



(12) 发明专利

(10) 授权公告号 CN 103198147 B

(45) 授权公告日 2016.05.11

(21) 申请号 201310137603.1

审查员 刘申

(22) 申请日 2013.04.19

(73) 专利权人 上海顺凯信息技术股份有限公司
地址 200002 上海市黄浦区虎丘路5号3楼

(72) 发明人 潘华 戴加东 金淼 褚伟洪
易爱华

(74) 专利代理机构 上海申蒙商标专利代理有限
公司 31214

代理人 徐小蓉

(51) Int. Cl.

G06F 17/30(2006.01)

(56) 对比文件

CN 102053610 A, 2011.05.11,

CN 102147597 A, 2011.08.10,

JP 4832609 B1, 2011.12.07,

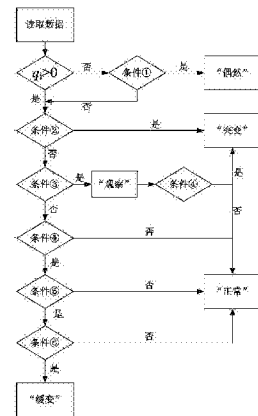
权利要求书2页 说明书5页 附图2页

(54) 发明名称

自动化监测异常数据的判别及处理方法

(57) 摘要

本发明属于自动监测领域，具体涉及一种自动化监测异常数据的判别及处理方法，该方法将监测数据中的异常数据细分为偶然数据、突变数据以及缓变数据，其中偶然数据由工程中的非正常施工因素造成，一旦确定，用监测数据的平均值替代；而突变数据与缓变数据是由工程中的正常施工因素引起的，一旦确定，应高度重视，采取相应的处理方式。本发明的优点是，相对于一般数据采集系统，采用本发明方法的自动化数据采集系统能够有效地对异常数据进行判别、分类，并进行相应的处理，真正实现自动化监测，合理指导施工。



1. 一种自动化监测异常数据的判别方法, 涉及由数据采集系统所采集的监测数据, 所述监测数据基于时间序列, 其特征在于将所述监测数据中的异常数据分类为偶然数据、突变数据和缓变数据, 以所述监测数据的发布时间点前推N个时间序列所对应的监测数据为基点 j_i , 建立统计的数据模型, 并定义符号因子 $q_i = \frac{j_{i+N} - j_i}{j_i - j_{i-1}}$; (1) 所述偶然数据的判别方法至少包括如下步骤: 当计算得所述符号因子 $q_i < 0$ 时, 则计算出所述数据模型中的数据平均值 S 和标准差 σ , 所述数据模型共有 $M+N+1$ 个监测数据点, 其中参数 M 、 N 根据具体工程取值; 以所述数据平均值 S 为中心线, 以 $S+3\sigma$ 为上边界, 以 $S-3\sigma$ 为下边界; 若所述基点 j_i 超出了所述的上边界或下边界, 则所述基点 j_i 被判定为所述偶然数据; 若所述基点 j_i 处于所述上边界和所述下边界之内, 则暂时保留该点以作进一步判定; (2) 所述突变数据的判定方法至少包括如下步骤: (a) 当计算得所述符号因子 $q_i > 0$ 时, 若再满足: $|j_i - j_{i-1}| \geq n_2 \times \Delta r$ 或 $|j_i - j_{i-1}| \geq n_3 \times \Delta z$, 则所述基点 j_i 为所述突变数据; 其中所述 n_2 、 n_3 为经验参数, 所述 Δr 为所述监测数据的日变量报警值, 所述 Δz 为所述监测数据的累计变量报警值; (b) 若计算未满足(a)中的条件, 则进行如下判断, 若满足 $|j_{i+N} - j_{i-1}| \geq n_2 \times \Delta r$ 或 $|j_{i+N} - j_{i-1}| \geq n_3 \times \Delta z$, 且同时满足 $|j_i - j_{i+d}| \geq n_4 \times \Delta r$ 或 $|j_i - j_{i+d}| \geq n_5 \times \Delta z$, 则将所述基点 j_i 标识为观察数据; 其中所述 n_4 、 n_5 为经验参数, 所述 j_{i+d} 为相对于基点 j_i 的24小时之前的监测数据; (c) 在(b)中所述观察数据的基础上, 对所述数据模型的单调性进行判断, 若所述符号因子 $q_i > 0$, 且 $q_{i+1} > 0$, $q_{i+2} > 0 \dots \dots q_{i+N-1} > 0$, 说明所述数据模型中的所述监测数据呈单调性, 则判定所述基点 j_i 为所述突变数据; (3) 所述缓变数据的判定方法至少包括如下步骤: 首先计算判断所述数据模型中的所述监测数据单调性, 满足单调性的所述监测数据, 若再满足 $n_2' \times \Delta r < |j_i - j_{i-1}| < n_2 \times \Delta r$ 且 $n_3' \times \Delta z < |j_i - j_{i-1}| < n_3 \times \Delta z$, 则进一步判别, 若符合下列两种情况之一, 则所述基点 j_i 判定为缓变数据; 其中所述 n_2' 、 n_3' 为经验参数; 情况一: $n_2' \times \Delta r < |j_{i+N} - j_{i-1}| < n_2 \times \Delta r$ 且 $n_3' \times \Delta z < |j_{i+N} - j_{i-1}| < n_3 \times \Delta z$ 情况二: $n_4' \times \Delta r < |j_i - j_{i+d}| < n_4 \times \Delta r$ 且 $n_5' \times \Delta z < |j_i - j_{i+d}| < n_5 \times \Delta z$, n_4' 、 n_5' 为经验参数。

