



(12) 发明专利申请

(10) 申请公布号 CN 119248560 A

(43) 申请公布日 2025. 01. 03

(21) 申请号 202411764839.2

G06Q 10/20 (2023.01)

(22) 申请日 2024.12.04

G06Q 50/40 (2024.01)

(71) 申请人 青岛民航凯亚系统集成有限公司

地址 266000 山东省青岛市城阳区棘洪滩

街道春阳路西端动车小镇科技馆230

室

(72) 发明人 姜璐璐 孟宪强 陈晓 薛玲祥

王晓辉 赵滨

(74) 专利代理机构 北京云嘉湃富知识产权代理

有限公司 11678

专利代理师 张亮

(51) Int. Cl.

G06F 11/07 (2006.01)

G06F 16/17 (2019.01)

G06F 16/18 (2019.01)

权利要求书2页 说明书8页 附图3页

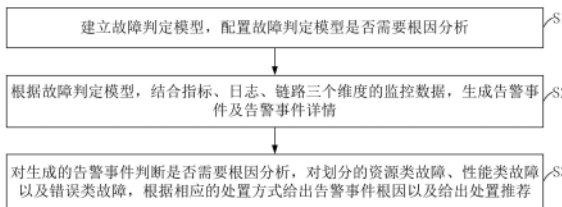
(54) 发明名称

一种机场数据服务接口故障分析方法及系

统

(57) 摘要

本发明属于机场数据处理技术领域,公开了一种机场数据服务接口故障分析方法及系统。该方法针对机场数据服务接口的运行特点,建立故障判定模型,分为资源类、性能类、错误类三种类型,并配置故障判定模型是否需要根因分析;根据故障判定模型,结合指标、日志、链路三个维度的监控数据,生成告警事件及告警事件详情;对生成的告警事件判断是否需要根因分析;对划分的资源类故障、性能类故障以及错误类故障,根据相应的处置方式给出告警事件根因以及给出处置推荐。本发明通过机器学习算法预测性能发展趋势,降低性能类故障的发生几率;通过知识图谱进行性能类和错误类故障根因分析,帮助运维人员及时发现问题、高效处理问题,排除风险隐患。



1. 一种机场数据服务接口故障分析方法,其特征在于,该方法包括:

S1,建立故障判定模型,配置故障判定模型是否需要根因分析;

S2,根据故障判定模型,结合指标、日志、链路三个维度的监控数据,生成告警事件及告警事件详情;

S3,对生成的告警事件判断是否需要根因分析,对划分的资源类故障、性能类故障以及错误类故障,根据相应的处置方式给出告警事件根因以及给出处置推荐。

2. 根据权利要求1所述的机场数据服务接口故障分析方法,其特征在于,在步骤S1中,所述故障判定模型分为资源类、性能类、错误类;

资源类针对服务器资源的使用情况进行故障判定,该类模型的判定要素为指标数据、比较符、阈值;

性能类针对服务接口的响应时长进行故障判定,该类型的判定要素为响应时长、比较符、阈值;

错误类针对服务的错误日志进行故障判定,该类型的判定要素为异常关键字。

3. 根据权利要求1所述的机场数据服务接口故障分析方法,其特征在于,在步骤S2中,指标维度监控数据,用于资源类故障的判定,当指标监控数据满足资源类故障判定模型的要素时,生成告警事件及告警事件详情;

日志维度监控数据,用于错误类故障的判定,当日志监控数据满足错误类故障判定模型的要素时,生成告警事件及告警事件详情;

链路维度监控数据,用于性能类故障的判定,当链路监控数据满足性能类故障判定模型的要素时,生成所述告警事件及告警事件详情。

4. 根据权利要求3所述的机场数据服务接口故障分析方法,其特征在于,针对性能类故障判定模型,判定要素中的阈值分固定阈值和动态阈值,固定阈值采用人工维护的方式进行配置,动态阈值采用机器学习算法通过历史时序数据自动计算获得。

5. 根据权利要求4所述的机场数据服务接口故障分析方法,其特征在于,动态阈值的计算方法以机场日常运行中接口的性能的变化规律为基础,引入机器学习算法,通过回归预测算法生成接口性能曲线,并依据此曲线判断数据服务接口的性能状态;

其中,接口性能曲线通过回归预测算法对历史数据进行收集、计算,具体包括:

(1) 数据样本采样:定时从Skywalking中抽取接口调用性能数据,将数据发送到Kafka消息中间件;

(2) 样本处理:利用流式计算引擎调Storm将大规模、实时的数据样本进行分析和计算,为降低误差,对异常表现的数据进行修正,将调用量过少的样本数据删除,调用量相似的数据按照相应的耗费时长计算平均值,再将调用量为0的数据修正为平均值,最后将预处理后的样本数据存至时间序列数据库openTSDB;

