(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 110347547 A
(43)申请公布日 2019.10.18

(21)申请号 201910448226.0

(22)申请日 2019.05.27

(71)申请人 中国平安人寿保险股份有限公司
地址 518000 广东省深圳市福田区益田路5033号平安金融中心14、15、16、41、44、45、46层

(72)发明人 石晓龙

(74)专利代理机构 深圳市赛恩倍吉知识产权代理有限公司 44334
代理人 杨毅玲

(51)Int.Cl.
G06F 11/30(2006.01)
G06F 21/55(2013.01)
G06F 16/35(2019.01)

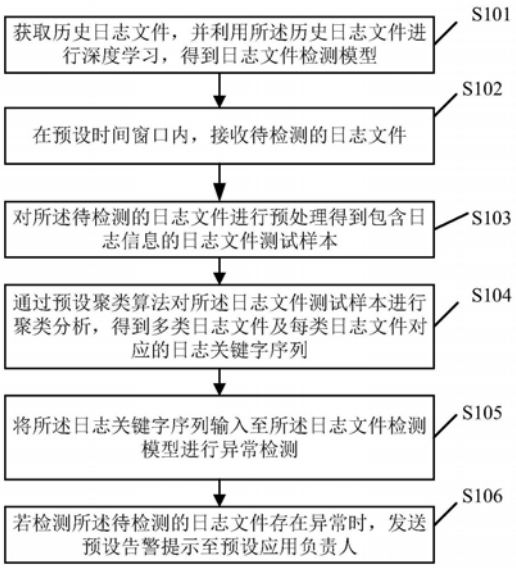
权利要求书2页 说明书10页 附图2页

(54)发明名称

基于深度学习的日志异常检测方法、装置、终端及介质

(57)摘要

本发明实施例提供一种基于深度学习的日志异常检测方法,利用历史日志文件进行深度学习,得到日志文件检测模型;在预设时间窗口内,接收待检测的日志文件;对待检测的日志文件进行预处理得到日志文件测试样本;对日志文件测试样本进行聚类分析,得到多类日志文件及每类日志文件对应的日志关键字序列;将日志关键字序列输入至日志文件检测模型进行异常检测;若存在异常时,发送预设告警提示至预设应用负责人。本发明实施例还提供一种基于深度学习的日志异常检测装置、终端以及计算机可读存储介质。本发明涉及日志监控,利用本发明实施例,可借助深度学习,自动检测日志文件是否存在异常,提高了日志告警的效率。



CN 110347547 A

1.一种基于深度学习的日志异常检测方法,其特征在于,所述基于深度学习的日志异常检测方法包括:

获取历史日志文件,并利用所述历史日志文件进行深度学习,得到日志文件检测模型;

在预设时间窗口内,接收待检测的日志文件;

对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本;

通过预设聚类算法对所述日志文件测试样本进行聚类分析,得到多类日志文件及每类日志文件对应的日志关键字序列;

将所述日志关键字序列输入至所述日志文件检测模型进行异常检测;

若检测所述待检测的日志文件存在异常时,发送预设告警提示至预设应用负责人。

2.根据权利要求1所述的基于深度学习的日志异常检测方法,其特征在于,所述利用所述历史日志文件进行深度学习,得到日志文件检测模型包括:

对所获取的所有历史日志文件中的每一个日志文件进行预处理得到包含日志信息的日志文件训练样本;

通过预设聚类算法对日志文件训练样本进行聚类分析,得到多类日志文件及每类日志文件对应的日志关键字序列;

将所述日志文件训练样本及对应的日志关键字序列输入预设的长短期记忆网络中进行训练,得到日志文件检测模型。

3.根据权利要求2所述的基于深度学习的日志异常检测方法,其特征在于,所述对所获取的所有历史日志文件中的每一个日志文件进行预处理包括:

按照预设规则对所述历史日志文件中的每一个日志文件进行切片或采样,其中,所述预设规则包括日志文件的时间序列、日志文件类型以及日志文件的来源信息。

4.根据权利要求1所述的基于深度学习的日志异常检测方法,其特征在于,所述日志文件检测模型的输入为预设时间窗口内的每类日志文件对应的日志关键字序列,输出为所有的预设日志关键字在所述关键字序列之后出现的概率向量。

5.根据权利要求4所述的基于深度学习的日志异常检测方法,其特征在于,所述将所述日志关键字序列输入至所述日志文件检测模型进行异常检测包括:

检测接收到待检测的日志文件的输出日志关键字是否为在所述日志关键字序列之后出现的概率最大的预设日志关键字;

若所述输出日志关键字不为在所述日志关键字序列之后出现概率最大的预设日志关键字,则判定所述待检测的日志文件存在异常。

6.根据权利要求5所述的基于深度学习的日志异常检测方法,其特征在于,在所述判定所述待检测的日志文件存在异常之前,所述方法还包括:

判断所述待检测的日志文件存在的异常是否为训练阶段已经出现的异常;

若判断结果为所述待检测的日志文件存在的异常为训练阶段已经出现的异常,则将报警提示发送给预设应用负责人;

