



(12) 发明专利

(10) 授权公告号 CN 114459574 B

(45) 授权公告日 2023.09.26

(21) 申请号 202210126103.7

G06F 17/18 (2006.01)

(22) 申请日 2022.02.10

(56) 对比文件

(65) 同一申请的已公布的文献号  
申请公布号 CN 114459574 A

CN 111159251 A, 2020.05.15

CN 108776683 A, 2018.11.09

CN 110658487 A, 2020.01.07

(43) 申请公布日 2022.05.10

CN 110806546 A, 2020.02.18

CN 113408070 A, 2021.09.17

(73) 专利权人 电子科技大学  
地址 611731 四川省成都市高新区(西区)  
西源大道2006号

EP 3540544 A1, 2019.09.18

US 2021257074 A1, 2021.08.19

专利权人 中国航发四川燃气涡轮研究院

WO 2020008339 A1, 2020.01.09

(72) 发明人 李晓瑜 向文嘉 何子睿 胡世杰  
陆超 吴锋

姚雨虹; 杨小兵. 基于多指标加权分析的航空客户价值预测模型. 中国计量大学学报. 2020, (第01期), 全文.

(74) 专利代理机构 成都华风专利事务所(普通合伙) 51223

李星南; 施展; 亢中苗; 苏卓. 基于孤立森林算法和BP神经网络算法的电力运维数据清洗方法. 电气应用. 2018, (第16期), 18-23.

专利代理师 张巨箭

审查员 付友昱

(51) Int. Cl.

G01F 25/10 (2022.01)

G06F 16/215 (2019.01)

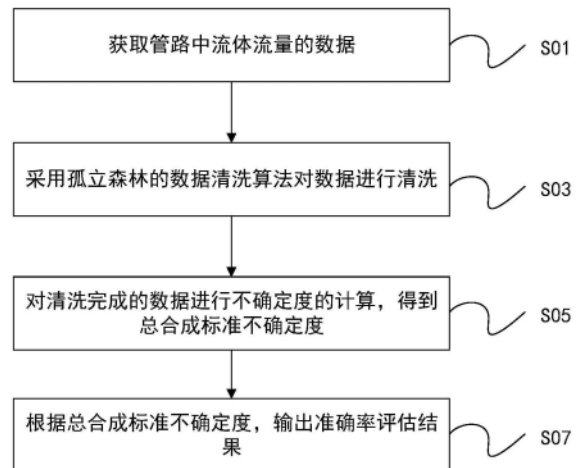
权利要求书3页 说明书11页 附图1页

(54) 发明名称

一种高速流体流量测量准确率自动化评估方法、装置和存储介质

(57) 摘要

本发明公开了一种高速流体流量测量准确率自动化评估方法、装置和存储介质,方法包括以下步骤:获取管路中流体流量的数据;采用孤立森林的数据清洗算法对数据进行清洗;对清洗完成的数据进行不确定度的计算,得到总合成标准不确定度;根据总合成标准不确定度,输出准确率评估结果。本发明提供了从测量数据搜集到输出准确率结果的全过程方案,在评估高速流体流量测量准确率上减少了用户的干预程度,一定程度上缓解了相关人员的工作负担,用户只需关注最后结果,提高了工作效率。



1. 一种高速流体流量测量准确率自动化评估方法,其特征在于:包括以下步骤:  
 获取管路中流体流量的数据;  
 采用孤立森林的数据清洗算法对数据进行清洗;  
 对清洗完成的数据进行不确定度的计算,得到总合成标准不确定度;  
 根据总合成标准不确定度,输出准确率评估结果;

所述采用孤立森林的数据清洗算法对数据进行清洗,包括:

S1:首先,从数据X中随机选择n个点作为子样本,放入一棵孤立树的根节点,其中X为一组测量数据数列,最大值为max,最小值为min;

S2:随机获得子样本中的一个数据x,定义:

$$f_i(x) = T\left[\left(x - \frac{x_i + x_{i+1}}{2}\right) / \frac{x_{i+1} - x_i}{2}\right]$$

其中,i为各数据序号, $x_i$ 和 $x_{i+1}$ 表示子样本中临近的不同测量数据,函数f(x)表示序号i的约束结果,T(u)是当前计算的约束方程且需同时满足的定义有:

$$\begin{cases} T(u) \geq 0; \\ T(u) = T(-u); \\ \int_1^{-1} T(u) = 1; \\ -1 < u < 1; \end{cases}$$

定义 $F_i(x) = \int_{min}^{max} f_i(x)$ ,  $P_i = \frac{F_i(x)}{s}$ ,  $s = \sum F_i(x)$ ;引入k段长度的次方,即:

$$g_i(x) = |x_{i+1} - x_i|^k f_i(x)$$

$$c = \sum_{i=1}^n P_i + \frac{\sum_1^{n-1} g_i(x)}{s}$$

至此计算出乘法系数c;

S3:在当前子样本范围内再随机产生一个值,并使得该值位于当前子样本空间中最大值与最小值之间;将该值与c相乘,结果形成切割点p;

S4:切割点p产生了一个超平面,将空间数据划分成了两个子空间;依次比较数据段中数据的大小,小于p的放在左边的子空间,即放在对当前节点的左分支;大于等于p的放在右边的子空间,即放在对当前节点的右分支;

S5:不断重复S3和S4,不断构造新的叶子节点,直到叶子节点上只有一个数据即无法再继续切割或树已经生长到了所设定的高度1;

S6:之后的处理过程按照孤立森林的基本算法,逐步筛选出异常数值,并进行删除;

所述对清洗完成的数据进行不确定度的计算,得到总合成标准不确定度,包括:

进行A类不确定度的计算;

进行B类不确定度的计算;

将计算所得的A类不确定度和B类不确定度的结果进行合成,计算总合成标准不确定度;

所述进行A类不确定度的计算,包括:

设 $x_i$ 为流量计/流量传感器每一时间段测得的瞬时流量数值,u(A)为该流量计/流量传

传感器的A类不确定度分析结果,则有:

$$u_{(A)} = \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \frac{1}{n} \sum_{i=1}^n x_i)^2}}{\sqrt{n}}$$