(3) 异常检测:异常检测模块考虑机场实际运营情况,包括流控期间、重保期间、航班大面积延误、暑运春运高峰、机场设备故障保障异常情况;利用L2正则化算法对预处理后的样本数据进行预测分析,对数据0-1标准化,利用公式进行线性变换,使结果落到[0,1]区间;

$$x' = \frac{x - \min}{\max - \min}$$

$$L = \|Y - XW\|^2 + \lambda \|W\|^2$$

式中, x' 为标准化后的数据, x 为原始数据, max 为原始数据最大值, min 为原始数据最小值, L 为损失系数, Y 为目标变量, X 为特征矩阵, W 为模型权重, λ 为正则化参数;

以删减部分无效信息、降低部分精度获得回归系数更为符合实际需求,回归系数越大,数据的整体增长趋势就越大;通过所得的回归系数,按照大小把数据服务接口的发展趋势进行划分,主动判断数据服务接口的性能状态,在性能故障发生前提前进行干预。

6. 根据权利要求3所述的机场数据服务接口故障分析方法,其特征在于,资源类故障处置方式中,针对不同类型资源指标,执行相应的指标下钻查询方法,调用Prometheus的查询API,获取详细数据作为故障根因。

7. 根据权利要求3所述的机场数据服务接口故障分析方法,其特征在于,性能类故障处置方式中,如果性能类故障发生,依托调用链,结合系统拓扑,定位链路中高耗时节点,将高耗时节点作为起点,按错误类故障处置流程追溯下游拓扑,并分析异常,获取异常因果链路,取链路终点作为告警事件根因。

8. 根据权利要求3所述的机场数据服务接口故障分析方法,其特征在于,错误类故障处置方法包括:获取事件相关资产,匹配故障案例库,计算知识图谱,获取故障事件的所有可疑路径,从时间维度计算相关事件,筛选根因路径,再从语义维度计算根因路径的相关度,筛选根因路径,最终生成根因路径,给出处置推荐。

9. 根据权利要求8所述的机场数据服务接口故障分析方法,其特征在于,错误类故障根因定位过程包括:

第1步,根据系统图谱,获取故障事件的所有可疑路径;系统图谱由各业务系统的物理及逻辑组成,包含系统所在的物理环境、相关的逻辑组件及其关系,通过CMDB库、数据链路、网络发现技术进行系统图谱的构建;

第2步,从时间维度分析每条路径上的节点事件相关性,缩小路径范围;

第3步,根据运维知识图谱,从语义维度分析每条路径上上下游节点事件相关性,精确定位根因路径,确定故障根因。

10. 一种机场数据服务接口故障分析系统,其特征在于,实施如权利要求1-9任意一项所述机场数据服务接口故障分析方法,该系统包括:

故障模型定义模块(1),用于定义资源类、性能类、链路类故障判定模型实例,定义故障判定模型要素信息,配置判定模型是都需要根因分析;

告警事件生成模块(2),用于告警事件生成,并生成告警事件详情;性能类故障动态阈值曲线生成;

根因分析模块(3),用于生成故障根因分析所需的系统图谱和运维知识图谱;对划分的资源类故障、性能类故障以及错误类故障,根据相应的处置方式给出告警事件根因以及给出处置推荐。

一种机场数据服务接口故障分析方法及系统

技术领域

[0001] 本发明属于机场数据处理技术领域,尤其涉及一种机场数据服务接口故障分析方法及系统。

背景技术

[0002] 机场内部信息化系统较多,例如A-CDM系统、地面服务系统、飞行区管理系统等,涉及生产、服务、安全和管理等方面,各信息系统间需进行数据的交互和共享,同时还需接入外部系统数据,例如空管数据、航空公司数据等,数据服务接口常用于各应用系统之间数据传输和存储,一旦出现异常则会导致机场数据接收不及时甚至中断,影响机场正常运营,导致旅客投诉,因此发明一种高效智能的数据服务接口故障分析方法尤为必要。

[0003] 发明专利(公开号CN113032238A,公开日20210625)公开了一种基于应用知识图谱的实时根因分析方法,通过构建应用知识图谱,实现对运维对象KPI指标的实时检测和根因分析。所用核心方法包括:多指标异常检测方法、跨层的应用知识图谱异常推理、故障链剪枝;基于相似性算法的实例级根因分析,最终能够定位到系统故障的根本原因。以上方法从应用服务和数据库、中间件等资源类故障入手,未涉及错误类、性能类故障的分析,不够全面且故障定位效率不高。

