

基于最大测点正常率与 GPU 并行加速的不良数据辨识方法

方 睿¹, 董树锋¹, 唐坤杰¹, 朱承治², 裴 活², 宋永华³

(1. 浙江大学电气工程学院, 浙江省杭州市 310027;

2. 国网浙江省电力有限公司, 浙江省杭州市 310007; 3. 澳门大学电机及电脑工程系, 澳门)

摘要: 基于测量不确定度的概念, 以测点正常率最大(MNMR)为目标的电力系统抗差状态估计方法具有较好的不良数据辨识能力。然而, 该模型求解困难, 已有研究对该模型进行了近似等效, 并采用现代内点法进行求解, 但存在因近似而辨识效果降低的问题。为此, 基于 MNMR 状态估计模型, 采用杂交变异粒子群算法, 提出一种基于图形处理器(GPU)并行加速的不良数据辨识算法。该算法不对 MNMR 模型进行近似等效, 根据 GPU 并行计算架构特点, 设计了粗粒度和细粒度结合的并行加速策略。算例结果表明, 所提的算法对不良数据的误检率和漏检率较低, 具有较好的不良数据辨识能力, 且计算时间短, 加速效率高, 能够满足实际运行需求。

关键词: 数据辨识; 状态估计; 测点正常率; 图形处理器; 并行计算

0 引言

电力系统状态估计是现代能量管理系统(energy management system, EMS)的重要组成部分, 为 EMS 中的高级应用提供数据支撑。但实际运行过程中, 量测采样数据中会偶然出现一些不良数据, 可能使得状态估计结果偏离实际值, 降低估计结果的可靠性。目前广泛使用最小二乘(WLS)估计^[1], 但该方法抗差能力较差, 易受不良数据的影响。针对抗差问题, 许多学者提出了多种抗差状态估计准则, 包括加权最小绝对值(WLAV)准则^[2]、非二次(non-quadratic)准则^[3]、最小中位数平方(LMS)准则和最小截平方(LTS)准则等^[4]。这些估计方法从残差理论出发, 尽量减少不良数据对状态估计结果的影响, 但没有从本质上解决不良数据的影响问题, 常常作为不良数据不能完全得到辨识时的抗差估计手段。事实上, 成功辨识出不良数据后, WLS 估计属于无偏估计, 计算速度快, 结果具有较好的统计特性。

目前对不良数据辨识的研究一类主要采用残差搜索法或建立在残差概念的基础上^[5-8]。残差越小能够表征测点的估计值离量测值较近, 但不能给出任何测点与真值之间的关系, 反而可能因此偏离测点的真值, 其结果未必合理。另一类不良数据辨识

的方法是量测量突变检测法^[9-10], 这类方法利用相邻采样时刻的量测信息, 能解决残差淹没和残差污染问题, 但要求采样时间内网络结构不发生变化, 同时前一时刻测点的不良数据已被辨识出。

近年来, 有学者提出以正常率最大为目标的状态估计方法^[11-14], 文献^[15-16]在此基础上进行了改进, 引入异常测点相容性校验和正常测点滤波。文献^[17-19]以测点正常率最大(maximum normal measurement rate, MNMR)方法为基础, 提出了最大正常率最小偏差度(maximum normal-rate least deviation, MNLD)的状态估计多目标模型。这类方法基于测量不确定度理论^[20], 以测点的状态估计结果为中心, 得到各个测点估计值的置信区间。与残差理论不同, MNMR 方法能够通过真值落在量测区间内的概率将测点的量测值与真值联系起来, 且通过评价指标能够更好地对估计结果进行考核。同时, 该方法以测点正常率最大为目标, 认为落在置信区间外的测点为异常测点, 可以在估计的同时较好地识别不良数据, 不受残差淹没和残差污染的影响。但该方法建立的模型非凸, 同时不是连续可微, 只能将模型近似等效, 求解较为困难且容易陷入局部最优, 辨识效果变差。随着电力系统的不断发展, 系统规模逐渐增大, 在线状态估计和不良数据辨识需要满足实时性的要求, 并行求解技术^[21]和高性能计算框架^[22]得以运用。

为此, 本文从不良数据辨识的效果和速度角度出发, 采用文献^[11]中的状态估计模型, 提出一种基

收稿日期: 2018-10-29; 修回日期: 2019-04-25。

上网日期: 2019-06-11。

国家电网公司科技项目(52110418000M)。

于图形处理器(GPU)并行加速的不良数据辨识方法。该方法采用杂交变异粒子群优化(particle swarm optimization, PSO)算法进行求解^[23],并利用统一计算设备架构(CUDA)进行 GPU 并行加速。最后,通过算例分析对比辨识效果和计算效率,验证本文算法的有效性。

1 MNMR 电力系统状态估计算法

1.1 数学模型

本节将简要介绍 MNMR 算法的数学模型和求解思路。MNMR 算法以测点正常率最大为目标,文献[11]中给出了测点正常率的概念。根据扩展不确定度的概念,测点量测值的概率分布决定了其合成标准不确定度,进而决定了其扩展不确定度的值,与量测装置本身的精度有关。若量测真值落在区间 $[Z_i - U_i, Z_i + U_i]$ 中的概率为 p ,可以得到:

$$P(|h_i(\bar{x}) - Z_i| \leq U_i) = p \quad (1)$$

式中: \bar{x} 为状态估计真值; $h_i(\bar{x})$ 为测点 i 在状态 \bar{x} 下的量测函数; Z_i 为测点 i 的量测值; U_i 为测点 i 在置信概率 p 下的扩展不确定度值,实际运行中 p 取某个大于0.5的值,如0.997或0.945。