分别计算A类不确定度分析结果,结果保存为 $u_{i(A)}$ ,其中 $i=1,2,3,\dots,n$ 为数据的数量;

所述进行B类不确定度的计算包括:

每个流量计/流量传感器的测量没有相关性,据此B类不确定度的定义如下:

$$u_{(B)} = \sqrt{m_1^2 v_1^2 + m_2^2 v_2^2 + m_3^2 v_3^2 + \dots + m_n^2 v_n^2}$$

其中, $m_1, m_2, \dots, m_n$ 为准确率系数, $v_1, v_2, \dots, v_n$ 为不同种类的流量计/流量传感器的流量测量不确定度分析方程;其中,准确率系数的计算方式如下:

$$m_j = \frac{\partial x}{\partial z_j}$$

其中 $z_j$ 为流量计或流量传感器的不同参数;

而 $v_1, v_2, \dots, v_n$ 为该参数的不确定度评估结果; $v_1, v_2, \dots, v_n$ 适用于: $v = k_p \times U_p$ ,式中 $U_p$ 为该流量计/流量传感器的拓展不确定度,通过查询可得, $k_p$ 为覆盖因子;当拓展不确定度的值查询不到时, $v_1, v_2, \dots, v_n$ 适用于: $v = \frac{\Delta}{\sqrt{3}}$ ,其中 $\Delta$ 为该参数的估计误差;

所述将计算所得的A类不确定度和B类不确定度的结果进行合成,计算总合成标准不确定度,包括:

将每一流量计/流量传感器的A类、B类不确定度分析结果进行合成;

$$u_i = \sqrt{u_{i(A)}^2 + u_{i(B)}^2}$$

$u_i$ 为各流量计/流量传感器的测量标准不确定度;

考虑高速流体的流量测量标准不确定度应取所有流量计/流量传感器标准不确定度的平均值,即:

$$u_{all} = \frac{\sum_{i=1}^n u_i}{n}$$

$u_{all}$ 为管道内流体流量的最终标准不确定度;

为避免个别测量故障导致的影响,在 $\exists u_i > u_{all} \times 2.34, i = 1, 2, 3, \dots$ 时,舍弃该 $u_i$ 值,建立新的标准不确定度列表 $u_i', i = 1, 2, 3, \dots, n'$ ,并取标准不确定度列表中的最小值 $u_{min}$ ,最大值 $u_{max}$ ,由下式最终算得总合成标准不确定度:

$$u'_{all} = 0.45 \times \frac{\sum_{i=1}^{n'} u_i' - u_{min} - u_{max}}{n - 2} + 0.55 \times \sum_{i=1}^n \frac{\frac{u_i}{q_i}}{\sum_{i=1}^n \frac{u_i}{q_i}} u_i$$

$u'_{all}$ 为总合成标准不确定度, $q_i$ 为每个流量计/流量传感器测得的平均流量数值,2.34为可替换为其他值的经验常数;

所述根据总合成标准不确定度,输出准确率评估结果,包括:

通过标准不确定度,计算出当前全部测量数据统一的相对标准不确定度;相对标准不确定度的计算方式为:

$$ur = u'_{all} / Q$$

式中, $u'_{all}$ 为总合成标准不确定度, $Q$ 为管道中的流体流量平均值;

将相对标准不确定度转化以百分号形式表达的准确率:

$$R = (1 - 10^3 \times ur) \times 100\%$$

其中, $R$ 表示以%形式表示的准确率。

2. 根据权利要求1所述的一种高速流体流量测量准确率自动化评估方法,其特征在于:所述获取管路中流体流量的数据为接收来自服务器传输的管路中流体流量的数据;

所述服务器的流体流量的数据来源为,安装在管路中的流量计或流量传感器初步采集得到的、由通过通用串行总线各自传输至数据采集卡进行数据的集成、数据采集卡汇总数据后通过本地局域网络并采用RTP和SRTP协议将数据上传的;

所述服务器的流体流量的数据存储于数据库。

3. 一种高速流体流量测量准确率自动化评估装置,包括存储器和处理器,所述存储器上存储有可在所述处理器上运行的计算机指令,其特征在于:所述处理器运行所述计算机指令时执行如权利要求1或2所述的一种高速流体流量测量准确率自动化评估方法的步骤。

4. 一种存储介质,其上存储有计算机指令,其特征在于:所述计算机指令运行时执行如权利要求1或2所述的一种高速流体流量测量准确率自动化评估方法的步骤。

## 一种高速流体流量测量准确率自动化评估方法、装置和存储介质

### 技术领域

[0001] 本发明涉及工业生产监测领域,尤其涉及一种高速流体流量测量准确率自动化评估方法、装置和存储介质。

### 背景技术

[0002] 在工业生产场景和科研实验场景下,存在许多在密闭空间(如管道等)进行高速流体流量测量的需求。目前的流体流量测量往往采取压差测量、流速测量等间接方法获取当前空间流体的状态,从而计算出流体的流量。此外,如空气流量已有空气流量计等电子测量设备直接获取流量情况。流量计作为目前测量流体流量的主要工具,为在密闭管路中测量各种流体的流量发挥了重要作用。鉴于流量测量的方法较多、计算方式的不同、使用场景的不一等,流量计种类繁多,有不同的适用条件,体现了其专用性和局限性。在高速流体的流量测量方面,目前缺乏一种广泛认可的通用流量测量工具,业内往往使用满足自身需求的专业或定制型测量工具。某些苛刻标准下测得数据的可信程度往往需要量化数值来表征准确率,准确率达标方可被采信。

[0003] 影响高速流体流量测量准确性的原因之一在于测量误差,其对流量测量准确率的影响程度已在业界基本形成共识,并往往使用不确定度这一概念来进行评估和量化。有测量即有误差,误差是当前测量值与真实值的差值,是不可避免的。作为仪器和测试学科的重要分支,误差理论已经经过了百年的发展,不确定度理论就是误差理论的升级与延伸。目前国内的流体流量准确率评估的规则已经明确,相关人员在计算流量测量准确率时,往往需要在计算工具辅助下进行手工计算,费时费力,甚至最后计算结果的正确与否也不能保证。在准确率计算的过程中,由于人工计算的局限性,往往还会省略一些繁琐困难的步骤,形成“计算误差的误差”。

