

面向配电网边缘计算装置的两阶段 PMU 数据压缩方法

习伟¹, 李鹏², 李鹏¹, 姚浩², 陈军健², 杨家辉¹, 于浩¹

(1. 天津大学 智能电网教育部重点实验室, 天津市 南开区 300072;

2. 南方电网数字电网研究院有限公司, 广东省 广州市 510700)

A two-stage PMU data compression method for edge computing devices of distribution networks

XI Wei¹, LI Peng², LI Peng¹, YAO Hao², CHEN Junjian², YANG Jiahui¹, YU Hao¹

(1. Key Laboratory of the Ministry of Education on Smart Power Grids, Tianjin University, Nankai District, Tianjin 300072, China;

2. Digital Grid Research Institute of China Southern Power Grid, Guangzhou 510700, Guangdong Province, China)

ABSTRACT: Due to the high upload rate of phasor measurement units for distribution networks, massive data is generated and puts significant strain on communication and storage systems. Considering the limited computing and storage resources of edge computing devices in the distribution networks, a data compression method of phasor measurements based on exception-swing door trending and exponential-Golomb coding is proposed. First, the swing door trending criterion is modified. By the improved exception-swing door trending algorithm, the original phasor measurements are lossily compressed. Then, the differential coding and normalization are applied to the compressed data as preprocessing. The 0-order exponential-Golomb coding is used to further reduce the data redundancy losslessly. Finally, the method is tested based on practical phasor measurements. Compared with the separate exception-swing door trending algorithm and general-purpose compression tools, the method achieves a higher compression ratio within the predefined precision requirement.

KEY WORDS: distribution networks; synchronous phasor measurements; data compression; exception-swing door trending; exponential-Golomb coding; edge computing

摘要: 针对配电网高上送速率的 PMU 装置产生的海量数据对通信系统和存储系统造成巨大压力的问题, 考虑到配电网边缘计算装置计算与存储资源有限, 提出了基于过滤旋转门和指数哥伦布编码的配电网边缘侧 PMU 数据压缩方法。首先, 对旋转门压缩判据进行改进, 利用改进的过滤旋转门压

缩算法对原始 PMU 数据进行有损压缩。其次, 对有损压缩数据进行差分编码和归一化处理, 利用 0 阶指数哥伦布编码方法进一步无损压缩, 降低数据冗余。最后, 利用实际的 PMU 数据, 对所提出的方法进行验证。与过滤旋转门压缩算法和通用压缩工具相比, 该方法在预先设定的精度要求下, 实现了更高的压缩比。

关键词: 配电网; 同步相量量测; 数据压缩; 过滤旋转门; 指数哥伦布编码; 边缘计算

0 引言

各种先进的信息化与数字化技术在电网中的应用极大地加速了电网的数字化转型, 云-边协同架构成为数字电网建设与运行中的一种新模式。其中, 在配电层面, 利用边缘计算技术实现海量数据的就地采集、分析、处理以及在此基础上的智能分析与决策, 是充分释放数据资源价值、突破数据大规模通信传输瓶颈、提高快速多变环境下决策快速性的重要手段^[1]。但与此同时, 配电网边缘计算装置的计算与存储资源有限, 并需要同时承载保护、控制、监测等多类型电网业务, 对单一业务占用资源的约束更加严格和关键, 这对边缘计算环境下的电网业务逻辑和算法设计提出了更高要求^[2]。

近年来, 以同步相量测量 (phasor measurement unit, PMU) 为代表的新型量测技术应用愈发广泛^[3], 并已开始应用于配电网。PMU 利用全球定位系统给每帧数据打上精确时标, 保证了配电网电压相量、电流相量^[4]等数据获取的同步性^[5]; 数据上送速率可达到 100 帧/s^[6], 保证了数据获取的实时性。上述

国家重点研发计划项目 (编号: 2020YFB0906000, 2020YFB0906001)。

Project Supported by the National Key Research and Development Program of China(No. 2020YFB0906000, 2020YFB0906001).

