



(12)发明专利申请

(10)申请公布号 CN 109040257 A

(43)申请公布日 2018.12.18

(21)申请号 201810900428.X

(22)申请日 2018.08.09

(71)申请人 湖南江河机电自动化设备股份有限公司

地址 410000 湖南省长沙市长沙高新技术
开发区岳麓西大道588号芯诚科技园
第5幢

(72)发明人 杨锋 陈述平 贺广武 陈建

(51)Int.Cl.

H04L 29/08(2006.01)

G06F 17/30(2006.01)

G06K 9/62(2006.01)

G08B 21/18(2006.01)

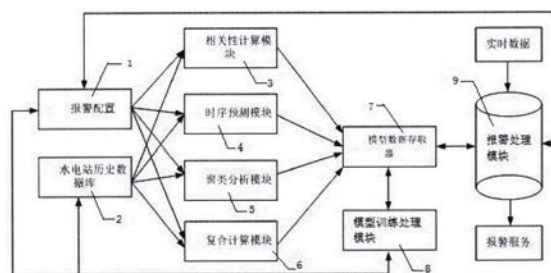
权利要求书1页 说明书5页 附图3页

(54)发明名称

一种基于大数据平台的水电站智能报警系统

(57)摘要

本发明涉及信号报警附属装置的技术领域，特别是涉及一种基于大数据平台的水电站智能报警系统，其无须复杂配置及复杂的设备建模，减少报警数目及误报，提高报警精度，解决模型和算法无法覆盖；包括报警配置、水电站历史数据库、相关性计算模块、时序预测模块、聚类分析模块、复合计算模块、模型数据存取器、模型训练处理模块和报警处理模块，相关性计算模块、时序预测模块、聚类分析模块和复合计算模块均与报警配置电连接，且相关性计算模块、时序预测模块、聚类分析模块和复合计算模块均与水电站历史数据库电连接，相关性计算模块、时序预测模块、聚类分析模块和复合计算模块均与模型数据存取器电连接，报警配置和报警处理模块电连接。



1. 一种基于大数据平台的水电站智能报警系统,其特征在于,包括报警配置(1)、水电站历史数据库(2)、相关性计算模块(3)、时序预测模块(4)、聚类分析模块(5)、复合计算模块(6)、模型数据存取器(7)、模型训练处理模块(8)和报警处理模块(9),相关性计算模块(3)、时序预测模块(4)、聚类分析模块(5)和复合计算模块(6)均与报警配置(1)电连接,且相关性计算模块(3)、时序预测模块(4)、聚类分析模块(5)和复合计算模块(6)均与水电站历史数据库(2)电连接,相关性计算模块(3)、时序预测模块(4)、聚类分析模块(5)和复合计算模块(6)均与模型数据存取器(7)电连接,模型训练处理模块(8)和报警处理模块(9)均与模型数据存取器(7)电连接,报警配置(1)和水电站历史数据库(2)均与模型训练处理模块(8)电连接,报警配置(1)和报警处理模块(9)电连接;

报警配置(1)主要包括报警相关的配置,如报警点描述、是否语音报警及是否短信微信告警以及是否需要复合计算等,如果是复合计算则还可以配置相关的计算公式,此外,报警配置(1)还可以配置模型训练处理算法和报警点在各模块的计算权重等信息;

水电站历史数据库(2)可以采用Apache HBase或Apache HDFS等存储水电站历年数据,用以后续的各种模块分析,分析的模块采用插件的方式,可以动态修改;

相关性计算模块(3)包括相关分析计算组件,包括协方差计算和相关系数计算等,用于找出和报警配置(1)点相关的其它关联点,然后使用这些关联点应用一元或多元回归等算法分析找出最佳参数及误差信息保存到模型数据存取器(7);

时序预测模块(4)包括ARIMA(差分整合移动平均自回归模型)以及Facebook的Prophet和LSTM等时序数据预测算法组件,根据历史数据预测未来一段时间的数据值,并将这些信息保存到模型数据存取器(7);