因此,对于任意一个测点 i ,若满足式(2)则认为正常测点。

$$|h_i(x) - Z_i| \leq U_i \quad (2)$$

式中: x 为状态估计计算结果; $h_i(x)$ 为测点 i 在状态 x 下的量测函数。

对于测点 i ,定义测点在估计结果 x 下的相对偏移 d_i 为:

$$d_i = \frac{h_i(x) - Z_i}{U_i} \quad (3)$$

根据式(2)可知,对于正常测点满足 $|d_i| \leq 1$,对于异常测点满足 $|d_i| > 1$,因此,根据“投票”思想可以得到测点的评价函数为:

$$g(d_i) = \begin{cases} 0 & |d_i| \leq 1 \\ 1 & |d_i| > 1 \end{cases} \quad (4)$$

以测点正常率最大为目标建立状态估计模型,也就是让全体测点的评价函数值之和最小。因此 MNMR 状态估计模型可以表示为:

$$\begin{cases} \min_x \sum_{i=1}^m g(d_i) \\ \text{s.t.} & d_i = \frac{h_i(x) - Z_i}{U_i} \\ & g(x) = 0 \\ & l(x) \leq 0 \end{cases} \quad (5)$$

式中: m 为测点总数; $g(x)$ 为潮流约束,如零功率注入节点的注入功率约束; $l(x)$ 为实际物理约束,如

线路潮流容量限制等。

1.2 模型特点

式(5)中求解的 MNMR 模型抗差性强,无论测点估计值偏离量测值多远,在目标函数中的大小都为1,具有较好的不良数据辨识能力。理想情况下,假设异常测点总数为 q ,那么满足:

$$\min_x \sum_{i=1}^m g(d_i) \approx q \quad (6)$$

通过求解该模型,可以将异常测点辨识出来。但由式(4)可以看出,目标函数不是处处连续可导,难以进行优化求解。文献[11]基于 sigmoid 函数对式(4)进行了近似,使其连续可导,得到的评价函数 $f(d_i)$ 具有以下特性。

$$f(d_i) \approx \begin{cases} 0 & |d_i| \leq 1 \\ 1 & |d_i| > 1 + \lambda \end{cases} \quad (7)$$

式中: λ 为大于0的常数。

根据式(7)特性,文献[11]中将 $f(d_i)$ 定义为:

$$\begin{cases} f(d_i) = \delta(d_i) + \delta(-d_i) \\ \delta(d_i) = \frac{1}{1 + e^{\frac{2r}{\lambda-1}[-d_i + (1 + \frac{\lambda-1}{2})]}} \end{cases} \quad (8)$$

式中: r 一般取大于2的自然数。

文献[11]采用近似后的 $f(d_i)$ 函数对求解模型进行了简化,而后采用现代内点法进行求解。但使用这种方式存在以下几个问题。

1)式(7)所示的近似模型无法辨别部分可疑测点。对于所有满足 $1 < |d_i| \leq 1 + \lambda$ 的测点,都无法判断这些测点是否为正常测点。如果将该近似模型用于不良数据辨识,会存在误检和漏检问题。

2)该近似模型中出现了两个参数 r 和 λ ,虽然文献[11]中大致进行了一个定量的分析,但对于不同的估计场景,仍需要进行参数调节,方法通用性受到限制。

3)该模型为非凸模型,近似后采用现代内点法进行求解,容易出现不收敛的情况,且计算结果会陷入局部最优,只能得到测点正常率较大的近似优化结果。

由此可以看出,目前采用的近似等效模型加上内点法的求解模式不能较好地处理 MNMR 状态估计模型。从状态估计的角度来看,这种求解模式可以求出一个测点正常率偏好的状态估计结果,但某些情况下求解出的结果并不是令人满意的。而从不良数据辨识的角度来看,在未简化式(4)的情况下, MNMR 算法从原理上来说具有很好的不良数据辨识能力,适合从不良数据辨识的角度对该模型加以运用。

1.3 求解思路

1) 由于式(4)所示的测点评价函数含有间断点,不是处处连续可导,如内点法等需要对函数进行求导而优化模型的方法难以直接运用,因此,本文考虑使用杂交变异 PSO 人工智能算法进行求解。这种方法能够跳出局部最优解,属于一类随机搜索方法,对于 MNMR 问题中出现的非连续优化问题能够很好地适应,在迭代搜索的过程中辨识出不良数据。

2) 使用杂交变异 PSO 算法,在迭代过程中计算量极其庞大。各个粒子在每次迭代中都要计算所有的量测函数值,且每次迭代会有多个粒子,迭代次数也会达到成千上万次。对于较大规模的系统,直接使用串行的算法计算速度是难以忍受的。因此,本文采用基于 GPU 并行加速的思想对该方法进行加速,设计粗粒度和细粒度的并行加速策略,使计算时间达到实际要求。

2 基于杂交变异 PSO 算法的优化模型求解

2.1 杂交变异 PSO 算法

标准 PSO 算法^[24]是一种模拟鸟群觅食行为的人工智能算法,该算法在计算过程中容易陷入局部最优。文献[23]在此基础上提出了一种杂交变异的 PSO 算法。在粒子迭代过程中加入遗传算法中的繁殖和变异思想,随机选中种群中的粒子进行两两杂交,再随机选中粒子进行高斯变异,进一步进行速度和位置的更新。子粒子的位置更新公式为:

$$x_{id,1}' = p_d x_{id,1} + (1 - p_d) x_{id,2} \quad (9)$$

$$x_{id,2}' = p_d x_{id,2} + (1 - p_d) x_{id,1} \quad (10)$$

式中: p_d 为 0~1 间的随机数; $x_{id,1}$ 和 $x_{id,2}$ 为第 i 组的 2 个父粒子第 d 维的位置;上标“'”表示子粒子。

子粒子的速度更新公式为:

$$v_{i,1}' = \frac{v_{i,1} + v_{i,2}}{|v_{i,1} + v_{i,2}|} |v_{i,1}| \quad (11)$$

$$v_{i,2}' = \frac{v_{i,1} + v_{i,2}}{|v_{i,1} + v_{i,2}|} |v_{i,2}| \quad (12)$$