优势使 PMU 可为运行人员提供更丰富的配电网运行信息,有效提升了配电网的运行状态感知水平^[7],为故障定位^[8]、状态估计^[9]等应用提供了重要的数据支撑。因此,PMU 功能被作为配电网边缘计算的核心业务功能之一。然而,由于其多类型、高频次量测特征,PMU 采集到的数据量将急剧增长,给其在边缘侧的就地处理和存储带来了巨大挑战。文献[10]指出,包含 20 种量测信息、数据上送速率为 60Hz 的 PMU,一天产生的数据量可达 100GB。这极大地增加了 PMU 业务在配电网边缘计算环境下的部署和应用难度。

数据压缩是指在保留有价值信息的前提下,通过对数据的处理或重新组织以缩减数据规模,提高其传输、存储和处理效率^[11],是解决边缘计算环境下 PMU 数据规模过大问题的一种重要方案。目前,针对 PMU 数据的压缩方法主要包括有损压缩和无损压缩两大类。其中,有损压缩是不可逆的,即经过压缩的数据与源数据存在一定的误差,但压缩比相对无损压缩较高,常用方法包括小波压缩算法、JPEG 图像压缩算法、傅里叶压缩算法等^[12]。例如,文献[10]利用 PMU 数据固有的相关性,提出了一种两阶段压缩算法,第一阶段采用主成分分析,第二阶段采用离散余弦变换,降低了多个 PMU 在空间和时间上的数据冗余;文献[13]将过滤压缩算法和旋转门压缩算法结合,从而兼具了二者的优点,可同时满足历史数据和实时数据的压缩需求;文献[14]提出了一种基于奇异值分解的智能配电系统有损数据压缩方法,能够显著减少通信网络传输的数据量,并能准确地重构原始数据;文献[15]将电力数据转化为图像,采用高斯滤波下采样方法对图像进行压缩,结合通道注意机制与空间注意机制,构建了一种残差双重注意机制网络,实现了对压缩数据的高精度重构;文献[16]针对配电网异构数据,提出了基于张量 Tucker 分解的配电网大数据压缩方法;文献[17]提出一种自适应联合重构算法,并应用到以分布式压缩感知为边缘算法的云边协同框架中,在云端对边缘上传的字典原子和测量值进行分析,实现电能质量数据的压缩存储。

无损压缩是可逆的,即在压缩过程中不损失数据精度,经过压缩的数据解压后与源数据相同,但压缩比通常较低,常用方法包括 Huffman 编码、算术编码、哥伦布编码等^[12]。目前,无损压缩算法主要应用于图像和视频压缩领域,在 PMU 数据压缩中的应用仍处于探索阶段。例如,文献[18]提出了一种新的基于分层预测和上下文自适应算术编码

的无损彩色图像压缩算法,降低了彩色图像编码数据大小;文献[19]基于模糊逻辑方法计算视频相邻帧的块间及块内相似性,以确定时间预测和空间预测的选择,最后通过相似度估计哥伦布编码的最优参数,实现了视频数据的无损压缩;文献[20]研究了风速和风能的关系以及数据间的时空相关性,结合香农熵理论,提出了一种用于压缩风电厂数据的无损数据压缩方法;文献[21]提出了一种结合差分编码和异或编码的变长编码框架,用于智能电网数据的无损压缩。

总体来看,对于 PMU 数据压缩问题,现有方法大多基于独立的有损压缩或无损压缩算法,在压缩比和重构精度方面仍有进一步提升空间,在边缘计算环境下的适用能力也有待探索。为此,本文首先在过滤旋转门压缩算法的基础上,采用改进的旋转门压缩判据提高了重构精度;并进一步将其与指数哥伦布无损编码算法相结合,在同等精度下实现了更高的压缩比。基于实际配电网 PMU 数据的测试结果表明,本文提出的方法具有良好的压缩效果,且计算量和内存占用较小,适合在边缘侧应用。

1 基于改进过滤旋转门的 PMU 数据压缩

过滤旋转门压缩算法可分为过滤压缩和旋转门压缩两个阶段^[13]。在每轮压缩中,首先读取一个压缩周期的原始数据,然后进行过滤压缩判断,接着对过滤压缩的输出进行改进的旋转门压缩判断,即可实现基于改进过滤旋转门的 PMU 数据压缩。

1.1 过滤压缩思路

在过滤压缩阶段,首先从数据流中读取一个压缩周期的 PMU 数据,然后按照数据点是否在过滤压缩框内,对数据点进行取舍。设压缩周期序数变量为 β 、压缩算法数据读入控制变量为 n_c 、PMU 数据上送速率为 f 、过滤压缩算法最大传输间隔为 T_{\max} 。则每个压缩周期需要读取的 PMU 数据个数 n_T 的计算公式如下:

$$n_T = \begin{cases} \frac{T_{\max}}{\Delta t} + 2, & \beta = 1 \\ \frac{T_{\max}}{\Delta t}, & \beta > 1 \text{ and } n_c = 0 \\ \frac{T_{\max}}{\Delta t} - 1, & \beta > 1 \text{ and } n_c = 1 \end{cases} \quad (1)$$

$$\Delta t = 1/f \quad (2)$$

式中,压缩周期序数变量 $\beta \in \mathbb{N}$,起始时 $\beta = 1$ 即第一个压缩周期, Δt 为 PMU 数据上送间隔, n_c 的取值由旋转门压缩过程决定。

