基于机器学习的智能运维预警系统

陈重韬，盛红雷，张实君，来　骥
国网冀北电力有限公司 信息通信分公司，北京 100000

摘　要：提出一种基于机器学习的智能运维预警系统，通过机器学习的方式，从采集的指标数据中学习各指标的特征与规律，通过对新采集的指标数据进行分析，检测其中的异常数据，并能够通过学习个性化的预警规则，实现智能预警的功能。

关键词：机器学习；智能运维；在线检测；个性预警

1　引言

随着网络的快速发展，出现了需要为大量用户服务的网络系统。这些网络系统具有分布在各个地址的大量计算机(服务器)或计算资源，而这些计算机或计算资源通常构建为集群的方式来为用户服务。随着提供服务的计算机或计算资源越来越多，对这些计算机或计算资源的各项指标进行监控并在出现故障时及时、准确地进行预警是非常重要的问题。以数据中心系统为例，需要对数据中心的计算机与计算资源的各项指标进行监控，以此发现数据中心系统发生的异常状况，使运维人员能够及早排除故障，保证系统的稳定运行。目前的监控方法主要为运维人员人工查看运维指标或采取设定固定阈值的方式对指标进行监控。人工查看运维指标除了需要大量人力之外，也极易在大量数据中出现遗漏的异常情况，并且当数据量增涨到一定程度之后，依靠人工查看的方式也变得不可行。而设定固定阈值进行监控的方法，要求对每一种指标都设置合理的阈值，当指标数量巨大时，该方法显得不太可行。除此之外，设定固定阈值的方式也只能对符合简单规则的异常情况进行预警，在复杂的实际生产环境中，极易产生大量误报。

2　系统框架与实现

运维，指运行维护，与研发、测试、系统管理同为互联网产品技术四大支撑，其核心目标是将交付的业务软件和硬件基础设施高效合理的整合，转换为可持续提供高质量服务的产品，同时最大限度降低服务运行的成本，保障服务运行的安全。运维的一个技术任务是提供服务故障管理。本文所提出的运维预警系统，针对设备或服务提供指标检测及预警，是运维服务故障管理中的一部分。

以数据中心为应用场景作为例子对运维预警系统进行介绍。数据中心是多方协作的特定设备网络，用来在 internet 网络基础设施上传递、加速、展示、计算、存储数据信息。针对数据中心，需要对各种指标进行监控，比如网络流量、存储空间、各个机器的 CPU、内存、网卡等指标数据进行监控。系统架构示意图如图 1 所示。

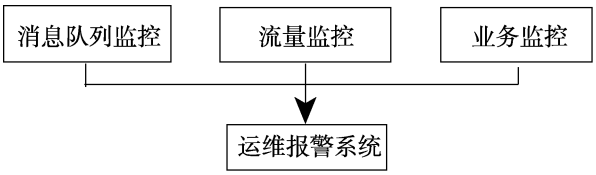

图 1　智能运维预警系统架构

数据中心是由很多计算资源、网络资源、存储资源构成的系统。在云计算中，计算资源(computing resource)，主要是指由设备或虚拟机提供计算能力的资源。在数据中心中，运维预警系统部署在数据监控设备的下游，如图 1 的例子，运维预警系统接收来自消息队列监控、流量监控、业务监控的监控数据，获取到待检测及预警的指标数据流。