若判断结果为所述待检测的日志文件存在的异常为未出现过的日志文件异常,则将所述未出现过的日志文件异常输出至第三方进行判断,其中,所述第三方包括用户、云端装置及用户与云端装置的组合。

7.根据权利要求6所述的基于深度学习的日志异常检测方法,其特征在于,所述方法还

包括：

获取在预设时间间隔内所述第三方输出的所有异常检测反馈信息；

根据所述异常检测反馈信息调整所述日志文件检测模型的参数，更新所述日志文件检测模型。

8.一种基于深度学习的日志异常检测装置，其特征在于，所述装置包括：

模型训练模块，用于获取历史日志文件，并利用所述历史日志文件进行深度学习，得到日志文件检测模型；

日志接收模块，用于在预设时间窗口内，接收待检测的日志文件；

预处理模块，用于对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本；

聚类分析模块，用于通过预设聚类算法对所述日志文件测试样本进行聚类分析，得到多类日志文件及每类日志文件对应的日志关键字序列；

异常检测模块，用于将所述日志关键字序列输入至所述日志文件检测模型进行异常检测；

告警提示模块，用于若检测所述待检测的日志文件存在异常时，发送预设告警提示至预设应用负责人。

9.一种终端，其特征在于，所述终端包括处理器，所述处理器用于执行存储器中存储的计算机程序时实现如权利要求1-7任意一项所述的基于深度学习的日志异常检测方法。

10.一种计算机可读存储介质，所述计算机可读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如权利要求1-7任意一项所述的基于深度学习的日志异常检测方法。

基于深度学习的日志异常检测方法、装置、终端及介质

技术领域

[0001] 本发明涉及云监控领域，尤其涉及一种基于深度学习的日志异常检测方法、基于深度学习的日志异常检测装置、终端以及计算机可读存储介质。

背景技术

[0002] 在当今高速发展的信息社会，人们已经离不开计算机等智能设备，随着计算机技术的成熟，运行于智能设备的应用程序虽然越来越可靠，但是实际运行过程中又难免会出现错误的情况。因此，工程技术人员通常需要通过日志系统，记录应用程序的运行状态及操作内容，以备工程技术人员查看、作为调试通信设备的依据。日志系统以日志的形式记录应用程序的各种运行状态和操作信息，生成日志文件，日志文件通常存储在本地存储设备内。

[0003] 随着数据处理、分析能力的提升，以及机器学习等人工智能技术的成熟，检测日志文件异常的技术层出不穷。现有技术在网络安全领域中通过对各类日志文件分析，进行异常发现、安全检测，进而保护应用程序运行安全。但是由于攻击方式、攻击手法具有多样性、不可预测性，基于先验知识检测手法难以应对新的异常。

发明内容

[0004] 鉴于此，有必要提供一种基于深度学习的日志异常检测方法、基于深度学习的日志异常检测装置、终端以及计算机可读存储介质，能够自动检测日志文件是否存在异常，并且实现对检测模型自动更新，提高了日志异常检测的效率。

[0005] 本发明实施例第一方面提供一种基于深度学习的日志异常检测方法，所述基于深度学习的日志异常检测方法包括：

[0006] 获取历史日志文件，并利用所述历史日志文件进行深度学习，得到日志文件检测模型；

[0007] 在预设时间窗口内，接收待检测的日志文件；

[0008] 对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本；

[0009] 通过预设聚类算法对所述日志文件测试样本进行聚类分析，得到多类日志文件及每类日志文件对应的日志关键字序列；

[0010] 将所述日志关键字序列输入至所述日志文件检测模型进行异常检测；

[0011] 若检测所述待检测的日志文件存在异常时，发送预设告警提示至预设应用负责人。

[0012] 进一步的，在本发明实施例提供的上述基于深度学习的日志异常检测方法中，所述利用所述历史日志文件进行深度学习，得到日志文件检测模型包括：

[0013] 对所获取的所有历史日志文件中的每一个日志文件进行预处理得到包含日志信息的日志文件训练样本；

[0014] 通过预设聚类算法对日志文件训练样本进行聚类分析，得到多类日志文件及每类日志文件对应的日志关键字序列；

[0015] 将所述日志文件训练样本及对应的日志关键字序列输入预设的长短期记忆网络中进行训练，得到日志文件检测模型。

[0016] 进一步的，在本发明实施例提供的上述基于深度学习的日志异常检测方法中，所述对所获取的所有历史日志文件中的每一个日志文件进行预处理包括：

[0017] 按照预设规则对所述历史日志文件中的每一个日志文件进行切片或采样，其中，所述预设规则包括日志文件的时间序列、日志文件类型以及日志文件的来源信息。

[0018] 进一步的，在本发明实施例提供的上述基于深度学习的日志异常检测方法中，所述日志文件检测模型的输入为预设时间窗口内的每类日志文件对应的日志关键字序列，输出为所有的预设日志关键字在所述关键字序列之后出现的概率向量。