发明内容

[0004] 为克服相关技术中存在的问题,本发明公开实施例提供了一种机场数据服务接口故障分析方法及系统。针对数据及时、高效传输是机场正常生产运营的基石,一旦发生故障将会产生不良影响,轻则航班延误、旅客积压,重则引发安全问题。本发明目的在于提出一种机场数据服务接口故障分析方法,通过机器学习算法检测异常,通过知识图谱进行故障根因分析,帮助运维人员及时发现问题、高效处理问题,排除风险隐患,避免因数据问题影响机场正常生产,从而提升旅客服务质量。

[0005] 所述技术方案如下:机场数据服务接口故障分析方法,包括:

S1,建立故障判定模型,配置故障判定模型是否需要根因分析;

S2,根据故障判定模型,结合指标、日志、链路三个维度的监控数据,生成告警事件及告警事件详情;

S3,对生成的告警事件判断是否需要根因分析,对划分的资源类故障、性能类故障以及错误类故障,根据相应的处置方式给出告警事件根因以及给出处置推荐。

[0006] 在步骤S1中,所述故障判定模型分为资源类、性能类、错误类;

资源类针对服务器资源的使用情况进行故障判定,该类模型的判定要素为指标数据、比较符、阈值;

性能类针对服务接口的响应时长进行故障判定,该类型的判定要素为响应时长、比较符、阈值;

错误类针对服务的错误日志进行故障判定,该类型的判定要素为异常关键字。

[0007] 在步骤S2中,指标维度监控数据,用于资源类故障的判定,当指标监控数据满足资源类故障判定模型的要素时,生成告警事件及告警事件详情;

日志维度监控数据,用于错误类故障的判定,当日志监控数据满足错误类故障判定模型的要素时,生成告警事件及告警事件详情;

链路维度监控数据,用于性能类故障的判定,当链路监控数据满足性能类故障判定模型的要素时,生成所述告警事件及告警事件详情。

[0008] 进一步,针对性能类故障判定模型,判定要素中的阈值分固定阈值和动态阈值,固定阈值采用人工维护的方式进行配置,动态阈值采用机器学习算法通过历史时序数据自动计算获得。

[0009] 进一步,动态阈值的计算方法以机场日常运行中接口的性能的变化规律为基础,引入机器学习算法,通过回归预测算法生成接口性能曲线,并依据此曲线判断数据服务接口的性能状态;

其中,接口性能曲线通过回归预测算法对历史数据进行收集、计算,具体包括:

(1)数据样本采样:定时从Skywalking中抽取接口调用性能数据,将数据发送到Kafka消息中间件;

(2)样本处理:利用流式计算引擎调Storm将大规模、实时的数据样本进行分析和计算,为降低误差,对异常表现的数据进行修正,将调用量过少的样本数据删除,调用量相似的数据按照相应的耗费时长计算平均值,再将调用量为0的数据修正为平均值,最后将预处理后的样本数据存至时间序列数据库openTSDB;

(3)异常检测:异常检测模块考虑机场实际运营情况,包括流控期间、重保期间、航班大面积延误、暑运春运高峰、机场设备故障保障异常情况;利用L2正则化算法对预处理后的样本数据进行预测分析,对数据0-1标准化,利用公式进行线性变换,使结果落到[0,1]区间;

$$x' = \frac{x - \min}{\max - \min}$$

$$L = \|Y - XW\|^2 + \lambda \|W\|^2$$

式中, x' 为标准化后的数据, x 为原始数据, $\max$ 为原始数据最大值, $\min$ 为原始数据最小值, L 为损失系数, Y 为目标变量, X 为特征矩阵, W 为模型权重, λ 为正则化参数;

以删减部分无效信息、降低部分精度获得回归系数更为符合实际需求,回归系数越大,数据的整体增长趋势就越大;通过所得的回归系数,按照大小把数据服务接口的发展趋势进行划分,主动判断数据服务接口的性能状态,在性能故障发生前提前进行干预。

[0010] 进一步,资源类故障处置方式中,针对不同类型资源指标,执行相应的指标下钻查询方法,调用Prometheus的查询API,获取详细数据作为故障根因。

[0011] 进一步,性能类故障处置方式中,如果性能类故障发生,依托调用链,结合系统拓扑,定位链路中高耗时节点,将高耗时节点作为起点,按错误类故障处置流程追溯下游拓扑,并分析异常,获取异常因果链路,取链路终点作为告警事件根因。

[0012] 进一步,错误类故障处置方法包括:获取事件相关资产,匹配故障案例库,计算知识图谱,获取故障事件的所有可疑路径,从时间维度计算相关事件,筛选根因路径,再从语

义维度计算根因路径的相关度,筛选根因路径,最终生成根因路径,给出处置推荐。

[0013] 进一步,错误类故障根因定位过程包括:

第1步,根据系统图谱,获取故障事件的所有可疑路径;系统图谱由各业务系统的物理及逻辑组成,包含系统所在的物理环境、相关的逻辑组件及其关系,通过CMDB库、数据链路、网络发现技术进行系统图谱的构建;