式中: $v_{i,1}$ 和 $v_{i,2}$ 为第 i 组的 2 个父粒子的位置向量。

然后随机选中部分粒子进行高斯变异,变异公式为:

$$x_i'' = x_i (1 + \text{Gaussian}(\sigma)) \quad (13)$$

式中: x_i 为选中的第 i 个粒子的位置向量;上标“''”表示变异粒子; $\text{Gaussian}(\sigma)$ 表示满足位置参数为 0、尺度参数为 σ 的高斯分布随机变量值。

若选择各节点电压和相角作为状态变量,则节点电压和相角会作为粒子的位置值。上述杂交过程

就是选出部分粒子,两两进行状态变量的线性叠加。而变异过程就是选出部分粒子,对状态变量的每一维进行服从高斯分布的随机变异。至此,在迭代过程中能够拥有跳出局部最优解的能力。

2.2 优化模型求解

使用杂交变异 PSO 算法对式(5)进行求解与传统的 PSO 优化步骤大体相同,不同之处如下。

1) 更新粒子的位置和速度后,采用式(9)至式(13)随机选出部分粒子进行上述杂交变异过程。

2) 由于式(5)存在约束条件,计算适应度值时,若粒子不满足约束,则根据模型的约束条件计算粒子的“约束偏移距离”^[25],并在此基础上加上一个较大的常数作为惩罚值,保证可行粒子优先选择、距离可行域最近的粒子次优先选择的原则。

3) 在进行收敛性判断时,可以实现设定一个期望的评价函数和,若适应度值已小于该阈值,则退出迭代循环。

3 GPU 并行加速策略

3.1 CUDA 并行计算

CUDA 是一种 GPU 上的通用计算技术。在该架构中,CPU 负责派生出运行在 GPU 上的“内核”程序,然后 GPU 负责将内核程序分配到 GPU 硬件上进行处理^[26]。通过这种方式,能够将大量相同的计算任务交给 GPU 处理实现并行化,提升程序运行速度。与中央处理器(CPU)计算架构不同,GPU 的设计针对的是解决执行大量离散而不相关任务的系统,通过将问题拆分为大量可独立运行的子问题来实现高效并行工作。

CUDA 并行计算模型如图 1 所示。

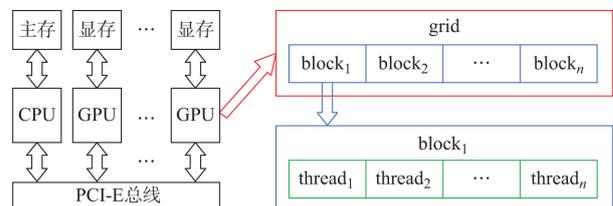


图 1 CUDA 并行计算模型
Fig.1 Parallel computation model of CUDA

CPU 和 GPU 必须通过 PCI-E 总线进行通信。由于总线带宽远远小于 CPU 和 GPU 的时钟频率,因此主存与显存之间的通信非常耗时,应该避免频繁的数据交互。在 GPU 上,每一个内核函数进行调用时都会占用一个线程网格(grid),该线程网格可以分为多个线程块(block),每个线程块内含有数量相同的线程(thread)。在一块 GPU 上含有多个

流处理器簇(streaming multiprocessor, SM),线程块中的线程会被组成为线程束(wrap)供 SM 进行调度和运行,从而实现多线程的并行执行。

3.2 并行性分析

GPU 通过开启大量的线程实现大规模并行化操作,在同一内核中,不同的线程对不同的数据执行相同的操作。因此,对于问题颗粒度分解得越细,越能充分发挥 GPU 的并行计算作用。

在本文所解决的问题中,采用杂交变异 PSO 算法求解式(5)模型,主要分为以下步骤。

步骤 1:各粒子更新自身速度和位置,该步骤是可并行的。

步骤 2:各粒子计算自身适应度值,该步骤是可并行的,主要表现为各粒子计算各个测点的量测函数 $h_l(x)$ 能够并行。

步骤 3:全体粒子更新个体历史最佳位置和种群历史最佳位置,该步骤也是可并行的。

这 3 个子步骤之间需要按顺序执行。为了保证 GPU 运算的高效,需要将该问题拆分为尽可能细的子问题,从而在同一时刻让尽可能多的线程进行计算,提高计算效率。

3.3 细粒度并行方法

细粒度的并行方法是让 GPU 的每一个线程对应处理目标问题的每一个最小任务单元。对于本问题,就是让 GPU 的每一个线程计算解集中的每个维度,从而能够开启大量的线程提高并行度。细粒度的并行方法在以上 3 个步骤中均有体现。

步骤 1 是对各粒子位置和速度的更新,其细粒度的并行方式如图 2(a)所示,其中, p_1 至 p_n 为粒子, n 为粒子总数, d 为粒子维度。由于各粒子每一维的位置和速度值在更新过程中都是独立的,每一个线程对应计算一维数据的更新值,相互不受干扰,达到最细粒度的并行。更新完粒子位置和速度后,会挑选出部分线程,实现式(9)至式(13)的杂交变异过程。在该步骤中,线程数与 nd 相等。

步骤 2 是计算各粒子的适应度值,主要是计算式(5)中的目标函数,其细粒度的并行方式如图 2(b)所示, m_l 表示第 l 类量测的测点总数。以节点注入功率量测为例对此进行解释,该量测可以表示为^[27]:

$$P_i = \sum_{j=1}^N v_i v_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \quad (14)$$

式中: P_i 为节点 i 的有功注入; N 为节点个数; v_i 和 v_j 分别为节点 i 和 j 的电压幅值; θ_{ij} 为节点 i 和 j 的相角差; G_{ij} 和 B_{ij} 分别为节点导纳矩阵对应元素的实部和虚部。

