



(12) 发明专利

(10) 授权公告号 CN 110609759 B

(45) 授权公告日 2021.09.14

(21) 申请号 201810623804.5

(22) 申请日 2018.06.15

(65) 同一申请的已公布的文献号
申请公布号 CN 110609759 A

(43) 申请公布日 2019.12.24

(73) 专利权人 华为技术有限公司
地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72) 发明人 李健 黄增军 胡敏 谢博 程康

(74) 专利代理机构 北京同达信恒知识产权代理有限公司 11291

代理人 冯艳莲

(51) Int. Cl.

G06F 11/07 (2006.01)

G06K 9/62 (2006.01)

(56) 对比文件

CN 103746831 A, 2014.04.23

CN 103746831 A, 2014.04.23

CN 107770797 A, 2018.03.06

CN 102136949 A, 2011.07.27

CN 103346912 A, 2013.10.09

CN 108156037 A, 2018.06.12

WO 2017008197 A1, 2017.01.19

US 2017180509 A1, 2017.06.22

CN 106941423 A, 2017.07.11

审查员 杨帆

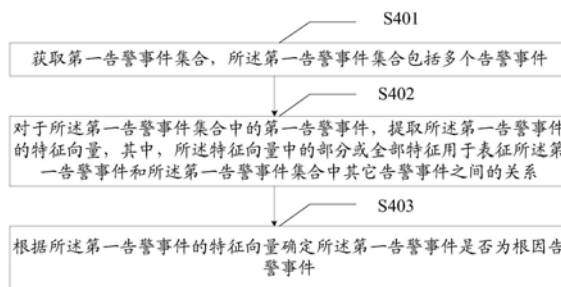
权利要求书3页 说明书14页 附图4页

(54) 发明名称

一种故障根因分析的方法及装置

(57) 摘要

一种故障根因分析的方法及装置,用以解决现有技术中故障根因识别的准确度低的问题。包括:故障根因分析装置获取第一告警事件集合,所述第一告警事件集合包括多个告警事件;对于所述第一告警事件集合中的第一告警事件,提取所述第一告警事件的特征向量,其中,所述特征向量中的部分或全部特征用于表征所述第一告警事件和所述第一告警事件集合中其它告警事件之间的关系,根据所述第一告警事件的特征向量确定所述第一告警事件是否为根因告警事件。本申请根据第一告警事件和其它告警事件之间的关系的特征向量来判断第一告警事件是否为根因告警事件,并提高故障根因识别的准确度。



1. 一种故障根因分析的方法,其特征在于,所述方法包括:

获取第一告警事件集合,所述第一告警事件集合包括多个告警事件;

对于所述第一告警事件集合中的第一告警事件,提取所述第一告警事件的特征向量,其中,所述特征向量中的部分或全部特征用于表征所述第一告警事件和所述第一告警事件集合中其它告警事件之间的关系;所述特征向量包括关联属性、时序关系、数量关系、资源关系中的至少一个,告警严重等级、告警名称、事件类型中的至少一个;其中,所述关联属性用于表征所述第一告警事件在所属的第一告警事件集合中与其他告警事件的关联关系,所述数量关系用于表征所述第一告警事件集合中包括第一告警事件在内的与第一告警事件相同的所有告警事件的个数的总和占第一告警事件集合中全部告警事件数量的比值,所述资源关系用于表征所述第一告警事件集合中第一告警事件与其它告警事件间的传播关系;

根据所述第一告警事件的特征向量,确定所述第一告警事件是否为根因告警事件。

2. 如权利要求1所述的方法,其特征在于,根据所述第一告警事件的特征向量确定所述第一告警事件是否为根因告警事件,具体包括:

将所述第一告警事件的特征向量输入到预先设置的分类模型中,以确定所述第一告警事件是否为根因告警事件,其中,所述分类模型用于根据输入的告警事件的特征向量,判断所述输入的告警事件是否为根因告警事件。

3. 如权利要求1或2所述的方法,其特征在于,所述获取第一告警事件集合,具体包括:

将多个告警事件聚合为至少一个告警事件集合,将所述至少一个告警事件集合中的一个告警事件集合作为所述第一告警事件集合。

4. 如权利要求3所述的方法,其特征在于,所述将多个告警事件聚合为至少一个告警事件集合,具体包括:

将多个告警事件按照设定规则聚合为至少一个告警集合,其中,所述设定规则包括时间关联度、拓扑关联度、文本相似度中的至少一个,其中,所述时间关联度用于表征各告警事件在时间上的关联关系,拓扑关联度用于表征各告警事件在物理位置上的关联关系,文本相似度用于表征各告警事件用字符串表示时,字符串的相似程度。

5. 如权利要求4所述的方法,其特征在于,所述特征向量还包括告警上下文,其中,所述告警上下文用于表征所述第一告警事件集合对应的故障类型。

6. 如权利要求2所述的方法,其特征在于,所述分类模型是根据至少一个训练告警事件集合,按照预设分类算法训练确定的,其中,所述至少一个训练告警事件集合中的每条训练告警事件的信息包括标签和特征向量,所述标签用于表征所述每条训练告警事件是否为根因告警事件;

对于所述训练告警事件集合中的任一训练告警事件,所述训练告警事件的关联属性用于表征所述训练告警事件在所述训练告警事件集合中与其他告警事件的关联关系,所述数量关系用于表征所述训练告警事件集合中包括所述训练告警事件在内的与所述训练告警事件相同的所有训练告警事件的个数的总和占所述训练告警事件集合中全部训练告警事件数量的比值,所述资源关系用于表征所述训练告警事件集合中所述训练告警事件与其它告警事件间的传播关系。

7. 如权利要求6所述的方法,其特征在于,所述训练告警事件集合中的每条训练告警事件的特征向量包括告警严重等级、告警名称、事件类型中的至少一个、以及关联属性、时序

关系、数量关系、资源关系中的至少一个。

8. 如权利要求7所述的方法,其特征在于,至少一个训练告警事件集合为多个训练告警事件集合,所述多个训练告警事件集合中的每个训练告警事件集合中的每条训练告警事件的特征向量还包括告警上下文,其中,所述每条训练告警事件的告警上下文用于表征所述每条训练告警事件所在的训练告警事件集合对应的故障类型。

9. 如权利要求6所述的方法,其特征在于,确定所述第一告警事件是否为根因告警事件之后,该方法还包括:

接收到经过所述分类模型判断并人工矫正后的所述第一告警事件集合中每条告警事件的标签,以及所述每条告警事件的特征向量,将所述标签以及所述每条告警事件的特征向量保存到用于训练所述分类模型的数据库中,其中,所述标签用于表征所述每条告警事件是否为根因告警事件。

10. 如权利要求6所述的方法,其特征在于,若所述预设分类算法为随机森林,则所述确定所述第一告警事件是否为根因告警,具体包括:

对于所述第一告警事件,若判断为根因告警的决策树与全部决策树的比值大于或等于设定阈值,则确定为根因告警,若判断为根因告警的决策树与全部决策树的比值小于所述设定阈值,则确定为非根因告警。

11. 一种装置,其特征在于,所述装置包括:

获取单元,用于获取第一告警事件集合,所述第一告警事件集合包括多个告警事件;

处理单元,用于对于所述第一告警事件集合中的第一告警事件,提取所述第一告警事件的特征向量,其中,所述特征向量中的部分或全部特征用于表征所述第一告警事件和所述第一告警事件集合中其它告警事件之间的关系;所述特征向量包括关联属性、时序关系、数量关系、资源关系中的至少一个,以及告警严重等级、告警名称、事件类型中的至少一个;其中,所述关联属性用于表征所述第一告警事件在所属的第一告警事件集合中与其他告警事件的关联关系,所述数量关系用于表征所述第一告警事件集合中包括第一告警事件在内的与第一告警事件相同的所有告警事件的个数的总和占第一告警事件集合中全部告警事件数量的比值,所述资源关系用于表征所述第一告警事件集合中第一告警事件与其它告警事件间的传播关系;

所述处理单元,还用于根据所述第一告警事件的特征向量,确定所述第一告警事件是否为根因告警事件。

12. 如权利要求11所述的装置,其特征在于,所述处理单元具体用于:

将所述第一告警事件的特征向量输入到预先设置的分类模型中,以确定所述第一告警事件是否为根因告警事件,其中,所述分类模型用于根据输入的告警事件的特征向量,判断所述输入的告警事件是否为根因告警事件。

