方面，本发明实施例提供了一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现上述任一种告警分类方法。

一方面，本发明实施例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在计算机可读存储介质上的计算机程序，所述计算机程序包括程序指令，当所述程序指令被计算机执行时，使所述计算机执行上述任一种告警分类方法的步骤。

本发明实施例中，根据网络设备的告警数据，计算获得告警的告警特征向量，对于复杂内容类型的告警，可以更加准确地表示告警的特征，解决了复杂内容的告警的特征提取困难的问题，基于该告警特征向量和分类模型来确定告警的类别，可以很大程度上提高告警分类的准确性，并且，根据网络设备的告警数据确定出的告警特征向量，与告警类别相关，因此，在训练过程中也不需要预先知道非常多的训练样本，减少了人工预先对告警类

别进行标注的成本，降低了复杂性。

附图说明

图 1 为本发明实施例中告警分类方法的应用场景示意图；

图 2 为本发明实施例中告警分类方法的流程图；

图 3 为本发明实施例中一种告警分类方法的流程示意图；

图 4 为本发明实施例中告警分类装置结构示意图；

图 5 为本发明实施例中电子设备结构示意图。

具体实施方式

下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，并不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

为便于对本发明实施例的理解，下面先对几个概念进行简单介绍：

日志告警数据：网络设备按照时间顺序生成的日志中提取的告警数据。

贝叶斯分类器：朴素贝叶斯分类器是 NLP 常用的一种分类器。从数学角度来说，分类问题可做如下定义：

已知集合： $C' = \{y_1, y_2, \dots, y_n\}$ 和 $I = \{x_1, x_2, \dots, x_m, \dots\}$ ，确定映射规则 $y = f(x)$ ，使得任意 $x_i \in I$ 有且仅有一个 $y_j \in C'$ 使得 $y_j = f(x_i)$ 成立。

其中 C' 叫做类别集合，其中每一个元素是一个类别，而 I 叫做项集合，其中每一个元素是一个待分类项， f 叫做分类器。分类算法的任务就是构造分类器 f 。

贝叶斯分类器的基础是贝叶斯定理。这个定理解决的是已知某条件概率，如何得到两个事件交换后的概率，也就是在已知 $P(A|B)$ 的情况下如何求得 $P(B|A)$ 。具体包括： $P(A|B)$ 表示事件 B 已经发生的前提下，事件 A 发生的概率，叫做事件 B 发生下事件 A 的条件概率。其基本求解公式为：

$$P(A|B) = \frac{P(AB)}{P(B)}$$

由于可以很容易直接得出 $P(A|B)$ ， $P(B|A)$ 则很难直接得出，但通常更关心 $P(B|A)$ ，贝叶斯定理给出了从 $P(A|B)$ 得到 $P(B|A)$ 的方法。即

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

参考图 1 所示，其为本发明实施例中告警分类方法的应用场景示意图。包括网络设备和网管系统，其中，网络设备通过在网络运营过程中，例如，为用户设备提供服务的过程中生成对应的日志，网管系统可以通过网络设备的日志，提取日志中的告警数据，进而确定告警数据的分类，以便于对告警进行分类处理。其中，网络设备可以是任何能够提供互联网服务的设备；用户设备可以是智能手机、平板电脑、便携式个人计算机、智能电视等

任何智能设备，网管系统可以为关联网络设备的网管设备。

网络设备与网管系统之间通过互联网相连，实现相互之间的通信，可选地，上述互联网使用标准通信技术和/或协议。互联网通常为因特网、但也可以是任何网络，包括但不限于局域网（Local Area Network, LAN）、城域网（Metropolitan Area Network, MAN）、广域网（Wide Area Network, WAN）、移动、有线或者无线网络、专用网络或者虚拟专用网络的任何组合。

需要说明的是，本发明实施例中提供的告警分类方法可以由网管系统实施，也可以由网络设备中设置的告警分类装置实施，也可以由单独的告警分类装置实施，在此不做限定。

本发明各个实施例中，以告警分类方法用于图1所示的应用场景图为例进行示意性说明。值得说明的是，本发明实施例中的应用场景架构图是为了更加清楚地说明本发明实施例中的技术方案，并不构成对本发明实施例提供的技术方案的限制，对于其它的应用场景架构和业务应用，本发明实施例提供的技术方案对于类似的问题，同样适用。

对金融行业的网络运维而言，例如银行等金融机构的网络运维，现有的技术方案都是通过运维人员配置告警关键字来实现对告警信息的分类。每一个分类都有对应的关键字。严格地利用关键字去匹配告警信息。大型金融公司若设置了多个IDC，为了规避风险，需要多厂商的网络设备，设备上架之后，由于多厂商带来的复杂性，具体的告警分类需要先进行设备的上架操作，检查新设备是否是新的品牌型号，检查品牌型号的告警关键字是否发生了变化，最后按需配置告警关键字。这样，仅依赖预先设置的告警关键字，对告警的内容特征提取是远远不够的，很容易造成分类不准确，而目前对于这些不同网络设备生成的日志产生的不同类型的告警数据，需要分别设置关键字，告警分类需要投入运维开发的人力，告警的配置系统需要运维开发的人力长期维护与更新，新的厂商新的型号新的软件版本甚至更换了网络管理系统的供应商都可能都需要重新配置，极大的增加了运维的成本。

基于上述问题，参阅图2所示，为本发明实施例中，告警分类方法的流程图，该方法包括：

步骤201：获取网络设备产生的告警数据；

具体的， $I = \{x_1, x_2, \dots, x_m, \dots\}$ 可以表示为告警数据集， $x_1, \dots, x_m$ 为从网络设备的日志中抓取的告警数据。