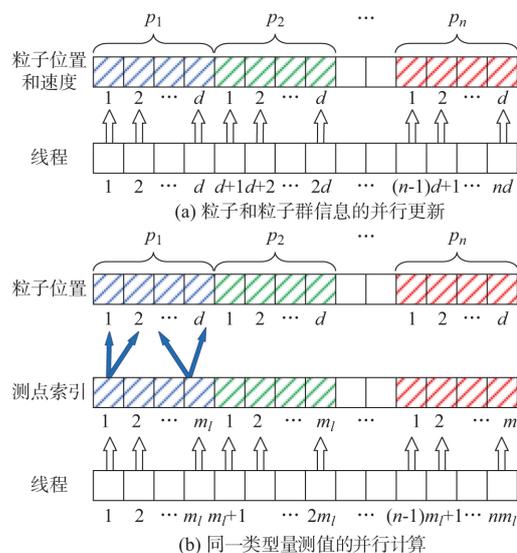


图 2 细粒度并行策略

Fig.2 Fine-grained parallel strategy

在计算节点注入功率的内核程序中,所有线程只对节点注入功率测点进行计算。各线程通过测点索引获取与该测点相关的粒子位置值,然后对该测点进行评价函数值的计算。

显然,式(5)中的目标函数是粒子所有测点相对偏移的评价函数,而不同类型量测函数的计算公式不同,式(14)只是其中一种。因此,应对编写不同的内核程序。在同一个内核程序中,所有线程对该量测类型下各粒子的每一个测点进行计算。最终,各线程计算出的测点评价函数值通过原子操作相加至对应粒子的适应度值中。在该步骤中,各内核线程数与 nm_l 相等。

步骤 3 是对个体历史最佳位置和种群历史最佳位置的更新,其细粒度的并行方式与步骤 1 大致相同。不同之处在于,若粒子在本次迭代中所得到的适应度值没有比历史最优最好,则该粒子的对应线程不会进行更新。

3.4 粗粒度并行方法

粗粒度的并行方法主要针对步骤 2。由于 3.3 节 3 个步骤中,步骤 2 的计算量很大,占计算总时间的绝大部分。从细粒度的并行策略中可以看出,需要计算粒子在目前状态下,各个测点的评价函数值总和。但在该步骤中,对于多种不同的量测类型,由于同一个内核函数只能实现一种量测函数值的计算,采用细粒度的并行方法各线程只能同时计算同一类型量测的测点评价函数值。所以,对于不同的量测类型,需要编写对应的内核函数。考虑并行执行各类型量测的内核函数,可以实现粗粒度的并行,如图 3 所示。

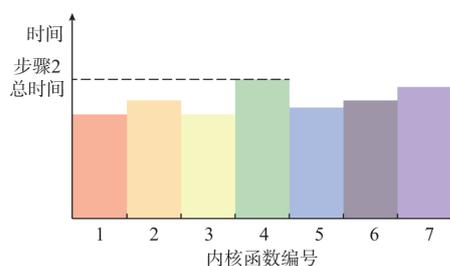


图3 粗粒度并行策略
Fig.3 Coarse-grained parallel strategy

GPU 在同一个时刻可以运行多个相互无关的流,一个流是一个工作队列。将节点电压、节点注入有功和无功功率、线路流入和流出的有功和无功功率共 7 种量测类型对应的 7 种不同的内核函数放入不同的流队列中,使用 GPU 进行并行流计算,多内核并行执行,实现网格级别粗粒度的并行。对于每一个内核函数,其作用是计算各粒子在该内核所对应的量测类型下所有测点的评价函数和。该步骤计算完后需要进行各个流的同步,总计算时间为各内核计算时间的最大值。

4 本文算法求解步骤

综上,本文所提出的基于 GPU 并行加速的最大测点正常率不良数据辨识方法的算法步骤如下。

1) 根据输入数据,形成节点导纳矩阵和各测点扩展不确定度值。

2) 初始化粒子群,在初值附近设置粒子的位置区间、速度区间,在区间范围内随机初始化各粒子的位置和速度,计算各粒子适应度值,若该粒子不满足约束,则该适应度值需要根据 2.2 节中的描述计算适应度值,惩罚值选为当前系统测点总数。

3) 根据适应度值设置粒子的个体最优位置、个体最优适应度值、全局最优位置和全局最优适应度值。

4) 将主机内存数据拷贝至 GPU 显存,并设置迭代次数 $k=0$ 和最大迭代次数 k_{max} 。

5) 在 GPU 上进行迭代步骤,具体如下。

① 启动粒子更新自身速度和位置的内核程序,选出部分粒子根据式(9)至式(13)进行杂交变异,并同步 GPU 线程。

② 启动计算各类型量测数据的测点评价内核程序,将各个内核程序设置到不同的流队列中并启动,最后同步 GPU 的所有流。

③ 根据式(5)中的等式约束和不等式约束,若粒子不满足约束,则步骤②中计算出的适应度值还需额外加上惩罚值。

④ 启动更新个体最优位置、个体最优适应度值、全局最优位置和全局最优适应度值的内核程序,并同步 GPU 线程。

⑤ 在 CPU 上令 $k=k+1$,若 k 还未达到 k_{max} ,则跳转到步骤①,否则跳出步骤⑤。

⑥ 将全局最优位置和全局最优适应度值从 GPU 显存拷贝至主机内存中。根据全局最优位置计算式(4),所有满足 $g(d_i)=1$ 的测点即为异常测点,则可以得到异常测点集合。

5 算例分析

为了验证本文方法的有效性,采用多个算例对本文算法进行测试。其中,算例 1 至算例 6 是 IEEE 标准算例,算例 7、算例 8 和算例 9 是某些实际输电网算例,各算例对应的节点数分别为 14, 30, 39, 57, 118, 300, 598, 6 386, 10 045。所有算例均在 64 bit 的 Windows 10 操作系统下测试,CPU 型号为 Intel Core i7-7700K,运行内存为 32 GB;GPU 型号为 NVIDIA GeForce GTX1080,支持 CUDA9.2 标准。