13. 如权利要求11或12所述的装置,其特征在于,所述获取单元具体用于:

将多个告警事件聚合为至少一个告警事件集合,将所述至少一个告警事件集合中的一个告警事件集合作为所述第一告警事件集合。

14. 如权利要求13所述的装置,其特征在于,所述获取单元具体用于:

将多个告警事件按照设定规则聚合为至少一个告警集合,其中,所述设定规则包括时间关联度、拓扑关联度、文本相似度中的至少一个,其中,所述时间关联度用于表征各告警

事件在时间上的关联关系,拓扑关联度用于表征各告警事件在物理位置上的关联关系,文本相似度用于表征各告警事件用字符串表示时,字符串的相似程度。

15. 如权利要求11所述的装置,其特征在于,所述特征向量还包括告警上下文,其中,所述告警上下文用于表征所述第一告警事件所属的第一告警事件集合对应的故障类型。

16. 如权利要求12所述的装置,其特征在于,所述分类模型是根据至少一个训练告警事件集合,按照预设分类算法训练确定的,其中,所述至少一个训练告警事件集合中的每条训练告警事件的信息包括标签和特征向量,所述标签用于表征所述每条训练告警事件是否为根因告警事件;

对于所述训练告警事件集合中的任一训练告警事件,所述训练告警事件的关联属性用于表征所述训练告警事件在所述训练告警事件集合中与其他告警事件的关联关系,所述数量关系用于表征所述训练告警事件集合中包括所述训练告警事件在内的与所述训练告警事件相同的所有训练告警事件的个数的总和占所述训练告警事件集合中全部训练告警事件数量的比值,所述资源关系用于表征所述训练告警事件集合中所述训练告警事件与其它告警事件间的传播关系。

17. 如权利要求16所述的装置,其特征在于,所述训练告警事件集合中的每条训练告警事件的特征向量包括告警严重等级、告警名称、事件类型中的至少一个、以及关联属性、时序关系、数量关系、资源关系中的至少一个。

18. 如权利要求17所述的装置,其特征在于,至少一个训练告警事件集合为多个训练告警事件集合,所述多个训练告警事件集合中的每个训练告警事件集合中的每条训练告警事件的特征向量还包括告警上下文,其中,所述每条训练告警事件的告警上下文用于表征所述每条训练告警事件所在的训练告警事件集合对应的故障类型。

19. 如权利要求16所述的装置,其特征在于,确定所述第一告警事件是否为根因告警事件之后,所述获取单元还用于:

接收到经过所述分类模型判断并人工矫正后的所述第一告警事件集合中每条告警事件的标签,以及所述每条告警事件的特征向量,将所述标签以及所述每条告警事件的特征向量保存到用于训练所述分类模型的数据库中,其中,所述标签用于表征所述每条告警事件是否为根因告警事件。

20. 如权利要求16所述的装置,其特征在于,若所述预设分类算法为随机森林,则所述处理单元具体用于:

对于所述第一告警事件,若判断为根因告警的决策树与全部决策树的比值大于或等于设定阈值,则确定为根因告警,若判断为根因告警的决策树与全部决策树的比值小于所述设定阈值,则确定为非根因告警。

21. 一种装置,其特征在于,包括处理器、存储器和通信接口;

所述存储器,存储有计算机程序;

所述处理器,用于调用并执行所述存储器中存储的计算机程序,通过所述通信接口来实现如权利要求1至中10任一项所述的方法。

22. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有计算机指令,当所述指令在计算机上运行时,使得计算机执行如权利要求1至10任一项所述的方法。

一种故障根因分析的方法及装置

技术领域

[0001] 本申请涉及信息技术领域,尤其涉及一种故障根因分析的方法及装置。

背景技术

[0002] 随着网络规模的扩大,各类网络设备的增多,这些网络设备会产生海量的各类型的告警,给运营支撑系统(operations support system,OSS)带来了很大的压力,原因在于这些海量告警中存在大量的无效告警,重复告警等不需要关注的告警,而运维人员无法从海量的告警中准确识别出故障根因,排障效率低,导致了大量的派单,浪费人力,运维成本高。

[0003] 在现有技术中,为了提高排障效率,减少派单,利用告警事件的数据,例如告警事件的类型名称,判断该告警事件是否为根因,但准确度低。

[0004] 综上所述,提高故障根因识别的准确度是目前需要解决的问题。

发明内容

[0005] 本申请提供一种故障根因分析的方法及装置,用以解决现有技术中故障根因识别的准确度低的问题。

[0006] 第一方面,本申请提供了一种故障根因分析的方法,所述方法包括:

[0007] 获取第一告警事件集合,所述第一告警事件集合包括多个告警事件;对于所述第一告警事件集合中的第一告警事件,提取所述第一告警事件的特征向量,其中,所述特征向量中的部分或全部特征用于表征所述第一告警事件和所述第一告警事件集合中其它告警事件之间的关系;根据所述第一告警事件的特征向量,确定所述第一告警事件是否为根因告警事件。

[0008] 通过上述方法,故障根因分析装置从多个维度提取所述第一个告警集合中第一告警事件的特征向量,使用第一告警事件和其它告警事件之间的关系的特征向量来判断所述第一告警事件是否为根因告警事件,提高了故障根因识别的准确度。

[0009] 在一种可能的设计中,根据所述第一告警事件的特征向量确定所述第一告警事件是否为根因告警事件,具体包括:将所述第一告警事件的特征向量输入到预先设置的分类模型中,以确定所述第一告警事件是否为根因告警事件,其中,所述分类模型用于表征特征向量和根因告警事件之间的关系。

[0010] 通过该方法,采用预先设置的分类模型判断第一告警事件是否为根因告警事件,通过机器学习的方法进一步提高了故障根因识别的准确度。

[0011] 在一种可能的设计中,所述获取第一告警事件集合,具体包括:将多个告警事件聚合为至少一个告警事件集合,将所述至少一个告警事件集合中的一个告警事件集合作为所述第一告警事件集合。

[0012] 通过该方法,确定出如何获取第一告警事件集合,是实现本申请的方法不可缺少的步骤。

[0013] 在一种可能的设计中,所述将多个告警事件聚合为至少一个告警事件集合,具体包括:

[0014] 将多个告警事件按照设定规则聚合为至少一个告警集合,其中,所述设定规则包括时间关联度、拓扑关联度、文本相似度中的至少一个,其中,所述时间关联度用于表征各告警事件在时间上的关联关系,拓扑关联度用于表征各告警事件在物理位置上的关联关系,文本相似度用于表征各告警事件用字符串表示时,字符串的相似程度。

[0015] 在一种可能的设计中,所述特征向量包括告警严重等级、告警名称、事件类型中的至少一个、以及关联属性、时序关系、数量关系、资源关系中的至少一个,其中,所述关联属性用于表征所述第一告警事件在所属的第一告警事件集合中与其他告警事件的关联关系,所述数量关系用于表征所述第一告警事件集合中包括第一告警事件在内的与第一告警事件相同的所有告警事件的个数的总和占第一告警事件集合中全部告警事件数量的比值,所述资源关系用于表征所述第一告警事件集合中第一告警事件与其它告警事件间的传播关系。

[0016] 在一种可能的设计中,所述特征向量还包括告警上下文,其中,所述告警上下文用于表征所述第一告警事件所属的第一告警事件集合对应的故障类型。

[0017] 在一种可能的设计中,所述分类模型是根据至少一个训练告警事件集合,按照预设分类算法训练确定的,其中,所述至少一个训练告警事件集合中的每条训练告警事件的信息包括标签和特征向量,所述标签用于表征所述每条训练告警事件是否为根因告警事件;

[0018] 对于所述训练告警事件集合中的任一训练告警事件,所述训练告警事件的关联属性用于表征所述训练告警事件在所述训练告警事件集合中与其他告警事件的关联关系,所述数量关系用于表征所述训练告警事件集合中包括所述训练告警事件在内的与所述训练告警事件相同的所有训练告警事件的个数的总和占所述训练告警事件集合中全部训练告警事件数量的比值,所述资源关系用于表征所述训练告警事件集合中所述训练告警事件与其它告警事件间的传播关系。