[0004] 随着科技的发展和工业的进步,如今流量测量已由传感器来负责采集,所采集的时间间隔短、测量位置多,以致最后的数据量过于庞大。如要手工计算或半自动化计算,工作量极其巨大。如果是高速流体的流量测量,传感器采集频率更是频繁,利用计算机软件技术实现高速流体的流量测量准确率自动化计算具有极其重要的意义。目前国内缺少计算测量准确率的软件工具,而国外已有测量数值的准确率计算器或一些自动化计算方法,但 these 方法仍然有无法迁移到高速流体流量测量场景,并且未实际打通全部计算流程,未真正实现准确率评估的自动化。

[0005] 因此,提供一种高速流体流量测量准确率自动化评估方法、装置和存储介质,属于本领域亟待解决的问题。

### 发明内容

[0006] 本发明的目的在于克服现有技术的不足,提供一种高速流体流量测量准确率自动化评估方法、装置和存储介质。

[0007] 本发明的目的是通过以下技术方案来实现的：

[0008] 本发明的第一方面，提供一种高速流体流量测量准确率自动化评估方法，包括以下步骤：

[0009] 获取管路中流体流量的数据；

[0010] 采用孤立森林的数据清洗算法对数据进行清洗；

[0011] 对清洗完成的数据进行不确定度的计算，得到总合成标准不确定度；

[0012] 根据总合成标准不确定度，输出准确率评估结果。

[0013] 进一步地，所述获取管路中流体流量的数据为接收来自服务器传输的管路中流体流量的数据；

[0014] 所述服务器的流体流量的数据来源为，安装在管路中的流量计或流量传感器初步采集得到的、由通过通用串行总线各自传输至数据采集卡进行数据的集成、数据采集卡汇总数据后通过本地局域网络并采用RTP和SRTP协议将数据上传的；

[0015] 所述服务器的流体流量的数据存储于数据库。

[0016] 进一步地，所述采用孤立森林的数据清洗算法对数据进行清洗，包括：

[0017] S1：首先，从数据X中随机选择n个点作为子样本，放入一棵孤立树的根节点，其中X为一组测量数据数列，最大值为max，最小值为min；

[0018] S2：随机获得子样本中的一个数据x，定义：

$$[0019] \quad f_i(x) = T\left[\left(x - \frac{x_i + x_{i+1}}{2}\right) / \frac{x_{i+1} - x_i}{2}\right]$$

[0020] 其中，i为各数据序号， $x_i$ 和 $x_{i+1}$ 表示子样本中临近的不同测量数据，函数f(x)表示序号i的约束结果，T(u)是当前计算的约束方程且需同时满足的定义有：

$$[0021] \quad \begin{cases} T(u) \geq 0; \\ T(u) = T(-u); \\ \int_1^{-1} T(u) = 1; \\ -1 < u < 1; \end{cases}$$

[0022] 定义 $F_i(x) = \int_{\min}^{\max} f_i(x)$ ， $P_i = \frac{F_i(x)}{s}$ ， $s = \sum F_i(x)$ ；引入k段长度的次方，即：

$$[0023] \quad g_i(x) = |x_{i+1} - x_i|^k f_i(x)$$

$$[0024] \quad c = \sum_{i=1}^n P_i + \frac{\sum_1^{n-1} g_i(x)}{s}$$

[0025] 至此计算出乘法系数c；

[0026] S3：在当前子样本范围内再随机产生一个值，并使得该值位于当前子样本空间中最大值与最小值之间；将该值与c相乘，结果形成切割点p；

[0027] S4：切割点p产生了一个超平面，将空间数据划分成了两个子空间；依次比较数据段中数据的大小，小于p的放在左边的子空间，即放在对当前节点的左分支；大于等于p的放在右边的子空间，即放在对当前节点的右分支；

[0028] S5：不断重复S3和S4，不断构造新的叶子节点，直到叶子节点上只有一个数据即无法再继续切割或树已经生长到了所设定的高度1；

[0029] S6:之后的处理过程按照孤立森林的基本算法,逐步筛选出异常数值,并进行删除。

[0030] 进一步地,所述对清洗完成的数据进行不确定度的计算,得到总合成标准不确定度,包括:

[0031] 进行A类不确定度的计算;

[0032] 进行B类不确定度的计算;

[0033] 将计算所得的A类不确定度和B类不确定度的结果进行合成,计算总合成标准不确定度。

[0034] 进一步地,所述进行A类不确定度的计算,包括:

[0035] 设 $x_i$ 为该流量计/流量传感器每一时间段测得的瞬时流量数值, $u(A)$ 为该流量计/流量传感器的A类不确定度分析结果,则有:

$$[0036] \quad u_{(A)} = \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \frac{1}{n} \sum_{i=1}^n x_i)^2}}{\sqrt{n}}$$

[0037] 分别计算A类不确定度分析结果,结果保存为 $u_{i(A)}$ ,其中 $i=1,2,3,\dots,n$ 为数据的数量。

[0038] 进一步地,所述进行B类不确定度的计算,包括:

[0039] 每个流量计/流量传感器的测量没有相关性,据此B类不确定度的定义如下:

$$[0040] \quad u_{(B)} = \sqrt{m_1^2 v_1^2 + m_2^2 v_2^2 + m_3^2 v_3^2 + \dots + m_n^2 v_n^2}$$

[0041] 其中, $m_1, m_2, \dots, m_n$ 为准确率系数, $v_1, v_2, \dots, v_n$ 为不同种类的流量计/流量传感器的流量测量不确定度分析方程;其中,准确率系数的计算方式如下:

$$[0042] \quad m_j = \frac{\partial x}{\partial z_j}$$

[0043] 其中 $z_j$ 为流量计或流量传感器的不同参数;