运维预警系统用于对网络系统中设备或计算资源的指标数据进行检测和异常预警。如图 2 所示，该系统包括：离线模型训练模块 1、在线检测模块 2 和预警模块 3。

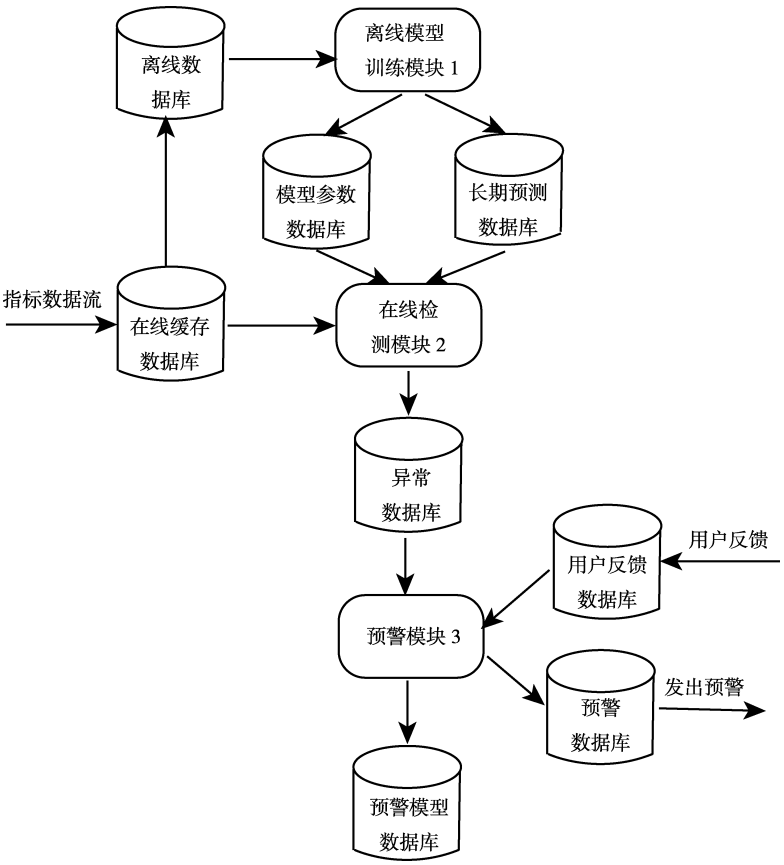

图 2　智能运维预警系统结构

将图 1 中的监控数据以数据流的形式输入到运维预警系统中。对于新输入的数据，在线检测模块 2 会使用各种检测算法对其中的异常数据进行检测。在线检测模块 2 中各检测算法所需的参数等数据由离线模型训练模块 1 计算获得。对于检测出来的异常数据，会通过预警模块 3，结合用户的反馈信息对用户进行预警。其主要的结构模块以及实现功能为：

（1）离线模型训练模块 1

主要用于基于机器学习对在线检测模块 2 所使用的检测算法所需的参数进行更新，以及基于历史数据的分析结果对指标数据的分布(趋势)进行预测。

可见，离线模型训练模块 1 主要有以下两种功能：

①通过机器学习的方式更新在线检测模块 2 各检测算法所需的参数。

具体的参数，是依据所使用的算法而变化的。例如，对于 GEN-ESD 算法、基于 LOWESS 平滑的分位数统计方法、基于 RPCA(数据降维)的检测算法、阈值判断法等。

②通过对历史数据分析，对指标数据未来的长期分布进行预测。

针对数据中心这个应用场景，具体的指标主要有机器的 CPU、内存、网卡监控数据、服务的 QPS(每秒查询率)、请求延时、网站登录信息等。

预测的内容可以包括两部分：一是各指标在不同时刻的取值的分布情况，二是具有周期性的指标在未来几个周期内的期望数值。

可以设置三个数据库协同离线模型训练模块 1 进行数据的存储。参考图 2，它们是离线数据库、模型参数数据库、长期预测数据库。离线模型训练模块会将相关数据存储于离线数据库、模型参数数据库以及长期预测数据库这三种数据库中。

（a）离线数据库用于存储离线模型训练所需的一定时间窗口内的运维指标数据；

（b）模型参数数据库用于存储在线检测模块 2 各检测算法所需的参数；

（c）长期预测数据库用于存储在离线模型训练模块 1 中对运维指标进行长期分布预测的结果。

上述设置离线数据库、模型参数数据库、长期预测数据库这三个数据库的方式是优选方式，实际上，可以采取其他方式实现，例如，仅设置一个数据库进行分块存储不同数据，或者，在离线模型训练模块 1 所在设备上对其所需或处理结果的数据进行分类存储。