步骤202：提取所述告警数据中的有效语义词，并根据所述告警数据的有效语义词确定所述告警数据对应的告警特征向量；

其中，每个有效语义词对应所述告警数据的告警特征向量中的一个特征元素；

具体的，在该场景中，可以先确定告警数据的有效语句作为告警信息的进一步提取，之后对有效语句进行分词，以提取所述告警数据中的有效语义词。其中，有效语义词可以为与告警分类相关的词汇，也可以通过确定与告警分类无关的词汇确定有效语义词。在具体实施过程中，提取所述告警数据中的有效语义词的方法可以有多种。

一种可能的实施方式，可以为将有效语句中除去时间、语句中的介词等与告警分类无关的词汇，剩余的词汇作为有效语义词。

举例来说，比如告警数据中一个有效语句为：VPN xxx is disconnected。可以确定与告警分类无关的词汇为“is”。进而，可以确定有效语义词为“VPN”，“xxx”，“disconnected”。

另一种实施例中，考虑到现实生产环境中对实时性要求比较高的情况，需要数据经过的回路需要尽可能短，以提高实效性。上述提取所述告警数据中的有效语义词花费的时间

不利于提高实效性，特别是告警数据中的有效语句比较长的情况，因此，本发明实施例中提供另一种实施方式，可以将提取所述告警数据中的有效语义词的预先处理动作删除，直接对整个有效语句进行直接分类。为了降低预处理步骤的省略导致的与告警分类无关的词汇对预测结果的影响。提高预测的准确率，一种可能的实现方式，可以在告警数据中的有效语句中加入标记的有效语义词，该有效语义词可以根据该有效语句中确定出的有效语义词确定。例如，端口 DOWN 告警类别的训练数据是:interface xxx is down。

可以将有效语句改为: Interface xxx is down interface。其中，xxx 可以表示端口号，这样可以降低“is”这类与告警分类无关的词汇的频率。让训练数据中重要特征元素的概率更高，并降低与告警分类无关的词汇的概率，能够有效减少与告警分类无关的词汇对预测结果的影响。

对于本发明的应用场景来说，类别集合就是告警的类别，以网络设备告警为例，告警可以分为端口 down、板卡异常、OSPF 异常、BGP 异常等类别。

步骤 203: 根据所述告警数据的告警特征向量和预先训练的分类模型，确定所述告警数据的告警类别；

其中，所述分类模型为根据历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别确定的；所述历史告警数据在各告警类别下的预估概率是根据所述历史告警数据的每个特征元素在每个告警类别的条件概率确定的。

所述分类模型的训练样本，可以人工打标签或者直接获取生产过往的历史告警数据作为训练样本。例如，训练数据可以为:“interface xxx is down”，其标签为端口 DOWN 告警类别。

本发明实施例以朴素贝叶斯的分类为例，说明本发明实施例中的分类模型，当然，也可以通过其他分类模型，实现本发明实施例的分类，在此不做限定。例如，分类模型可以为支持向量机（Support Vector Machine, SVM）分类模型，也可以为逻辑回归（Logistic Regression）、决策树、神经网络、邻近算法或 K 最近邻（k-NearestNeighbor, KNN）分类模型等，本发明实施例中并不进行限制。

如图 3 所示，本发明实施例提供一种朴素贝叶斯分类方法的流程，包括：

步骤 301、确定待分类项 x，结合上述实施例，待分类项 x 可以为网络设备产生的历史告警数据，也可以是新获取的告警数据。历史告警数据的标签可以为预先设置的，也可以为训练过程中设置的。

具体的，还可以包括提取所述告警数据中的有效语义词，并根据所述告警数据的有效语义词确定所述告警数据对应的告警特征向量；

具体的，告警特征向量可以表示为 $x = \{a_1, a_2, \dots, a_m\}$ ，而每个告警特征向量包括特征元素 $a_1, \dots, a_m$ 。

步骤 302、确定训练数据。

在具体实施过程中，训练数据可以为预先设置的历史告警数据及历史告警数据的标签；也可以针对训练或运行分类器后的校验结果，重新生成的新的告警学习数据。

步骤 303、确定告警类别集 $C' = \{y_1, y_2, \dots, y_n\}$ ，确定各类别下各个特征元素的预估概率。

对于本发明的应用场景来说，类别集合就是告警类别，以网络设备告警为例，告警可以分为端口 down、板卡异常、OSPF 异常、BGP 异常等类别。

具体的, 可以表示为 $P(y_1|x)$, $P(y_2|x)$, ..., $P(y_n|x)$ 。

一种可能的实现方式, 在具体实施过程中, 可以包括:

步骤一、获取所述网络设备的历史告警数据集; 所述历史告警数据集包括多个历史告警数据及各历史告警数据对应的告警类别;

步骤二、针对每个历史告警数据, 提取所述历史告警数据中的有效语义词, 将所述有效语义词作为所述历史告警数据的告警特征向量的特征元素;

步骤三、针对同一告警类别, 确定每个特征元素在所述告警类别的条件概率;

步骤四、针对每个历史告警数据, 根据所述历史告警数据的告警特征向量中的每个特征元素在每个告警类别下的条件概率确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率;

具体的, 条件概率可以包括:

$$P(a_1|y_1), P(a_2|y_1), \dots, P(a_m|y_1); P(a_1|y_2), P(a_2|y_2), \dots, P(a_m|y_2); \dots; P(a_1|y_n), P(a_2|y_n), \dots, P(a_m|y_n)$$