[0019] 在一种可能的设计中,所述训练告警事件集合中的每条训练告警事件的特征向量包括告警严重等级、告警名称、事件类型中的至少一个、以及关联属性、时序关系、数量关系、资源关系中的至少一个。

[0020] 在一种可能的设计中,至少一个训练告警事件集合为多个训练告警事件集合,所述多个训练告警事件集合中的每个训练告警事件集合中的每条训练告警事件的特征向量还包括告警上下文,其中,所述每条训练告警事件的告警上下文用于表征所述每条训练告警事件所在的训练告警事件集合对应的故障类型。

[0021] 在一种可能的设计中,确定所述第一告警事件是否为根因告警事件之后,该方法还包括:接收到经过所述分类模型判断并人工矫正后的所述第一告警事件集合中每条告警事件的标签,以及所述每条告警事件的特征向量,将所述标签以及所述每条告警事件的特征向量保存到用于训练所述分类模型的数据库中,其中,所述标签用于表征所述每条告警事件是否为根因告警事件。

[0022] 通过该方法,使用人工矫正后的所述第一告警事件的标签进行分类模型训练,可以提高分类模型的准确度。

[0023] 在一种可能的设计中,若所述预设分类算法为随机森林,则所述确定所述第一告警事件是否为根因告警,具体包括:对于所述第一告警事件,若判断为根因告警的决策树与全部决策树的比值大于或等于设定阈值,则确定为根因告警,若判断为根因告警的决策树与全部决策树的比值小于所述设定阈值,则确定为非根因告警。

[0024] 第二方面,本申请提供一种装置,所述装置包括:获取单元,用于获取第一告警事件集合,所述第一告警事件集合包括多个告警事件;处理单元,用于对于所述第一告警事件集合中的第一告警事件,提取所述第一告警事件的特征向量,其中,所述特征向量中的部分或全部特征用于表征所述第一告警事件和所述第一告警事件集合中其它告警事件之间的关系;所述处理单元,还用于根据所述第一告警事件的特征向量,确定所述第一告警事件是否为根因告警事件。

[0025] 在一种可能的设计中,所述处理单元具体用于:将所述第一告警事件的特征向量输入到预先设置的分类模型中,以确定所述第一告警事件是否为根因告警事件,其中,所述分类模型用于表征特征向量和根因告警事件之间的关系。

[0026] 在一种可能的设计中,所述获取单元具体用于:将多个告警事件聚合为至少一个告警事件集合,将所述至少一个告警事件集合中的一个告警事件集合作为所述第一告警事件集合。

[0027] 在一种可能的设计中,所述获取单元具体用于:将多个告警事件按照设定规则聚合为至少一个告警集合,其中,所述设定规则包括时间关联度、拓扑关联度、文本相似度中的至少一个,其中,所述时间关联度用于表征各告警事件在时间上的关联关系,拓扑关联度用于表征各告警事件在物理位置上的关联关系,文本相似度用于表征各告警事件用字符串表示时,字符串的相似程度。

[0028] 在一种可能的设计中,所述特征向量包括告警严重等级、告警名称、事件类型中的至少一个、以及关联属性、时序关系、数量关系、资源关系中的至少一个,其中,所述关联属性用于表征所述第一告警事件在所属的第一告警事件集合中包括第一告警事件在内的与第一告警事件相同的所有告警事件的个数的总和占第一告警事件集合中全部告警事件数量的比值,所述资源关系用于表征所述第一告警事件集合中第一告警事件与其它告警事件间的传播关系。

[0029] 在一种可能的设计中,所述特征向量还包括告警上下文,其中,所述告警上下文用于表征所述第一告警事件所属的第一告警事件集合对应的故障类型。

[0030] 在一种可能的设计中,所述分类模型是根据至少一个训练告警事件集合,按照预设分类算法训练确定的,其中,所述至少一个训练告警事件集合中的每条训练告警事件的信息包括标签和特征向量,所述标签用于表征所述每条训练告警事件是否为根因告警事件;

[0031] 对于所述训练告警事件集合中的任一训练告警事件,所述训练告警事件的关联属性用于表征所述训练告警事件在所述训练告警事件集合中与其他告警事件的关联关系,所述数量关系用于表征所述训练告警事件集合中包括所述训练告警事件在内的与所述训练告警事件相同的所有训练告警事件的个数的总和占所述训练告警事件集合中全部训练告警事件数量的比值,所述资源关系用于表征所述训练告警事件集合中所述训练告警事件与其它告警事件间的传播关系。

[0032] 在一种可能的设计中,所述训练告警事件集合中的每条训练告警事件的特征向量包括告警严重等级、告警名称、事件类型中的至少一个、以及关联属性、时序关系、数量关系、资源关系中的至少一个。

[0033] 在一种可能的设计中,至少一个训练告警事件集合为多个训练告警事件集合,所述多个训练告警事件集合中的每个训练告警事件集合中的每条训练告警事件的特征向量还包括告警上下文,其中,所述每条训练告警事件的告警上下文用于表征所述每条训练告警事件所在的训练告警事件集合对应的故障类型。

[0034] 在一种可能的设计中,确定所述第一告警事件是否为根因告警事件之后,所述获取单元还用于:接收到经过所述分类模型判断并人工矫正后的所述第一告警事件集合中每条告警事件的标签,以及所述每条告警事件的特征向量,将所述标签以及所述每条告警事件的特征向量保存到用于训练所述分类模型的数据库中,其中,所述标签用于表征所述每条告警事件是否为根因告警事件。

[0035] 在一种可能的设计中,若所述预设分类算法为随机森林,则所述处理单元具体用于:

[0036] 对于所述第一告警事件,若判断为根因告警的决策树与全部决策树的比值大于或等于设定阈值,则确定为根因告警,若判断为根因告警的决策树与全部决策树的比值小于所述设定阈值,则确定为非根因告警。

[0037] 第三方面,本申请实施例还提供了一种装置,包括处理器、存储器和通信接口,所述存储器用于存储计算机程序,所述处理器用于读取所述存储器中存储的计算机程序并实现第一方面、第一方面的任意一种设计提供的方法。

[0038] 第四方面,本申请还提供了一种计算机可读存储介质,用于存储为执行上述第一方面、第一方面的任意一种设计的功能所用的计算机软件指令,其包含用于执行上述第一方面、第一方面的任意一种设计的方法所设计的程序。

[0039] 第五方面,本申请提供一种包含指令的计算机程序产品,当其在计算机上运行时,使得计算机执行第一方面和/或第一方面的任一种设计所述的方法。

附图说明

[0040] 图1为本申请提供的一种告警事件示意图;

[0041] 图2为本申请提供的一种资源关系树示意图;

[0042] 图3为本申请提供的一种系统架构图;

[0043] 图4为本申请提供的一种故障根因分析的方法流程图;

[0044] 图5为本申请提供的另一种故障根因分析的方法流程图;

[0045] 图6为本申请提供的随机森林示意图;

[0046] 图7为本申请提供的一种的装置示意图;

[0047] 图8为本申请提供的一种的硬件结构图。

具体实施方式

[0048] 下面将结合附图对本申请实施例作进一步地详细描述。

[0049] 本申请实施例提供一种故障根因分析的方法及装置,用于解决现有技术中故障根

因识别的准确度低的问题。其中,方法和设备是基于同一发明构思的,由于方法及设备解决问题的原理相似,因此设备与方法的实施可以相互参见,重复之处不再赘述。

[0050] 随着网络规模的扩大,路由器,交换机,防火墙,负载均衡设备,接入网关设备等各类网络设备增多,这些网络设备会上报海量告警给OSS,但这些海量告警中存在大量的无效告警,重复告警等不需要关注的告警,运维人员无法从海量的告警中准确识别出故障根因,导致了大量的派单,浪费人力,运维成本高。在现有技术中,利用告警事件的数据,例如告警事件的类型名称,判断该告警事件是否为根因,准确度低,因此,如何提高故障根因识别的准确度是目前需要解决的问题。