按照 PMU 数据时标顺序,定义上一个被保存的数据点为 ξ_{save} 、过滤压缩阶段被保存的数据点集

合为 ξ_{temp} 、当前新数据点为 ξ_{curr} 、当前新数据点的前一个数据点为 ξ_{prev} ，则当前新数据点是否被压缩的判断依据如下：

$$\Delta T < T_{\text{max}} \quad (3)$$

$$\begin{cases} |\xi_{\text{curr}} - \xi_{\text{save}}| \leq \delta_{\text{ExcDev}}, & \xi_{\text{temp}} = \emptyset \\ |\xi_{\text{curr}} - \xi_{\text{temp}}(s)| \leq \delta_{\text{ExcDev}}, & \xi_{\text{temp}} \neq \emptyset \end{cases} \quad (4)$$

式中， ΔT 为 ξ_{curr} 与 ξ_{save} 的时间差； δ_{ExcDev} 用于控制过滤压缩算法最大误差， $\xi_{\text{temp}}(s)$ 为 ξ_{temp} 集合中最后一个元素。若连续2个数据点均满足式(3)和式(4)，即数据点均在过滤压缩框内，则前一个数据点被舍弃，否则将这两个数据点归入被保存的数据点集合 ξ_{temp} 中（第一个数据点不参与运算），这一机制决定了过滤压缩的最大误差不超过 $2\delta_{\text{ExcDev}}$ 。

以图1为例，点A为起点建立过滤压缩框，从点B开始进行过滤压缩判断。对点B，由于点B、C均在当前过滤压缩框内，故点B被舍弃；对点C，由于点D不在当前过滤压缩框内，故点C、D均被保留；下一步，以点D为起点建立新的过滤压缩框，重复上述判断过程，直至该压缩周期末尾数据点，过滤压缩过程结束，此时点F为 ξ_{curr} ，点E为 ξ_{prev} 。

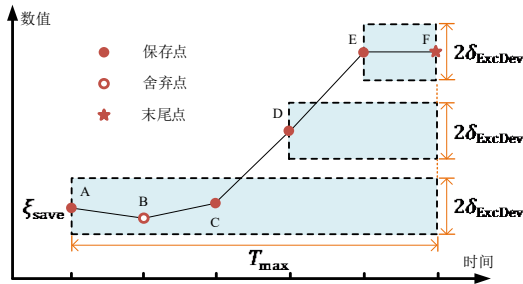


图1 过滤压缩算法示意图

Fig. 1. Diagram of the exception compression method

1.2 改进的旋转门压缩方法

旋转门压缩方法通过判断数据点是否在旋转门压缩框内，对数据点进行取舍。传统旋转门压缩方法侧重于对算法效率的提升，以保证应用的实时性^[13]。但在配电网边缘计算应用中，降低本地的数据存储规模是数据压缩的重要目的。此时，选取的 T_{max} 一般较大， ξ_{temp} 集合中数据点较多，传统旋转门压缩方法的误差可能超过 δ ，造成精度下降：

$$\delta = 2(\delta_{\text{ExcDev}} + \delta_{\text{CompDev}}) \quad (5)$$

为此，本文在动态旋转门压缩算法的基础上^[22]，提出一种旋转门压缩的改进判据，以提高数据压缩精度。具体改进判据公式如下：

$$k_{\text{up}} = \max \left(\frac{\xi_{\text{SDT}}(i) - \xi_{\text{SDT}}^{\text{up}}(j)}{t_{\text{SDT}}(i) - t_{\text{SDT}}(j)} \right) \quad (6)$$

$$k_{\text{down}} = \min \left(\frac{\xi_{\text{SDT}}(i) - \xi_{\text{SDT}}^{\text{down}}(j)}{t_{\text{SDT}}(i) - t_{\text{SDT}}(j)} \right) \quad (7)$$

$$k_{\text{up}} \leq k_{\text{down}} \quad (8)$$

$$i \leq (s + 2) \quad (9)$$

式中， ξ_{SDT} 为 $(s + 3)$ 维向量，由 ξ_{save} 、 ξ_{temp} 、 ξ_{prev} 按照时标顺序组合得到； $\xi_{\text{SDT}}^{\text{up}}(j)$ 与 $\xi_{\text{SDT}}^{\text{down}}(j)$ 是与 $\xi_{\text{SDT}}(j)$ 具有相同时标、距离为 δ_{CompDev} 的上下两个支点；对 ξ_{SDT} 集合中数据点按时标顺序进行判断，并通过以下3种方式进行处理：