[0019] 进一步的，在本发明实施例提供的上述基于深度学习的日志异常检测方法中，所述将所述日志关键字序列输入至所述日志文件检测模型进行异常检测包括：

[0020] 检测接收到待检测的日志文件的输出日志关键字是否为在所述日志关键字序列之后出现的概率最大的预设日志关键字；

[0021] 若所述输出日志关键字不为在所述日志关键字序列之后出现概率最大的预设日志关键字，则判定所述待检测的日志文件存在异常。

[0022] 进一步的，在本发明实施例提供的上述基于深度学习的日志异常检测方法中，在所述判定所述待检测的日志文件存在异常之前，所述方法还包括：

[0023] 判断所述待检测的日志文件存在的异常是否为训练阶段已经出现的异常；

[0024] 若判断结果为所述待检测的日志文件存在的异常为训练阶段已经出现的异常，则将报警提示发送给预设应用负责人；

[0025] 若判断结果为所述待检测的日志文件存在的异常为未出现过的日志文件异常，则将所述未出现过的日志文件异常输出至第三方进行判断，其中，所述第三方包括用户、云端装置及用户与云端装置的组合。

[0026] 进一步的，在本发明实施例提供的上述基于深度学习的日志异常检测方法中，所述方法还包括：

[0027] 获取在预设时间间隔内所述第三方输出的所有异常检测反馈信息；

[0028] 根据所述异常检测反馈信息调整所述日志文件检测模型的参数，更新所述日志文件检测模型。

[0029] 本发明实施例第二方面还提供一种基于深度学习的日志异常检测装置，所述装置包括：

[0030] 模型训练模块，用于获取历史日志文件，并利用所述历史日志文件进行深度学习，得到日志文件检测模型；

[0031] 日志接收模块，用于在预设时间窗口内，接收待检测的日志文件；

[0032] 预处理模块，用于对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本；

[0033] 聚类分析模块，用于通过预设聚类算法对所述日志文件测试样本进行聚类分析，得到多类日志文件及每类日志文件对应的日志关键字序列；

[0034] 异常检测模块，用于将所述日志关键字序列输入至所述日志文件检测模型进行异常检测；

[0035] 告警提示模块，用于若检测所述待检测的日志文件存在异常时，发送预设告警提示至预设应用负责人。

[0036] 本发明实施例第三方面还提供一种终端，所述终端包括处理器，所述处理器用于执行存储器中存储的计算机程序时实现上述任意一项所述的基于深度学习的日志异常检测方法。

[0037] 本发明实施例第四方面还提供一种计算机可读存储介质，所述计算机可读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意一项所述的基于深度学习的日志异常检测方法。

[0038] 本发明实施例提供一种基于深度学习的日志异常检测方法、基于深度学习的日志异常检测装置、终端以及计算机可读存储介质，获取历史日志文件，并利用所述历史日志文件进行深度学习，得到日志文件检测模型；在预设时间窗口内，接收待检测的日志文件；对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本；通过预设聚类算法对所述日志文件测试样本进行聚类分析，得到多类日志文件及每类日志文件对应的日志关键字序列；将所述日志关键字序列输入至所述日志文件检测模型进行异常检测；若检测所述待检测的日志文件存在异常时，发送预设告警提示至预设应用负责人。利用本发明实施例，可借助深度学习，自动检测日志文件是否存在异常，且神经网络具备高度学习和自适应能力，能够自动调整模型参数，从而更新检测模型，提高了日志告警的效率。

附图说明

[0039] 图1是本发明第一实施方式提供的基于深度学习的日志异常检测方法的流程图。

[0040] 图2是本发明一实施方式的终端的结构示意图。

[0041] 图3是图2所示的终端的示例性的功能模块图。

[0042] 主要元件符号说明

[0043]

终端	1
存储器	10

[0044]

显示屏	20
处理器	30
基于深度学习的日志异常检测装置	100
模型训练模块	101
日志接收模块	102
预处理模块	103
聚类分析模块	104
异常检测模块	105
告警提示模块	106

[0045] 如下具体实施方式将结合上述附图进一步说明本发明实施例。

具体实施方式

[0046] 为了能够更清楚地理解本发明实施例的上述目的、特征和优点，下面结合附图和具体实施方式对本发明进行详细描述。需要说明的是，在不冲突的情况下，本申请的实施方式中的特征可以相互组合。

[0047] 在下面的描述中阐述了很多具体细节以便于充分理解本发明实施例，所描述的实施方式仅仅是本发明一部分实施方式，而不是全部的实施方式。基于本发明中的实施方式，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方式，都属于本发明实施例保护的范围。

[0048] 除非另有定义，本文所使用的所有的技术和科学术语与属于本发明实施例的技术领域的技术人员通常理解的含义相同。本文中在本发明的说明书中所使用的术语只是为了描述具体的实施方式的目的，不是旨在于限制本发明实施例。

[0049] 图1是本发明第一实施方式提供的基于深度学习的日志异常检测方法的流程图。所述基于深度学习的日志异常检测方法可以应用于终端，所述终端可以是例如智能手机、笔记本电脑、台式/平板电脑、智能手表等智能设备。如图1所示，所述基于深度学习的日志异常检测方法可以包括如下步骤：