其中, a_m 为 x 告警特征向量中的特征元素, $x = \{a_1, a_2, \dots, a_m\}$ 为所述历史告警数据的告警特征向量, y_i 为第 i 个告警类别; $P(x|y_i)$ 为告警类别为 y_i 出现告警特征向量 x 的概率; $P(a_j|y_i)$ 为第 a_j 个特征元素在第 i 个告警类别的条件概率;

因此, 可以确定:

$$P(y_i|x) = \frac{P(x|y_i)P(y_i)}{P(x)}$$

因为 $P(x)$ 分母对于所有类别 y_i (i 为任意值) 为常数, 因为我们只要将分子最大化即可确定 $P(y_i|x)$ 的最大值。由于各特征元素是条件独立的, 因此, 所述历史告警数据的告警特征向量在每个告警类别下的预估概率, 可以根据以下公式确定:

$$P(x|y_i)P(y_i) = P(a_1|y_i)P(a_2|y_i) \dots P(a_m|y_i)P(y_i) = P(y_i) \prod_{j=1}^m P(a_j|y_i)$$

其中, $x = \{a_1, a_2, \dots, a_m\}$ 为所述历史告警数据的告警特征向量, a_m 为 x 告警特征向量中的特征元素, y_i 为第 i 个告警类别; $P(x|y_i)$ 为告警类别为 y_i 出现告警特征向量 x 的概率; $P(y_i)$ 为第 i 个告警类别在历史告警数据集中的发生概率; $\prod_{j=1}^m P(a_j|y_i)$ 为告警类别为 y_i 出现每个特征元素的条件概率的乘积。

此处, 可以假设每个故障告警发生的概率是一样的, 接着对每个特征元素计算针对每个类别的条件概率。具体的, 可以依照词频统计实现。每条训练数据中的告警的每个有效语义词都做词频统计。

比如训练数据中类别为 VPN 异常的告警: “VPN xxx is disconnected”。

去掉与分类无关的词汇 “is”。“VPN”, “xxx”, “disconnected” 在 VPN 异常出现的频率都是 1/3。

例如, 针对所有训练数据, 可以确定出 VPN 在类别为 “VPN 异常” 的概率的统计值。类似的, 可以确定出所有有效词汇在类别为 “VPN 异常” 的概率的统计值作为每个特征元素针对每个类别的条件概率。

需要注意的是当 $P(a_i|y_j)=0$ 时, 当某个类别下某个特征元素没有出现时, 就是产生这种

现象,这会令分类器质量大大降低。为了解决这个问题,一种可能的实现方式,可以包括:

若确定所述历史告警数据的告警特征向量中的第一特征元素在第一告警类别下的条件概率为零,则将所述历史告警数据的告警特征向量中的所述第一特征元素在所述第一告警类别的条件概率增加预定偏置项;所述第一告警类别为所述告警类别中的任一告警类别;所述第一特征元素为所述历史告警数据的告警特征向量中的任一特征元素。

具体的,预定偏置项可以为1。

步骤五、根据各历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别进行所述分类模型的训练;具体的,训练所有训练数据,直至分类器的准确率达到预设阈值。

步骤304、确定预测的分类。

以 $P(x|y_k)P(y_k)$ 最大为例,则待分类项 x 的所属类别为 y_k 。

具体的,若

$P(y_k|x) = \max\{P(y_1|x), P(y_2|x), \dots, P(y_n|x)\}$, 则确定 $x \in y_k$ 。

步骤305、根据训练数据的标签,校验所述分类的预测值。若确定校验通过,则执行步骤306,否则,针对该训练数据生成新的学习数据,执行步骤302;

具体的,可以包括:根据各历史告警数据在各告警类别下的预估概率,确定各历史告警数据的预测分类;

根据所述预测分类及所述各历史告警数据对应的告警类别,调整所述分类模型的参数,以使所述预测分类的置信度大于预设阈值。

步骤306:确定待分类项的分类结果。

需要说明的是,上述校验过程,还可以在分类过程中执行,即如果训练数据十分有限,分类器可能并不能在新的告警第一次出现时就正确分类。那么就需要人工矫正。假设有一条从未出现的、类别是B的告警数据,被分类器判断成了A类,那么就需要人工加入该告警数据的学习数据,例如,通常就是此告警的类似告警,重新生成分类器,进而加强分类器的准确率。

基于相同的发明构思,如图4所示,本发明另一个实施例提供了一种告警分类装置,包括:

收发单元401,用于获取网络设备产生的告警数据;

处理单元402,用于提取所述告警数据中的有效语义词,并根据所述告警数据的有效语义词确定所述告警数据对应的告警特征向量;其中,每个有效语义词对应所述告警数据的告警特征向量中的一个特征元素;根据所述告警数据的告警特征向量和预先训练的分类模型,确定所述告警数据的告警类别;所述分类模型为根据历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别确定的;所述历史告警数据在各告警类别下的预估概率是根据所述历史告警数据的每个特征元素在每个告警类别的条件概率确定的。

一种可能的实现方式,收发单元401,还用于:

获取所述网络设备的历史告警数据集;所述历史告警数据集包括多个历史告警数据及各历史告警数据对应的告警类别;

处理单元402,还用于:针对每个历史告警数据,提取所述历史告警数据中的有效语义词,将所述有效语义词作为所述历史告警数据的告警特征向量的特征元素;针对同一告

警类别, 确定每个特征元素在所述告警类别的条件概率; 针对每个历史告警数据, 根据所述历史告警数据的告警特征向量中的每个特征元素在每个告警类别下的条件概率确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率; 根据各历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别进行所述分类模型的训练。

一种可能的实现方式, 处理单元 402, 具体用于:

根据各历史告警数据在各告警类别下的预估概率, 确定各历史告警数据的预测分类; 根据所述预测分类及所述各历史告警数据对应的告警类别, 调整所述分类模型的参数, 以使所述预测分类的置信度大于预设阈值。