2. 根据权利要求1所述的一种自动化监测异常数据的判别方法, 其特征在于所述经验参数 n_2 和 n_3 可分别取0.5和0.1。

3. 根据权利要求1所述的一种自动化监测异常数据的判别方法, 其特征在于所述经验参数 n_4 和 n_5 可分别取2和0.4。

4. 根据权利要求1所述的一种自动化监测异常数据的判别方法, 其特征在于所述经验参数 n_2' 和 n_3' 可分别取0.2和0.05。

5. 根据权利要求1所述的一种自动化监测异常数据的判别方法, 其特征在于所述经验参数 n_4' 和 n_5' 可分别取0.3和0.08。

6. 一种权利要求1述及的自动化监测异常数据的处理方法, 其特征在于所述处理方法

至少包括如下步骤:当所述基点 j_i 被判定为所述偶然数据后,将该点剔除,并采用所述数据平均值取代该点。

7.一种权利要求1或3所述的自动化监测异常数据的处理方法,其特征在于所述处理方法至少包括如下步骤:当所述基点 j_i 被判定为所述突变数据后,延迟发布时间,观察所述监测数据变化趋势,同时自动加密监测频率;之后检查仪器可靠性;其后结合现场施工情况综合判别,发布报警信息。

8.一种权利要求1或4所述的自动化监测异常数据的处理方法,其特征在于所述处理方法至少包括如下步骤:当所述基点 j_i 被判定为所述缓变数据后,延迟发布时间,观察所述监测数据变化趋势;之后若所述监测数据变化趋势收敛,则进行平滑处理;若所述监测数据变化趋势继续发展,则结合现场施工情况综合判别,决定是否需要报警。

自动化监测异常数据的判别及处理方法

技术领域

[0001] 本发明属于自动监测领域,具体涉及一种自动化监测异常数据的判别及处理方法。

背景技术

[0002] 自动化监测系统具有高精度、连续性、实时性的特征,已被广泛应用于工程测量的多个方面。在自动化监测的实施过程中,不可避免地会出现异常数据,究其原因,主要分为两类:

[0003] 非正常施工因素:仪器突发故障、仪器性能不稳定(包括数据漏采为“0”)、断电、测点破坏或被遮挡等,该类异常数据应在第一时间判定出,并予以剔除;

[0004] 正常施工因素:该类异常数据能客观反映整个工程的状态,是业主、施工方等单位重点关心的对象,应谨慎处理,确立合理的报警机制。

[0005] 目前,还没有对异常数据进行自动化判别的方法形成,也没有相应的处理方法。

发明内容

[0006] 本发明的目的是根据上述现有技术的不足之处,提供一种自动化监测异常数据的判别及处理方法,该方法通过建立统计的数据模型并计算符号因子,以对异常数据进行判别、分类,并进行相应的处理,实现真正的自动化监测,合理指导施工。

[0007] 本发明目的实现由以下技术方案完成:

[0008] 一种自动化监测异常数据的判别方法,涉及由数据采集系统所采集的监测数据,所述监测数据基于时间序列,其特征将所述监测数据中的异常数据分类为偶然数据、突变数据和缓变数据,以所述监测数据的发布时间点前推N个时间序列的时间点为基点 j_i ,