[0051] 以下,对本申请中的部分用语进行解释说明,以便使本领域技术人员理解。

[0052] 1) 告警事件集合,也可以称为situation,是根据时间关联度,拓扑关联度,文本相似度中的至少一个维度,对一个可能的故障对应的一系列告警事件进行聚合得到的,举例说明:假设原始告警事件的集合为 $A=[a_1, a_2, \dots, a_n]$,聚合后的所有situation记为 $S, S=\{s_1:[a_1, \dots, a_k], \dots, s_i:[a_p, \dots, a_q], \dots, s_r:[a_m, \dots, a_n]\}$,其中 r 是situation的数量, $1 < k, p, q, m < n$,即每一个situation s_i 都是一系列告警事件的集合,其中,所述告警事件集合可以是通过聚合的方式得到的也可以是人工确定的。

[0053] 2) 至少一个,可以为一个,也可以为多个,本申请对其不做限定。

[0054] 3) 分类模型,用于表征所述特征向量和根因告警事件之间的关系,根据输入的告警事件的特征向量,判断该告警事件是否为根因告警事件。其中,所述分类模型是按照预设分类算法根据至少一个训练告警事件集合确定的,所述训练告警事件集合中的告警事件是离线获取的,所述分类模型可以按照设定时间进行更新,更新时使用的数据库中包括的在线实时告警事件可以是经过分类模型判断并人工矫正后的告警事件。

[0055] 4) 告警严重等级,用于表征告警事件的等级,可以用中文或者英文的字符串表示,例如,用中文表示时可以为紧急、重要、次要、提醒、未知5个等级,在数据处理时,可以将中文字符串处理为对应的特征,由于告警严重等级间存在递进的关系,需要进行单独编码,因此紧急、重要、次要、提醒、未知可以分别处理成5,4,3,2,1等告警严重等级特征,告警事件在上报时携带告警严重等级,假设,一个situation中有4条告警事件,告警严重等级分别为紧急、重要、重要、重要,则该4条告警事件的告警严重等级特征分别为5,4,4,4,编码也可以为其它形式,本申请对其不做限定。

[0056] 5) 告警名称,用于表征告警事件的属性,可以表征具体的故障原因,可选的,告警名称还会对应一个恢复操作建议。其中,告警名称可以用离散的中文或者英文的字符串表示,例如,告警名称可以为ETH_LOS、MPLS_TUNNEL_LOCV、ETH_APS_LOST、TUNNEL_DOWN、整机用户数降到最小阈值等;由于告警名称间都是独立的,因此在编码时采用独热(one-hot)编码,假设,一个situation中的告警名称有ETH_LOS、MPLS_TUNNEL_LOCV、ETH_APS_LOST、TUNNEL_DOWN、整机用户数降到最小阈值,进行one-hot编码后,这5个告警名称分别对应的编码为00001,00010,00100,01000,10000,即告警名称特征,编码也可以为其它形式,本申请对其不做限定。

[0057] 6) 事件类型,用于表征告警事件的类型,具体的,可以分为硬件告警、环境告警、线路告警、业务告警、安全告警等。也可以同通过one-hot编码对事件类型进行编码,假设,一个situation中的事件类型有硬件告警、环境告警、线路告警、业务告警、安全告警,进行

one-hot编码后,这5个事件类型分别对应的编码为00001,00010,00100,01000,10000,即事件类型特征,编码也可以为其它形式,本申请对其不做限定。

[0058] 7) 关联属性,用于表征告警事件在一个situation中与其他告警事件的关联关系,具体可以为根因告警事件、非根因告警事件也可能无属性,各告警事件的关联属性中所标记的是否为根因告警事件,是根据该告警事件和其他告警事件的关联关系确定的,例如告警事件A引起告警事件B,则告警事件A对于告警事件B就是根因告警事件,则将告警事件A的关联属性标记为根因告警事件,将告警事件B的关联属性标记为非根因告警事件。当告警事件为根因告警事件时,标记为P,当告警事件为非根因告警事件时标记为C。假设一个situation中关联属性标记分别为P、C和空,进行one-hot编码后,这3个关联属性分别对应的编码为001,010,100,即关联属性特征,编码也可以为其它形式,本申请对其不做限定。

[0059] 8) 时序关系,用于表征告警事件在一个situation中发生的时间顺序,还可以表征各告警事件之间在时间上的距离关系,具体的,首先根据告警事件发生时间对situation中所有的告警事件进行排序,将告警事件发生时间构造为向量 $T=[t_1, t_2, \dots, t_i, \dots, t_n]$, $1 \leq i \leq n$,由于不同的situation持续时间不一样,因此需要进行归一化处理,即 t_i 时刻的告警事件所对应的时序特征 $T_i = (t_i - \min(T)) / (\max(T) - \min(T))$ 。例如,一个situation中有5条告警事件,发生时间分别为10:01,10:02,10:02,10:02,10:08,则这5个告警事件对应的时序关系特征分别为0.0,0.14,0.14,0.14,1。

[0060] 9) 数量关系,用于表征一个situation中相同告警事件的数量占situation中全部告警事件数量的比值,具体的,situation中有 N_i 个 i 告警事件,记 i 告警事件的次数为 N_i ,situation中告警事件总数有 N 个,则 i 告警事件的数量关系特征为 C_i , $C_i = N_i / N$ 。在具体实现时,可以将来自同一告警源的具有相同告警名称的告警事件看做相同告警事件。

[0061] 10) 资源关系,用于表征一个situation中各个告警事件间的传播关系,在垂直方向上,都是server层告警事件引起其承载的client层告警事件,举例说明,如图1所示,图1中1,2,3,4,5,6,7,8,9代表网元,a,b,c,d,e为发生的告警事件,将告警事件与有向图无环图G相结合,可以得到如图2所示的该situation中告警事件资源关系树,位于树根的告警事件是根因告警事件,标记为1,叶子节点标记为0。例如,一个situation中有5个告警事件,通过资源关系树判断后,处在最底层的告警只有第一个告警,其他告警位于更上层次,那么这5个告警的特征分别为1,0,0,0,0,其中,所示有向图无环图G为根据网络拓扑和存量,将网元、单板、端口间的物理和逻辑连接关系进行表示的图,本申请实施例中对1和0所表示的内容不做限定。

[0062] 11) 告警上下文,用于表征告警事件所属的situation对应的故障类型,例如,线路类故障、传输故障、网元脱管、掉线群障等,编码时也可以采用独热one-hot编码。

[0063] 12) 多个,是指两个或两个以上。

[0064] 本申请实施例中,对于采用one-hot编码的特征,假设该特征具有 m 个可能的取值,则经过one-hot编码后,该特征就变成了 m 个二元特征,采用one-hot编码,可以解决分类器不方便处理属性数据的问题,同时也在一定程度上起到了扩充特征的作用,在进行数据处理时,One-hot编码需要针对每个特征先进行编码,编码后再进行拼接。假设,告警名称、事件类型、关联属性这3个特征每个可能的取值的数量分别为 a, b, c ,则经过one-hot编码后,一条告警的特征向量维度从3维变成了 $a+b+c$ 维。

[0065] 另外,需要理解的是,在本申请的描述中,“第一”、“第二”等词汇,仅用于区分描述的目的,而不能理解为指示或暗示相对重要性,也不能理解为指示或暗示顺序。

[0066] 本申请实施例中,故障根因分析装置采用告警事件聚合的方式将至少一个告警事件聚合为至少一个告警事件集合,然后从多个维度提取所述至少一个告警集合中每条告警事件的特征向量,将所述每条告警事件的特征向量输入到预先设置的通过机器学习确定的分类模型中,确定所述每条告警事件是否为根因告警事件。可以减少专家人力投入,提高故障根因识别的准确度。

[0067] 在本申请实施例中,故障根因分析装置可以为一个单独的服务器,也可以是网管设备的一个软件模块,具体的,故障根因分析装置所在的系统架构可以图3所示,所述故障根因分析装置也可以称为根因识别装置,路由器,交换机,防火墙,负载均衡设备,接入网关设备等各类网元设备产生各种类型的告警事件(也可以称为告警日志、原始告警实例),所述告警事件的数据如表1所示,表1仅仅是一个简单的例子,并没有包括告警事件的全部特征,在线实时告警可以从数据总线获取。

[0068] 表1

关联属性	告警严重等级	告警名称	告警源	网元类型	定位信息	发生时间	事件类型
-	紧急	NE_NOT_LOGI N	Ne1	PTN xx	##	2017/10/15 18:04:25	安全告警
-	紧急	NE_NOT_LOGI N	Ne1	PTN xx	##	2017/10/15 18:04:25	安全告警
非根因告警	重要	MPLS_PW_LO CV	NEx	PTN xx	##	2017/10/15 18:04:26	通信质量
根因报警	重要	ETH_LOS	NEx	PTN xx	##	2017/10/15 18:04:26	通信质量