在标准系统算例中,假设各测点的量测值服从高斯分布,将计算所得的潮流数据添加 2% 的高斯噪声得到试验所用的量测生数据,得到扩展不确定度值。通过对量测值改变符号、置 0 或者加减 20% 等手段来模拟制造不良数据^[16]。选取节点电压和相角作为状态变量,粒子群总数设置为 $10 + 2\sqrt{a}$ ^[28],电压位置区间设置为 $[0.95, 1.05]$,相角位置区间设置为 $[-1.57, 1.57]$,粒子杂交概率设置为 0.5,变异概率设置为 0.2,变异参数 σ 设置为位置区间长度的 10%,最大迭代次数为 10 000 次。在本节中,粒子群初值通过文献[11]中所采用的近似方法计算得出。

5.1 相同算例的多不良数据辨识度分析

为了观察系统中不良数据的占比对本文算法的影响,以算例 1 为基础,设置 1%~6% 的不良数据,将本文算法效果与文献[11]的算法进行对比。其中,根据文献[11]中的描述,参数 r 取 2, λ 在第一阶段迭代取 54.059 5,第二阶段取 2。算例 1 中一共模拟生成了 107 个测点。具体计算结果详见附录 A 表 A1。

从表 A1 中可以看出,在 IEEE 14 节点系统(算例 1)下,本文所采用的算法在各情况下都能实现不良数据的精确检索,没有发生误检和漏检现象。文献[11]算法在不良数据较少(小于 4%)时效果较好,没有漏检测点,误检测点也保持在较低水平;在不良数据较多(大于 4%)时解的合理性开始下降,

出现了漏检测点且误检测点也有上升。算例结果表明,本文算法在不良数据占比较少和较多的情况下都能实现不良数据的精确检索。

5.2 不同算例的不良数据辨识度分析

为了测试不同规模的算例下本文算法对于不良数据辨识的影响,设置 3% 的不良数据,将本文算法效果与文献[11]的算法进行对比。其中,参数 r 和 λ 取值与文献[11]中相同。计算结果如表 1 和图 4 所示。

表 1 各 IEEE 测试系统不良数据辨识结果对比
Table 1 Identification result comparison of bad data identification results in IEEE test systems

算例	不良数据个数	本文算法		文献[11]算法	
		误检个数	漏检个数	误检个数	漏检个数
1	3	0	0	2	0
2	7	2	0	12	1
3	7	1	0	17	1
4	12	8	0	10	2
5	31	5	0	18	6
6	63	22	2	66	17

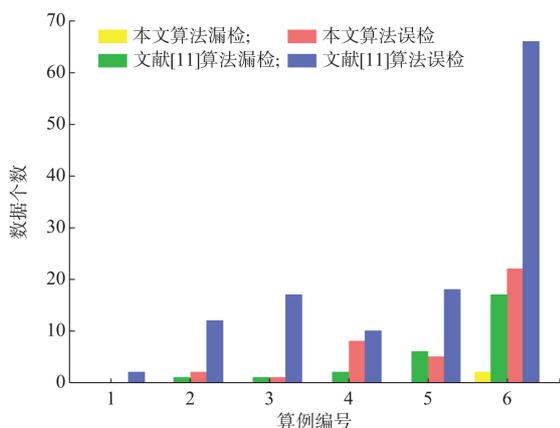


图 4 IEEE 各节点系统不良数据辨识结果
Fig.4 Bad data identification results of IEEE systems node

从表 1 和图 4 可以看出,文献[11]算法由于对评价函数曲线进行了近似,获得的结果中有较多的测点属于可疑测点,无法判断其是否正常。因此,随着系统规模的增大,该算法的误检个数和漏检个数都有上升,效果逐渐变差。且在测试过程中,使用文献[11]算法进行了多次的参数调节才取得了这些相对较好的结果。而本文算法的漏检测点个数都保持在较低的水平,对于在含有 2 000 个测点以上的 IEEE 300 节点系统(算例 6)中,本文算法仅有 2 个漏检测点。同时,本文算法的误检测点个数也能够保持较低水平,不到测点总数的 1%,剔除后对系统可观性几乎无影响。算例结果表明,本文算法可以

得到更为合理不良数据辨识结果。

5.3 粒子个数分析

为了测试不同粒子数对于本文算法计算结果的影响,设置 3% 的不良数据,测试漏检的个数,具体计算结果详见附录 A 表 A2。

从附录 A 表 A2 中可以看出,当算例规模较小时,较少的粒子数可以满足辨识要求;当算例规模增大时,粒子数目越多辨识效果越好,但计算规模也会相应增加。对于较大规模算例,若粒子数目较少,算法在初值的基础上难以搜索到更优的解。粒子群总数设置为 $10+2\sqrt{d}$ 时,基本能够满足不良数据辨识的要求,同时能够保证计算时间要求。

5.4 并行加速比分析

为了测试本文算法的加速效果,在各算例中设置 3% 的不良数据,采用串行算法和本文算法进行比较,验证并行算法的有效性。计算结果如表 2 所示,将计算时间以对数坐标为纵轴,绘制对比图如图 5 所示。

表 2 串并行算法计算时间比较
Table 2 Comparison of calculation time between serial and parallel algorithms

算例	并行算法用时/s	串行算法用时/s	加速时间/s	加速比
1	7.77	0.64	-7.12	0.08
2	8.02	1.36	-6.67	0.17
3	6.90	1.38	-5.52	0.20
4	8.10	2.93	-5.16	0.36
5	8.34	4.45	-3.89	0.53
6	9.30	28.68	19.38	3.08
7	8.85	383.03	374.18	43.26
8	53.68	7 069.93	7 016.25	131.71
9	92.52	12 495.28	12 402.76	135.06