[0050] S101：获取历史日志文件，并利用所述历史日志文件进行深度学习，得到日志文件检测模型。

[0051] 在本实施方式中，获取历史日志文件，并利用所述历史日志文件进行深度学习，得到日志文件检测模型，所述日志文件检测模型可以基于LSTM(长短期记忆网络模型)为初始模型进行训练得到的。所述历史日志文件包括已进行异常检测的日志文件与未进行异常检测的日志文件。所述利用所述历史日志文件进行深度学习，得到日志文件检测模型包括：对所获取的所有历史日志文件中的每一个日志文件进行预处理得到包含日志信息的日志文

件训练样本;通过预设聚类算法对日志文件训练样本进行聚类分析,得到多类日志文件及每类日志文件对应的日志关键字序列;将所述日志文件训练样本及对应的日志关键字序列输入预设的长短期记忆网络中进行训练,得到日志文件检测模型。

[0052] 其中,所述对所获取的所有历史日志文件中的每一个日志文件进行预处理包括:按照预设规则对所述历史日志文件中的每一个日志文件进行切片或采样,其中,所述预设规则包括日志文件的时间序列、日志文件类型以及日志文件的来源信息。所述按照预设规则对所述历史日志文件中的每一个日志文件进行切片或采样包括:按照日志文件的时间序列对日志文件进行切片或采样;或者,按照日志文件的类型对日志文件进行切片或采样;或者,按照日志文件的来源信息对日志文件进行切片或采样。

[0053] 具体的,以预设规则为日志文件的时间序列为例,所述按照日志文件的时间序列对日志文件进行切片或采样。假设目前获取的历史日志文件的数量为10,按照日志文件的时间序列先后顺序分别为日志1、日志2、…日志10。对日志文件进行切片或采样的处理是通过云节点进行的,所述云节点之间可以相互通信。假设目前可用于对日志文件切片或采样进行底层处理的云节点的数量为5个,即第一至第五云节点。可以将日志1与日志2作为一个日志数据切片发送给第一云节点,将日志3与日志4作为一个日志数据切片发送给第二云节点,将日志5与日志6作为一个日志数据切片发送给第三云节点,将日志7与日志8作为一个日志数据切片发送给第四云节点,将日志9与日志10作为一个日志数据切片发送给第五云节点。

[0054] 所述通过预设聚类算法对日志文件训练样本进行聚类分析,得到多类日志文件及每类日志文件对应的日志关键字序列。所述预设聚类算法为终端用户预先设置的。所述预设聚类算法可以包括:K-Means(K均值)聚类算法、均值漂移聚类算法、基于密度的聚类算法、用高斯混合模型的最大期望聚类算法、凝聚层次聚类算法以及图团体检测算法。按照所述预设聚类算法对日志文件训练样本进行分类处理,以生成相似日志记录的组(也即生成多类日志文件),每类日志文件对应日志关键字序列,通过聚类算法实现将整个日志文件训练样本转换为一个离散事件序列。

[0055] S102:在预设时间窗口内,接收待检测的日志文件。

[0056] 在本实施方式中,在预设时间窗口内,接收待检测的日志文件。所述预设时间窗口为终端用户预先设置的时间窗口,所述预设时间窗口可以预设为1分钟、5分钟或者更长的时间。当业务访问量大时,预设时间窗口可以根据具体情况相应缩短,反之可延长所述预设时间窗口。所述待检测的日志文件可以源自任何日志产生源位置,例如数据库管理系统、数据库应用、中间件、硬件日志、操作系统日志、应用日志、应用服务器日志、数据库服务器日志以及监视系统或应用的行为的任何其他类型的日志。

[0057] S103:对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本。

[0058] 在本实施方式中,所述对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本包括:按照预设规则对所述待检测的日志文件中的每一个日志文件进行切片或采样,所述预设规则包括日志文件的时间序列、日志文件类型以及日志文件的来源信息。

[0059] S104:通过预设聚类算法对所述日志文件测试样本进行聚类分析,得到多类日志

文件及每类日志文件对应的日志关键字序列。

[0060] 在本实施方式中，通过预设聚类算法对所述日志文件测试样本进行聚类分析，得到多类日志文件及每类日志文件对应的日志关键字序列。所述预设聚类算法为终端用户预先设置的。所述预设聚类算法包括：K-Means(K均值)聚类算法、均值漂移聚类算法、基于密度的聚类算法、用高斯混合模型的最大期望聚类算法、凝聚层次聚类算法以及图团体检测算法。按照所述预设聚类算法对日志文件测试样本进行分类处理，以生成相似日志记录的组(也即生成多类日志文件)，每类日志文件对应日志关键字序列，通过聚类算法实现将整个日志文件测试样本转换为一个离散事件序列。

[0061] 在聚类分析过程中，每创建一个新的聚类，系统会分配给新的聚类一个编号作为标识。聚类分析可以自动将相似度高的数据划分到同一个类中，而不同类的数据对象之间的相似度很小。在所述日志文件进行预处理后，对所述日志文件进行聚类分析，能够最大化提取保存有效的数据信息，保证有用信息的不丢失。