1) 若对 $j = 1$ 且 $i \in \{j + 1, \dots, s + 2\}$ 中的所有点，式(6)至式(9)均成立，则改进判据成立， ξ_{temp} 集合中的数据点全部被舍弃，将 ξ_{prev} 保存，更新为 ξ_{save} ，置 $n_c = 0$ ，并传递 ξ_{save} 、 ξ_{curr} 及其时标和 k_{up} 、 k_{down} 到下一个压缩周期；

2) 若对 $i = q$ 且 $q \in \{j + 1, \dots, s + 2\}$ 中的某点，式(8)不成立，式(9)成立，则保留 $\xi_{\text{SDT}}(q - 1)$ ，并令 $j = q - 1$ ，对 $i \in \{j + 1, \dots, s + 2\}$ 重新判断；

3) 若式(9)不成立，则旋转门压缩过程结束，置 $n_c = 1$ ，传递 ξ_{save} 、 ξ_{prev} 、 ξ_{curr} 及其时标和 k_{up} 、 k_{down} 到下一个压缩周期。上述过程可总结为图2所示的计算流程。

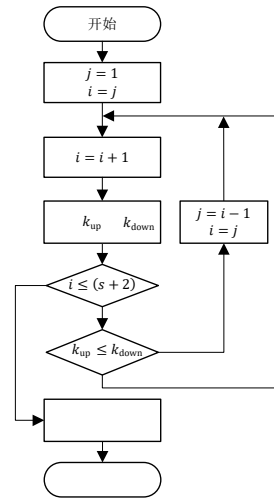


图2 旋转门压缩算法流程图

Fig. 2. Flowchart of the modified swing door trending compression method

进一步地，由图3的示例进行说明。已知点B在过滤压缩中被舍弃，点A为 ξ_{save} ，点C、D属于在过滤压缩中保留的 ξ_{temp} 集合，点E为 ξ_{prev} ，点F为 ξ_{curr} ；点A至点F组成 ξ_{SDT} 集合。初始时，点A作为旋转门起始点；点C完全满足改进判据，点D仅不满足式(8)，故点C被保存并作为新的旋转门起始点；点D、E完全满足改进判据，点F不满足式(9)，故点D被舍弃，点E、F的取舍由下一个压缩周期决定。

的压缩与重构是无损的,在改进的过滤旋转门压缩阶段,数据的压缩与重构是有损的。因此,本文在数据重构时,首先进行指数哥伦布解码,得到过滤旋转门算法压缩后的数据,再通过线性插值的方法得到完整的数据。

3.1 指数哥伦布解码

指数哥伦布编码具有对称的码元结构,即在编码 1 前,先编码 g 个 0,编码 1 后,再编码 g 位的组内偏移量 $offset$ 。因此,解码步骤如下:

- 1) 按位读取二进制数据,直到读取到位 1 结束,读取到的位 0 的个数即为原数据所在的组号 g ;
- 2) 继续读取 g 位二进制数据,将其转换为十进制数据,即为 $offset$;

- 3) 利用式 (15) 反解出 ξ_{EG} 。

由于在进行指数哥伦布编码前,对数据进行了预处理,所以在得到 ξ_{EG} 之后,还需进行预处理的反变换,步骤如下:

- 1) 根据 ξ_{EG} 的奇偶性,计算 $\Delta\xi_{int}$:

$$\Delta\xi_{int} = \begin{cases} \frac{(\xi_{EG}+1)}{2}, & \xi_{EG} \text{ 是奇数} \\ -\frac{\xi_{EG}}{2}, & \xi_{EG} \text{ 是偶数} \end{cases} \quad (16)$$

- 2) 对 $\Delta\xi_{int}$,根据原小数位数 α ,计算差分值 $\Delta\xi$,由于差分序列第一个数据值为真值,根据式 (17) 和式 (18) 即可逐次求出预处理前的数据 ξ :

$$\Delta\xi = \Delta\xi_{int} \times 10^{-\alpha} \quad (17)$$

$$\xi_2 = \Delta\xi + \xi_1 \quad (18)$$

其中, ξ_2 为当前数据点, ξ_1 为上一个数据点, $\Delta\xi$ 为 ξ_2 与 ξ_1 的差值,逐次计算过程从第二个数据点开始。

3.2 线性插值数据重构

在本文改进的过滤旋转门压缩算法中,无论是在过滤压缩阶段还是在旋转门压缩阶段,其压缩思想都是在一定的容许误差限值下,用两个数据点作为端点构成的线段去代替原始数据点连成的折线。因此,采用线性插值的方法对指数哥伦布解码后的数据序列进行重构。