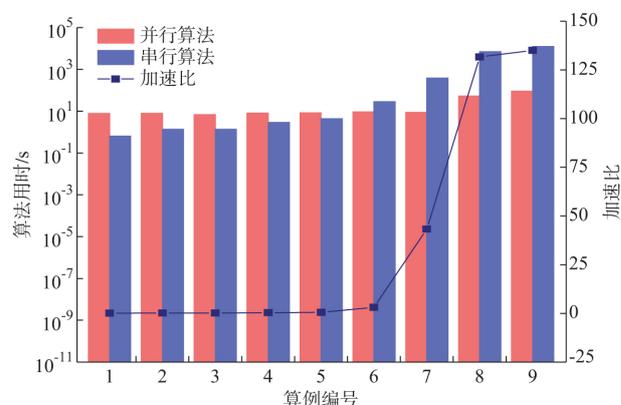


图 5 所提出的算法加速效果分析
Fig.5 Acceleration effect analysis of proposed algorithm

从表 2 可以看出,在系统规模较小时,并行算法运行时间较长,反而落后于串行算法。然而,当系统

节点数达到 300 以上时,并行算法的计算效率有了非常大的提升。从图 5 可以看出,节点数小于 300 时,并行算法的计算时间比串行算法慢 1 个数量级左右;节点数大于 300 时,并行算法的计算时间比串行算法快 1~2 个数量级。从并行计算原理上可以对这个现象进行解释。当系统规模较小时,主存与显存之间的通信、线程调度等步骤耗时占比较大,使得采用并行算法反而效果更差;当系统规模变大时,问题被分解为大量的线程块交给 GPU 的 SM 进行处理,使得整体求解时间变化不大;当系统规模进一步变大时,GPU 中的各 SM 都始终保持忙碌状态,一块 SM 中有多个线程块等待其调度和运行,求解时间逐渐上升。随着系统规模的增大,并行算法能够拥有更大的加速比,计算效率更高。对于万级节点系统,加速比达到了 135.06,能在 1.5 min 左右完成计算,符合实际要求,加速效果较好。

5.5 内存占用测试

大规模高性能计算时,内存占用是算法不可忽略的重要性能因素。为了测试本文算法的资源占用情况,在各算例中对本文算法进行内存占用和显存占用测试,具体计算结果详见附录 A 表 A3。

从附录 A 表 A3 中可以看出,对于万级节点系统,本文算法的内存占用峰值仅为 332 MB,显存占用峰值仅为 231 MB。总体而言,本文算法对内存和显存容量的要求较低。

6 结语

本文以 MNMR 状态估计算法为基础,提出了一种基于 GPU 并行加速的量测不良数据辨识方法,该方法具有以下特点。

1) 算法没有对 MNMR 模型进行近似等效,辨识过程中不会出现可疑测点,无须进行参数调节,且不存在收敛性问题,全局搜索能力强。

2) 算法有较好的不良数据辨识能力,能够保持较低的误检率和极低的漏检率。

3) 算法采用 CUDA 并行计算架构,通过粗粒度和细粒度的并行策略进行加速,加速后的算法计算时间短,加速效率高,能够满足实际运行需求。

4) 算法的内存和显存占用峰值较小,在普通的个人计算机上即可进行计算。

算例结果表明,本文算法对于不良数据的辨识效果好,计算速度快,资源占用低。在后续的研究中,可以考虑在辨识前进行预处理,尝试过滤正常测点的噪声误差,降低本方法的误检率。

附录见本刊网络版 (<http://www.aeps-info.com/aeps/ch/index.aspx>)。

参考文献

- [1] SCHWEPPE F C, WILDES J. Power system static-state estimation: Part one exact model[J]. IEEE Transactions on Power Apparatus & Systems, 1970, 89(1): 125-130.
- [2] CELIK M K, ABUR A, MELIOPOULOS A P, et al. A robust WLAV state estimator using transformations [J]. IEEE Transactions on Power Systems, 1992, 7(1): 106-113.
- [3] BALDICK R, CLEMENTS K A, PINJO-DZIGAL Z, et al. Implementing nonquadratic objective functions for state estimation and bad data rejection [J]. IEEE Transactions on Power Systems, 1997, 12(1): 376-382.
- [4] MILI L, CHENIAE M G, ROUSSEUW P J. Robust state estimation of electric power systems [J]. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 2002, 41(5): 349-358.
- [5] 卢志刚,程慧琳,冯磊,等.基于证据融合理论的多不良数据辨识 [J]. 电网技术, 2012, 36(1): 123-128.
LU Zhigang, CHENG Huilin, FENG Lei, et al. Multi-bad data identification based on evidence fusion theory [J]. Power System Technology, 2012, 36(1): 123-128.
- [6] 孙国强,卫志农,周封伟.改进迭代自组织数据分析的不良数据辨识 [J]. 中国电机工程学报, 2006, 26(11): 162-166.
SUN Guoqiang, WEI Zhinong, ZHOU Fengwei. The application of ISODATA to bad data detection and identification based on genetic algorithms [J]. Proceedings of the CSEE, 2006, 26(11): 162-166.
- [7] SINGH D, PANDEY J P, CHAUHAN D S. Topology identification, bad data processing, and state estimation using fuzzy pattern matching [J]. IEEE Transactions on Power Systems, 2005, 20(3): 1570-1579.
- [8] 卢志刚,王浩锐,孙继凯.基于灵敏度分析的数据最优筛选与不良数据辨识 [J]. 电网技术, 2011, 35(2): 38-42.
LU Zhigang, WANG Haorui, SUN Jikai. Optimal data screening and bad data identification based on sensitive analysis [J]. Power System Technology, 2011, 35(2): 38-42.
- [9] 卫志农,张云岗,郑玉平.基于量测量突变检测的新方法 [J]. 中国电机工程学报, 2002, 22(6): 35-38.
WEI Zhinong, ZHANG Yungang, ZHENG Yuping. The improvement of measurement suddenly-change detection method [J]. Proceedings of the CSEE, 2002, 22(6): 35-38.
- [10] 张海波,李林川.电力系统状态估计的混合不良数据检测方法 [J]. 电网技术, 2001, 25(10): 17-20.
ZHANG Haibo, LI Linchuan. A hybrid approach for detection of bad data in power system state estimation [J]. Power System Technology, 2001, 25(10): 17-20.
- [11] HE Guangyu, DONG Shufeng, QI Junjian, et al. Robust state estimator based on maximum normal measurement rate [J]. IEEE Transactions on Power Systems, 2011, 26(4): 2058-2065.
- [12] 何光宇,董树锋.基于测量不确定度的电力系统状态估计(一)