[0044] 而 $v_1, v_2, \dots, v_n$ 为该参数的不确定度评估结果; $v_1, v_2, \dots, v_n$ 适用于: $v=k_p \times U_p$ 式中 $U_p$ 为该流量计/流量传感器的拓展不确定度,通过查询可得, $k_p$ 为覆盖因子;当拓展不确定度的值查询不到时, $v_1, v_2, \dots, v_n$ 适用于: $v = \frac{\Delta}{\sqrt{3}}$ ,其中 $\Delta$ 为该参数的估计误差。

[0045] 进一步地,所述将计算所得的A类不确定度和B类不确定度的结果进行合成,计算总合成标准不确定度,包括:

[0046] 将每一流量计/流量传感器的A类、B类不确定度分析结果进行合成。

$$[0047] \quad u_i = \sqrt{u_{i(A)}^2 + u_{i(B)}^2}$$

[0048]  $u_i$ 为各流量计/流量传感器的测量标准不确定度;

[0049] 考虑高速流体的流量测量标准不确定度应取所有流量计/流量传感器标准不确定度的平均值,即:

$$[0050] \quad u_{all} = \frac{\sum_{i=1}^n u_i}{n}$$

[0051]  $u_{all}$ 为管道内流体流量的最终标准不确定度;



[0052] 为避免个别测量故障导致的影响,在 $\exists u_i > u_{all} \times 2.34, i = 1, 2, 3, \dots$ 时,舍弃该 $u_i$ 值,建立新的标准不确定度列表 $u'_i, i = 1, 2, 3, \dots, n'$ ,并取标准不确定度列表中的最小值 $u_{min}$ ,最大值 $u_{max}$ ,由下式最终算得总合成标准不确定度:

$$[0053] \quad u'_{all} = 0.45 \times \frac{\sum_{i=1}^{n'} u'_i - u_{min} - u_{max}}{n - 2} + 0.55 \times \sum_{i=1}^n \frac{\frac{u_i}{q_i}}{\sum_{i=1}^n \frac{u_i}{q_i}} u_i$$

[0054]  $u'_{all}$ 为总合成标准不确定度, $q_i$ 为每个流量计/流量传感器测得的平均流量数值,2.34为可替换为其他值的经验常数。

[0055] 进一步地,所述根据总合成标准不确定度,输出准确率评估结果,包括:

[0056] 通过标准不确定度,计算出当前全部测量数据统一的相对标准不确定度;相对标准不确定度的计算方式为:

$$[0057] \quad ur = u'_{all} / Q$$

[0058] 式中, $u'_{all}$ 为总合成标准不确定度, $Q$ 为管道中的流体流量平均值;

[0059] 将相对标准不确定度转化以百分号形式表达的准确率:

$$[0060] \quad R = (1 - 10^3 \times ur) \times 100\%$$

[0061] 其中, $R$ 表示以%形式表示的准确率。

[0062] 本发明的第二方面,提供一种高速流体流量测量准确率自动化评估装置,包括存储器和处理器,所述存储器上存储有可在所述处理器上运行的计算机指令,所述处理器运行所述计算机指令时执行所述的一种高速流体流量测量准确率自动化评估方法的步骤。

[0063] 本发明的第三方面,提供一种存储介质,其上存储有计算机指令,所述计算机指令运行时执行所述的一种高速流体流量测量准确率自动化评估方法的步骤。

[0064] 本发明的有益效果是:

[0065] (1) 在本发明的一示例性实施例中,首先实现待处理的流体流量测量数据的转存(优选为从服务器到计算器);采用简单的统计分析和基于孤立森林的数据清洗算法,用于检出异常数值;不确定度评估的步骤开始读入数据,计算流量测量的准确率;随后,不确定度评估工具计算完成,通过相对标准不确定度来转化成准确率。因此,提供了从测量数据搜集到输出准确率结果的全过程方案,在评估高速流体流量测量准确率上减少了用户的干预程度,一定程度上缓解了相关人员的工作负担,用户只需关注最后结果,提高了工作效率。

[0066] (2) 在本发明的一示例性实施例中,采用孤立森林的数据清洗算法,并具体公开了该算法的具体步骤,该算法是孤立森林原算法的改进算法,应用概率密度的思想,降低了数据清洗环节的时间复杂度,提升了数据处理效率。同时,相较原算法,该算法在计算结果准确率上有一定提高,并保留了原算法的直观性和孤立森林快速随机构建的优点。

[0067] (3) 在本发明的一示例性实施例中,通过分别将A类不确定度计算和B类不确定度计算,并进行合成,提出了针对具体业务场景的新不确定度合成计算方法,提升了不确定度评估精度,最后得到不确定度的结果,使得数据分析更加准确。

[0068] (4) 在本发明的一示例性实施例中,B类不确定度的计算针对高速流体流量测量的具体业务场景进行了特殊定义,该计算方式是在国家标准上的延伸和优化,计算结果符合要求、贴近具体情况并且更加准确。



## 附图说明

[0069] 图1为本发明一示例性实施例提供的一种高速流体流量测量准确率自动化评估方法的流程示意图；

[0070] 图2为本发明一示例性实施例提供的相关硬件的部署说明和整个方法的系统拓扑架构示意图。

## 具体实施方式

[0071] 下面结合附图对本发明的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，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例，都属于本发明保护的范围。

[0072] 在本发明的描述中，需要说明的是，属于“中心”、“上”、“下”、“左”、“右”、“竖直”、“水平”、“内”、“外”等指示的方向或位置关系为基于附图所述的方向或位置关系，仅是为了便于描述本发明和简化描述，而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作，因此不能理解为对本发明的限制。

[0073] 在本发明的描述中，需要说明的是，除非另有明确的规定和限定，属于“安装”、“相连”、“连接”应做广义理解，例如，可以是固定连接，也可以是可拆卸连接，或一体地连接；可以是机械连接，也可以是电连接；可以是直接相连，也可以通过中间媒介间接相连，可以是两个元件内部的连通。对于本领域的普通技术人员而言，可以根据具体情况理解上述术语在本发明中的具体含义。