（2）在线检测模块 2

主要用于接收待检测的指标数据，并基于离线模型训练模块 1 对指标数据的分布的预测结果，利用预置的检测算法对指标数据进行检测。

在线检测模块接收针对数据中心所采集的指标数据流，使用多种检测算法对不同的异常情况进行检测。例如，可以通过 kafka(一种高吞吐量的分布式发布订阅消息系统)采集数据，然后需要进行异常检测的指标通过 kafka 接入运维预警系统，而且，还可以设定在系统中进行数据登记，只对登记过的数据才能进行检测。

对于不同的指标可以根据指标各自的特性，组合各种检测算法对其进行检测。例如，假设对某种指标，某种算法检测出的异常总是被用户反馈为不感兴趣，那么在之后的异常检测中将不再继续使用该检测算法进行检测。又例如，如果某种算法对某种类型的指标的检测效果非常好(通过用户反馈获知效果)，那么将对之前没有使用该算法的该类指标增加该算法进行检测。其检测算法为：对具有周期性的数据进行 STL 分解，对残余项通过 GEN-ESD 算法与分位数统计量进行异常检测。

设置两个数据库协同在线检测模块 2 进行数据的存储。参考图 2，它们是在线数据缓存数据库和异常数据库。在线检测模块 2 会将相关数据存储于在线数据缓存数据库和异常数据库中。其中：

①在线数据缓存数据库用于存储在线检测模块 2 各检测算法所需的一定时间窗口内的运维指标数据；

②异常数据库用于存储在线检测模块 2 检测出的异常数据记录，以及异常记录对应的指标序列的瞬时特征描述信息；特征描述信息包括但不限于：周期性信息、分布信息、自相关信息、偏度信息、峰值信息。

上述设置线数据缓存数据库、异常数据库的方式是优选方式，实际上，可以采取其他方式实现，例如，仅设置一个数据库进行分块存储不同数据，或者，在在线检测模块 2 所在设备上对其所需或处理结果的数据进行分类存储。

（3）预警模块 3

预警模块 3 主要用于针对在线检测模块 2 的检测结果，基于预置的预警规则确定是否预警，以及，在预警后接收用户反馈，并根据用户反馈更新预警规则。

可见，预警模块 3 主要包括以下两个功能:

①对在线检测模块 2 检测出的异常进行实时预警过滤，根据预置的预警规则判断是否进行预警。例如可以通过短信、邮件、内部 app 的形式发至用户；用户一般为运维人员或相关业务人员。

②根据用户反馈，以及序列的特征描述，学习用户的兴趣，并更新预警规则。其中，用户反馈是用户对预警作出的反馈，用户一般为运维人员或相关业务人员，用户可以通过访问 API(应用程序编程接口)的方式或网页的形式对预警作出反馈。预警模块 3 可以使用半监督学习算法根据用户反馈学习用户是否对某类预警信息感兴趣。其中，半监督学习(Semi-Supervised Learning，SSL)是模式识别和机器学习领域研究的重点问题，是监督学习与无监督学习相结合的一种学习方法。它主要考虑如何利用少量的标注样本和大量的未标注样本进行训练和分类的问题。主要分为半监督分类，半监督回归，半监督聚类和半监督降维算法。

预警模块 3 会将相关数据存储于预警数据库、用户反馈数据库以及预警模型数据库中。其中：

①预警数据库用于存储最终发出预警的数据的相关信息；

②用户反馈数据库用于存储用户对预警的反馈信息，如该预警是否为误报，是否对该类异常感兴趣等；

③预警模型数据库用于存储预警模块 3 所使用的机器学习算法的模型参数。

上述设置预警数据库、用户反馈数据库以及预警模型数据库的方式是优选方式，实际上，可以采取其他方式实现，例如，仅设置一个数据库进行分块存储不同数据，或者，在预警模块 3 所在设备上对其所需或处理结果的数据进行分类存储。