- 结果评价[J].电力系统自动化,2009,33(19):21-24.
HE Guangyu, DONG Shufeng. Power system static-state estimation based on uncertainty of measurement: Part one result evaluation[J]. Automation of Electric Power Systems, 2009, 33(19): 21-24.
- [13] 何光宇,董树锋.基于测量不确定度的电力系统状态估计(二)方法研究[J].电力系统自动化,2009,33(20):32-36.
HE Guangyu, DONG Shufeng. Power system static-state estimation based on uncertainty of measurement: Part two a new method[J]. Automation of Electric Power Systems, 2009, 33(20): 32-36.
- [14] 何光宇,董树锋.基于测量不确定度的电力系统状态估计(三)算法比较[J].电力系统自动化,2009,33(21):28-31.
HE Guangyu, DONG Shufeng. Power system static-state estimation based on uncertainty of measurement: Part three algorithms compared [J]. Automation of Electric Power Systems, 2009, 33(21): 28-31.
- [15] WANG Bin, HE Guangyu, DONG Shufeng. An improved algorithm for state estimator based on maximum normal measurement rate[J]. IEEE Transactions on Power Systems, 2013, 28(4): 4879-4880.
- [16] 常乃超,王彬,何光宇,等.以测点正常率最大为目标的状态估计改进算法[J].电力系统自动化,2014,38(11):62-67.
CHANG Naichao, WANG Bin, HE Guangyu, et al. An improved algorithm for state estimator based on maximum normal measurement rate[J]. Automation of Electric Power Systems, 2014, 38(11): 62-67.
- [17] 陈艳波,谢瀚阳,王金丽,等.基于不确定测度的电力系统抗差状态估计(一)理论基础[J].电力系统自动化,2018,42(1):8-15.DOI:10.7500/AEPS20161213006.
CHEN Yanbo, XIE Hanyang, WANG Jinli, et al. Uncertain measure based robust state estimation of power system: Part one theoretical principle[J]. Automation of Electric Power Systems, 2018, 42(1): 8-15. DOI: 10.7500/AEPS 20161213006.
- [18] 陈艳波,谢瀚阳,王鹏,等.基于不确定测度的电力系统抗差状态估计(二)模型方法[J].电力系统自动化,2018,42(2):26-33. DOI:10.7500/AEPS20161213007.
CHEN Yanbo, XIE Hanyang, WANG Peng, et al. Uncertain measure based robust state estimation of power system: Part two model and solution[J]. Automation of Electric Power Systems, 2018, 42(2): 26-33. DOI: 10.7500/AEPS 20161213007.
- [19] 陈艳波,谢瀚阳,王鹏,等.基于不确定测度的电力系统抗差状态估计(三)算法对比[J].电力系统自动化,2018,42(3):8-13. DOI:10.7500/AEPS20161213008.
CHEN Yanbo, XIE Hanyang, WANG Peng, et al. Uncertain measure based robust state estimation of power system: Part three algorithm comparison [J]. Automation of Electric Power Systems, 2018, 42(3): 8-13. DOI: 10.7500/AEPS 20161213008.
- [20] BIPM, IEC, IFCC, et al. Guide to the expression of uncertainty in measurement[S]. 1993.
- [21] 陈荃韡,龚成明,赵晋泉,等.并行稀疏系统直接求解库 SuperLU_MT 在状态估计中的应用[J].电力系统自动化,2017,41(3):83-88.DOI:10.7500/AEPS20160607008.
CHEN Quanwei, GONG Chengming, ZHAO Jinquan, et al. Application of parallel sparse system direct solver library SuperLU_MT in state estimation[J]. Automation of Electric Power Systems, 2017, 41(3): 83-88. DOI: 10.7500/AEPS20160607008.
- [22] 郭健,周京阳,李强,等.高性能在线分析计算现状与协同计算关键技术[J].电力系统自动化,2018,42(3):149-159. DOI:10.7500/AEPS20170320007.
GUO Jian, ZHOU Jingyang, LI Qiang, et al. Current status of high-performance on-line analysis computation and key technologies for cooperating computation[J]. Automation of Electric Power Systems, 2018, 42(3): 149-159. DOI: 10.7500/AEPS20170320007.
- [23] HIGASHI N, IBA H. Particle swarm optimization with Gaussian mutation[C]// Proceedings of the 2003 IEEE Swarm Intelligence Symposium (SIS 03), April 26, 2003, Indiana, USA: 72-79.
- [24] KENNEDY J, EBERHART R. Particle swarm optimization [C]// Proceedings of ICNN'95-International Conference on Neural Networks, November 27-December 1, 1995, Perth, Australia: 1942-1948.
- [25] PULIDO G T, COELLO C C. A constraint-handling mechanism for particle swarm optimization[C]// Proceedings of the 2004 Congress on Evolutionary Computation, June 19-23, 2004, Portland, USA.
- [26] COOK S.CUDA 并行程序设计[M].北京:机械工业出版社,2014.
COOK S. CUDA programming: a developer's guide to parallel computing with GPUs [M]. Beijing: China Machine Press, 2014.
- [27] 于尔铿.电力系统状态估计[M].北京:水利电力出版社,1985.
YU Erkeng. Power system state estimation [M]. Beijing: China Water & Power Press, 1985.
- [28] JAMES K, MAURICE C. PSO version 2006[EB/OL]. (2006-02-27) [2018-04-22]. http://www.particleswarm.info/Standard_PSO_2006.c.