聚类分析模块(5),该历史数据进行无监督机器学习算法将数据分为不同的类别,如机组运行在不同功率等,并计算同一个类中的均值、众数、中位数、四分位数、标准差、方差等统计信息,并将这些信息保存到模型数据存取器(7);

复合计算模块(6)主要用于一些特殊处理,系统提供常用的函数,如求绝对值ABS、平均值AVG以及某点对应时间点的数据BEFORE,此外还支持用户自定义函数及二次开发,复合计算模块(6)可用于多个点的复杂计算后合成一个新的报警点或无法用其它模块解决的报警计算等,并将这些信息保存到模型数据存取器(7);

模型数据存取器(7)保存上述各种模块的计算模型,可采用关系型数据库或NoSQL甚至文件系统存储,供实时数据的报警处理模块(9)读取使用;

模型训练处理模块(8)根据报警配置(1)及模型数据存取器(7)中相关性计算等四个模块模型参数,并应用决策树、随机森林、AdaBoost和逻辑回归等相关算法,通过训练历史数据得到最优的算法及其参数,将这些信息保存到模型数据存取器(7);

报警处理模块(9)读取模型数据存取器(7)中的模型计算及参数数据,根据报警配置(1)的相关信息及实时数据,分别使用相关性计算模块(3)、时序预测模块(4)、聚类分析模块(5)及复合计算模块(6),并应用模型训练处理模块(8)的最优算法计算是否报警。

一种基于大数据平台的水电站智能报警系统

技术领域

[0001] 本发明涉及信号报警附属装置的技术领域,特别是涉及一种基于大数据平台的水电站智能报警系统。

背景技术

[0002] 众所周知,传统的水电站报警包括模拟量报警和开关量报警,报警形式单一、缺少点之间的关联分析,一旦有操作则会存在大量过程刷屏报警,而这些报警大都不是有效报警,导致有效报警被大量无效的报警掩盖,为此,一些水电站监控厂家提出了智能报警的技术方案,主要通过人工建立业务模型方式定义报警、综合报警以及建立面向对象的设备状态,参照IEC61850或IEC61970标准建模,而上述智能报警针对单个电站进行建模,需要大量的人工梳理业务,以人工建立业务模型方式定义报警为例,需要根据该电站的控制流程和经验或测试时延进行报警配置,而对象报警则需要建立统一的水电厂标准化数据模型,这种方式只适合大型水电站,此外,由于历史原因,通用的大型水电站统一模型很难建立,而中小型电站则更难建立模型。

发明内容

[0003] 为解决上述技术问题,本发明提供一种无须复杂配置及复杂的设备建模,减少报警数目及误报,提高报警精度,解决模型和算法无法覆盖的基于大数据平台的水电站智能报警系统。

[0004] 本发明的一种基于大数据平台的水电站智能报警系统,包括报警配置、水电站历史数据库、相关性计算模块、时序预测模块、聚类分析模块、复合计算模块、模型数据存取器、模型训练处理模块和报警处理模块,相关性计算模块、时序预测模块、聚类分析模块和复合计算模块均与报警配置电连接,且相关性计算模块、时序预测模块、聚类分析模块和复合计算模块均与水电站历史数据库电连接,相关性计算模块、时序预测模块、聚类分析模块和复合计算模块均与模型数据存取器电连接,模型训练处理模块和报警处理模块均与模型数据存取器电连接,报警配置和水电站历史数据库均与模型训练处理模块电连接,报警配置和报警处理模块电连接;

[0005] 报警配置主要包括报警相关的配置,如报警点描述、是否语音报警及是否短信微信告警以及是否需要复合计算等,如果是复合计算则还可以配置相关的计算公式,此外,报警配置还可以配置模型训练处理算法和报警点在各模块的计算权重等信息;

[0006] 水电站历史数据库可以采用Apache HBase或Apache HDFS等存储水电站历年数据,用以后续的各种模块分析,分析的模块采用插件的方式,可以动态修改;

[0007] 相关性计算模块包括相关分析计算组件,包括协方差计算和相关系数计算等,用于找出和报警配置点相关的其它关联点,然后使用这些关联点应用一元或多元回归等算法分析找出最佳参数及误差信息保存到模型数据存取器;