[0074] 在本申请使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本申请。在本申请和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0075] 应当理解，尽管在本申请可能采用术语第一、第二、第三等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本申请范围的情况下，第一信息也可以被称为第二信息，类似地，第二信息也可以被称为第一信息。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。此外，属于“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性。

[0076] 此外，下面所描述的本发明不同实施方式中所涉及的技术特征只要彼此之间未构成冲突就可以相互结合。

[0077] 参见图1，图1示出了本发明的一示例性实施例提供的一种高速流体流量测量准确率自动化评估方法，包括以下步骤：

[0078] S01: 获取管路中流体流量的数据；

[0079] S03: 采用孤立森林的数据清洗算法对数据进行清洗；

[0080] S05: 对清洗完成的数据进行不确定度的计算，得到总合成标准不确定度；

[0081] S07: 根据总合成标准不确定度，输出准确率评估结果。

[0082] 具体地，在该示例性实施例中，首先实现待处理的流体流量测量数据的转存(优选为从服务器到计算器)；采用简单的统计分析和基于孤立森林的数据清洗算法，用于检出异

常数值；不确定度评估的步骤开始读入数据，计算流量测量的准确率；随后，不确定度评估工具计算完成，通过相对标准不确定度来转化成准确率。

[0083] 因此，本示例性实施例提供了从测量数据搜集到输出准确率结果的全过程方案，在评估高速流体流量测量准确率上减少了用户的干预程度，一定程度上缓解了相关人员的工作负担，用户只需关注最后结果，提高了工作效率。

[0084] 更优地，在一示例性实施例中，步骤S01中所述获取管路中流体流量的数据为接收来自服务器传输的管路中流体流量的数据；

[0085] 所述服务器的流体流量的数据来源为，安装在管路中的流量计或流量传感器初步采集得到的、由通过通用串行总线各自传输至数据采集卡进行数据的集成、数据采集卡汇总数据后通过本地局域网络并采用RTP和SRTP协议将数据上传的；

[0086] 所述服务器的流体流量的数据存储于数据库。

[0087] 具体地，在该示例性实施例中，安装在直管道上的若干流量计（包括支持数据采集传输的压差流量计、电磁流量计、涡街流量计等）或其他流量传感器就行高速流体流量的初步测量，即流量信息采集。每个流量计/流量传感器采集的数据通过通用串行总线，各自传输至数据采集卡，进行数据的集成。

[0088] 数据采集卡汇总数据后，通过本地局域网络，采用RTP（实时传输协议，Real-time Transport Protocol）和SRTP（安全实时传输协议，Secure Real-time Transport Protocol）协议将数据上传至服务器。SRTP是建立在RTP基础上的安全传输，以确保数据的完整性、真实性和可靠性，使用SRTP协议要求服务器对接收到的测量数据进行初步的检查，检查方式依照SRTP相关标准。服务器将接收到的测量数据进行存储至数据库，数据库的类型为时序型数据库，本实例采用IoTDB数据库方案。

[0089] 之后，计算机通过本地局域网，采用TCP/IP协议从服务器的数据库中获取测量数据。获取的数据暂时保存在计算机的本地磁盘上，待所述方法所有流程结束后，自动释放和删除这些数据。本步骤需要用户许可才可执行。

[0090] 图2提供了本发明相关硬件的部署说明和整个方法的系统拓扑架构。本发明需要的关键设施为高速流体管道、若干的流量计/流量传感器、数据采集卡、服务器和计算机。此外，作为传输介质，还需要通用串行总线、网线等，其他的配件如上述关键设施必需的配件或其它非必要组成零件，这里不作说明。由图2可知，一个管道上可安装多个流量计/流量传感器，这些流量计/流量传感器是独立工作的。每个流量计/流量传感器将独自测得的流量数据通过通用串行总线传输给数据采集卡，由数据采集卡将所有数据进行集成。数据采集卡、服务器、计算机接入了同一本地局域网络，这三个设备可按照不同的通信协议进行信息沟通。数据采集卡和服务器的通信协议为RTP和SRTP，服务器和计算机之间的通信协议为TCP/IP。服务器可根据实际情况进行增添和扩容。服务器内部建设有数据库，负责储存海量的实时流量测量数据。用户可操作计算机，控制计算机开始下载数据并自动开始数据的处理和流量准确率的评估流程，最后从计算机上得知最后的准确率评估结果。

[0091] 而在又一示例性实施例中，可以在服务器中执行所述方法的数据处理步骤。

[0092] 更优地，在一示例性实施例中，步骤S03中所述采用孤立森林的数据清洗算法对数据进行清洗，包括：

[0093] 孤立森林算法的主要原理已是公有技术，这里不作赘述。孤立森林的主要思想是

通过创建随机构建的孤立树森林来检测一组数据中的异常值,孤立树递归地分割子空间,直到一个数据子节点被孤立或者达到树的高度限制。在原本算法中,有以下公式定义:

$$[0094] \quad s(x, n) = 2 \frac{E(h(x))}{C(n)}$$

$$[0095] \quad C(n) = 2H(n-1) - \frac{2(n-1)}{n}$$

[0096] 其中,  $s(x, n)$  为数据  $x$  的异常值;  $E(h(x))$  为数据点  $x$  在孤立树中的路径长度  $h(x)$  的期望;  $C(n)$  为  $n$  个数据点产生路径长度的平均值, 并用以标准化  $h(x)$ ;  $H(n)$  为调和数, 可以被估计为确切的某一数值, 如  $\ln(n) + 0.5772156649$ 。

[0097] 本发明就孤立森林中孤立树的建立过程进行重新定义。该过程的输入为:  $X$ -一组测量数据数列,  $e$ -当前树的高度,  $l$ -树的高度限制,  $k$ - $k$ 段长度的次方; 输出为孤立树  $iTree$ 。

[0098] 首先, 从数据  $X$  中随机选择  $n$  个点作为子样本, 放入一棵孤立树的根节点。

[0099] 其次, 已知子样本的个数为  $n$ , 最大值为  $\max$ , 最小值为  $\min$ 。随机获得子样本中的一个数据  $x$ , 定义:

$$[0100] \quad f_i(x) = T\left[\left(x - \frac{x_i + x_{i+1}}{2}\right) / \frac{x_{i+1} - x_i}{2}\right]$$