一种可能的实现方式, 处理单元 402, 具体用于根据以下公式确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率:

$$P(x|y_i)P(y_i) = P(a_1|y_i) P(a_2|y_i) \dots P(a_m|y_i)P(y_i) = P(y_i) \prod_{j=1}^m P(a_j|y_i)$$

其中, $x = \{a_1, a_2, \dots, a_m\}$ 为所述历史告警数据的告警特征向量, a_m 为 x 告警特征向量中的特征元素, y_i 为第 i 个告警类别; $P(x|y_i)$ 为告警类别为 y_i 出现告警特征向量 x 的概率; $P(y_i)$ 为第 i 个告警类别在历史告警数据集中的发生概率; $P(a_j|y_i)$ 为第 a_j 个特征元素在第 i 个告警类别的条件概率; $\prod_{j=1}^m P(a_j|y_i)$ 为告警类别为 y_i 出现每个特征元素的条件的概率的乘积。

一种可能的实现方式, 处理单元 402, 还用于:

若确定所述历史告警数据的告警特征向量中的第一特征元素在第一告警类别下的条件概率为零, 则将所述历史告警数据的告警特征向量中的所述第一特征元素在所述第一告警类别的条件概率增加预定偏置项; 所述第一告警类别为所述告警类别中的任一告警类别; 所述第一特征元素为所述历史告警数据的告警特征向量中的任一特征元素。

基于相同的发明构思, 如图 5 所示, 本发明实施例中, 一种电子设备的结构示意图。

本发明实施例提供了一种电子设备, 其中, 该电子设备可以为服务器或其它计算机装置等, 该电子设备可以包括处理器 710 (Center Processing Unit, CPU)、存储器 720、输入设备 730 和输出设备 740 等, 输入设备 730 可以包括键盘、鼠标、触摸屏等, 输出设备 740 可以包括显示设备, 如液晶显示器 (Liquid Crystal Display, LCD)、阴极射线管 (Cathode Ray Tube, CRT) 等。

存储器 720 可以包括只读存储器 (ROM) 和随机存取存储器 (RAM), 并向处理器 710 提供存储器 720 中存储的程序指令和数据。在本发明实施例中, 存储器 720 可以用于存储上述实施例中的告警分类方法的程序。

处理器 710 通过调用存储器 720 存储的程序指令, 处理器 710 用于按照获得的程序指令执行:

提取所述告警数据中的有效语义词, 并根据所述告警数据的有效语义词确定所述告警数据对应的告警特征向量; 其中, 每个有效语义词对应所述告警数据的告警特征向量中的一个特征元素; 根据所述告警数据的告警特征向量和预先训练的分类模型, 确定所述告警数据的告警类别; 所述分类模型为根据历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别确定的; 所述历史告警数据在各告警类别下的预估概率是根据所述历史告警数据的每个特征元素在每个告警类别的条件概率确定的。

一种可能的实现方式，处理器 710，还用于：针对每个历史告警数据，提取所述历史告警数据中的有效语义词，将所述有效语义词作为所述历史告警数据的告警特征向量的特征元素；针对同一告警类别，确定每个特征元素在所述告警类别的条件概率；针对每个历史告警数据，根据所述历史告警数据的告警特征向量中的每个特征元素在每个告警类别下的条件概率确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率；根据各历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别进行所述分类模型的训练。

一种可能的实现方式，处理器 710，具体用于：

根据各历史告警数据在各告警类别下的预估概率，确定各历史告警数据的预测分类；根据所述预测分类及所述各历史告警数据对应的告警类别，调整所述分类模型的参数，以使所述预测分类的置信度大于预设阈值。

一种可能的实现方式，处理器 710，具体用于根据以下公式确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率：

$$P(x|y_i)P(y_i) = P(a_1|y_i) P(a_2|y_i) \dots P(a_m|y_i)P(y_i) = P(y_i) \prod_{j=1}^m P(a_j|y_i)$$

其中， $x = \{a_1, a_2, \dots, a_m\}$ 为所述历史告警数据的告警特征向量， a_m 为 x 告警特征向量中的特征元素， y_i 为第 i 个告警类别； $P(x|y_i)$ 为告警类别为 y_i 出现告警特征向量 x 的概率； $P(y_i)$ 为第 i 个告警类别在历史告警数据集中的发生概率； $P(a_j|y_i)$ 为第 a_j 个特征元素在第 i 个告警类别的条件概率； $\prod_{j=1}^m P(a_j|y_i)$ 为告警类别为 y_i 出现每个特征元素的条件概率的乘积。

一种可能的实现方式，处理器 710，还用于：

若确定所述历史告警数据的告警特征向量中的第一特征元素在第一告警类别下的条件概率为零，则将所述历史告警数据的告警特征向量中的所述第一特征元素在所述第一告警类别的条件概率增加预定偏置项；所述第一告警类别为所述告警类别中的任一告警类别；所述第一特征元素为所述历史告警数据的告警特征向量中的任一特征元素。

基于相同的发明构思，本发明实施例提供了一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现上述告警分类方法的步骤。

基于相同的发明构思，本发明实施例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在计算机可读存储介质上的计算机程序，所述计算机程序包括程序指令，当所述程序指令被计算机执行时，使所述计算机执行上述告警分类方法的步骤。

本领域内的技术人员应明白，本发明的实施例可提供为方法、系统、或计算机程序产品。因此，本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本发明是参照根据本发明实施例的方法、设备（系统）、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器

以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

尽管已描述了本发明的优选实施例，但本领域内的技术人员一旦得知了基本创造性概念，则可对这些实施例作出另外的变更和修改。所以，所附权利要求意欲解释为包括优选实施例以及落入本发明范围的所有变更和修改。