[0008] 时序预测模块包括ARIMA(差分整合移动平均自回归模型)以及Facebook的

Prophet和LSTM等时序数据预测算法组件,根据历史数据预测未来一段时间的数据值,并将这些信息保存到模型数据存取器;

[0009] 聚类分析模块,该历史数据进行无监督机器学习算法将数据分为不同的类别,如机组运行在不同功率等,并计算同一个类中的均值、众数、中位数、四分位数、标准差、方差等统计信息,并将这些信息保存到模型数据存取器;

[0010] 复合计算模块主要用于一些特殊处理,系统提供常用的函数,如求绝对值ABS、平均值AVG以及某点对应时间点的数据BEFORE,此外还支持用户自定义函数及二次开发,复合计算模块可用于多个点的复杂计算后合成一个新的报警点或无法用其它模块解决的报警计算等,并将这些信息保存到模型数据存取器;

[0011] 模型数据存取器保存上述各种模块的计算模型,可采用关系型数据库或NoSQL甚至文件系统存储,供实时数据的报警处理模块读取使用;

[0012] 模型训练处理模块根据报警配置及模型数据存取器中相关性计算等四个模块模型参数,并应用决策树、随机森林、AdaBoost和逻辑回归等相关算法,通过训练历史数据得到最优的算法及其参数,将这些信息保存到模型数据存取器;

[0013] 报警处理模块读取模型数据存取器中的模型计算及参数数据,根据报警配置的相关信息及实时数据,分别使用相关性计算模块、时序预测模块、聚类分析模块及复合计算模块,并应用模型训练处理模块的最优算法计算是否报警。

[0014] 与现有技术相比本发明的有益效果为:该系统可以与水电站监控系统等通信,通过对历史数据相关性分析、时序预测和聚类统计分析进行建模,无须复杂配置及复杂的设备建模,并通过决策树、随机森林等算法聚合多个模型结果,解决单个模型处理的精度不够,减少报警数目及误报,提高报警精度,对模型和算法无法处理的特殊异常通过复合计算处理,解决模型和算法无法覆盖的特殊报警,相关性计算模块工作时,以一元回归分析为例,假设报警点#101和点#201、点#102及点#211相关,其一元线性回归函数如下:报警点#101= $a_0+a_1*\text{点}\#201+a_2*\text{点}\#102+a_3*\text{点}\#211$,通过历史数据计算最优的 a_0 、 a_1 、 a_2 、 a_3 值及误差范围,连同关联点信息及误差信息保存到模型数据存取器,时序预测模块工作时,以ARIMA算法为例,通过报警点的历史数据,可以计算得到最优的自回归(AR)的项数 p 、差分(I)的系数 d 、移动平均(MA)的项数 q 以及历史预测误差范围,并将这些信息保存到模型数据存取器,复合计算模块工作时,以计算数据点100#在十秒的变化小于百分之五并且开关点201#处于合闸状态为例,其符合计算为: $ABS(\text{点}\#100-\text{BEFORE}(10, \text{点}\#100)) * 100 / \text{BEFORE}(10, \text{点}\#100) < 0.05 \text{ AND } (\text{点}\#201 = 1)$,复合计算模块将用户配置的复合计算伪代码编译成可执行代码,将这些信息保存到模型数据存取器,报警处理模块工作时,假设报警点#1的模型训练处理最优算法为线性回归,且相关的参数如下:相关性计算参数0.2,时序预测计算参数0.3,聚类分析参数0.5,复合计算参数为0,要求报警计算条件为0.7,那么:如果计算该报警点#1的相关性计算模型计算为异常(即为1),时序预测计算为异常,聚合计算计算为正常,那么计算报警值为 $1*0.2+1*0.3=0.5$ 小于报警条件0.7即不报警,如果计算该报警点#1的相关性计算模型计算为异常(即为1),时序预测计算为正常,聚合计算计算为异常,那么计算报警值为 $1*0.2+1*0.5=0.7$ 符合报警条件即报警,如告警则向报警服务发送相关报警信息,与此同时,介绍一种模型训练处理模块工作时模型生成训练的方法,包括如下步骤:步骤0、根据报警配置读取历史数据;步骤1、分析报警点和其它数据之间的相关性,尝试