第2步,从时间维度分析每条路径上的节点事件相关性,缩小路径范围;

第3步,根据运维知识图谱,从语义维度分析每条路径上上下游节点事件相关性,精确定位根因路径,确定故障根因。

[0014] 本发明的另一目的在于提供一种机场数据服务接口故障分析系统,该系统实施所述机场数据服务接口故障分析方法,该系统包括:

故障模型定义模块,用于定义资源类、性能类、链路类故障判定模型实例,定义故障判定模型要素信息,配置判定模型是都需要根因分析;

告警事件生成模块,用于告警事件生成,并生成告警事件详情;性能类故障动态阈值曲线生成;

根因分析模块,用于生成故障根因分析所需的系统图谱和运维知识图谱;对划分的资源类故障、性能类故障以及错误类故障,根据相应的处置方式给出告警事件根因以及给出处置推荐。

[0015] 结合上述的所有技术方案,本发明所具备的有益效果为:本发明将机场数据接口故障进行分类处理,通过机器学习算法预测性能发展趋势,降低性能类故障的发生几率;通过知识图谱进行性能类和错误类故障根因分析,帮助运维人员及时发现问题、高效处理问题,排除风险隐患,避免因数据问题影响机场正常生产,从而提升旅客服务质量。本发明有效降低了机场数据接口的故障频率,并提高了故障解决的效率,减少了对技术专家的依赖。解决了机场运维人员在数据接口服务的运维过程中遇到的故障难预测、难解决问题。

附图说明

[0016] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本公开的实施例,并与说明书一起用于解释本公开的原理;

图1是本发明实施例提供的机场数据服务接口故障分析方法流程图;

图2是本发明实施例提供的接口性能预测过程示意图;

图3是本发明实施例提供的机场数据服务接口故障分析系统示意图;

图中:1、故障模型定义模块;2、告警事件生成模块;3、根因分析模块。

具体实施方式

[0017] 为使本发明的上述目的、特征和优点能够更加明显易懂,下面结合附图对本发明的具体实施方式做详细的说明。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。但是本发明能够以很多不同于在此描述的其他方式来实施,本领域技术人员可以在不违背本发明内涵的情况下做类似改进,因此本发明不受下面公开的具体实施的限制。

[0018] 本发明的创新点在于:在机场数据服务的运维作业中,以智能系统化的方式取代传统人工分析、解决问题的过程,通过机器学习算法预测性能发展趋势,降低性能类故障的

发生几率;通过知识图谱进行性能类和错误类故障根因分析,帮助运维人员及时发现问题、高效处理问题,排除风险隐患。

[0019] 实施例1,如图1所示,本发明实施例提供的机场数据服务接口故障分析方法包括:

S1,建立故障判定模型,配置故障判定模型是否需要根因分析;

示例性的,资源类针对服务器资源的使用情况进行故障判定,如CPU使用率、内存使用率、剩余磁盘空间等,该类模型的判定要素为指标数据+比较符+阈值,如CPU使用率>90%;性能类针对服务接口的响应时长进行故障判定,该类型的判定要素为响应时长+比较符+阈值,如XX接口响应时长>3秒;错误类针对服务的错误日志进行故障判定,该类型的判定要素为异常关键字,如“OutOfMemoryError”“NullPointerException”等;

S2,根据故障判定模型,结合指标、日志、链路三个维度的监控数据,生成告警事件及告警事件详情;

示例性的,指标维度监控数据,来源于Prometheus,用于资源类故障的判定,当指标监控数据满足资源类故障判定模型的要素时,生成所述告警事件及告警事件详情,如Server1服务器的CPU使用率指标符合“CPU使用率>90%”的判定要素时,生成告警事件“Server1服务器CPU使用率较高”,同时生成告警事件详情“Server1服务器CPU使用率95%”;

日志维度监控数据,来源于Loki,用于错误类故障的判定,当日志监控数据满足错误类故障判定模型的要素时,生成所述告警事件及告警事件详情,如Service1服务的日志存在“OutOfMemoryError”的判定要素时,生成告警事件“Service1服务的日志存在OutOfMemoryError错误”,同时生成告警事件详情“Service1服务错误日志:
java.lang.OutOfMemoryError:Java heap space

```
at java.util.ArrayList.grow(ArrayList.java:242)
```

```
at java.util.ArrayList.add(ArrayList.java:411)
```

```
at com.example.MemoryLeakExample.addToList(MemoryLeakExample.java:23)”;
```

链路维度监控数据,来源于Skywalking,用于性能类故障的判定,当链路监控数据满足性能类故障判定模型的要素时,生成所述告警事件及告警事件详情,如Service1服务的getList()接口响应时长符合“>3秒”的判定要素时,生成告警事件“Service1服务的getList()接口响应较慢”,同时生成告警事件详情“Service1服务的getList()接口总耗时4500ms,其中setp1耗时1000ms,setp2耗时3500ms”。