显然，本领域的技术人员可以对本发明实施例进行各种改动和变型而不脱离本发明实施例的精神和范围。这样，倘若本发明实施例的这些修改和变型属于本发明权利要求及其等同技术的范围之内，则本发明也意图包含这些改动和变型在内。

权 利 要 求

1、一种告警分类方法，其特征在于，包括：

获取网络设备产生的告警数据；

提取所述告警数据中的有效语义词，并根据所述告警数据的有效语义词确定所述告警数据对应的告警特征向量；其中，每个有效语义词对应所述告警数据的告警特征向量中的一个特征元素；

根据所述告警数据的告警特征向量和预先训练的分类模型，确定所述告警数据的告警类别；所述分类模型为根据历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别确定的；所述历史告警数据在各告警类别下的预估概率是根据所述历史告警数据的每个特征元素在每个告警类别的条件概率确定的。

2、如权利要求 1 所述的方法，其特征在于，所述分类模型为根据历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别确定的，包括：

获取所述网络设备的历史告警数据集；所述历史告警数据集包括多个历史告警数据及各历史告警数据对应的告警类别；

针对每个历史告警数据，提取所述历史告警数据中的有效语义词，将所述有效语义词作为所述历史告警数据的告警特征向量的特征元素；

针对同一告警类别，确定每个特征元素在所述告警类别的条件概率；

针对每个历史告警数据，根据所述历史告警数据的告警特征向量中的每个特征元素在每个告警类别下的条件概率确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率；

根据各历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别进行所述分类模型的训练。

3、如权利要求 2 所述的方法，其特征在于，所述根据各历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别进行所述分类模型的训练，包括：

根据各历史告警数据在各告警类别下的预估概率，确定各历史告警数据的预测分类；

根据所述预测分类及所述各历史告警数据对应的告警类别，调整所述分类模型的参数，以使所述预测分类的置信度大于预设阈值。

4、如权利要求 1-3 任一项所述的方法，其特征在于，所述针对每个历史告警数据，根据所述历史告警数据的告警特征向量中的每个特征元素在每个告警类别下的条件概率确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率，为根据以下公式确定：

$$P(x|y_i)P(y_i) = P(a_1|y_i) P(a_2|y_i) \dots P(a_m|y_i)P(y_i) = P(y_i) \prod_{j=1}^m P(a_j|y_i)$$

其中， $x = \{a_1, a_2, \dots, a_m\}$ 为所述历史告警数据的告警特征向量， a_m 为 x 告警特征向量中的特征元素， y_i 为第 i 个告警类别； $P(x|y_i)$ 为告警类别为 y_i 出现告警特征向量 x 的概率； $P(y_i)$ 为第 i 个告警类别在历史告警数据集中的发生概率； $P(a_j|y_i)$ 为第 a_j 个特征元素在第 i 个告警类别的条件概率； $\prod_{j=1}^m P(a_j|y_i)$ 为告警类别为 y_i 出现每个特征元素的条

件概率的乘积。

5、如权利要求4所述的方法，其特征在于，所述方法还包括：

若确定所述历史告警数据的告警特征向量中的第一特征元素在第一告警类别下的条件概率为零，则将所述历史告警数据的告警特征向量中的所述第一特征元素在所述第一告警类别的条件概率增加预定偏置项；所述第一告警类别为所述告警类别中的任一告警类别；所述第一特征元素为所述历史告警数据的告警特征向量中的任一特征元素。

6、一种告警分类装置，其特征在于，包括：

收发单元，用于获取网络设备产生的告警数据；

处理单元，用于提取所述告警数据中的有效语义词，并根据所述告警数据的有效语义词确定所述告警数据对应的告警特征向量；其中，每个有效语义词对应所述告警数据的告警特征向量中的一个特征元素；根据所述告警数据的告警特征向量和预先训练的分类模型，确定所述告警数据的告警类别；所述分类模型为根据历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别确定的；所述历史告警数据在各告警类别下的预估概率是根据所述历史告警数据的每个特征元素在每个告警类别的条件概率确定的。

7、如权利要求6所述的装置，其特征在于，所述收发单元，还用于：

获取所述网络设备的历史告警数据集；所述历史告警数据集包括多个历史告警数据及各历史告警数据对应的告警类别；

所述处理单元，还用于：针对每个历史告警数据，提取所述历史告警数据中的有效语义词，将所述有效语义词作为所述历史告警数据的告警特征向量的特征元素；针对同一告警类别，确定每个特征元素在所述告警类别的条件概率；针对每个历史告警数据，根据所述历史告警数据的告警特征向量中的每个特征元素在每个告警类别下的条件概率确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率；根据各历史告警数据在各告警类别下的预估概率和各历史告警数据对应的告警类别进行所述分类模型的训练。

8、如权利要求7所述的装置，其特征在于，所述处理单元，具体用于：

根据各历史告警数据在各告警类别下的预估概率，确定各历史告警数据的预测分类；根据所述预测分类及所述各历史告警数据对应的告警类别，调整所述分类模型的参数，以使所述预测分类的置信度大于预设阈值。

9、如权利要求6-8任一项所述的装置，其特征在于，所述处理单元，具体用于根据以下公式确定所述历史告警数据的告警特征向量在每个告警类别下的预估概率：

$$P(x|y_i)P(y_i) = P(a_1|y_i)P(a_2|y_i) \dots P(a_m|y_i)P(y_i) = P(y_i) \prod_{j=1}^m P(a_j|y_i)$$