模块内定义的各种模型,如线性回归模型,计算模型相关参数,找出最优模型及相关参数、误差信息等保存到模型数据存取器;步骤2、采用各种时序预测算法,如ARIMA、LSTM等时序数据预测算法,根据历史数据找出最优模型及相关参数、误差信息等保存到模型数据存取器;步骤3、对历史数据进行聚类分类,并计算分类中的统计信息数据;保存模型的相关参数和统计信息数据到模型数据存取器;步骤4、判断是否存在复合计算配置,如存在则步骤5,否则直接跳至步骤6;步骤5、将配置的复合计算公司编译成可执行代码,将这些信息保存到模型数据存取器;步骤6、根据报警配置及模型数据存取器中相关模块模型参数,并应用决策树、随机森林、AdaBoost、逻辑回归等相关算法,通过训练历史数据得到最优的算法及其参数,将这些信息保存到模型数据存取器,介绍一种报警处理模块工作时告警判断的方法,包括如下步骤:步骤0、加载模型及相关数据;步骤1、获取实时数据及模型关联数据,如从Kafka或Redis等读取实时数据点及关联点数据;步骤2、使用大数据流处理组件,如Apache Storm和Apache Spark Streaming等,应用模型计算判断是否产生告警,如产生告警进入步骤3,否则进入步骤1;步骤3、发送告警到相应的告警服务,并进入步骤1。

附图说明

[0015] 图1是本发明的结构示意图;

[0016] 图2是模型训练处理模块工作的流程图;

[0017] 图3是报警处理模块工作的流程图;

[0018] 附图中标记:1、报警配置;2、水电站历史数据库;3、相关性计算模块;4、时序预测模块;5、聚类分析模块;6、复合计算模块;7、模型数据存取器;8、模型训练处理模块;9、报警处理模块。

具体实施方式

[0019] 下面结合附图和实施例,对本发明的具体实施方式作进一步详细描述。以下实施例用于说明本发明,但不用来限制本发明的范围。

[0020] 如图1至图3所示,本发明的一种基于大数据平台的水电站智能报警系统,包括报警配置1、水电站历史数据库2、相关性计算模块3、时序预测模块4、聚类分析模块5、复合计算模块6、模型数据存取器7、模型训练处理模块8和报警处理模块9,相关性计算模块、时序预测模块、聚类分析模块和复合计算模块均与报警配置电连接,且相关性计算模块、时序预测模块、聚类分析模块和复合计算模块均与水电站历史数据库电连接,相关性计算模块、时序预测模块、聚类分析模块和复合计算模块均与模型数据存取器电连接,模型训练处理模块和报警处理模块均与模型数据存取器电连接,报警配置和水电站历史数据库均与模型训练处理模块电连接,报警配置和报警处理模块电连接;

[0021] 报警配置主要包括报警相关的配置,如报警点描述、是否语音报警及是否短信微信告警以及是否需要复合计算等,如果是复合计算则还可以配置相关的计算公式,此外,报警配置还可以配置模型训练处理算法和报警点在各模块的计算权重等信息;

[0022] 水电站历史数据库可以采用Apache HBase或Apache HDFS等存储水电站历年数据,用以后续的各种模块分析,分析的模块采用插件的方式,可以动态修改;

[0023] 相关性计算模块包括相关分析计算组件,包括协方差计算和相关系数计算等,用

于找出和报警配置点相关的其它关联点,然后使用这些关联点应用一元或多元回归等算法分析找出最佳参数及误差信息保存到模型数据存取器;

[0024] 时序预测模块包括ARIMA(差分整合移动平均自回归模型)以及Facebook的Prophet和LSTM等时序数据预测算法组件,根据历史数据预测未来一段时间的数据值,并将这些信息保存到模型数据存取器;

[0025] 聚类分析模块,该历史数据进行无监督机器学习算法将数据分为不同的类别,如机组运行在不同功率等,并计算同一个类中的均值、众数、中位数、四分位数、标准差、方差等统计信息,并将这些信息保存到模型数据存取器;