[0020] S3,对生成的告警事件判断是否需要根因分析,对划分的资源类故障、性能类故障以及错误类故障,根据相应的处置方式给出告警事件根因以及给出处置推荐;

示例性的,判断告警事件是否需要根因分析,此因素继承自该告警事件产生来源的故障判定模型配置,由人工进行配置,如不需要根因分析则就此结束,如需根因分析则进行下述根因分析流程。

[0021] 在步骤S3中,如果是资源类故障,则根据不同的资源类型设置固定的下钻查询方式,定位故障原因;

在机场实际运行环境中,该类型的故障处置方法比较简单、速度较快,按资源类型设置固定的下钻查询方式,下钻查询通过调用Prometheus的查询API,针对不同的资源类型,执行相应的查询表达式以获取所需的详细数据,如服务器CPU类告警时,从告警事件信息中获取服务器id和告警产生的时间,进一步查询服务器信息,确定服务器类型,根据服务

器类型确定查询表达式,如windows服务器的查询表达式为“win_proc_Percent_Processor_Time”,按照告警时间和所属服务器组装查询CPU占用前10进程的表达式为“topk(10,sort_desc(win_proc_Percent_Processor_Time{assetCode='Server1',instance!='Idle'}))&time=1728956026.897”,所得的进程作为告警事件根因。

[0022] 如果是性能类故障,定位链路中高耗时节点,将高耗时节点作为起点,按错误类流程追溯下游拓扑,并与案例库相匹配,分析异常,获取异常因果链路定位故障原因;

在机场实际运行环境中,性能类故障相对发生频率较高,传统的性能分析通过调用耗时、人工画图来判断,无法根据机场实际运营状态自动分析数据服务接口性能,将故障实际发生前就能有效的进行性能优化处置,无法将性能恶化等不同场景进行有效的分析迭代。本方案以历史接口调用数据为基础,挖掘接口性能变化规律,引入机器学习算法,通过回归预测模型,自动识别接口性能恶化趋势,建立接口性能分析的闭环,提升服务运营团队在服务接口性能治理上的工作效率,消除人工分析投入,只保留结果验证工作。

[0023] 如果是错误类故障,获取事件相关资产,匹配故障案例库,计算知识图谱,获取故障事件的所有可疑路径,从时间维度计算相关事件,筛选根因路径,再从语义维度计算根因路径的相关度,筛选根因路径,最终生成根因路径,给出处置推荐。

[0024] 其中,在机场实际运行环境中,错误类故障是根因分析中最复杂的一类,由于故障系统结构的复杂性,故障事件的因果路径通常较为复杂。

[0025] 示例性的,性能类故障处置方式包括:

数据服务接口的性能类故障具备可预测性,以机场日常运行中接口的性能的变化规律为基础,通过机器学习算法,设置阈值,自动进行接口性能的分析,在达到阈值时进行告警,提醒运维人员及时处置,降低故障实际发生次数。

[0026] 记录数据接口每一次数据调用的消耗时间,汇集形成样本库,每隔十分钟进行抽样,通过回归预测算法判断接口的性能状态。如图2所示,接口性能预测过程示意图,先利用微服务架构,自动将数据接口实时的性能数据发送到消息中间件,然后对样本进行流式计算,对异常数据进行修正,将预处理后的样本数据存至时间序列数据库,最后通过异常检测模块对数据进行预测分析,从而得出数据接口的性能趋势;

具体如下:

(1) 数据样本采样:开发定时采集程序,定时从Skywalking中抽取接口调用性能数据,将数据发送到Kafka消息中间件;

(2) 样本处理:利用流式计算引擎调Storm将大规模、实时的数据样本进行预处理,为降低误差,对异常表现的数据例如耗费时长和调用量为0以及过少的样本数据进行修正,修正措施包括:①将调用量过少的样本数据删除;②调用量相似的数据按照相应的耗费时长计算平均值;③将调用量为0的数据修正为平均值。最后将预处理后的样本数据存至时间序列数据库openTSDB;

(3) 异常检测:异常检测模块要考虑机场实际运营情况,例如流控期间、重保期间、航班大面积延误、暑运春运高峰、机场设备故障保障异常等特殊情况。利用L2正则化算法对预处理后的样本数据进行预测分析,L2正则化算法是一种改良后的最小二乘法,该算法在原始损失函数中加上L2正则项,防止模型出现过拟合,避免了非满秩矩阵情况下矩阵不可逆、数值不稳定性增加、以及解的不确定性等求逆困难问题。数据0-1标准化,先利用公式进