其中， $x = \{a_1, a_2, \dots, a_m\}$ 为所述历史告警数据的告警特征向量， a_m 为 x 告警特征向量中的特征元素， y_i 为第 i 个告警类别； $P(x|y_i)$ 为告警类别为 y_i 出现告警特征向量 x 的概率； $P(y_i)$ 为第 i 个告警类别在历史告警数据集中的发生概率； $P(a_j|y_i)$ 为第 a_j 个特征元素在第 i 个告警类别的条件概率； $\prod_{j=1}^m P(a_j|y_i)$ 为告警类别为 y_i 出现每个特征元素的条件概率的乘积。

10、如权利要求9所述的装置，其特征在于，所述处理单元，还用于：

若确定所述历史告警数据的告警特征向量中的第一特征元素在第一告警类别下的条件概率为零，则将所述历史告警数据的告警特征向量中的所述第一特征元素在所述第一告

警类别的条件概率增加预定偏置项；所述第一告警类别为所述告警类别中的任一告警类别；所述第一特征元素为所述历史告警数据的告警特征向量中的任一特征元素。

11、一种电子设备，其特征在于，包括：

至少一个存储器，用于存储程序指令；

至少一个处理器，用于调用所述存储器中存储的程序指令，按照获得的程序指令执行上述权利要求 1-5 任一项所述的方法。

12、一种计算机可读存储介质，其上存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现上述权利要求 1-5 任一项所述的方法。

13、一种计算机程序产品，其特征在于，所述计算机程序产品包括存储在计算机可读存储介质上的计算机程序，所述计算机程序包括程序指令，当所述程序指令被计算机执行时，使所述计算机执行权利要求 1-5 任一项所述的方法的步骤。