[0026] 复合计算模块主要用于一些特殊处理,系统提供常用的函数,如求绝对值ABS、平均值AVG以及某点对应时间点的数据BEFORE,此外还支持用户自定义函数及二次开发,复合计算模块可用于多个点的复杂计算后合成一个新的报警点或无法用其它模块解决的报警计算等,并将这些信息保存到模型数据存取器;

[0027] 模型数据存取器保存上述各种模块的计算模型,可采用关系型数据库或NoSQL甚至文件系统存储,供实时数据的报警处理模块读取使用;

[0028] 模型训练处理模块根据报警配置及模型数据存取器中相关性计算等四个模块模型参数,并应用决策树、随机森林、AdaBoost和逻辑回归等相关算法,通过训练历史数据得到最优的算法及其参数,将这些信息保存到模型数据存取器;

[0029] 报警处理模块读取模型数据存取器中的模型计算及参数数据,根据报警配置的相关信息及实时数据,分别使用相关性计算模块、时序预测模块、聚类分析模块及复合计算模块,并应用模型训练处理模块的最优算法计算是否报警。

[0030] 本发明的一种基于大数据平台的水电站智能报警系统,其在工作时,该系统可以与水电站监控系统等通信,通过对历史数据相关性分析、时序预测和聚类统计分析进行建模,无须复杂配置及复杂的设备建模,并通过决策树、随机森林等算法聚合多个模型结果,解决单个模型处理的精度不够,减少报警数目及误报,提高报警精度,对模型和算法无法处理的特殊异常通过复合计算处理,解决模型和算法无法覆盖的特殊报警,相关性计算模块工作时,以一元回归分析为例,假设报警点#101和点#201、点#102及点#211相关,其一元线性回归函数如下:报警点#101= $a_0+a_1*\text{点}\#201+a_2*\text{点}\#102+a_3*\text{点}\#211$,通过历史数据计算最优的 a_0 、 a_1 、 a_2 、 a_3 值及误差范围,连同关联点信息及误差信息保存到模型数据存取器,时序预测模块工作时,以ARIMA算法为例,通过报警点的历史数据,可以计算得到最优的自回归(AR)的项数 p 、差分(I)的系数 d 、移动平均(MA)的项数 q 以及历史预测误差范围,并将这些信息保存到模型数据存取器,复合计算模块工作时,以计算数据点100#在十秒的变化小于百分之五并且开关点201#处于合闸状态为例,其符合计算为: $ABS(\text{点}\#100-\text{BEFORE}(10, \text{点}\#100)) * 100 / \text{BEFORE}(10, \text{点}\#100) < 0.05 \text{ AND } (\text{点}\#201 = 1)$,复合计算模块将用户配置的复合计算伪代码编译成可执行代码,将这些信息保存到模型数据存取器,报警处理模块工作时,假设报警点#1的模型训练处理最优算法为线性回归,且相关的参数如下:相关性计算参数0.2,时序预测计算参数0.3,聚类分析参数0.5,复合计算参数为0,要求报警计算条件为0.7,那么:如果计算该报警点#1的相关性计算模型计算为异常(即为1),时序预测计算为异常,聚合计算计算为正常,那么计算报警值为 $1*0.2+1*0.3=0.5$ 小于报警条件0.7即不报警,如果计算该报警点#1的相关性计算模型计算为异常(即为1),时序预测计算为正常,聚