在线检测预警部分是在接收到待检测的指标数据流之后实时启动工作：

（1）业务或运维人员将需要检测的数据打入

kafka，并对数据进行注册；

（2）在线检测模块从 kafka 接收数据流；

（3）在线检测模块将实时接收的数据流片段与缓存数据库中的历史数据进行拼接；

（4）在线检测模块将实时接收的数据流片段增加至离线数据库(或者另设置一个服务专门负责将实时数据保存至离线数据库中)；

（5）在线检测模块从模型参数数据库中获取相关模型的参数，从离线预测数据库中获得离线预测对应的预测值；

（6）在线检测模块根据拼接获得的数据序列、相关模型参数、离线预测值，利用各检测算法对序列进行异常检测；

（7）在线检测模块将异常检测结果保存至异常数据库，并通知预警模块对异常进行过滤；

（8）预警模块根据预警模型数据库对异常进行过滤；

（9）预警模块将过滤出的异常保存入预警数据库，并向用户员发送预警信息；

（10）当用户对预警信息进行反馈，预警模块会将反馈信息保存至用户反馈数据库中；

（11）预警模块根据预警的反馈在线更新预警模型与策略。

离线建模部分是定期(例如一天)运行：

（1）离线模型训练模块从离线数据库中读取历史数据；

（2）离线模型训练模块根据历史数据对模型参数进行更新，并对未来一天的数据情况进行预测；

（3）离线模型训练模块将更新后的模型参数保存至模型参数数据库，将预测结果保存至长期预测数据库。

运维预警系统，通过机器学习的方式，从采集的指标数据中学习各指标的特征与规律，通过对新采集的指标数据进行分析，检测其中的异常数据。此过程基本为自动化过程，基本不需要人工配置。在检测出异常数据之后，并不是简单地直接发出预警，而是根据不同指标的运维人员过去对不同预警的反馈，学习个性化的预警规则，从而可过滤运维人员不感兴趣的异常，以此减轻运维人员负担，降低误报率。

3 总结

本文提出的运维预警系统，通过机器学习的方式，从采集的指标数据中学习各指标的特征与规律，通过对新采集的指标数据进行分析，检测其中的异常数据。此过程基本为自动化过程，基本不需要人工配置。在检测出异常数据之后，并不是简单地直接发出预警，而是根据不同指标的运维人员过去对不同预警的反馈，学习个性化的预警规则，从而可过滤运维人员不感兴趣的异常，以此减轻运维人员负担，降低误报率。该运维预警系统，可以理解为是面向时间序列的指标的运维预警系统，其应用场景包括上述的数据中心、网络业务系统、网站系统等(以“网络系统”统一表示)，所监控的指标数据也因各场景的不同而不同。

参考文献：

[1] 李锐. 基于无线通信网络的智能化专家运维系统研究[J]. 电子技术与软件工程, 2016(12).

[2] 江务学. 基于结构优化递归神经网络的网络流量预测[J]. 西南大学学报: 自然科学版, 2016(2).

[3] 王浩, 吕云飞, 陈源宝, 等. 基于格兰杰因果关系贝叶斯网络的大规模无线局域网流量预测方法[J]. 电信科学, 2015(8).

[4] 陆晓燕, 魏晖. 网络运维智能化, 开启运维新价值[J]. 通信世界, 2015(21).

[5] 耿直. 大数据时代统计学面临的机遇与挑战[J]. 统计研究, 2014(1).

[6] 张雄灵, 杨贯中. 数据挖掘在河道洪水准确预测中的应用研究[J]. 计算机仿真, 2013(1).