建立统计的数据模型,并定义符号因子 $q_i = \frac{j_{i+1} - j_i}{j_i - j_{i-1}}$; (1)所述偶然数据的判别方法至少包

括如下步骤:当计算得所述符号因子 $q_i < 0$ 时,则计算出所述数据模型中的数据平均值S和标准差 σ ,所述数据模型共有M+N+1个监测数据点,其中参数M、N根据具体工程取值;以所述数据平均值S为中心线,以 $S+3\sigma$ 为上边界,以 $S-3\sigma$ 为下边界;若所述基点 j_i 超出了所述的上边界或下边界,则所述基点 j_i 被判定为所述偶然数据;若所述基点 j_i 处于所述上边界和所述下边界之内,则暂时保留该点以作进一步判定;(2)所述突变数据的判定方法至少包括如下步骤:(a)当计算得所述符号因子 $q_i > 0$ 时,若再满足: $|j_i - j_{i-1}| \geq n_2 \times \Delta r$ 或 $|j_i - j_{i-1}| \geq n_3 \times \Delta z$,则所述基点 j_i 为所述突变数据;其中所述 n_2 、 n_3 为经验参数,所述 Δr 为所述监测数据的日变量报警值,所述 Δz 为所述监测数据的累计变量报警值;(b)若计算未满足(a)中的条件,则进行如下判断,若满足 $|j_{i+N} - j_{i-1}| \geq n_2 \times \Delta r$ 或 $|j_{i+N} - j_{i-1}| \geq n_3 \times \Delta z$,且同时满足 $|j_i - j_{i+N}| \geq n_4 \times \Delta r$ 或 $|j_i - j_{i+N}| \geq n_5 \times \Delta z$,则将所述基点 j_i 标识为观察数据;其中所

述 n_4 、 n_5 为经验参数,所述 j_{i1d} 为相对于基点 j_i 的24小时之前的监测数据;(c)在(b)中所述观察数据的基础上,对所述数据模型的单调性进行判断,若所述符号因子 $q_i > 0$,且 $q_{i+1} > 0$, $q_{i+2} > 0 \dots \dots q_{i+N-1} > 0$,说明所述数据模型中的所述监测数据呈单调性,则判定所述基点 j_i 为所述突变数据;(3)所述缓变数据的判定方法至少包括如下步骤:首先计算判断所述数据模型中的所述监测数据单调性,满足单调性的所述监测数据,若再满足 $n_2' \times \Delta r < |j_i - j_{i-1}| < n_2 \times \Delta r$ 且 $n_3' \times \Delta z < |j_i - j_{i-1}| < n_3 \times \Delta z$,则进一步判别,若符合下列两种情况之一,则所述基点 j_i 判定为缓变数据;其中所述 n_2' 、 n_3' 为经验参数;情况一: $n_2' \times \Delta r < |j_{i+N} - j_{i-1}| < n_2 \times \Delta r$ 且 $n_3' \times \Delta z < |j_{i+N} - j_{i-1}| < n_3 \times \Delta z$ 情况二: $n_4' \times \Delta r < |j_i - j_{i1d}| < n_4 \times \Delta r$ 且 $n_5' \times \Delta z < |j_i - j_{i1d}| < n_5 \times \Delta z$ 。

[0009] 所述经验参数 n_2 和 n_3 可分别取0.5和0.1。

[0010] 所述经验参数 n_4 和 n_5 可分别取2和0.4。

[0011] 所述经验参数 n_2' 和 n_3' 可分别取0.2和0.05。

[0012] 所述经验参数 n_4' 和 n_5' 可分别取0.3和0.08。

[0013] 一种涉及上述自动化监测异常数据的处理方法,其特征在于所述处理方法至少包括如下步骤:当所述基点 j_i 被判定为所述偶然数据后,将该点剔除,并采用所述数据平均值取代该点。

[0014] 一种涉及上述自动化监测异常数据的处理方法,其特征在于所述处理方法至少包括如下步骤:当所述基点 j_i 被判定为所述突变数据后,延迟发布时间,观察所述监测数据变化趋势,同时自动加密监测频率;之后检查仪器可靠性;其后结合现场施工情况综合判别,发布报警信息。