合计算计算为异常,那么计算报警值为 $1*0.2+1*0.5=0.7$ 符合报警条件即报警,如告警则向报警服务发送相关报警信息,与此同时,介绍一种模型训练处理模块工作时模型生成训练的方法,包括如下步骤:步骤0、根据报警配置读取历史数据;步骤1、分析报警点和其它数据之间的相关性,尝试模块内定义的各种模型,如线性回归模型,计算模型相关参数,找出最优模型及相关参数、误差信息等保存到模型数据存取器;步骤2、采用各种时序预测算法,如ARIMA、LSTM等时序数据预测算法,根据历史数据找出最优模型及相关参数、误差信息等保存到模型数据存取器;步骤3、对历史数据进行聚类分类,并计算分类中的统计信息数据;保存模型的相关参数和统计信息数据到模型数据存取器;步骤4、判断是否存在复合计算配置,如存在则步骤5,否则直接跳至步骤6;步骤5、将配置的复合计算公司编译成可执行代码,将这些信息保存到模型数据存取器;步骤6、根据报警配置及模型数据存取器中相关模块模型参数,并应用决策树、随机森林、AdaBoost、逻辑回归等相关算法,通过训练历史数据得到最优的算法及其参数,将这些信息保存到模型数据存取器,介绍一种报警处理模块工作时告警判断的方法,包括如下步骤:步骤0、加载模型及相关数据;步骤1、获取实时数据及模型关联数据,如从Kafka或Redis等读取实时数据点及关联点数据;步骤2、使用大数据流处理组件,如Apache Storm和Apache Spark Streaming等,应用模型计算判断是否产生告警,如产生告警进入步骤3,否则进入步骤1;步骤3、发送告警到相应的告警服务,并进入步骤1。

[0031] 以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明技术原理的前提下,还可以做出若干改进和变型,这些改进和变型也应视为本发明的保护范围。