[0101] 其中,  $i$  为各数据序号,  $x_i$  和  $x_{i+1}$  表示子样本中临近的不同测量数据, 函数  $f(x)$  表示序号  $i$  的约束结果,  $T(u)$  是当前计算的约束方程且需同时满足的定义有:

$$[0102] \quad \begin{cases} T(u) \geq 0; \\ T(u) = T(-u); \\ \int_{-1}^{-1} T(u) = 1; \\ -1 < u < 1; \end{cases}$$

[0103] 定义  $F_i(x) = \int_{\min}^{\max} f_i(x)$ ,  $P_i = \frac{F_i(x)}{s}$ ,  $s = \sum F_i(x)$ ; 引入  $k$  段长度的次方 (一般按经验赋值), 即

$$[0104] \quad g_i(x) = |x_{i+1} - x_i|^k f_i(x)$$

$$[0105] \quad c = \sum_{i=1}^n P_i + \frac{\sum_{i=1}^{n-1} g_i(x)}{s}$$

[0106] 至此计算出乘法系数  $c$ , 参与后续过程。该步算法运用概率密度思想, 通过乘法系数  $c$  使后续算法的分离过程能够更早地孤立异常值, 使孤立树的路径更短, 从而减少循环, 提升效率; 同时在异常范围内分配较高概率密度, 能使得算法的异常值筛选更加准确。

[0107] 第三, 在当前子样本范围内再随机产生一个值, 并使得该值位于当前子样本空间中最大值与最小值之间。将该值与  $c$  相乘, 结果形成切割点  $p$ 。

[0108] 第四, 切割点产生了一个超平面, 将空间数据划分成了两个子空间。依次比较数据段中数据的大小, 小于  $p$  的放在左边的子空间, 即放在对当前节点的左分支; 大于等于  $p$  的放在右边的子空间, 即放在对当前节点的右分支。

[0109] 最后, 不断重复第三步和第四步, 不断构造新的叶子节点, 直到叶子节点上只有一个数据 (无法再继续切割) 或树已经生长到了所设定的高度  $l$ 。

[0110] 之后的处理过程按照孤立森林的基本算法, 逐步筛选出异常数值。本发明将这些异常数据进行删除。

[0111] 具体地,在该示例性实施例中,计算机将暂存数据进行数据处理,即数据清洗或转换,排除明显的故障数据、异常数据和不可靠数据,避免其影响之后的准确率判定。其中,该步骤中的数据清洗是必须的,数据转换根据计算机是否有相关要求来决定是否执行,如有,用户需要制定转换规则表,由程序读取后指导计算机完成数据转换。数据清洗的主要工作是检出测量数据中的异常数值,将其删除,采取的方法为确定异常数据项。数据转换的主要工作为根据转换规则表,转换测量数据的格式、量纲等基本数据形式。

[0112] 而在本示例性实施例的采用孤立森林的数据清洗算法中,在第二步应用了概率密度的思想,通过乘法系数调整切割点的选取值,从而减少算法第三步和第四步循环次数,降低了数据清洗环节的时间复杂度,提升了数据处理效率;同时算法重新分配了正常数据和离群数据的概率密度,提高了异常数据的区分能力,在计算结果准确率上有一定提高;算法是对原孤立树建立算法的优化改进,极大程度保留了原孤立森林算法的直观性和孤立森林快速随机构建的优点。

[0113] 更优地,在一示例性实施例中,步骤S05中所述对清洗完成的数据进行不确定度的计算,得到总合成标准不确定度,包括:

[0114] S0501:进行A类不确定度的计算;

[0115] S0503:进行B类不确定度的计算;

[0116] S0505:将计算所得的A类不确定度和B类不确定度的结果进行合成,计算总合成标准不确定度。

[0117] 具体地,在该示例性实施例中,通过分别将A类不确定度计算和B类不确定度计算,并进行合成,最后得到不确定度的结果,使得数据分析更加准确。

[0118] 更优地,在一示例性实施例中,步骤S0501中所述进行A类不确定度的计算,包括:

[0119] 设 $x_i$ 为该流量计/流量传感器每一时间段测得的瞬时流量数值, $u(A)$ 为该流量计/流量传感器的A类不确定度分析结果,则有:

$$[0120] \quad u_{(A)} = \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \frac{1}{n} \sum_{i=1}^n x_i)^2}}{\sqrt{n}}$$

[0121] 分别计算A类不确定度分析结果,结果保存为 $u_{i(A)}$ ,其中 $i=1,2,3,\dots,n$ 为数据的数量。

[0122] 更优地,在一示例性实施例中,步骤S0503中所述进行B类不确定度的计算,该方法在国家标准JJF1059.1-2012上针对本发明场景进行了改进,计算结果符合要求,是标准基础上的延伸和深化,通过把不同流量计/流量传感器的内部参数囊括到计算过程,而不是直接计算测得流量数据的B类不确定度,从而极大程度地考虑到各个因素对测量的影响,使高速流体流量测量B类不确定度计算的精度更高。包括:

[0123] 每个流量计/流量传感器的测量没有相关性,据此B类不确定度的定义如下:

$$[0124] \quad u_{(B)} = \sqrt{m_1^2 v_1^2 + m_2^2 v_2^2 + m_3^2 v_3^2 + \dots + m_n^2 v_n^2}$$

[0125] 其中, $m_1, m_2, \dots, m_n$ 为准确率系数, $v_1, v_2, \dots, v_n$ 为不同种类的流量计/流量传感器的流量测量不确定度分析方程;其中,准确率系数的计算方式如下:

$$[0126] \quad m_j = \frac{\partial x}{\partial z_j}$$

[0127] 其中 $z_j$ 为流量计或流量传感器的不同参数;