[0062] S105：将所述日志关键字序列输入至所述日志文件检测模型进行异常检测。

[0063] 在本实施方式中，将所述日志关键字序列输入至所述日志文件检测模型进行异常检测。所述日志文件检测模型的输入为预设时间窗口内的每类日志文件对应的日志关键字序列，输出为所有的预设日志关键字在所述日志关键字序列之后出现的概率向量。所述将所述日志关键字序列输入至所述日志文件检测模型进行异常检测包括：检测接收到待检测的日志文件的输出日志关键字是否为在所述日志关键字序列之后出现的概率最大的预设日志关键字；若所述输出日志关键字不为在所述日志关键字序列之后出现概率最大的预设日志关键字，则判定所述待检测的日志文件存在异常。

[0064] 作为可用于基于深度学习的日志异常检测方法的示例，考虑总是在日志文件的一致时间内以一致的概率出现的某种类型的日志关键字。例如，一段时间内所述日志关键字序列为{k1,k2,k3,k4,k5,k6}，读取日志的窗口为3，则输入序列和输出序列分别为{k1,k2,k3->k4}，{k2,k3,k4->k5}和{k3,k4,k5->k6}(以输入序列和输出序列为{k1,k2,k3->k4}为例，其中，{k1,k2,k3}为输入序列，{k4}为输出序列)。如果这种类型的日志关键字突然不以相同的概率出现在日志文件中(也即，对于输入序列为{k1,k2,k3}的序列来说，在该序列之后出现概率最大的输出序列应该为{k4}，但实际输出的日志关键字不为{k4})，则可以向应用负责人通知可能存在值得注意的异常。可以理解的是，相反的情况也可以构成日志文件的异常，例如，其中某种类型的日志关键字完全不或者只是很罕见地出现在日志文件中，但是突然以大得多的概率在一天中的错误时间出现在日志文件中，则可以向预设应用负责人通知可能存在系统问题或安全漏洞的异常。

[0065] 由于各种攻击手法、方式具有多样性、不可预测性，日志文件中可能会存在新的威胁，可能会出现在训练阶段没有训练的某种新的网络行为的日志文件。所述日志文件存在的异常可以包括训练阶段已经出现的异常以及未出现过的日志文件异常。本发明实施例还提供一种日志训练异常收集库，日志训练异常收集库中用于存放训练阶段已经出现的异常。在本实施方式中，所述方法还包括：通过调整日志文件检测模型的参数更新所述日志文件检测模型。对所述日志文件检测模型进行更新，以便适应随着时间推移而出现的新的日志异常。所述日志文件检测模型的参数包括模型的权重信息与偏置信息，所述权限信息与偏置信息可以在模型训练过程中进行调整。

[0066] 在所述判定所述待检测的日志文件存在异常之前，所述方法还包括：判断所述待检测的日志文件存在的异常是否为训练阶段已经出现的异常(也即判断所述待检测的日志文件存在的异常是否存在于日志训练异常收集库中)；若判断结果为所述待检测的日志文件存在的异常为训练阶段已经出现的异常，则将报警提示发送给预设应用负责人；若判断结果为所述待检测的日志文件存在的异常为未出现过的日志文件异常，则将所述未出现过的日志文件异常输出至第三方进行判断。其中，所述第三方包括用户(例如，让用户在基于对自身业务和安全问题的理解来处理日志文件异常的问题)、云端装置(例如，通过云端联动，获取云端其他用户处理的日志文件异常情况，根据云端其他用户处理的日志文件异常情况来处理当前日志文件异常的问题)及用户与云端装置的组合(例如，先将检测结果发送至云端，查看是否存在其他用户处理过这种日志文件异常，当查看到未存在其他用户处理过这种日志文件异常时，再将查看结果发送给用户，让用户在基于其对自身业务和安全问题的理解来处理当前日志文件异常的问题)。

[0067] 可以理解的是，未出现过的日志文件异常作为日志文件检测模型没有训练过的一种异常，其可能存在威胁，也可能不存在威胁。同理可知，作为对未出现过的日志文件异常的判断，异常检测反馈信息可能判断未出现过的日志文件特征存在威胁，也可能判断未出现过的日志文件特征不存在威胁。也就是说，当异常检测反馈信息判定未出现过的日志文件特征确实存在威胁时，更新日志文件检测模型，在下一次遇到该情况时，将其判定为训练阶段已经出现的异常，并向预设应用负责人发送报警提示；当异常检测反馈信息判定未出现过的日志文件特征不存在威胁时，也需更新检测模型，在下一次遇到该情况时，判定日志文件无异常。

[0068] 在所述第三方向所述基于深度学习的日志异常检测装置100发送异常检测反馈信息之后，所述方法还包括：获取在预设时间间隔内所述第三方输出的所有异常检测反馈信息；根据所述异常检测反馈信息调整所述日志文件检测模型的参数，更新所述日志文件检测模型。所述预定时间间隔可以为终端用户预先设置的，例如，所述预定时间间隔为3天。