方睿(1995—),男,硕士研究生,主要研究方向:状态估计、电力系统高性能计算。E-mail: garrybest@foxmail.com

董树锋(1982—),男,通信作者,博士,副教授,主要研究方向:状态估计、有源配电网分析。E-mail: dongshufeng@zju.edu.cn

唐坤杰(1994—),男,博士研究生,主要研究方向:电力系统高性能计算方法。E-mail: tangkunjie1994@163.com

(编辑 施冬敏)

(下转第 115 页 continued on page 115)

附录 A

表 A1 IEEE 14 节点系统不良数据辨识结果对比

Table A1 Comparison of bad data identification results of IEEE 14-bus system

不良数据个数	不良数据占比	本文算法		文献[11]算法	
		误检个数	漏检个数	误检个数	漏检个数
1	1%	0	0	2	0
2	2%	0	0	2	0
3	3%	0	0	2	0
4	4%	0	0	2	0
5	5%	0	0	5	1
6	6%	0	0	5	2

表 A2 不同粒子数下不良数据辨识结果对比

Table A2 Comparison of bad data identification results of different swarm sizes

算例	不良数据个数	粒子维度	漏检个数			
			粒子个数			
			20	$10 + \sqrt{d}$	$10 + 2\sqrt{d}$	$10 + d$
case 1	3	28	0	0	0	0
case 2	7	60	0	0	0	0
case 3	7	78	0	1	0	0
case 4	12	114	2	2	0	0
case 5	31	236	5	3	0	0
case 6	63	600	15	6	2	1

表 A3 算法内存占用

Table A3 Memory usage of the algorithm in this paper

算例	量测总数	内存占用峰值/MB	显存占用峰值/MB
case 1	107	179	115
case 2	226	183	115
case 3	248	181	115
case 4	406	183	117
case 5	1 053	229	118
case 6	2 003	234	118
case 7	2 666	247	119
case 8	19 197	316	175
case 9	27 252	332	231

[16] RUIZ P A, GOLDIS E, RUDKEVICH A M, et al. Security-constrained transmission topology control MILP formulation using sensitivity factors [J]. IEEE Transactions on Power Systems, 2017, 32(2): 1597-1605.

葛朝强(1975—),男,博士,高级工程师,主要研究方向:电力系统及其自动化、电力市场。

吴鑫(1986—),女,硕士,工程师,主要研究方向:电力系统及其自动化、电力市场。

(编辑 章黎)

徐帆(1984—),男,通信作者,高级工程师,主要研究方向:电力系统运行和优化。E-mail: xufansea@sina.com

Market Mechanism and Clearing Model of Inter-provincial Peak Regulation Ancillary Service for Regional Power Grid

XU Fan^{1,2}, GE Zhaoqiang³, WU Xin³, ZHU Minjian^{1,2}, TU Mengfu^{1,2}

(1. NARI Group Corporation (State Grid Electric Power Research Institute), Nanjing 211106, China;

2. State Key Laboratory of Smart Grid Protection and Control, Nanjing 211106, China;

3. East China Branch of State Grid Corporation of China, Shanghai 200120, China)

Abstract: Aiming at the problem of peak regulation caused by clean energy accommodation in the regional power grid, a market clearing model for peak regulation ancillary service is proposed. The market-oriented mechanism is used to compensate the units participating in peak regulation, which solves the difficulty of energy return under the current inter-provincial energy replacement mechanism. According to the capacity of available peak regulation and the limit of deep peak regulation, the unit declares the peak regulation capacity and compensation price in stages to participate in the market. The clearing results meet the unit operation constraints and network security constraints and have good performance in operation. The clearing price of peak regulation market has clear market guidance function, which means the peak regulation market can provide a clear price signal to attract more units to provide peak regulation capacity, and promote the effective accommodation of clean energy in the regional power grid. Case studies are conducted based on the simulated operational data of peak regulation auxiliary service market for the regional power grid to validate the effectiveness of the proposed model.

This work is supported by State Grid Corporation of China.

Key words: peak regulation market; ancillary service; deep peak regulation; market clearing

(上接第 93 页 continued from page 93)

Bad Data Identification Method Based on Maximum Normal Measurement Rate and GPU Parallel Acceleration

FANG Rui¹, DONG Shufeng¹, TANG Kunjie¹, ZHU Chengzhi², PEI Tian², SONG Yonghua³

(1. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China;

2. State Grid Zhejiang Electric Power Co. Ltd., Hangzhou 310007, China;

3. Department of Electrical and Computer Engineering, University of Macau, Macau, China)

Abstract: Based on the concept of measurement uncertainty, the robust state estimation method for power system with maximum normal measurement rate (MNMR) has good identification ability of bad data. However, the model is difficult to solve. In existing researches, the model is approximated and solved by the modern interior point method, but the problems such as lower identification effect due to approximation are existed. Therefore, based on the state estimation model of MNMR, a hybrid particle swarm optimization (PSO) algorithm with hybrid mutation is used to propose a bad data identification algorithm based on graphics processing unit (GPU) parallel acceleration. Without approximating the MNMR model and according to the architecture characteristics of the GPU parallel computation, the algorithm designs a parallel acceleration strategy combining coarse and fine granularity. The results of the case analysis show that the proposed algorithm has low false detection rate and missed detection rate for bad data, and has good identification ability of bad data, short calculation time and high acceleration efficiency, which could meet the actual operation requirements.

This work is supported by State Grid Corporation of China (No. 52110418000M).

Key words: data identification; state estimation; normal measurement rate; graphics processing unit (GPU); parallel computation