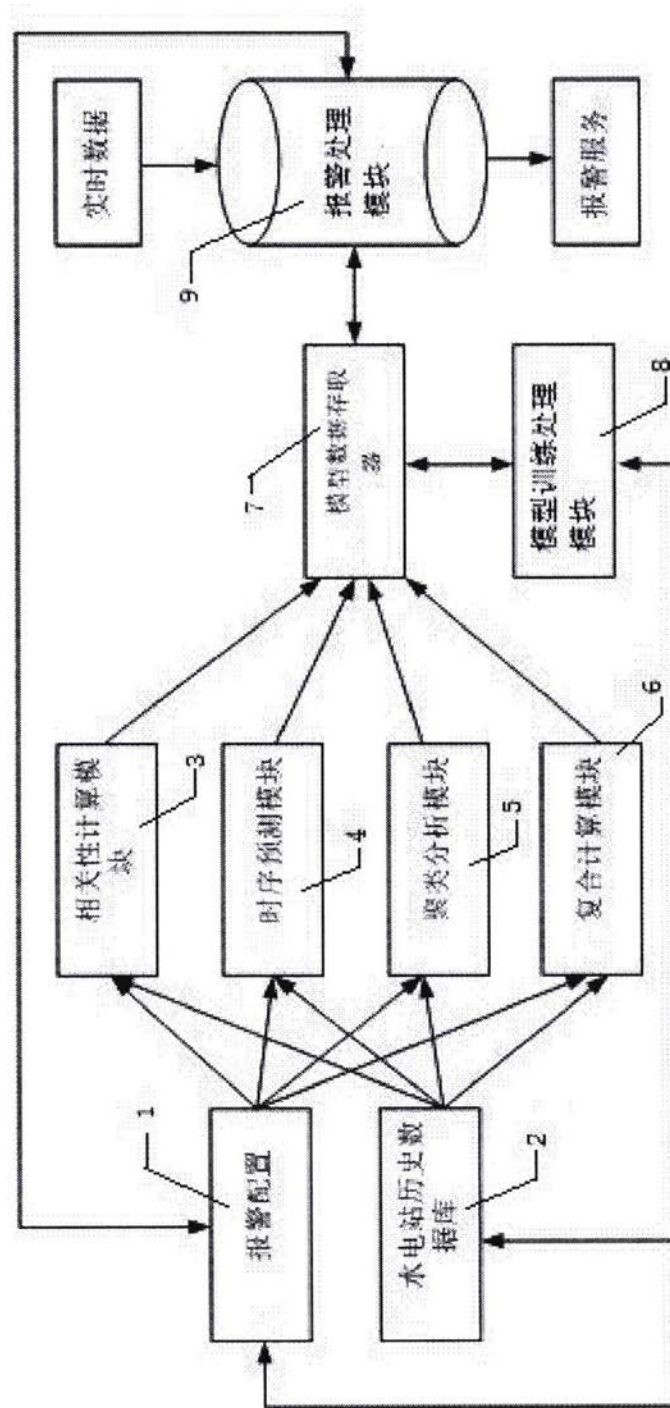


图1

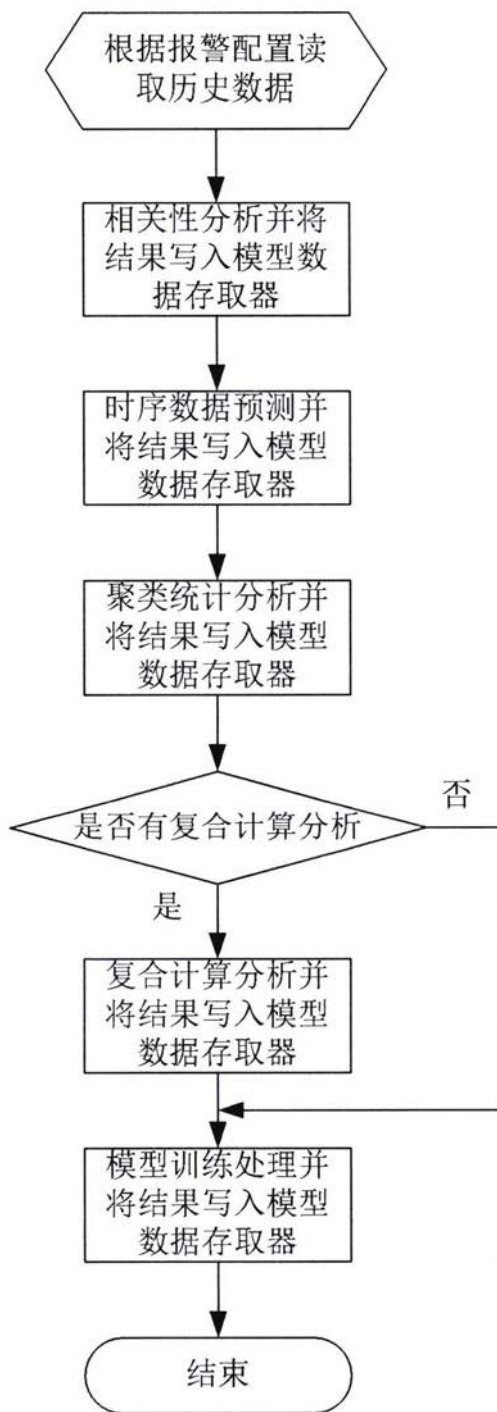


图2

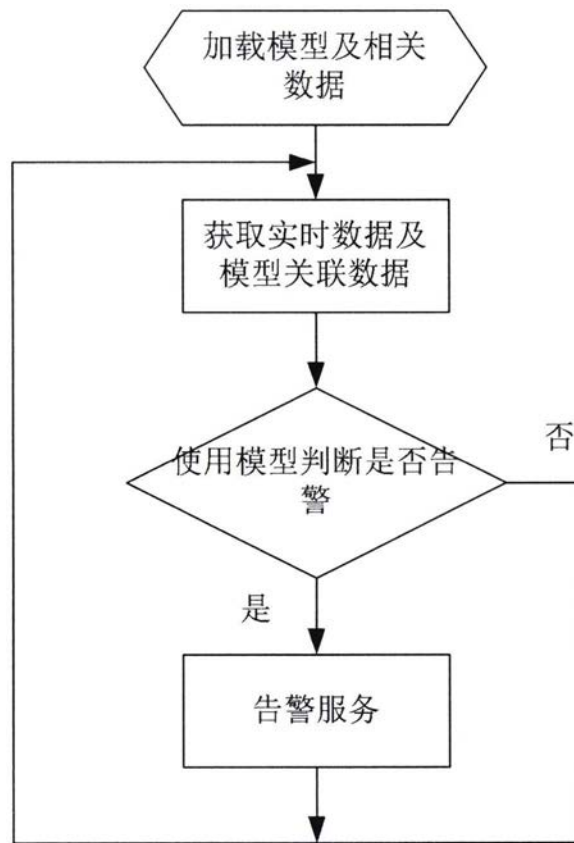


图3