每次参与线性插值的运算数据为有损压缩数据序列中相邻的两个数据及其对应的时标,分别记为 ξ_{head} 、 ξ_{end} 和 t_{head} 、 t_{end} 。当 t_{end} 与 t_{head} 的差为 Δt 时,说明两个数据的时标是相邻的,无需进行线性插值。当 t_{end} 与 t_{head} 的差为 $p\Delta t$ ($p > 1, p \in \mathbb{N}$) 时,说明两个数据间存在 $(p-1)$ 个数据被舍弃,需要通过线性插值的方法进行重构。设待恢复的数据点为 $\xi_{recon}(i)$,对应的时标为 $(t_{head} + i\Delta t)$, $i \in \{1, \dots, p-1\}$ 。 $\xi_{recon}(i)$ 的计算公式为:

$$\xi_{recon}(i) = \frac{\xi_{end} - \xi_{head}}{t_{end} - t_{head}} \times i \times \Delta t + \xi_{head} \quad (19)$$

特别地,对于相位数据而言,本文基于文献[26]提出的相位数据无损转换方法进行了预处理,消除了相位数据因达到 $\pm\pi$ 两侧范围而产生的跃变。因此,重构相位数据时,需要对线性插值得到的相位数据按照文献[26]中提出的相位数据恢复方法进行无损变换,使得重构后的相位数据与原始相位数据保持相同的数据格式。

4. 算例分析

4.1 算例介绍

为了验证本文两阶段 PMU 数据压缩方法的有效性和可行性,采用实际示范工程的配电网 PMU 量测数据对压缩算法性能进行测试。

现场 PMU 的上送速率为 50Hz,即相邻两帧数据的时间间隔 $\Delta t = 0.02s$,电压等级为 10kV。本文选取需要在边缘侧存储的 PMU 量测数据类型包括三相节点电压相量、三相支路电流相量、频率数据。选取连续 1 小时的数据进行压缩算法性能测试,每种量测类型数据包括 180000 个数据点。本文假定每个量测类型的 PMU 数据分别存储在一个文件中,编码方式为 UTF-8,每个文件内仅包含 1 小时该类型的量测数据,该量测类型数据的 ID 标识符及初始数据点对应的时刻标识在文件名中,其余数据点对应的时刻可根据 Δt 计算得到。

指数哥伦布编码得到的压缩数据文件为两个二进制文件,分别存储被压缩算法保留的量测数据及其时标索引数据。算例中的重构数据文件,格式与原始数据文件相同,编码方式为 UTF-8。重构数据精度的计算方式为原始数据文件与重构数据文件的对应数据间的绝对和相对误差。

4.2 压缩控制参数设置

对于 T_{max} 、 δ_{ExcDev} 和 $\delta_{CompDev}$ 3 个控制参数,考虑电网的实际运行特征和 PMU 电压、电流、频率量测数据的特点,兼顾存储空间和数据时效性,设定 $T_{max} = 1s$;每种数据类型的单点最大误差要求 δ 决定了控制参数 δ_{ExcDev} 与 $\delta_{CompDev}$ 和的最大取值。依据经验,对于电压有效值数据、电流有效值数据、频率数据,设定 δ_{ExcDev} 与 $\delta_{CompDev}$ 的比值为 1:2;特别地,由于配电网线路短,线路两端相角差较小^[3],因此将相角数据的单点最大误差设置为 0.04° ,并据此设置 δ_{ExcDev} 为 0 , $\delta_{CompDev}$ 为 0.02° 。表 1 给出了本文采用的压缩控制参数的具体取值,其中,脚标 ϕ 代表相数据,角标 RMS 代表有效值, P

代表相位。相电压基准值 $U_0=5.774\text{kV}$ ，支路电流基准值 $I_0=1000\text{A}$ 。

表 1 压缩参数设置

Table 1 Setting of the compression parameters

量测数据类型	$U_{\phi}^{(\text{RMS})}$	$U_{\phi}^{(\text{P})}$	$I_{\phi}^{(\text{RMS})}$	$I_{\phi}^{(\text{P})}$	f
单点最大误差	$0.05\% \times U_0$	0.04°	$0.5\% \times I_0$	0.04°	0.003Hz
δ_{ExcDev} 限值	0.48V	0°	0.83A	0°	0.0005Hz
δ_{CompDev} 限值	0.96V	0.02°	1.67A	0.02°	0.001Hz