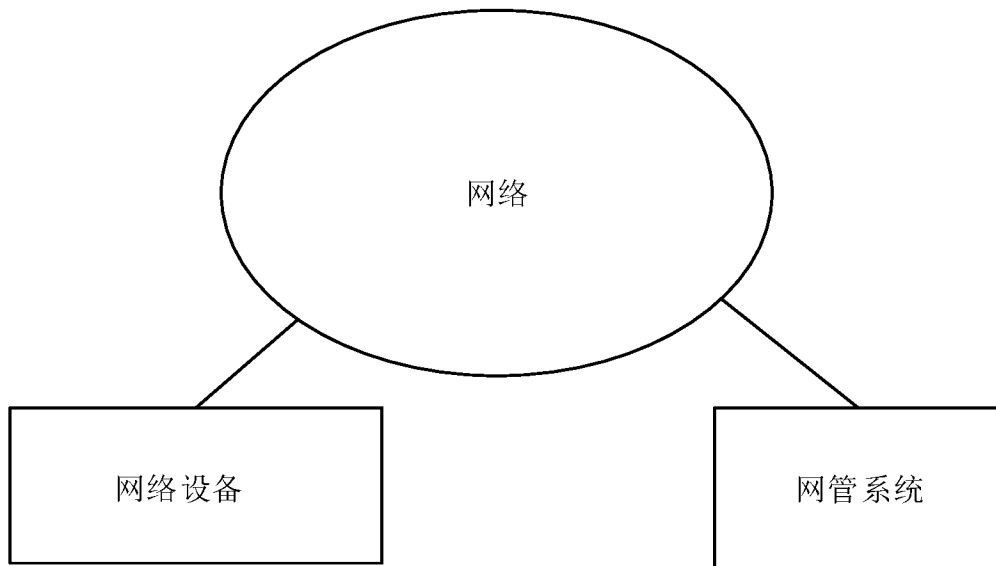


图 1

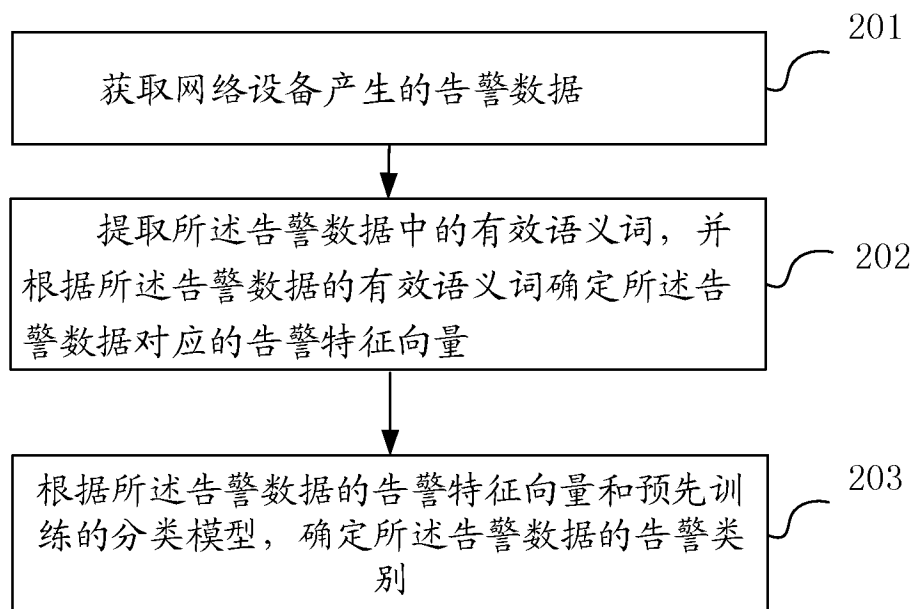


图 2

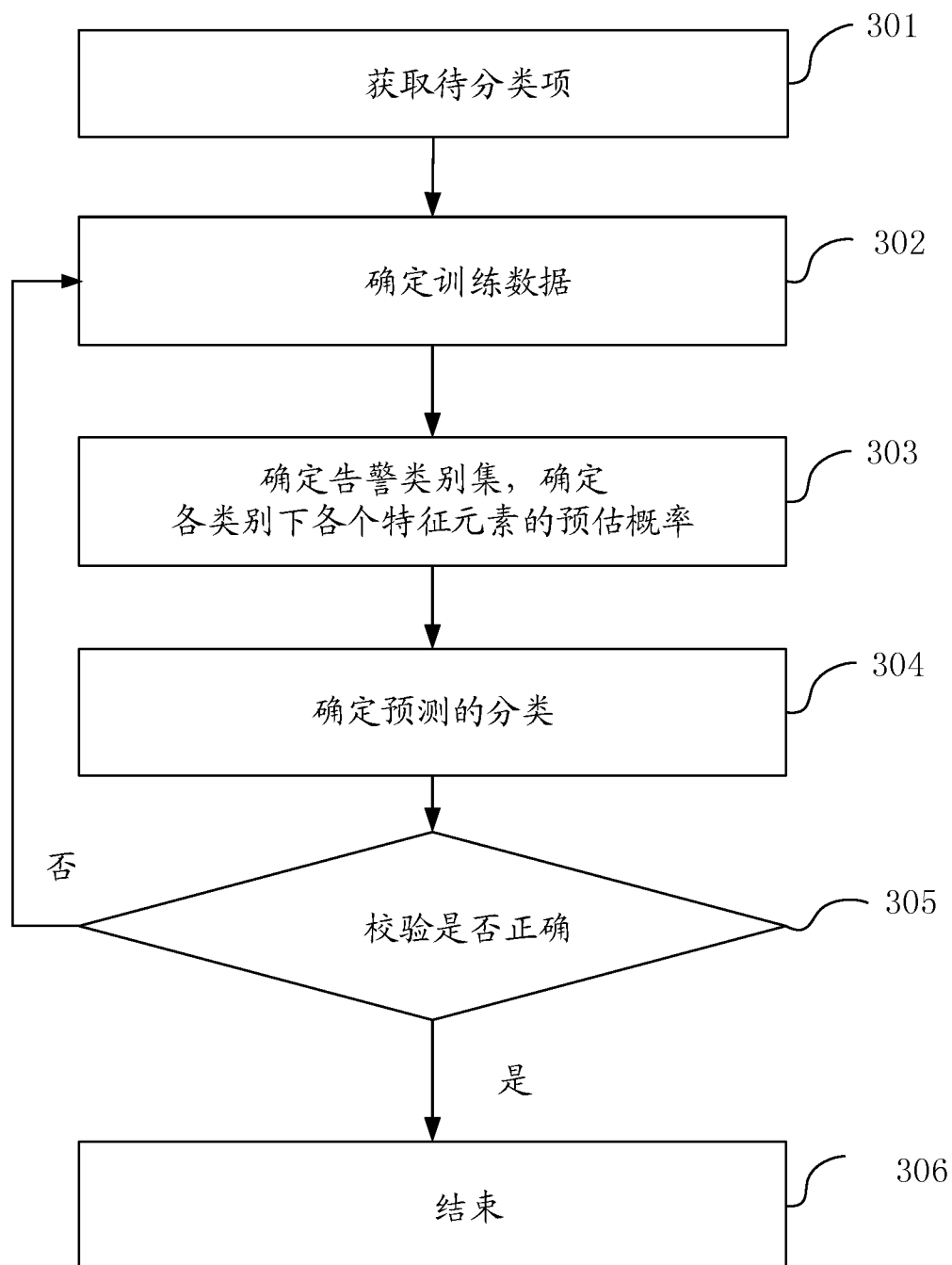


图 3

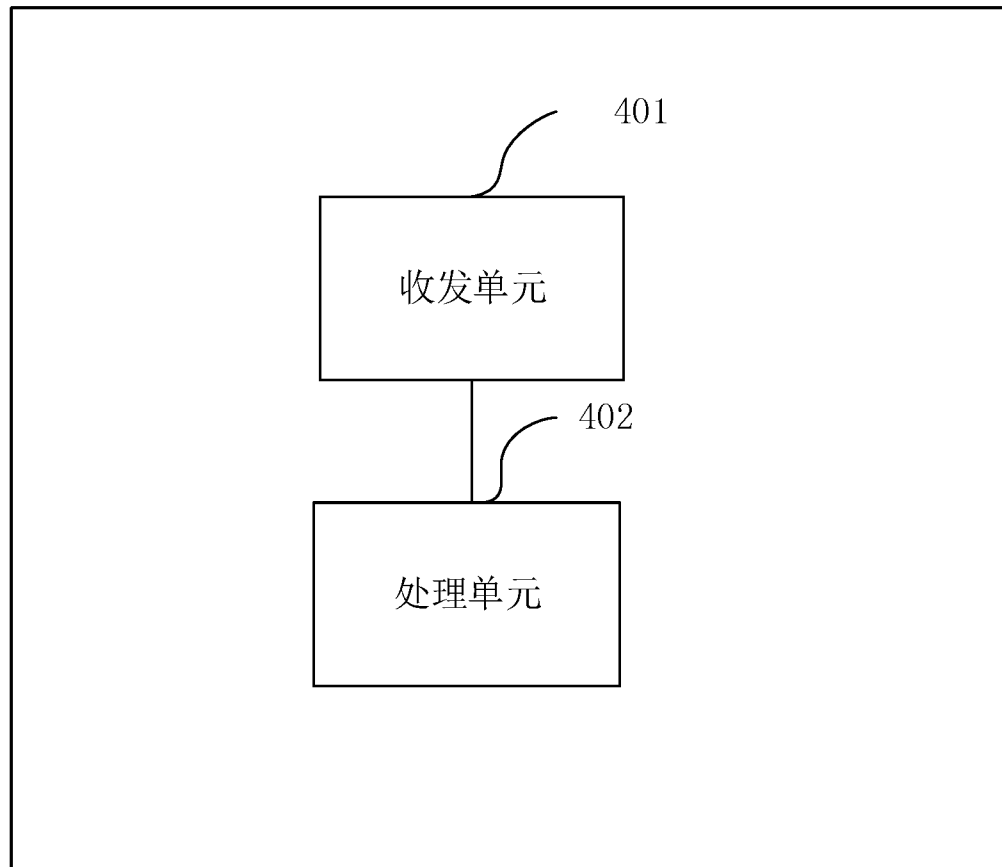


图 4

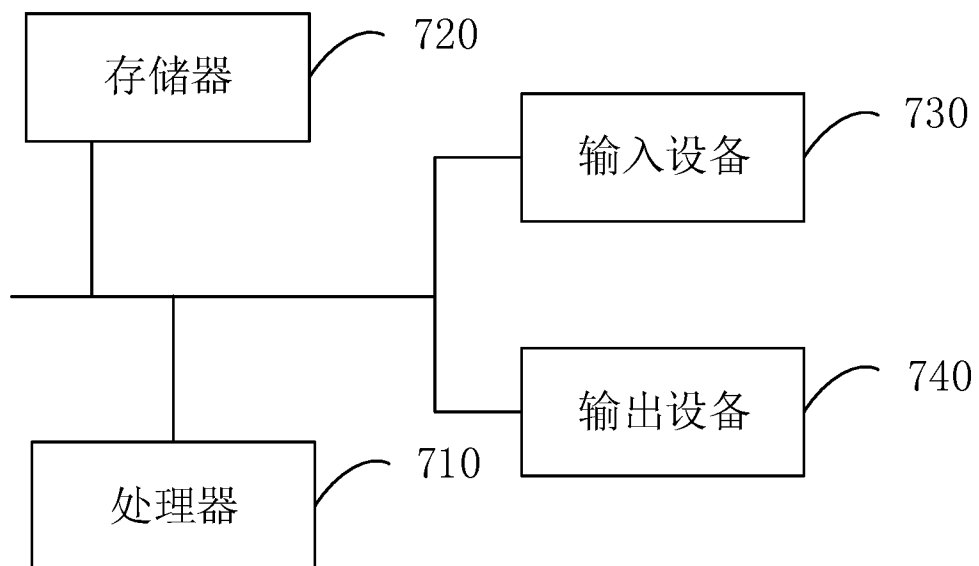


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/086878

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/35(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, IEEE, CNKI: 分类, 类别, 告警, 报警, 分词, 向量, 元素, 概率, 贝叶斯, sort, classify, emergency, warning, word, separation, vector, element, probability, Bayes

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110287316 A (SHENZHEN QIANHAI WEBANK CO., LTD.) 27 September 2019 (2019-09-27) claims 1-12, description paragraphs [0142]-[0143]	1-13
X	CN 108228622 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 29 June 2018 (2018-06-29) description paragraphs [0030]-[0071], paragraph [0128]	1-13
X	CN 109165294 A (ANHUI XUNFEI INTELLIGENT TECHNOLOGY CO., LTD.) 08 January 2019 (2019-01-08) description, paragraphs [0035]-[0057]	1-13
A	CN 107291723 A (ALIBABA GROUP HOLDING LIMITED) 24 October 2017 (2017-10-24) entire document	1-13
A	CN 108595634 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 28 September 2018 (2018-09-28) entire document	1-13



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

06 July 2020

Date of mailing of the international search report

22 July 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/CN2020/086878