[0015] 一种涉及上述自动化监测异常数据的处理方法,其特征在于所述处理方法至少包括如下步骤:当所述基点 j_i 被判定为所述缓变数据后,延迟发布时间,观察所述监测数据变化趋势;之后若所述监测数据变化趋势收敛,则进行平滑处理;若所述监测数据变化趋势继续发展,则结合现场施工情况综合判别,决定是否需要报警。

[0016] 本发明的优点是,相对于一般数据采集系统,采用本发明方法的自动化数据采集系统能够有效地对异常数据进行判别、分类,并进行相应的处理,真正实现自动化监测,合理指导施工,具体体现为:①结合趋势分析法,对大量的监测数据实现自动化判别与分类,无需人工进行干预;②实现了动态判别,体现了自动化监测的实时性;③能自动对异常数据进行分类,并对不同类型的异常数据采取不同的处理方式,为建立合理的报警机制提供了基础。

附图说明

[0017] 图1为本发明的流程示意图;

[0018] 图2为本发明中偶然数据分析的数据模型示意图;

[0019] 图3为本发明中突变、缓变数据分析的数据模型示意图;

[0020] 图4为本发明实施后的效果图。

具体实施方式

[0021] 以下结合附图通过实施例对本发明的特征及其它相关特征作进一步详细说明,以便于同行业技术人员的理解:

[0022] 如图1-4,图中标记1-5分别为:偶然数据1、突变数据2、突变数据3、偶然数据4、偶然数据5。

[0023] 实施例:如图1所示,本实施例具体涉及一种自动化监测异常数据的判别及处理方法,在本实施例中通过数据采集系统采集监测数据,并将监测数据中的异常数据细分为:偶然数据、突变数据和缓变数据。其中偶然数据由工程中的非正常施工因素引起,如仪器突发故障、仪器性能不稳定(包括数据漏采为“0”)、断电、测点破坏或被遮挡等,一旦确定,均予以剔除,并用相应的平均值替代;而突变数据和缓变数据则由工程中的正常施工因素所引起,应高度重视,并采取合理的处理方法。

[0024] 为方便如图1中所示的流程示意图绘制,对相关判别条件进行如下约定:

[0025] (1)“基点 j_i 超出以 $S+n_1 \times \Delta r$ 为上边界,以 $S-n_1 \times \Delta r$ 为下边界的界限”或者“基点 j_i 超出以 $S+3\sigma$ 为上边界,以 $S-3\sigma$ 为下边界的界限”等价于条件①(即进行偶然数据的判断);

[0026] (2)“ $|j_i - j_{i-1}| > n_2 \times \Delta r$ 或者 $|j_i - j_{i-1}| > n_3 \times \Delta z$ ”等价于条件②(即进行突变数据判断的关键条件之一);

[0027] (3)“ $|j_{i+N} - j_{i-1}| > n_2 \times \Delta r$ 或者 $|j_{i+N} - j_{i-1}| > n_3 \times \Delta z$ ”且“ $|j_i - j_{i+d}| > n_4 \times \Delta r$ 或者 $|j_i - j_{i+d}| > n_5 \times \Delta z$ ”等价于条件③(即进行突变数据判断的另一个关键条件)

[0028] (4)“ $q_i > 0$, 且 $q_{i+1} > 0$, $q_{i+2} > 0 \dots \dots q_{i+N-1} > 0$ ”等价于条件④(即“单调性”判断);

[0029] (5)“ $|j_i - j_{i-1}| > n_2 \times \Delta r$ 且 $|j_{i+N} - j_{i-1}| > n_3 \times \Delta z$ ”等价于条件⑤(即进行缓变数据判断的下限补充条件);

[0030] (6)“ $n_2 \times \Delta r < |j_{i+N} - j_{i-1}| < n_2 \times \Delta r$ 且 $n_3 \times \Delta z < |j_{i+N} - j_{i-1}| < n_3 \times \Delta z$ ”或者“ $n_4 \times \Delta r < |j_i - j_{i+d}| < n_4 \times \Delta r$ 且 $n_5 \times \Delta z < |j_i - j_{i+d}| < n_5 \times \Delta z$ ”等价于条件⑥(即进行缓变数据判断的关键条件)。

[0031] 如图1、2、3所示,本实施例中的自动化监测异常数据的判别及处理方法具体如下:

[0032] (一)偶然数据的判定及处理方法:

[0033] 判断总则:监测数据为一时间序列,以日变量报警值 Δr 和累计变量报警值 Δz 两个指标进行控制,以日变量报警值为首要控制指标,累计变量报警值为最终控制指标,报警值应由设计方提供或根据相关规范进行取值。

[0034] 以发布时间点前推 N 个时间序列为基点 j_i ,建立统计的数据模型(如图2),例如:数据采集系统每5分钟采集一次数据,取 $N = 6$,则实际数据发布的时间约延迟半小时,对一般情况而言,半小时的时间差不会造成工程风险。

[0035] 定义“符号因子” $q_i = \frac{j_{i+1} - j_i}{j_i - j_{i-1}}$,当满足条件 $q_i < 0$ 时,再进一步判断:

[0036] 计算出整个分析模型中的数据平均值 S 和标准差 σ ,共 $M+N+1$ 数据点,参数 M 、 N 根据具体工程调整,取 $M=10$ 、 $N=6$;其中参数 M 、 N 的取值依据具体为:依据 3σ 准则中对样本数量的要求, M 应不小于 10 ; N 的取值要综合考虑采样频率和延迟发布时间,一般延迟 30 分钟发布数据,如果采样频率为 5 分钟,那么 $N=6$ 。以 S 为中心线,以 $S+3\sigma$ 为上边界,以 $S-3\sigma$ 为下边界。若基点 j_i 超出了上述的界限,则可判断该点为“偶然数据”;若基点 j_i 并未超出了上述的界限,则暂时保留该点,并对该点开展进一步的判断;

[0037] 一旦被判定为“偶然数据”,需将该点剔除,并采用平均值 S 取代该点,从而方便后续点的判别。

[0038] (二)突变数据的判定及处理方法:

[0039] 具体数据模型如图3所示,满足下列两种情况之一,可将基点 j_i 判别为“突变数据”。

[0040] 情况一:当“符号因子” $q_i > 0$ 时,若 $|j_i - j_{i-1}| \geq n_2 \times \Delta r$ 或 $|j_i - j_{i-1}| \geq n_3 \times \Delta z$ (n_2 、 n_3 为经验参数,建议可分别取 $0.5, 0.1$),则基点 j_i 为“突变数据”;

[0041] 情况二:若不满足情况一的条件,则进行下面的判断,

[0042] 若存在: $|j_{i+N} - j_{i-1}| \geq n_2 \times \Delta r$ 或 $|j_{i+N} - j_{i-1}| \geq n_3 \times \Delta z$,且同时满足 $|j_i - j_{i+d}| \geq n_4 \times \Delta r$ 或 $|j_i - j_{i+d}| \geq n_5 \times \Delta z$ (n_4 、 n_5 为经验参数,建议可分别取 $2, 0.4$),则可将基点 j_i 标识为“观察数据”,提醒监测人员密切关注所采集的数据,其中 j_{i+d} 为相对于基点 j_i 24小时之前的数据。

[0043] 之后在“观察数据”的基础之上,再进行整个数据模型“单调性”判断;若 $q_i > 0$,且 $q_{i+1} > 0, q_{i+2} > 0 \dots \dots q_{i+N-1} > 0$,即数据分析模型中的判断因子均大于零,也就是说该数据模型中的监测数据是单调的,则可判定基点 j_i 为“突变数据”。

[0044] 一旦满足上述两种情况之一,即判定出现了“突变数据”,建议采取如下的处理措施:

[0045] 第一步:延迟发布时间,观察监测数据变化趋势,同时自动加密监测频率;

[0046] 第二步:检查仪器可靠性;

[0047] 第三步:结合现场施工情况综合判别,发布报警信息。

[0048] (三)缓变数据的判定及处理方法:

[0049] 首先应进行数据模型“单调性”的判断,具体判别方法同(二)中述及的“单调性”判断相同。在满足“单调性”的前提下,再进行以下判别(数据模型如图3所示):

[0050] 如果满足 $n_2' \times \Delta r < |j_i - j_{i-1}| < n_2 \times \Delta r$ 且 $n_3' \times \Delta z < |j_i - j_{i-1}| < n_3 \times \Delta z$,则进一步判别,若符合下列两种情况之一,则判定为“缓变数据”。(其中 n_2' 、 n_3' 为经验参数,建议可分别取 $0.2, 0.05$)