行线性变换,使结果落到[0,1]区间。

$$[0027] \quad x' = \frac{x - \min}{\max - \min}$$

$$L = \|Y - XW\|^2 + \lambda \|W\|^2$$

式中, x' 为标准化后的数据, x 为原始数据, $\max$ 为原始数据最大值, $\min$ 为原始数据最小值, L 为损失系数, Y 为目标变量, X 为特征矩阵, W 为模型权重, λ 为正则化参数;

以删减部分无效信息、降低部分精度来获得回归系数更为符合实际需求,回归系数越大,数据的整体增长趋势就越大。通过所得的回归系数,按照大小把数据服务接口的发展趋势进行划分,系数越大的代表接口性能恶化趋势越大,从而主动判断数据服务接口的性能状态,在性能故障发生前提前进行干预。

[0028] 如果性能类故障发生,依托调用链,结合系统拓扑,按照链路中span的执行时长取调用链中耗时最长的span,然后根据span的类型来定位慢节点,如该span为本地调用,则取当前节点为高耗时节点,如该span为跨节点调用,则取目标节点为高耗时节点,将高耗时节点作为起点,按错误类故障处置流程追溯下游拓扑,并分析异常,获取异常因果链路,取链路终点作为告警事件根因。

[0029] 示例性的,错误类故障处置方法包括:

错误类故障依托系统拓扑,追溯、分析异常节点下游节点的异常情况,获取关联度高的异常因果链路,取链路终点作为错误类故障告警事件根因。

[0030] 错误类故障根因定位过程分为以下三个步骤:

第1步,根据系统图谱,从异常节点出发,所有的可达路径作为故障事件的所有可疑路径;

第2步,从时间维度分析每条路径上的节点事件相关性,同时发生的节点事件认为具有相关性,缩小可疑路径范围;

第3步,根据运维知识图谱,从语义维度分析每条路径上上下游节点事件相关性,精确定位根因路径,确定故障根因,如路径中两个节点的事件符合运维知识图谱中“错误日志:”Too many connections”与“数据库可用连接数不足”的因果关系,则认为该路径的这两个节点存在强因果关系,对路径中的所有下游节点进行逐对计算,所有计算结果都为强因果关系时,该路径被认定为强因果根因路径。

[0031] 其中,第1步和第3步利用知识图谱技术,其中分为系统图谱与运维知识图谱两类应用,系统图谱指各业务系统的物理及逻辑组成,包含系统所在的物理环境、相关的逻辑组件及其关系,通过CMDB库、数据链路、网络发现等技术手段进行系统图谱的构建,CMDB中包含系统所涉及的软件、硬件及其部署关系,通过这部分信息生成系统的软、硬件节点及部署关系;数据链路中包含系统中各应用的调用关系,通过这部分信息生成系统的软件间调用关系;网络发现获取网络中的网络设备及网络跳路,补充图谱中的硬件网络关系。

[0032] 运维知识图谱是机场各业务系统共用的运维知识库,用于故障的识别与处置、故障影响范围的分析、异常事件的因果分析,其数据来源包含系统图谱、机场运维知识库、机场运维及开发专业知识、机场历史故障及案例数据等。知识图谱的构建过程如下:

(i) 数据源收集:收集与根因分析相关的各类数据,包括系统部署架构、系统故障

记录、机场运维日志、机场用户反馈、系统设计文档、系统测试报告等。

[0033] 数据来源应具有代表性和全面性,以确保构建的知识图谱能够覆盖根因分析所需的各个方面。

[0034] (ii) 知识抽取:从收集到的数据中提取实体、关系、属性等基本信息,形成结构化或半结构化的数据。采用自然语言处理(NLP)技术、机器学习算法或人工标注等方式进行知识抽取,首先,通过自然语言处理技术对文本进行预处理、实体识别、关系抽取和属性提取,利用规则或模型从文本中自动识别关键信息;其次,采用机器学习算法,在特征工程的基础上训练模型以识别和抽取文本中的实体、关系和属性,通过评估与优化提高模型性能;最后,辅以人工标注,通过人工对文本进行实体、关系和属性的标记,并经过质量控制和结果整合,确保数据的准确性和完整性。

[0035] (iii) 知识融合:将不同来源的知识进行融合,通过边界连接,将分散的知识图谱连接形成完整的大图,解决知识间的冲突和冗余问题。建立统一的知识表示体系,确保知识图谱的一致性和准确性。

[0036] (iv) 知识图谱存储:使用图数据库Neo4j存储知识图谱。存储时需要考虑数据的查询效率、可扩展性和维护成本等因素。

[0037] 本发明还提供了一种机场数据服务接口故障分析系统,该系统实施所述机场数据服务接口故障分析方法,该系统包括:

故障模型定义模块1,用于定义资源类、性能类、链路类故障判定模型实例,定义故障判定模型要素信息,配置判定模型是都需要根因分析;

告警事件生成模块2,用于告警事件生成,并生成告警事件详情;性能类故障动态阈值曲线生成;