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110287316	A	27 September 2019	None			
CN	108228622	A	29 June 2018	None			
CN	109165294	A	08 January 2019	None			
CN	107291723	A	24 October 2017	TW	201737118	A	16 October 2017
				WO	2017167067	A1	05 October 2017
CN	108595634	A	28 September 2018	None			

国际检索报告

国际申请号

PCT/CN2020/086878

A. 主题的分类

G06F 16/35 (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, IEEE, CNKI: 分类, 类别, 告警, 报警, 分词, 向量, 元素, 概率, 贝叶斯, sort, classify, emergency, warning, word, separation, vector, element, probability, Bayes

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 110287316 A (深圳前海微众银行股份有限公司) 2019年 9月 27日 (2019 - 09 - 27) 权利要求1-12, 说明书第[0142]-[0143]段	1-13
X	CN 108228622 A (平安科技深圳有限公司) 2018年 6月 29日 (2018 - 06 - 29) 说明书第[0030]-[0071]段, 第[0128]段	1-13
X	CN 109165294 A (安徽讯飞智能科技有限公司) 2019年 1月 8日 (2019 - 01 - 08) 说明书第[0035]-[0057]段	1-13
A	CN 107291723 A (阿里巴巴集团控股有限公司) 2017年 10月 24日 (2017 - 10 - 24) 全文	1-13
A	CN 108595634 A (腾讯科技深圳有限公司) 2018年 9月 28日 (2018 - 09 - 28) 全文	1-13

☐ 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 7月 6日

国际检索报告邮寄日期

2020年 7月 22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

任兴超

电话号码 86-(10)-53961375

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/086878

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	110287316	A	2019年 9月 27日	无	
CN	108228622	A	2018年 6月 29日	无	
CN	109165294	A	2019年 1月 8日	无	
CN	107291723	A	2017年 10月 24日	TW 201737118 A 2017年 10月 16日 WO 2017167067 A1 2017年 10月 5日	
CN	108595634	A	2018年 9月 28日	无	