[0051] 情况一: $n_2' \times \Delta r < |j_{i+N} - j_{i-1}| < n_2 \times \Delta r$ 且 $n_3' \times \Delta z < |j_{i+N} - j_{i-1}| < n_3 \times \Delta z$

[0052] 情况二： $n_4' \times \Delta r < |j_i - j_{i1d}| < n_4 \times \Delta r$ 且 $n_5' \times \Delta z < |j_i - j_{i1d}| < n_5 \times \Delta z$ (n_4' 、 n_5' 为经验参数,建议可分别取0.3,0.08)

[0053] 一旦判定出现了“缓变数据”,建议采取如下的处理措施:

[0054] 第一步:延迟发布时间,观察监测数据变化趋势;

[0055] 第二步:若监测数据的变化趋势收敛,则进行平滑处理;

[0056] 第三步:若监测数据的变化趋势继续发展,则结合现场施工情况综合判别,决定是否需要报警。

[0057] 上述八个经验参数主要是依据对自动化监测系统试运营期间监测数据的总结、反演得出。

[0058] 如图4所示为实施本发明方法之后的效果示意图,其中由非正常施工因素引起的异常数据为偶然数据1、偶然数据4和偶然数据5,由正常施工因素引起的异常数据为突变数据2和突变数据3;按照本发明提供的方法对数据进行判别,能够正确地剔除偶然数据1、4、5,同时能够合理地判别出由于施工原因引起的异常数据,即突变数据2、3。

[0059] 本实施例的有益效果是:相对于一般数据采集系统,采用本方法的自动化数据采集系统能够有效地对异常数据进行判别、分类,并进行相应的处理,真正实现自动化监测,合理指导施工。