[0128] 而 $v_1, v_2, \dots, v_n$ 为该参数的不确定度评估结果; $v_1, v_2, \dots, v_n$ 适用于: $v = k_p X U_p$ ,式中 $U_p$ 为该流量计/流量传感器的拓展不确定度,通过查询可得, $k_p$ 为覆盖因子;当拓展不确定度的值查询不到时, $v_1, v_2, \dots, v_n$ 适用于: $v = \frac{\Delta}{\sqrt{3}}$ ,其中 $\Delta$ 为该参数的估计误差。

[0129] 举例说明,如某一压差式流量计,通过测得压力,结合自身的孔口直径 $d$ 等不同参数,最后获得流量数据,其B类不确定度分析方程为:

$$[0130] \quad u_{(B)} = \sqrt{m_1^2 v_1^2 + m_2^2 v_2^2 + m_3^2 v_3^2 + m_4^2 v_4^2}$$

$$[0131] \quad m_1 = \frac{\partial x}{\partial C}$$

$$[0132] \quad m_2 = \frac{\partial x}{\partial d}$$

$$[0133] \quad m_3 = \frac{\partial x}{\partial D}$$

$$[0134] \quad m_4 = \frac{\partial x}{\partial p}$$

[0135] 式中, $C$ 为该流量计的流量系数, $d$ 为该流量计的孔口直径, $D$ 为该流量计的内部孔管长度, $p$ 为该流量计测得的瞬时流体压力。每个参数的偏导计算过程与流量计产品本身相关,不同产品计算过程可能不同,这里不进行阐述。另外,针对覆盖因子 $k_p$ ,本发明规定在置信水平95%下,该值为1.92。

[0136] 更优地,在一示例性实施例中,步骤S0505中所述将计算所得的A类不确定度和B类不确定度的结果进行合成,计算总合成标准不确定度,包括:

[0137] 将每一流量计/流量传感器的A类、B类不确定度分析结果进行合成。

$$[0138] \quad u_i = \sqrt{u_{i(A)}^2 + u_{i(B)}^2}$$

[0139]  $u_i$ 为各流量计/流量传感器的测量标准不确定度;

[0140] 考虑到每个流量计/流量传感器都是独立工作,其在概率模型上是基本符合相对独立的。在寻常情况,可考虑高速流体的流量测量标准不确定度应取所有流量计/流量传感器标准不确定度的平均值,即:

$$[0141] \quad u_{all} = \frac{\sum_{i=1}^n u_i}{n}$$

[0142]  $u_{all}$ 为管道内流体流量的最终标准不确定度;

[0143] 为避免个别测量故障导致的影响,在 $\exists u_i > u_{all} \times 2.34, i = 1, 2, 3, \dots$ 时(2.34为经验常数),舍弃该 $u_i$ 值,建立新的标准不确定度列表 $u_i', i = 1, 2, 3, \dots, n'$ ,并取标准不确定度列表中的最小值 $u_{min}$ ,最大值 $u_{max}$ ,由下式最终算得总合成标准不确定度:

$$[0144] \quad u'_{all} = 0.45 \times \frac{\sum_{i=1}^n u'_i - u_{min} - u_{max}}{n - 2} + 0.55 \times \sum_{i=1}^n \frac{\frac{u_i}{q_i}}{\sum_{i=1}^n \frac{u_i}{q_i}} u_i$$

[0145]  $u'_{all}$  为总合成标准不确定度,  $q_i$  为每个流量计/流量传感器测得的平均流量数值。

[0146] 更优地, 在一示例性实施例中, 步骤S07中所述根据总合成标准不确定度, 输出准确率评估结果, 包括:

[0147] 通过标准不确定度, 计算出当前全部测量数据统一的相对标准不确定度; 相对标准不确定度的计算方式为:

$$[0148] \quad ur = u'_{all} / Q$$

[0149] 式中,  $u'_{all}$  为总合成标准不确定度,  $Q$  为管道中的流体流量平均值;

[0150] 将相对标准不确定度转化以百分号形式表达的准确率:

$$[0151] \quad R = (1 - 10^3 \times ur) \times 100\%$$

[0152] 其中,  $R$  表示以 % 形式表示的准确率。

[0153] 最后可以通过显示屏幕对结果进行显示。

[0154] 与上述示例性实施例具有相同的发明构思, 本发明的又一示例性实施例提供一种高速流体流量测量准确率自动化评估装置, 包括存储器和处理器, 所述存储器上存储有可在所述处理器上运行的计算机指令, 所述处理器运行所述计算机指令时执行所述的一种高速流体流量测量准确率自动化评估方法的步骤。

[0155] 电子设备以通用计算设备的形式表现。电子设备的组件可以包括但不限于: 上述至少一个处理单元、上述至少一个存储单元、连接不同系统组件 (包括存储单元和处理单元) 的总线。

[0156] 其中, 所述存储单元存储有程序代码, 所述程序代码可以被所述处理单元执行, 使得所述处理单元执行本说明书上述“示例性方法”部分中描述的根据本发明各种示例性实施方式的步骤。例如, 所述处理单元可以执行如图1中所示方法。

[0157] 存储单元可以包括易失性存储单元形式的可读介质, 例如随机存取存储单元 (RAM) 3201和/或高速缓存存储单元, 还可以进一步包括只读存储单元 (ROM)。

[0158] 存储单元还可以包括具有一组 (至少一个) 程序模块的程序/实用工具, 这样的程序模块包括但不限于: 操作系统、一个或者多个应用程序、其它程序模块以及程序数据, 这些示例中的每一个或某种组合中可能包括网络环境的实现。

[0159] 总线可以为表示几类总线结构中的一种或多种, 包括存储单元总线或者存储单元控制器、外围总线、图形加速端口、处理单元或者使用多种总线结构中的任意总线结构的局域总线。

[0160] 电子设备也可以与一个或多个外部设备 (例如键盘、指向设备、蓝牙设备等) 通信, 还可与一个或者多个使得用户能与该电子设备交互的设备通信, 和/或与使得该电子设备能与一个或多个其它计算设备进行通信的任何设备 (例如路由器、调制解调器等等) 通信。这种通信可以通过输入/输出 (I/O) 接口进行。并且, 电子设备还可以通过网络适配器与一个或者多个网络 (例如局域网 (LAN), 广域网 (WAN) 和/或公共网络, 例如因特网) 通信。网络适配器通过总线与电子设备的其它模块通信。应当明白, 可以结合电子设备使用其它硬件和/或软件模块, 包括但不限于: 微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、