[0069] 所述根据所述异常检测反馈信息调整所述日志文件检测模型的参数包括自动调整参数与手动控制调整参数。具体的，日志文件检测模型自动调整参数包括：根据所述异常检测反馈信息自动对所述检测模型再次进行训练，自动更新所述日志文件检测模型的参数，从而更新所述日志文件检测模型，以便再次遇到该情况时能够做出准确的判断。所述手动控制调整参数包括：检测是否接收到终端用户输出的更新日志文件检测模型的指令；若检测到接收到终端用户输出的更新日志文件检测模型的指令，则根据所述异常检测反馈信息对所述检测模型再次训练，更新所述日志检测模型。

[0070] S106：若检测所述待检测的日志文件存在异常时，发送预设告警提示至预设应用负责人。

[0071] 在本实施方式中，若检测到所述待检测的日志文件存在异常时，发送预设告警提示至预设应用负责人。可以理解的是，不同的服务器对应不同的测试系统，不同的测试系统都对应预设应用负责人。也即所述预设应用负责人为终端用户预先设置的，能够解决其对应的测试系统所接收到的日志文件异常问题的负责人。将不同的测试系统对应不同的预设应用负责人，从而能够避免在异常信息处理的过程中由于应用负责人不擅长该领域而不能够准确分析部分异常信息或者分析日志文件的效率以及由于异常信息过多导致技术人员

漏掉关键日志文件的情况，能够更高效、准确的完成日志文件的分析。发送告警提示至预设应用负责人的方式包括但不限于邮件通知、短信通知、电话通知等形式。

[0072] 本发明实施例提供一种基于深度学习的日志异常检测方法，获取历史日志文件，并利用所述历史日志文件进行深度学习，得到日志文件检测模型；在预设时间窗口内，接收待检测的日志文件；对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本；通过预设聚类算法对所述日志文件测试样本进行聚类分析，得到多类日志文件及每类日志文件对应的日志关键字序列；将所述日志关键字序列输入至所述日志文件检测模型进行异常检测；若检测所述待检测的日志文件存在异常时，发送预设告警提示至预设应用负责人。利用本发明实施例，可借助深度学习，自动检测日志文件是否存在异常，且神经网络具备高度学习和自适应能力，能够调整模型参数，更新日志文件检测模型，提高了日志异常检测的效率。

[0073] 以上是对本发明实施例所提供的方法进行的详细描述。根据不同的需求，所示流程图中方块的执行顺序可以改变，某些方块可以省略。下面对本发明实施例所提供的终端进行描述。

[0074] 本发明实施例还提供一种终端，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述程序时实现上述任一实施方式中所述的基于深度学习的日志异常检测方法的步骤。

[0075] 图2是本发明一实施方式的终端的结构示意图，如图2所示，终端1包括存储器10，存储器10中存储有基于深度学习的日志异常检测装置100。所述的终端1可以是手机、平板电脑、个人数字助理等具有应用显示功能的终端。所述基于深度学习的日志异常检测装置100可以获取历史日志文件，并利用所述历史日志文件进行深度学习，得到日志文件检测模型；在预设时间窗口内，接收待检测的日志文件；对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本；通过预设聚类算法对所述日志文件测试样本进行聚类分析，得到多类日志文件及每类日志文件对应的日志关键字序列；将所述日志关键字序列输入至所述日志文件检测模型进行异常检测；若检测所述待检测的日志文件存在异常时，发送预设告警提示至预设应用负责人。利用本发明实施例，可借助深度学习，自动检测日志文件是否存在异常，提高了日志告警的效率。

[0076] 本实施方式中，终端1还可以包括显示屏20及处理器30。存储器10、显示屏20可以分别与处理器30电连接。

[0077] 所述的存储器10可以是不同类型存储设备，用于存储各类数据。例如，可以是终端1的存储器、内存，还可以是可外接于该终端1的存储卡，如闪存、SM卡(Smart Media Card，智能媒体卡)、SD卡(SecureDigital Card，安全数字卡)等。此外，存储器10可以包括高速随机存取存储器，还可以包括非易失性存储器，例如硬盘、内存、插接式硬盘，智能存储卡(Smart Media Card,SMC)，安全数字(Secure Digital,SD)卡，闪存卡(Flash Card)、至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。存储器10用于存储各类数据，例如，所述终端1中安装的各类应用程序(Applications)、应用上述基于深度学习的日志异常检测方法而设置、获取的数据等信息。

[0078] 显示屏20安装于终端1，用于显示信息。

[0079] 处理器30用于执行所述基于深度学习的日志异常检测方法以及所述终端1内安装

的各类软件，例如操作系统及应用显示软件等。处理器30包含但不限于处理器(Central Processing Unit，CPU)、微控制单元(Micro Controller Unit，MCU)等用于解释计算机以及处理计算机软件中的数据的装置。

[0080] 所述的基于深度学习的日志异常检测装置100可以包括一个或多个的模块，所述一个或多个模块被存储在终端1的存储器10中并被配置成由一个或多个处理器(本实施方式为一个处理器30)执行，以完成本发明实施例。例如，参阅图3所示，所述的基于深度学习的日志异常检测装置100可以包括模型训练模块101、日志接收模块102、预处理模块103、聚类分析模块104、异常检测模块105与告警提示模块106。本发明实施例所称的模块可以是完成一特定功能的程序段，比程序更适合于描述软件在处理器中的执行过程。