[0070] 网管设备接收各类网元设备发送出的告警事件,进行situation构建,将situation中的告警事件通过故障根因分析装置进行判断,确定告警事件是否为根因告警事件,若为根因则派单给OSS,本申请实施例中,situation构建也可以在故障根因分析装置中完成。

[0071] 下面结合附图4对上述系统架构中situation构建和故障根因分析装置进行判断告警事件是否为根因告警事件的方案进行具体说明。具体参见图4,为本申请提供了一种故障根因分析的方法流程图。该方法包括:

[0072] 步骤S401、故障根因分析装置获取第一告警事件集合,所述第一告警事件集合包括多个告警事件。

[0073] 步骤S402、故障根因分析装置提取所述第一告警集合中第一告警事件的特征向量,其中,所述特征向量中的部分或全部特征用于表征所述第一告警事件和所述第一告警事件集合中其它告警事件之间的关系。

[0074] 具体的,所述特征向量包括告警严重等级、告警名称、事件类型中的至少一个、以及关联属性、时序关系、数量关系、资源关系中的至少一个,可选的,所述特征向量还包括告警上下文。所述关联属性用于表征所述第一告警事件在所属的第一告警事件集合中与其他

告警事件的关联关系,所述数量关系用于表征所述第一告警事件集合中包括第一告警事件在内的与第一告警事件相同的所有告警事件的个数的总和占第一告警事件集合中全部告警事件数量的比值,所述资源关系用于表征所述第一告警事件集合中第一告警事件与其它告警事件间的传播关系,所述告警上下文用于表征所述第一告警事件集合对应的故障类型。

[0075] 步骤S403、根据所述第一告警事件的特征向量,确定所述第一告警事件是否为根因告警事件。

[0076] 具体的,故障根因分析装置将所述每条告警事件的特征向量输入到预先设置的分类模型中,其中,所述分类模型用于表征所述特征向量和根因告警事件之间的关系。

[0077] 该分类模型可以是根据至少一个训练告警事件集合,按照预设分类算法训练确定的,其中,所述至少一个训练告警事件集合中的每条训练告警事件的信息包括标签和特征向量,所述标签用于表征所述每条训练告警事件是否为根因告警事件。对于所述训练告警事件集合中的任一训练告警事件,所述训练告警事件的关联属性用于表征所述训练告警事件在所述训练告警事件集合中与其他告警事件的关联关系,所述数量关系用于表征所述训练告警事件集合中包括所述训练告警事件在内的与所述训练告警事件相同的所有训练告警事件的个数的总和占所述训练告警事件集合中全部训练告警事件数量的比值,所述资源关系用于表征所述训练告警事件集合中所述训练告警事件与其它告警事件间的传播关系。

[0078] 上述至少一个训练告警事件集合可以为多个训练告警事件集合,所述多个训练告警事件集合中的每个训练告警事件集合中的每条训练告警事件的特征向量还可以包括告警上下文,其中,所述每条训练告警事件的告警上下文用于表征所述每条训练告警事件所在的训练告警事件集合对应的故障类型。

[0079] 可以理解的是,在训练分类模型时用到的训练告警事件的特征向量和步骤S402中提取的第一告警事件的特征向量通常是一致的,如,在训练分类模型时用到的训练告警事件的特征向量包括告警严重等级、告警名称、事件类型、关联属性、时序关系和数量关系,相应地,在步骤S402中提取的第一告警事件的特征向量也包括告警严重等级、告警名称、事件类型、关联属性、时序关系和数量关系。

[0080] 本申请实施例中,故障根因分析装置从多个维度提取所述第一个告警集合中第一告警事件的特征向量,使用第一告警事件和其它告警事件之间的关系的特征向量来判断所述第一告警事件是否为根因告警事件,提高了故障根因识别的准确度。

[0081] 所述第一告警事件集合中的其他告警事件的处理方法与所述第一告警事件相同,本发明在此不一一赘述。

[0082] 在一种可能的设计中,步骤S401之前,该方法还包括:将多个告警事件聚合为至少一个告警事件集合,将所述至少一个告警事件集合中的一个告警事件集合作为所述第一告警事件集合。

[0083] 具体的,将多个告警事件按照设定规则聚合为至少一个告警事件集合,其中,所述设定规则包括时间关联度、拓扑关联度、文本相似度中的一个或多个,本申请对其不做限定,其中,所述时间关联度用于表征各告警事件在时间上的关联关系,举例说明,时间相近的告警事件关联度比较好,时间越接近的告警事件属于同一个故障的可能性大,应该被聚合到一个告警集合中;拓扑关联度用于表征各告警事件在物理位置上的关联关系,举例说

明,从空间拓扑看,物理上位置比较靠近的告警事件关联度比较好,越靠近的告警事件属于同一个故障的可能性更大,应该被聚合到一个告警集合中,文本相似度用于表征各告警事件用字符串表示时,字符串的相似程度,举例说明,将告警事件的告警名称、告警源、告警定位信息拼接为字符串,通过比较不同告警事件字符串的文本相似程度,认为文本相似度越高的告警事件属于同一个故障的可能性更大,应该被聚合到一个告警集合中。

[0084] 在一种可能的设计中,在步骤S404之后,还包括:接收到经过所述分类模型判断并人工矫正后的第一告警事件集合中每条告警事件的标签,以及所述每条告警事件的特征向量,将所述标签以及所述每条告警事件的特征向量保存到用于训练所述分类模型的数据库中,其中,所述标签用于表征所述每条告警事件是否为根因告警事件。

[0085] 举例说明,假设分类模型判断为根因告警事件,标记为1,派单后,运维人员发现判断错误,实际上是非根因告警事件,运维人员对其进行人工矫正,标记为0,本申请实施例中1和0可以为分别代表非根因告警事件和根因告警事件,本申请对其不做限定。

[0086] 本申请实施例中,采用的分类模型在对告警事件进行实时判断时可以根据不断更新的数据进行更新,也可以定期更新,不断更新的数据中增加的数据为运维人员矫正后的告警事件对应的标签和特征向量,最初的分类模型可以根据至少一个训练告警事件集合,按照预设分类算法训练确定的,训练告警事件集合中的告警事件可以从单机或者分布式数据库中获取,本申请对其不做限定。

[0087] 可选的,若所述第一告警事件集合只有一个告警事件,将该告警事件直接确定为根因告警事件,也可以将该告警事件中用于表征与其他告警事件之间的关系的特征设置为默认值,并按照前述步骤S401至S403的方法确定该告警事件是否为根因告警事件。

[0088] 下面通过一个完整的实施例对本申请提出的故障根因分析的方法进行详细的说明,具体如图5所示。

[0089] 步骤S501、接收各类网元设备产生各种类型的告警事件。

[0090] 步骤S502、根据时间关联度,拓扑关联度,文本相似度中的至少一个维度对将接收到的告警事件进行聚合,确定出至少一个告警事件集合。

[0091] 步骤S503、针对每一个告警事件集合,进行特征向量提取。

[0092] 举例说明,告警事件集合经过特征向量提取,并进行编码后,可以生产如下表2所示的特征向量矩阵,表2中f0表示告警严重等级,f11、f12、f13、f14的组合表示告警名称,f21、f22的组合表示事件类型,f31、f32、f33的组合表示关联属性,f4表示时序关系,f5表示数量关系,f6表示资源关系,f71、f72的组合表示告警上下文。

[0093] 表2

[0094]

f0	f11	f12	f13	f14	f21	f22	f31	f32	f33	f4	f5	f6	f71	f72
5	0	0	0	1	0	1	0	0	1	0	0.29	1	0	1
3	0	0	1	0	0	1	0	1	0	0.25	0.71	0	0	1
4	0	0	1	0	0	1	1	0	0	0.25	0.71	0	0	1
3	0	0	1	0	0	1	0	1	0	0.25	0.71	0	0	1
3	0	0	1	0	0	1	0	1	0	0.25	0.71	0	0	1
5	0	0	0	1	0	1	0	0	1	0	0.29	1	0	1
3	0	0	1	0	0	1	0	1	0	0.25	0.71	0	0	1