考虑到配电网各个应用场景对数据精度的需求不同，压缩控制参数的选择需要与应用场景相匹配。在面向配电网故障定位等应用场景时，可根据精度需求调整压缩控制参数，获得满足精度要求的 PMU 重构数据。反之，在数据精度要求较低场景下，通过改变压缩控制参数，可获得更大压缩比，实现更优的压缩性能。

4.3 数据压缩算法性能指标

本文采用压缩比(compression ratio, CR)和重构数据精度来反映数据压缩算法性能。其中，压缩比为原始数据文件大小与压缩文件总大小的比值：

$$\gamma_{\text{CR}} = \frac{F_{\text{Orig}}}{F_{\text{Comp}}} \quad (20)$$

式中， F_{Orig} 为原始数据文件的大小， F_{Comp} 为压缩文件的总大小，即压缩数据文件和时标索引数据文件的大小之和。

重构数据精度采用最大绝对误差(maximum absolute error, MAE)、平均绝对误差百分比(mean absolute error percentage, MAEP)和平均相对误差百分比(mean relative error percentage, MREP)3 个指标。其中，最大绝对误差用来衡量压缩算法对单个数据点精度造成的影响；平均绝对误差百分比用来衡量压缩算法对相位和频率数据整体精度造成的影响；平均相对误差百分比用来衡量压缩算法对电压和电流有效值数据整体精度造成的影响。具体的计算公式如下：

1) 最大绝对误差 MAE:

$$\gamma_{\text{MAE}} = \max_{1 \leq n \leq N} |\xi_n - \hat{\xi}_n| \quad (21)$$

2) 平均绝对误差百分比 MAEP:

$$\gamma_{\text{MAEP}} = \frac{1}{N} \sum_{n=1}^N |\xi_n - \hat{\xi}_n| \quad (22)$$

3) 平均相对误差百分比 MREP:

$$\gamma_{\text{MREP}} = \frac{1}{N} \sum_{n=1}^N \left| \frac{\xi_n - \hat{\xi}_n}{\xi_n} \right| \quad (23)$$

式中， ξ_n 为原始数据， $\hat{\xi}_n$ 为重构数据， N 为原始数据个数。

4.4 两阶段 PMU 数据压缩算法性能测试结果分析

两阶段 PMU 数据压缩算法性能测试结果如表 2 所示。其中， γ_{CR0} 为对原始数据经过改进过滤旋转门压缩后，原始数据文件与压缩文件总大小的比值，即改进过滤旋转门阶段的压缩比； γ_{CR} 为进一步完成指数哥伦布编码后，原始数据文件与压缩文件总大小的比值，即整个压缩算法总压缩比；电压或电流相量数据的脚标 A、B、C 分别代表各相。

从表 2 中可以得出，所有数据的最大绝对误差均不超过设定的最大单点误差，验证了改进的过滤旋转门压缩算法误差是受控的；所有数据的总压缩比均超过 8，三相节点电压相量或三相支路电流相量的压缩效果接近，表明了该算法能较好地适应各类型 PMU 数据的压缩。

对于相电压有效值数据，总压缩比超过 30，单点最大误差不超过 2.1V，平均相对误差控制在 0.01% 以内；对于频率数据，总压缩比超过 30，单点最大误差 0.002Hz，平均绝对误差为 0.0006Hz。可以得出，对于波动较小的原始数据，如电压相量、频率数据，有损压缩阶段压缩比较高，数据在经过差分预处理后，数据绝对值较小，符合指数哥伦布无损压缩编码特性，因此无损压缩阶段在保证相同数据重构精度的同时，进一步提高了数据压缩比。

对于相电流有效值数据，总压缩比超过 8，单点最大误差不超过 3.1A，平均相对误差控制在 0.1% 以内，这是由于电流有效值数据与负荷密切相关，不确定性较强，波动较大，在有损压缩阶段为了保证数据重构精度，压缩比较低，导致无损压缩后压缩比提升程度相对于电压相量、频率数据较小，总压缩比相对较低。

对于电压相位数据和电流相位数据，总压缩比分别为 16 和 12 左右，单点最大误差不超过 0.04° ，平均绝对误差不超过 0.008° 。电压相位数据相比电流相位数据平稳，在相同误差精度条件下，有损压缩阶段舍弃了更多的数据点，实现了更高的压缩比。总的来说，本文方法在压缩相位数据时，在保持较高精度的同时，仍能实现高效压缩。