[0081] 可以理解的是，对应上述基于深度学习的日志异常检测方法中的各实施方式，终端1可以包括图3中所示的各功能模块中的一部分或全部，各模块的功能将在以下具体介绍。需要说明的是，以上基于深度学习的日志异常检测方法的各实施方式中相同的名词相关名词及其具体的解释说明也可以适用于以下对各模块的功能介绍。为节省篇幅及避免重复起见，在此就不再赘述。

[0082] 模型训练模块101可以用于获取历史日志文件，并利用所述历史日志文件进行深度学习，得到日志文件检测模型。

[0083] 日志接收模块102可以用于在预设时间窗口内，接收待检测的日志文件。

[0084] 预处理模块103可以用于对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本。

[0085] 聚类分析模块104可以用于通过预设聚类算法对所述日志文件测试样本进行聚类分析，得到多类日志文件及每类日志文件对应的日志关键字序列。

[0086] 异常检测模块105可以用于将所述日志关键字序列输入至所述日志文件检测模型进行异常检测。

[0087] 告警提示模块106可以用于若检测所述待检测的日志文件存在异常时，发送预设告警提示至预设应用负责人。

[0088] 本发明实施例还提供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实施方式中的基于深度学习的日志异常检测方法的步骤。

[0089] 所述基于深度学习的日志异常检测装置/终端/计算机设备集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本发明实现上述实施方式方法中的全部或部分流程，也可以通过计算机程序来指令相关的硬件来完成，所述的计算机程序可存储于一计算机可读存储介质中，该计算机程序在被处理器执行时，可实现上述各个方法实施例的步骤。其中，所述计算机程序包括计算机程序代码，所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读存储介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM，Read-Only Memory)、随机存取存储器(RAM，Random Access Memory)、电载波信号、电信信号以及软件分发介质等。

[0090] 所称处理器可以是中央处理单元(Central Processing Unit，CPU)，还可以是其

他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等,所述处理器是所述基于深度学习的日志异常检测装置/终端的控制中心,利用各种接口和线路连接整个基于深度学习的日志异常检测装置/终端的各个部分。

[0091] 所述存储器用于存储所述计算机程序和/或模块,所述处理器通过运行或执行存储在所述存储器内的计算机程序和/或模块,以及调用存储在存储器内的数据,实现所述基于深度学习的日志异常检测装置/终端的各种功能。所述存储器可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序(比如声音播放功能、图像播放功能等)等;存储数据区可存储根据手机的使用所创建的数据(比如音频数据、电话本等)等。

[0092] 在本发明所提供的几个具体实施方式中,应该理解到,所揭露的终端和方法,可以通过其它的方式实现。例如,以上所描述的系统实施方式仅仅是示意性的,例如,所述模块的划分,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。

[0093] 对于本领域技术人员而言,显然本发明实施例不限于上述示范性实施例的细节,而且在不背离本发明实施例的精神或基本特征的情况下,能够以其他的具体形式实现本发明实施例。因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明实施例的范围由所附权利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明实施例内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。系统、装置或终端权利要求中陈述的多个单元、模块或装置也可以由同一个单元、模块或装置通过软件或者硬件来实现。

[0094] 以上实施方式仅用以说明本发明实施例的技术方案而非限制,尽管参照以上较佳实施方式对本发明实施例进行了详细说明,本领域的普通技术人员应当理解,可以对本发明实施例的技术方案进行修改或等同替换都不应脱离本发明实施例的技术方案的精神和范围。