5	0	1	0	0	1	0	0	0	1	0	0.5	0	1	0
5	1	0	0	0	0	1	0	1	0	0	0.5	0	1	0
5	0	1	0	0	1	0	0	0	1	0.1	0.5	0	1	0
5	1	0	0	0	0	1	0	1	0	0.1	0.5	0	1	0
5	0	1	0	0	1	0	0	0	1	1	0.5	0	1	0
5	1	0	0	0	0	1	0	1	0	1	0.5	0	1	0

[0095] 步骤S504、将上述表2中任一行特征向量输入到预先设置的分类模型中,进行根因识别。

[0096] 具体的,所述分类模型中采用的分类算法可以为SVM、GBDT、决策树、随机森林、贝叶斯分类器,神经网络,但不限于上述算法,本申请对其不做限定。

[0097] 以随机森林算法为例,在进行根因识别时,即计算判断所述任一行特征向量对应的告警事件为根因告警的决策树与全部决策树的比值,当比值大于或等于设定阈值,则确定告警事件为根因告警,若判断为根因告警的决策树与全部决策树的比值小于所述设定阈值,则确定告警事件为非根因告警。举例说明:假设有m棵决策树的随机森林分类模型对告警事件A进行识别,若有m1棵决策树判定告警事件A为根因且 $m1 > m/2$,随机森林分类模型的输出为此告警事件A是根因告警事件,随机森林决策树示意图如图6所示。

[0098] 具体的,所述随机森林算法为:1) 假定训练集样本总数为N,总特征为F;2) 从训练集通过有放回的随机重采样方式产生n个样本,形成单次的训练集;3) 对单次训练集,从F个特征中选取其中的k个特征,用建立决策树的方式获得最佳分割点,生成1颗CART决策树;4) 重复m次步骤2) 和步骤3),产生m棵决策树,其中,m是由外部配置输入;5) 将m棵决策树合并即构成随机森林的模型。

[0099] 其中,每一棵CART决策树的训练过程如下:1) 设结点的训练数据集为D,CART决策树是一个2叉树,训练CART决策树的过程就是将数据集D不断切分为2个分支,直到满足算法停止条件。切分点的选择通过计算Gini指数得出;2) 设定停止条件:结点中的样本个数小于预定阈值,或样本集的基尼指数小于预定阈值,或没有更多特征;3) 对于某特征 $f_i, f_i \in F$,记 f_i 所有的可能取值为A,a是其中的某个具体值。则可将数据集分为 $A=a$,和 $A \neq a$ 两个子集D1和D2,计算集合D的基尼指数:

$$[0100] \quad \text{Gini}(D, A) = \frac{|D_1|}{|D|} \text{Gini}(D_1) + \frac{|D_2|}{|D|} \text{Gini}(D_2);$$

[0101] 4) 遍历特征 f_i 的所有取值,找出Gini指数最大的值为此特征对应的切分点;5) 遍历所有的特征,找出Gini指数最大的特征,作为分裂特征。至此可以将数据集D划分为两个子集,也即完成了针对某个具体特征的最佳分割点的选择;6) 对上述两个子节点递归调用步骤3)、4)、5),直到满足停止条件;7) 生成CART决策树,将生成m棵CART决策树保存即得了离线训练后的分类模型。

[0102] 步骤S505、识别出根因告警事件后,进行派单,运维人员对识别出的根因告警事件进行校正并增加反馈标记。若分类模型判断正确则标记为1,若分类模型判断错误,则标记为0,将标记后的数值增加到表2的特征向量矩阵中,生成表3,表3中第一列即为反馈标记,生成的表3更新到数据库中,作为分类模型的训练告警事件集合。

[0103] 表3

[0104]

label	f0	f11	f12	f13	f14	f21	f22	f31	f32	f33	f4	f5	f6	f71	f72
1	5	0	0	0	1	0	1	0	0	1	0	0.29	1	0	1
0	3	0	0	1	0	0	1	0	1	0	0.25	0.71	0	0	1
0	4	0	0	1	0	0	1	1	0	0	0.25	0.71	0	0	1
0	3	0	0	1	0	0	1	0	1	0	0.25	0.71	0	0	1
0	3	0	0	1	0	0	1	0	1	0	0.25	0.71	0	0	1
1	5	0	0	0	1	0	1	0	0	1	0	0.29	1	0	1
0	3	0	0	1	0	0	1	0	1	0	0.25	0.71	0	0	1
1	5	0	1	0	0	1	0	0	0	1	0	0.5	0	1	0
0	5	1	0	0	0	0	1	0	1	0	0	0.5	0	1	0
1	5	0	1	0	0	1	0	0	0	1	0.1	0.5	0	1	0
0	5	1	0	0	0	0	1	0	1	0	0.1	0.5	0	1	0
1	5	0	1	0	0	1	0	0	0	1	1	0.5	0	1	0
0	5	1	0	0	0	0	1	0	1	0	1	0.5	0	1	0

[0105] 本申请实施例中,分类模型是根据至少一个训练告警事件集合,按照预设分类算法训练确定的,其中,所述至少一个训练告警事件集合中的每条训练告警事件的信息包括标签和特征向量,即表3所示的内容,所述标签用于表征所述每条训练告警事件是否为根因告警事件。

[0106] 本申请实施例中,训练分类模型时训练数据的获取只需普通的运维人员在日常工作中简单判断故障根因分析装置识别的根因告警事件是否正确即可,无需专家投入,降低了人力投入,并且具有通用性,无需针对不同的产品,组网重新设计开发,避免了重复投入的问题。结合不同的数据,提取多维度特征,保证了算法的准确度;并且通过不断的学习,得到新的模型,识别准确度还会逐渐提高。

[0107] 基于与方法实施例同样的发明构思,本申请还提供了一种装置示意图,如图7所示,所述装置包括:获取单元701,用于获取第一告警事件集合,所述第一告警事件集合包括多个告警事件;处理单元702,用于对于所述第一告警事件集合中的第一告警事件,提取所述第一告警事件的特征向量,其中,所述特征向量中的部分或全部特征用于表征所述第一告警事件和所述第一告警事件集合中其它告警事件之间的关系;所述处理单元702,还用于根据所述第一告警事件的特征向量,确定所述第一告警事件是否为根因告警事件。

[0108] 在一种可能的实现方式中,所述处理单元具体用于:将所述第一告警事件的特征向量输入到预先设置的分类模型中,以确定所述第一告警事件是否为根因告警事件,其中,所述分类模型用于表征特征向量和根因告警事件之间的关系。

[0109] 在一种可能的实现方式中,所述获取单元具体用于:将多个告警事件聚合为至少一个告警事件集合,将所述至少一个告警事件集合中的一个告警事件集合作为所述第一告警事件集合。

[0110] 可选的,所述获取单元具体用于:将多个告警事件按照设定规则聚合为至少一个告警集合,其中,所述设定规则包括时间关联度、拓扑关联度、文本相似度中的至少一个。

[0111] 示例性的,所述特征向量包括告警严重等级、告警名称、事件类型中的至少一个、以及关联属性、时序关系、数量关系、资源关系中的至少一个,其中,所述关联属性用于表征

所述第一告警事件在所属的第一告警事件集合中与其他告警事件的关联关系,所述数量关系用于表征所述第一告警事件集合中包括第一告警事件在内的与第一告警事件相同的所有告警事件的个数的总和占第一告警事件集合中全部告警事件数量的比值,所述资源关系用于表征所述第一告警事件集合中第一告警事件与其它告警事件间的传播关系。

[0112] 可选的,所述特征向量还包括告警上下文,其中,所述告警上下文用于表征所述第一告警事件所属的第一告警事件集合对应的故障类型。

[0113] 在一种可能的实现方式中,所述分类模型是根据至少一个训练告警事件集合,按照预设分类算法训练确定的,其中,所述至少一个训练告警事件集合中的每条训练告警事件的信息包括标签和特征向量,所述标签用于表征所述每条训练告警事件是否为根因告警事件;

[0114] 对于所述训练告警事件集合中的任一训练告警事件,所述训练告警事件的关联属性用于表征所述训练告警事件在所述训练告警事件集合中与其他告警事件的关联关系,所述数量关系用于表征所述训练告警事件集合中包括所述训练告警事件在内的与所述训练告警事件相同的所有训练告警事件的个数的总和占所述训练告警事件集合中全部训练告警事件数量的比值,所述资源关系用于表征所述训练告警事件集合中所述训练告警事件与其它告警事件间的传播关系。