为了直观地反映重构数据精度，以相电压有效值数据为例，绘制出 1 小时原始 A 相电压有效值数据、重构数据及其误差曲线，如图 4 所示。其中，原始数据在 6050V 左右上下波动，重构数据能够很好地跟随原始数据；误差曲线也表明原始数据与重构数据间误差是受控的。

表 2 PMU 数据压缩结果
Tabel 2 Results of the PMU data compression

数据类型	T_{\max}	δ_{ExcDev}	δ_{CompDev}	γ_{CR0}	γ_{CR}	γ_{MAE}	γ_{MREP}
$U_{\text{A}}^{(\text{RMS})}$	1 s	0.4 V	0.8 V	5.161	30.732	2.010 V	0.0076%
$U_{\text{B}}^{(\text{RMS})}$	1 s	0.4 V	0.8 V	5.787	34.200	1.960 V	0.0076%
$U_{\text{C}}^{(\text{RMS})}$	1 s	0.4 V	0.8 V	5.633	33.356	2.009 V	0.0077%
$I_{\text{A}}^{(\text{RMS})}$	1 s	0.8 A	1.6 A	2.001	8.884	3.077 A	0.0801%
$I_{\text{B}}^{(\text{RMS})}$	1 s	0.8 A	1.6 A	1.975	8.775	3.030 A	0.0797%
$I_{\text{C}}^{(\text{RMS})}$	1 s	0.8 A	1.6 A	1.979	8.758	3.024 A	0.0792%
数据类型	T_{\max}	δ_{ExcDev}	δ_{CompDev}	γ_{CR0}	γ_{CR}	γ_{MAE}	γ_{MAEP}
$U_{\text{A}}^{(\text{P})}$	1 s	0°	0.02°	1.512	16.357	0.0372°	0.0078°
$U_{\text{B}}^{(\text{P})}$	1 s	0°	0.02°	1.517	16.373	0.0372°	0.0079°
$U_{\text{C}}^{(\text{P})}$	1 s	0°	0.02°	1.523	16.422	0.0372°	0.0079°
$I_{\text{A}}^{(\text{P})}$	1 s	0°	0.02°	0.969	12.152	0.0367°	0.0073°
$I_{\text{B}}^{(\text{P})}$	1 s	0°	0.02°	0.983	12.256	0.0361°	0.0074°
$I_{\text{C}}^{(\text{P})}$	1 s	0°	0.02°	0.985	12.310	0.0367°	0.0074°
f	1 s	0.0005 Hz	0.001 Hz	3.317	39.599	0.002 Hz	0.0006 Hz

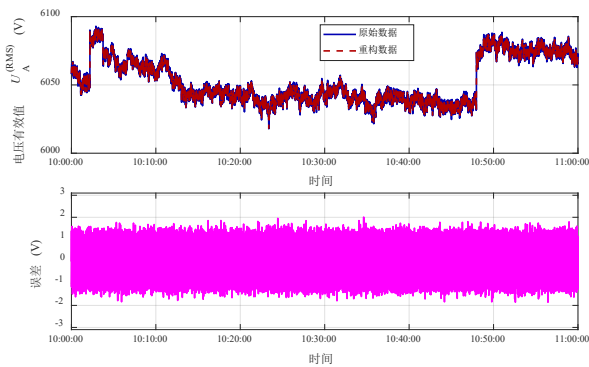


图 4 A 相电压有效值数据重构效果

Fig. 4. Reconstructed profile of A-phase RMS voltage

4.5 不同数据压缩算法性能测试结果对比

为了验证本文两阶段 PMU 数据压缩算法的性能优势,选取了 ZIP 和 RAR 两种通用的无损数据压缩工具,基于相同的 PMU 数据进行了测试,结果如图 5 所示。

从图 5 可以看出,对于电压有效值、电压相位、电流相位、频率等数据类型,本文提出的两阶段 PMU 数据压缩算法的压缩比均明显高于其他压缩工具,显示出较好的压缩性能;对于电流有效值数据,各方法压缩效果接近,原因是本文算法在压缩时需要额外保留索引数据文件,对于波动较大的电流有效值数据,索引数据文件相对较大,影响了压缩比。