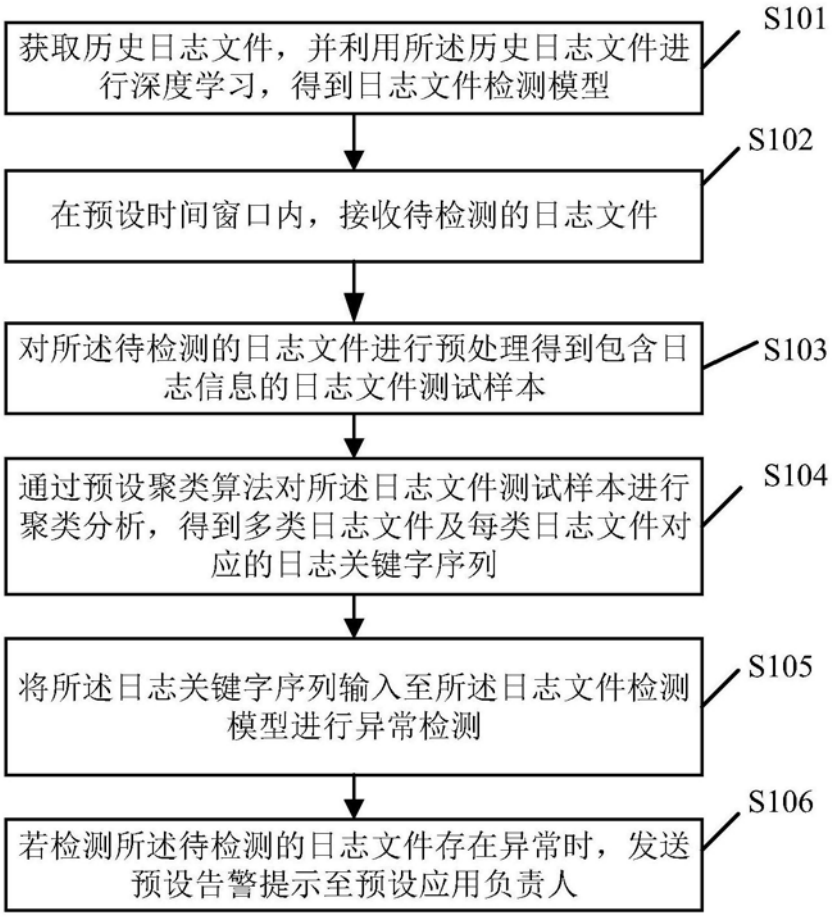

S101
获取历史日志文件，并利用所述历史日志文件进行深度学习，得到日志文件检测模型
S102
在预设时间窗口内，接收待检测的日志文件
S103
对所述待检测的日志文件进行预处理得到包含日志信息的日志文件测试样本
S104
通过预设聚类算法对所述日志文件测试样本进行聚类分析，得到多类日志文件及每类日志文件对应的日志关键字序列
S105
将所述日志关键字序列输入至所述日志文件检测模型进行异常检测
S106
若检测所述待检测的日志文件存在异常时，发送预设告警提示至预设应用负责人

图1

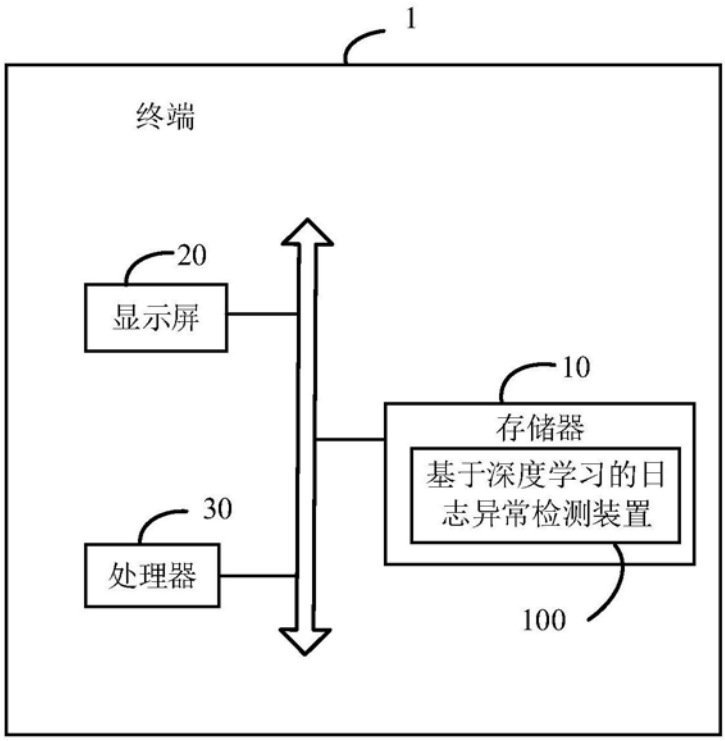

1
终端
20
显示屏
10
存储器
基于深度学习的日志异常检测装置
30
处理器
100

图2

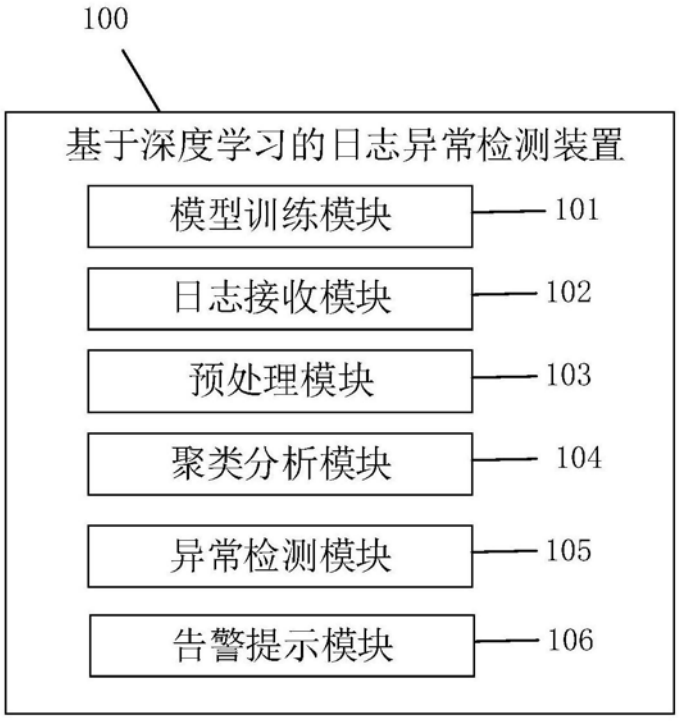

100
基于深度学习的日志异常检测装置
模型训练模块
101
日志接收模块
102
预处理模块
103
聚类分析模块
104
异常检测模块
105
告警提示模块
106

图3