[0115] 可选的,所述训练告警事件集合中的每条训练告警事件的特征向量包括告警严重程度、告警名称、事件类型中的至少一个、以及关联属性、时序关系、数量关系、资源关系中的至少一个。

[0116] 可选的,至少一个训练告警事件集合为多个训练告警事件集合,所述多个训练告警事件集合中的每个训练告警事件集合中的每条训练告警事件的特征向量还包括告警上下文,其中,所述每条训练告警事件的告警上下文用于表征所述每条训练告警事件所在的训练告警事件集合对应的故障类型。

[0117] 在一种可能的实现方式中,确定所述第一告警事件是否为根因告警事件之后,所述获取单元还用于:接收到经过所述分类模型判断并人工矫正后的所述第一告警事件集合中每条告警事件的标签,以及所述每条告警事件的特征向量,将所述标签以及所述每条告警事件的特征向量保存到用于训练所述分类模型的数据库中,其中,所述标签用于表征所述每条告警事件是否为根因告警事件。

[0118] 可选的,若所述预设分类算法为随机森林,则所述处理单元具体用于:对于所述第一告警事件,若判断为根因告警的决策树与全部决策树的比值大于或等于设定阈值,则确定为根因告警,若判断为根因告警的决策树与全部决策树的比值小于所述设定阈值,则确定为非根因告警。

[0119] 本申请实施例中对模块的划分是示意性的,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,另外,在本申请各个实施例中的各功能模块可以集成在一个处理器中,也可以是单独物理存在,也可以两个或两个以上模块集成在一个模块中。上述集成的模块既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。

[0120] 其中,集成的模块既可以采用硬件的形式实现时,如图8所示,一种装置可以包括处理器802。上述处理单元702对应的实体的硬件可以为处理器802。处理器802,可以是一个中央处理模块(英文:central processing unit,简称CPU),或者为数字处理模块等等。所

述设备还可以包括通信接口801(可以为收发器),上述获取单元701对应的硬件实体可以为通信接口801。该设备还可以包括:存储器803,用于存储处理器802执行的程序。存储器803可以是非易失性存储器,比如硬盘(英文:hard disk drive,缩写:HDD)或固态硬盘(英文:solid-state drive,缩写:SSD)等,还可以是易失性存储器(英文:volatile memory),例如随机存取存储器(英文:random-access memory,缩写:RAM)。存储器803是能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其他介质,但不限于此。

[0121] 处理器802用于执行存储器803存储的程序代码,具体用于执行图4、图5所示实施例所述的方法。可以参见图4、图5所示实施例所述的方法,本申请在此不再赘述。

[0122] 本申请实施例中不限定上述通信接口801、处理器802以及存储器803之间的具体连接介质。本申请实施例在图8中以存储器803、处理器802以及通信接口801之间通过总线804连接,总线在图8中以粗线表示,其它部件之间的连接方式,仅是进行示意性说明,并不引以为限。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示,图8中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。

[0123] 本发明实施例还提供了一种计算机可读存储介质,用于存储为执行上述处理器所需执行的计算机软件指令,其包含用于执行上述处理器所需执行的程序。

[0124] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请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,本申请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。

[0125] 本申请中,“和/或”,描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B的情况,其中A,B可以是单数或者复数。字符“/”一般表示前后关联对象是一种“或”的关系。

[0126] 本申请中,“至少一个”是指一个或者多个,“多个”是指两个或两个以上。“以下至少一项(个)”或其类似表达,是指的这些项中的任意组合,包括单项(个)或复数项(个)的任意组合。例如,“a,b,或c中的至少一项(个)”,或,“a,b,和c中的至少一项(个)”,均可以表示:a,b,c,a-b(即a和b),a-c,b-c,或a-b-c,其中a,b,c分别可以是单个,也可以是多个。

[0127] 应理解,在本申请的各种实施例中,上述各过程的序号的大小并不意味着执行顺序的先后,部分或全部步骤可以并行执行或先后执行,各过程的执行顺序应以其功能和内在逻辑确定,而不对本申请实施例的实施过程构成任何限定。

[0128] 本申请是参照根据本申请的方法、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。

[0129] 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。

[0130] 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。

[0131] 显然,本领域的技术人员可以对本申请进行各种改动和变型而不脱离本申请的精神和范围。这样,倘若本申请的这些修改和变型属于本申请权利要求及其等同技术的范围之内,则本申请也意图包含这些改动和变型在内。