根因分析模块3,用于生成故障根因分析所需的系统图谱和运维知识图谱;对划分的资源类故障、性能类故障以及错误类故障,根据相应的处置方式给出告警事件根因以及给出处置推荐。

[0038] 为进一步说明本发明实施例相关效果,进行如下实验:

验证环境包含3台服务器分别以Server1、Server2、Server3命名,包含2个应用服务节点和一个数据库分别以Service1、Service2、DB命名,Service1部署于Server1,Service2部署于Server2,DB部署于Server3,调用链路为Service1->Service2->DB,模拟场景为DB连接数耗尽导致Service1接口异常。

[0039] 实验过程如下:

1.配置故障判定模型:①应用日志ERROR异常模型;②数据库连接数不足异常模型,并配置①为需要根因分析。

[0040] 2.服务启动:依次启动DB、Service2、Service1,并通过jmeter工具持续调用Service1测试接口,此时接口一切正常。

[0041] 3.模拟故障:启动测试程序,将DB连接数占满,此时生成3个告警事件:①Service1应用日志存在ERROR异常,详情为“ERROR:call Service2 failed”;②Service2应用日志存在ERROR异常,详情为“ERROR:Could not create connection to database server”;③DB存在链接数不足异常,详情为“数据库可用连接数=0”。

[0042] 4.根因分析:系统对告警事件①进行根因分析,计算所得根因路径为①->②->③,

成功获取根因事件③。

[0043] 以上所述,仅为本发明较优的具体的实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,凡在本发明的精神和原则之内所作的任何修改、等同替换和改进等,都应涵盖在本发明的保护范围之内。