[0060] 首先,本方法结合趋势分析法,对大量的监测数据实现自动化判别与分类,无需人工进行干预。

[0061] 其次,本方法实现了动态判别,体现了自动化监测的实时性。

[0062] 还有,本方法能自动对异常数据进行分类,并对不同类型的异常数据采取不同的处理方式,为建立合理的报警机制提供了基础。

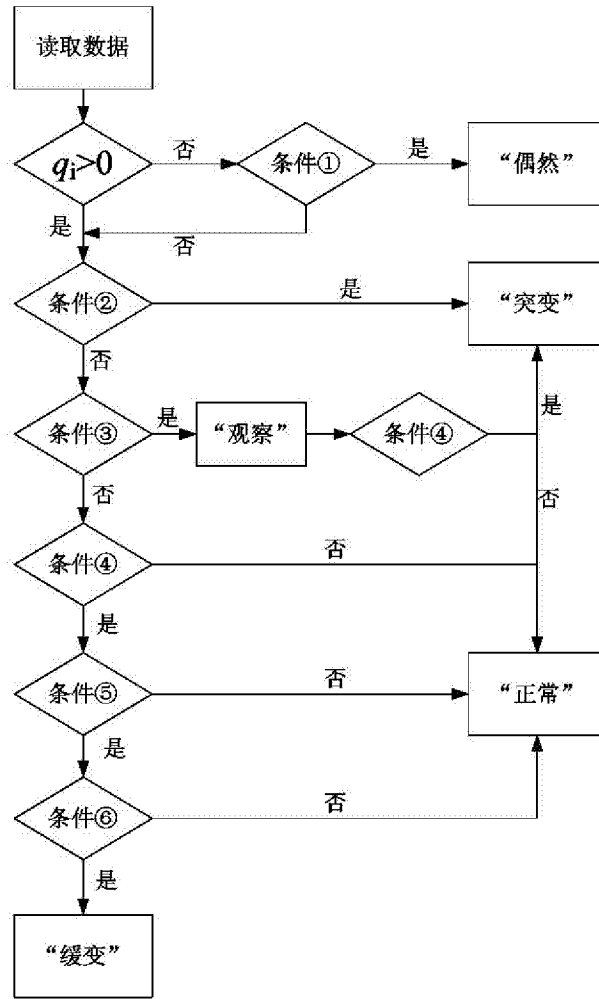


图1

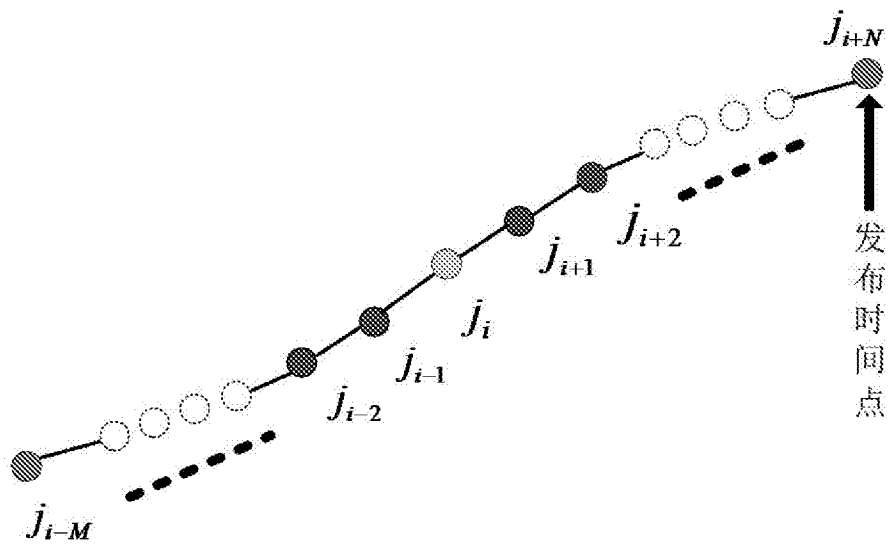


图2

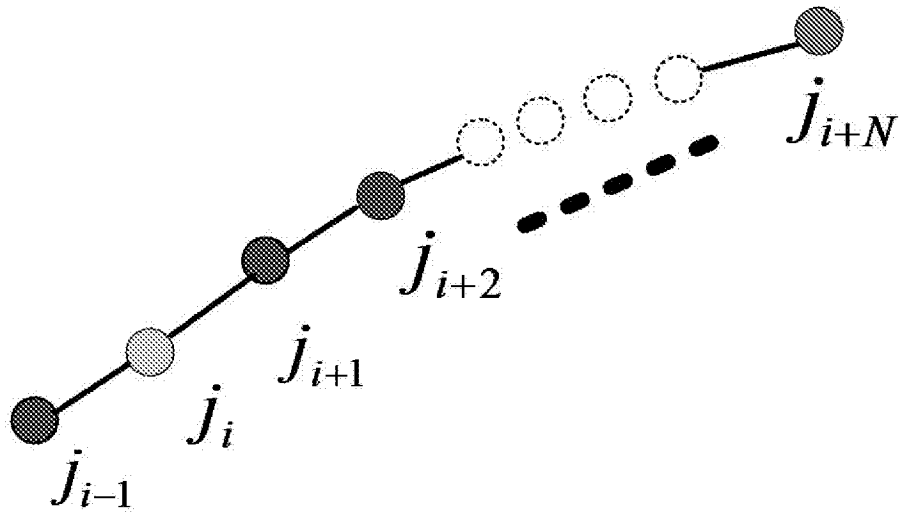


图3

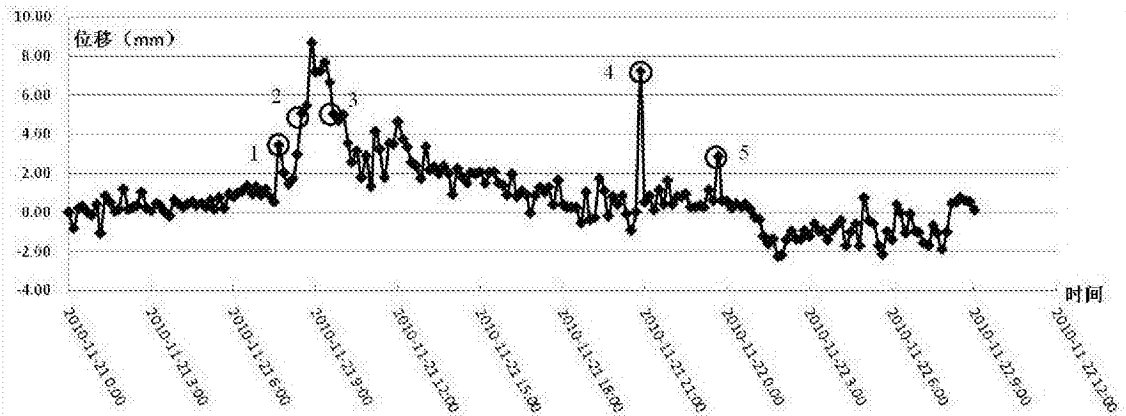


图4