RAID系统、磁带驱动器以及数据备份存储系统等。

[0161] 通过以上的方式的描述,本领域的技术人员易于理解,这里描述的示例实施方式可以通过软件实现,也可以通过软件结合必要的硬件的方式来实现。因此,根据本示例性实施例的技术方案可以以软件产品的形式体现出来,该软件产品可以存储在一个非易失性存储介质(可以是CD-ROM,U盘,移动硬盘等)中或网络上,包括若干指令以使得一台计算设备(可以是个人计算机、服务器、终端装置、或者网络设备等)执行根据本示例性实施例的方法。

[0162] 与上述示例性实施例具有相同的发明构思,本发明的又一示例性实施例提供一种存储介质,其上存储有计算机指令,所述计算机指令运行时执行所述的一种高速流体流量测量准确率自动化评估方法的步骤。

[0163] 基于这样的理解,本实施例的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品(程序产品)的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。

[0164] 所述程序产品可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以为但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

[0165] 计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了可读程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。可读信号介质还可以是可读存储介质以外的任何可读介质,该可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

[0166] 可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于无线、有线、光缆、RF等等,或者上述的任意合适的组合。

[0167] 可以以一种或多种程序设计语言的任意组合来编写用于执行本发明操作的程序代码,所述程序设计语言包括面向对象的程序设计语言—诸如Java、C++等,还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算设备上部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中,远程计算设备可以通过任意种类的网络,包括局域网(LAN)或广域网(WAN),连接到用户计算设备,或者,可以连接到外部计算设备(例如利用因特网服务提供商来通过因特网连接)。

[0168] 显然,上述实施例仅仅是为清楚地说明所作的举例,而并非对实施方式的限定,对于所属领域的普通技术人员来说,在上述说明的基础上还可以做出其他不同形式的变化或变动。这里无需也无法对所有的实施方式予以穷举。而由此所引申出的显而易见的变化或变动仍处于本发明创造的保护范围之内。



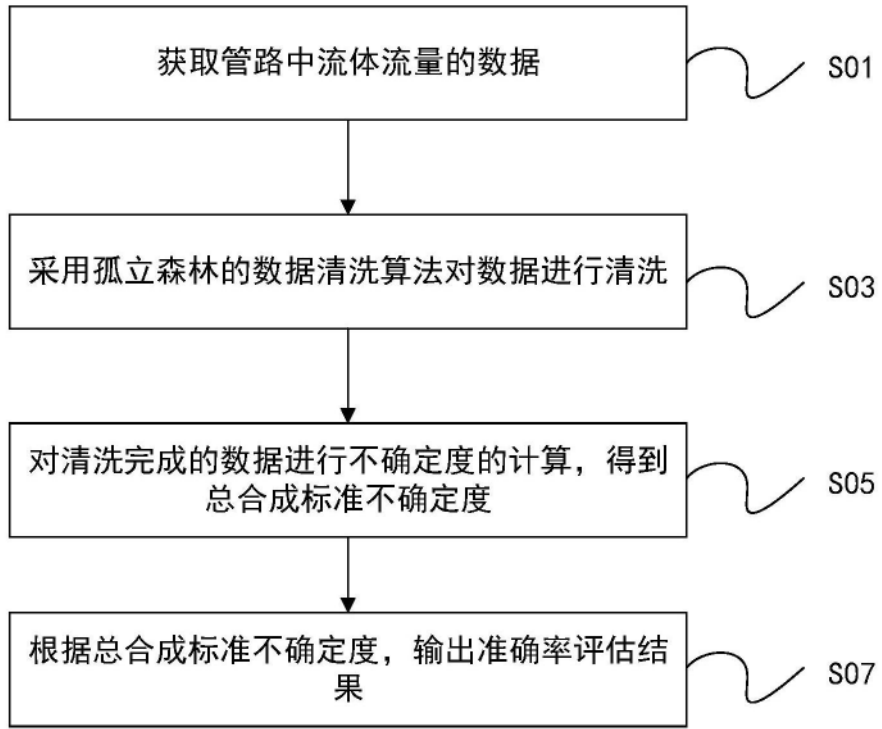


图1

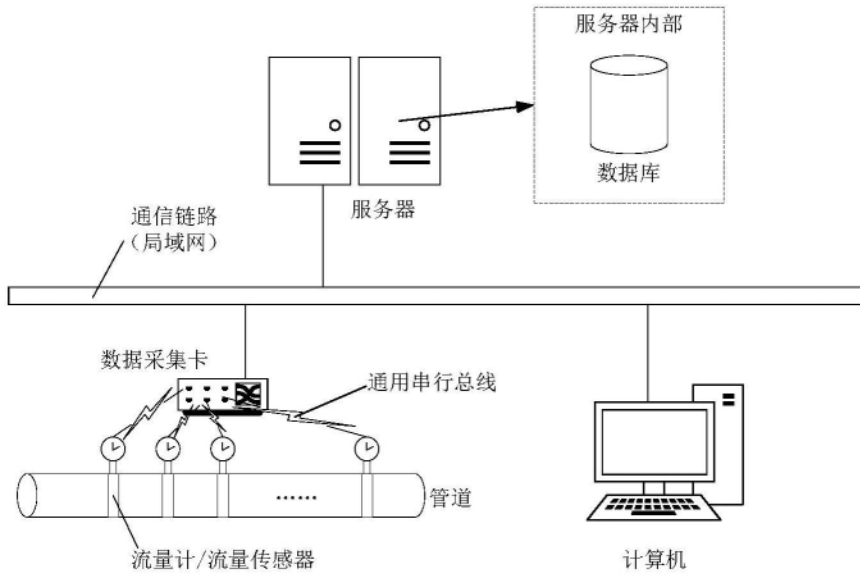


图2