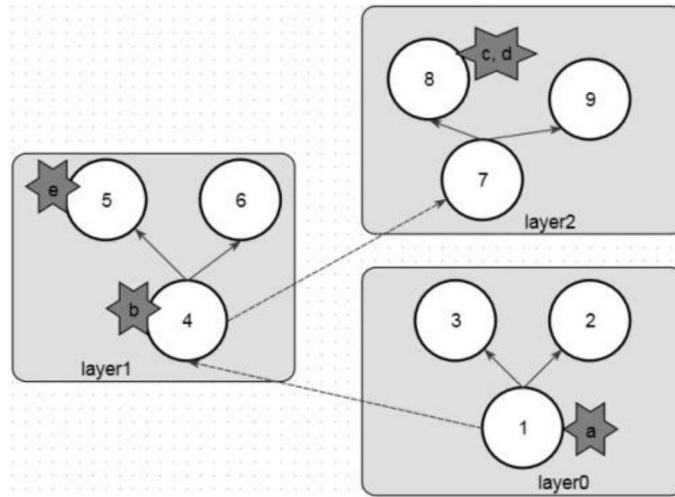


图1

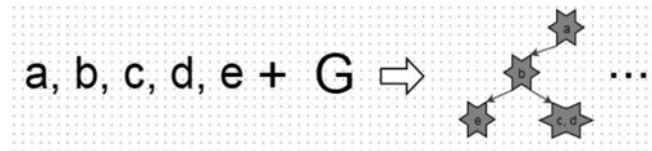


图2

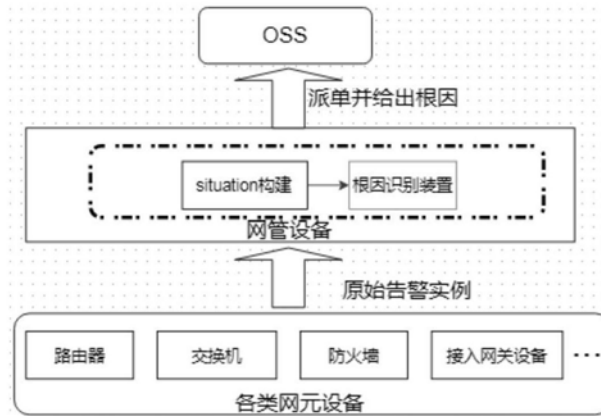


图3

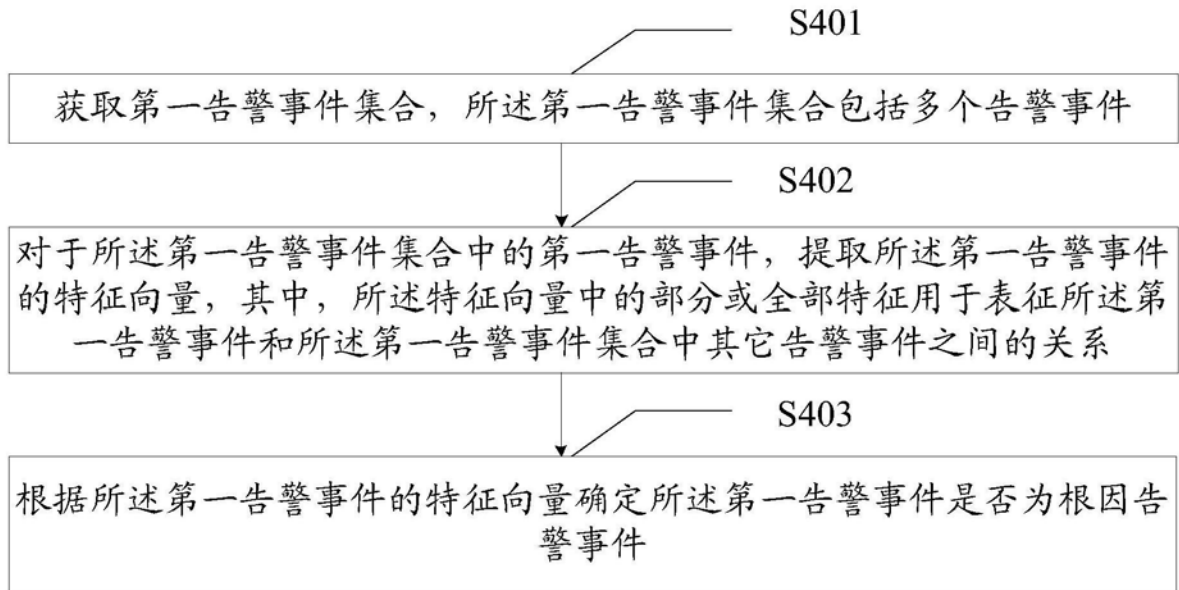


图4

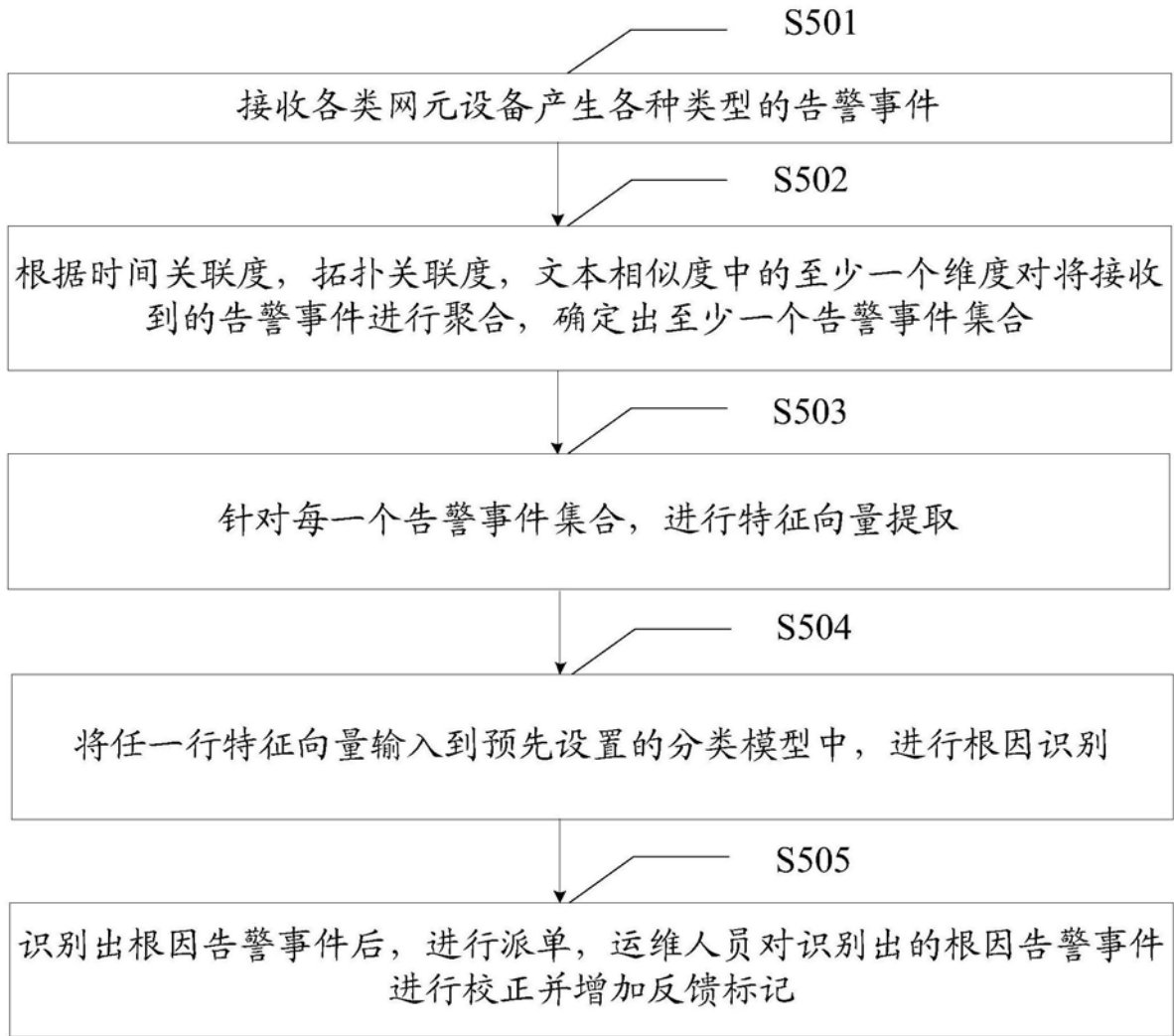


图5

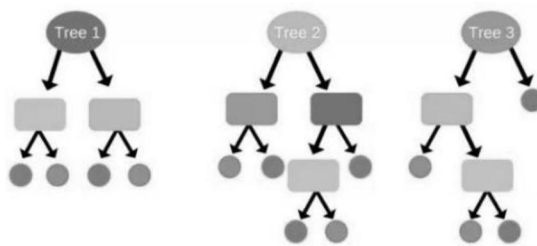


图6

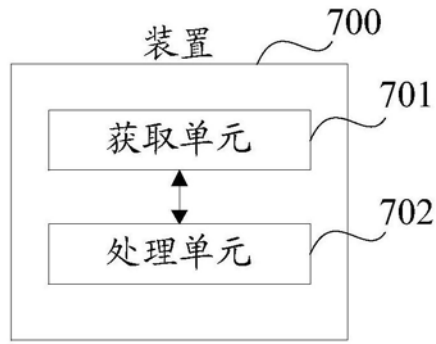


图7

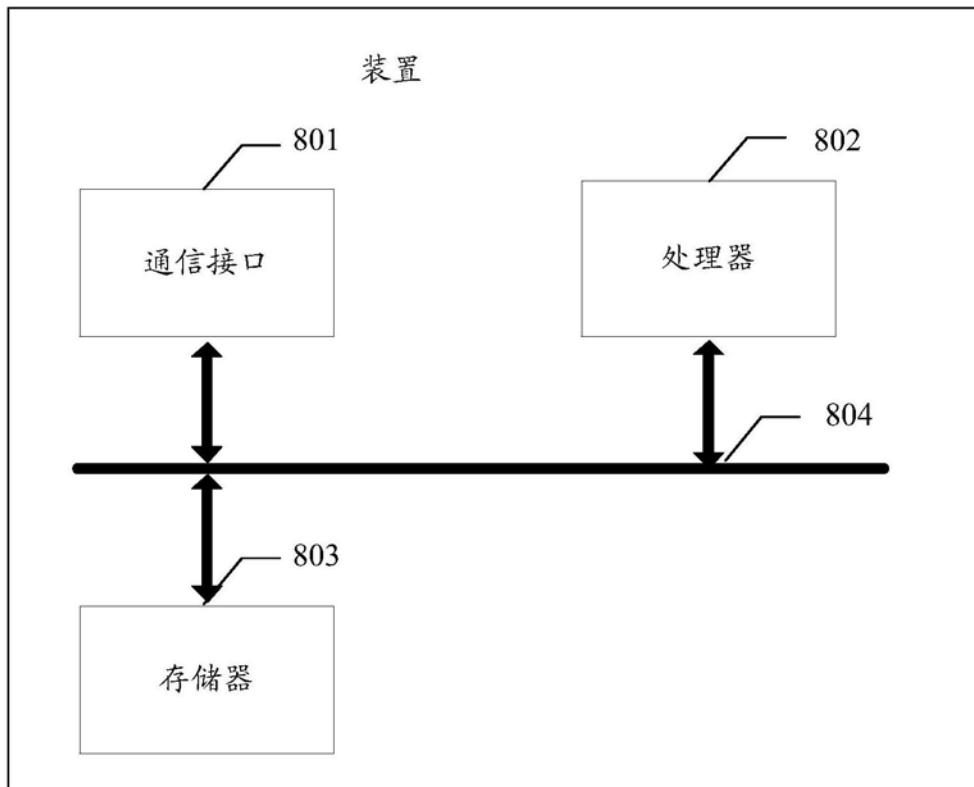


图8