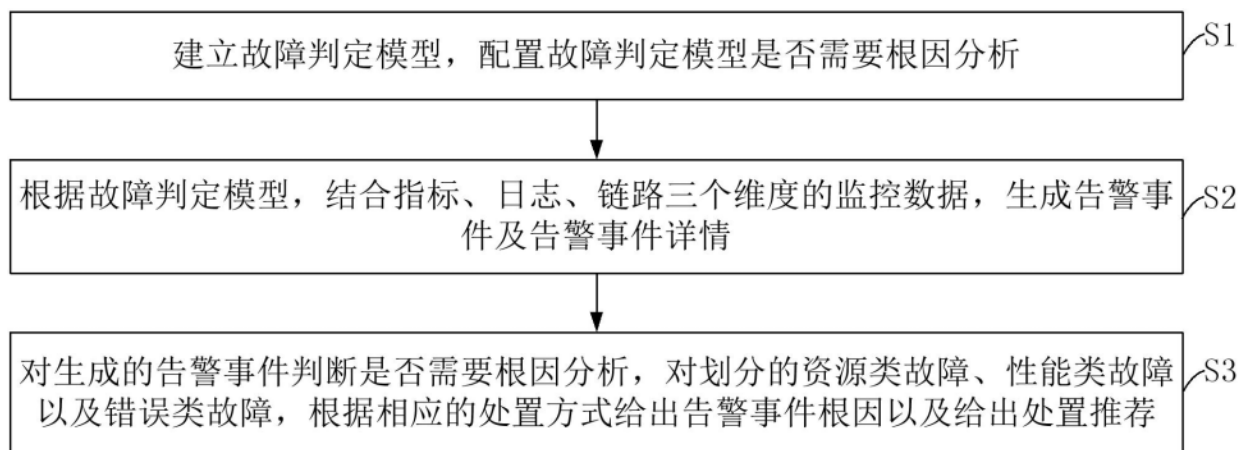


图1

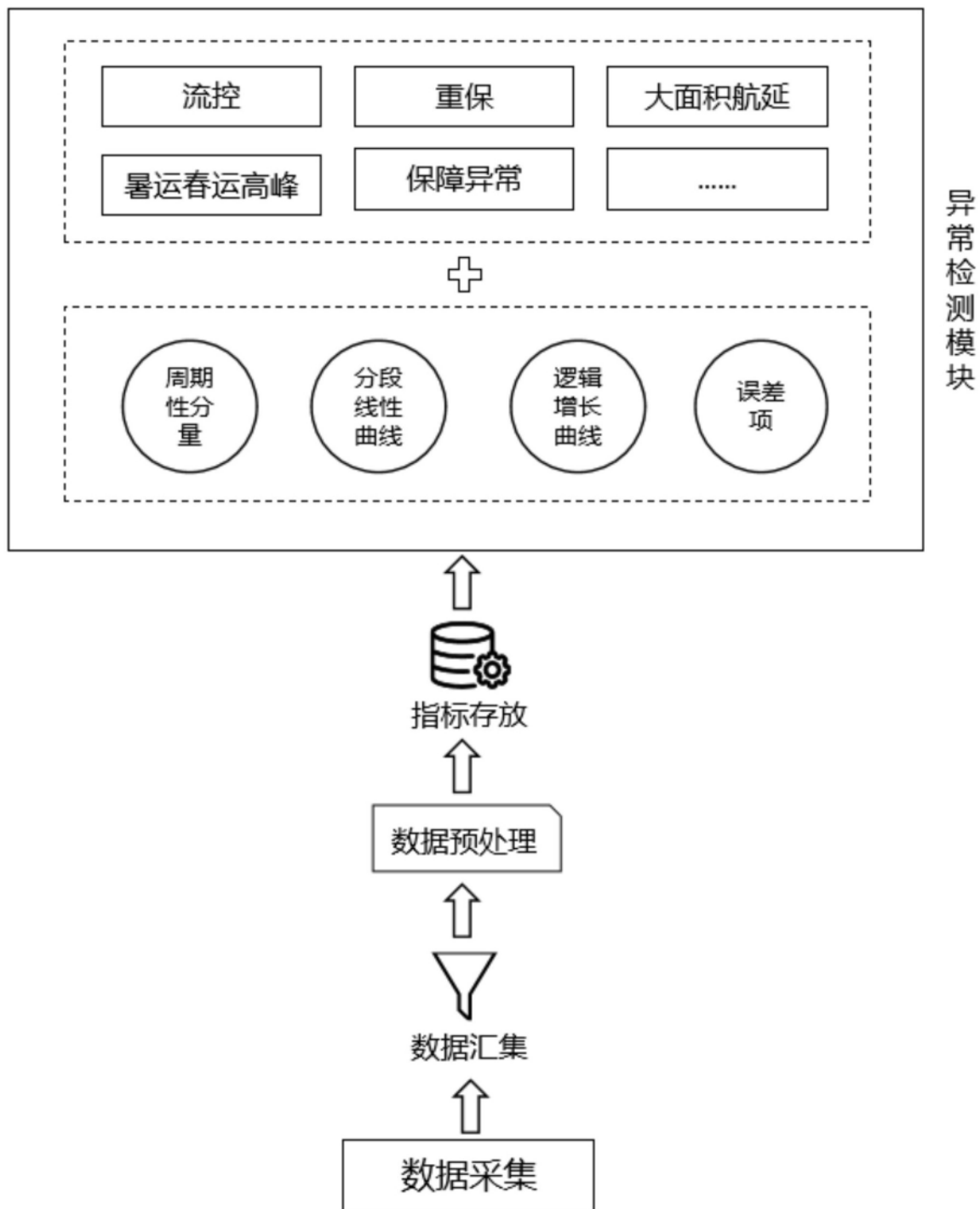


图2

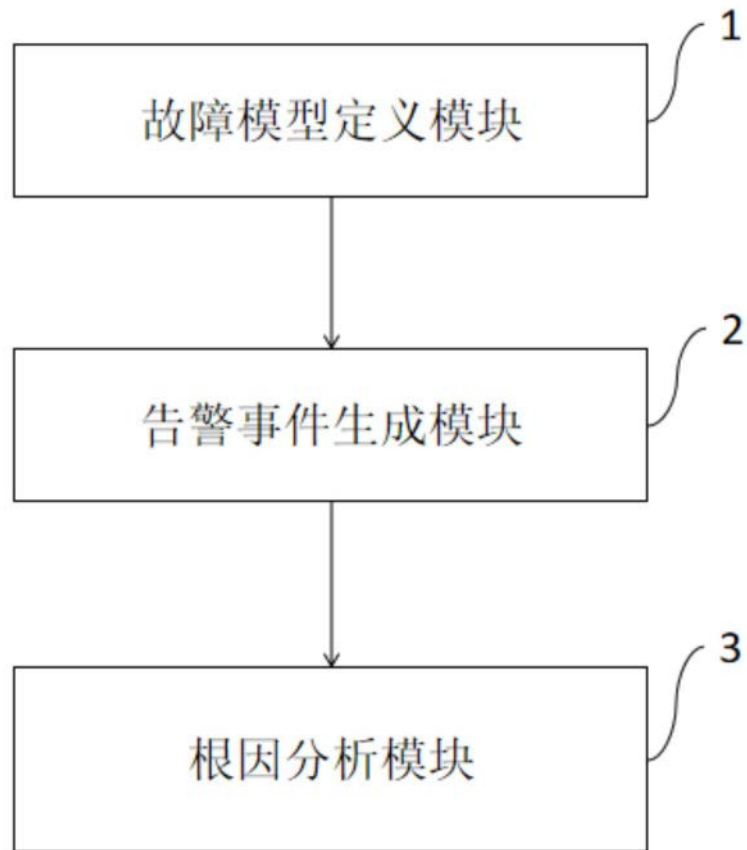


图3