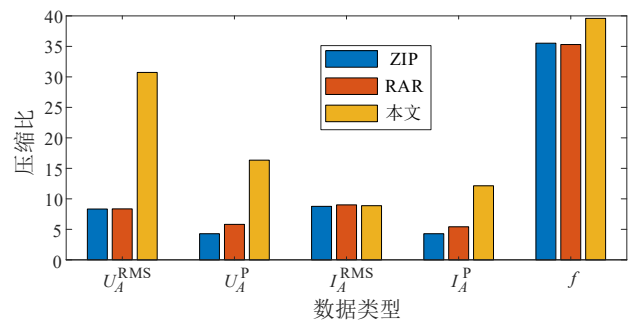


图 5 不同压缩方法结果对比

Fig. 5. Comparison of different compression methods

5. 结论

PMU 将成为配电网边缘计算中的一项重要业务功能,但由于其高上送速率的特点,产生了海量数据,带来了较大的通信和存储负担,与边缘计算装置的计算、存储资源有限性形成了矛盾。对此,本文提出了基于过滤旋转门和指数哥伦布编码的配电网 PMU 数据压缩方法,改进了旋转门压缩判据,提高了重构数据精度;将指数哥伦布编码与过滤旋转门压缩算法结合,在保持相同数据精度的同时,进一步提高了压缩比。算例测试表明,由于本文方法不涉及复杂的频域变换或迭代计算,计算量小、易于实施,可在配电网边缘计算装置等计算资源有限的环境下灵活部署,满足多样化场景下的数据压缩应用需求。

在未来工作方面,由于过滤旋转门压缩为一阶线性压缩,在配电网出现故障等导致 PMU 数据波动较大时,压缩比降低,这一问题需要在未来研究中解决。此外,本文在压缩算法控制参数的选择中,基于单点最大误差要求,将控制参数人为设定为常数。事实上,对于系统中不同位置 PMU 数据的压缩,应具备参数自适应调整能力,这将是未来工作的重点。

参考文献

- [1] SHI Weisong, CAO Jie, ZHANG Quan, et al. Edge computing: vision and challenges[J]. IEEE Internet of Things Journal, 2016, 3(5): 637-646.
- [2] 岑伯维, 蔡泽祥, 胡凯强, 等. 电力物联网边缘计算终端业务时序逻辑与计算负荷建模方法[J]. 电力系统自动化, 2021, 45(9): 107-112.
CEN Bowei, CAI Zexiang, HU Kaiqiang, et al. Modeling method for bussiness sequential logic and workload of edge computing terminal in electric internet of things[J]. Automation of Electric Power Systems, 2021, 45(9): 107-112.
- [3] 刘灏, 毕天姝, 徐全, 等. 配电网高精度同步相量测量技术方案与展望[J]. 电力系统自动化, 2020, 44(18): 23-29.
LIU Hao, BI Tianshu, XU Quan, et al. Scheme and prospect of high-precision synchrophasor measurement technology for distribution network[J]. Automation of Electric Power Systems, 2020, 44(18): 23-29.
- [4] 张道农, 于跃海, 时伯年, 等. 电力系统同步相量测量技术及应用[M]. 北京: 中国电力出版社, 2017.
ZHANG Daonong, YU Yuehai, SHI Bonian, et al. Synchronous phasor measurement technology and application in power system[M]. Beijing: China Electric Power Press, 2017.
- [5] 李依泽, 陆超, 王印峰, 等. 基于 Kalman 滤波与神经网络的高精度同步时钟算法[J]. 电网技术, 2019, 43(3): 777-783.
LI Yize, LU Chao, Wang Yinfeng, et al. A high accuracy synchronous clock algorithm based on kalman filter and neural network[J]. Power System Technology, 2019, 43(3): 777-783.
- [6] 王金丽, 韦春元, 段祥骏, 等. 含微型同步相量测量的智能配电网信息集成方法[J]. 电网技术, 2021, 45(1): 284-291.
WANG Jinli, WEI Chunyuan, DUAN Xiangjun, et al. Information integration method of intelligent distribution network with micro synchronous phasor measurement[J]. Power System Technology, 2021, 45(1): 284-291.
- [7] 刘喆林, 王成山, 李鹏, 等. 多源量测数据融合的配电网状态估计及应用[J]. 中国电机工程学报, 2021, 41(8): 2605-2615.
LIU Zhelin, WANG Chengshan, LI Peng, et al. State estimation of distribution networks based on multi-source measurement data and its applications[J]. Proceedings of the CSEE, 2021, 41(8): 2605-2615.
- [8] 于力, 焦在滨, 王晓鹏, 等. 基于 PMU 的中压配电网精确故障定位方法及关键技术[J]. 电力系统自动化, 2020, 44(18): 30-38.
YU Li, JIAO Zaibin, WANG Xiaopeng, et al. Accurate fault location scheme and key technology of medium-voltage distribution network with synchrophasor measurement units[J]. Automation of Electric Power Systems, 2020, 44(18): 30-38.
- [9] 何西, 涂春鸣, 于力. 配电网状态估计量测处理方法及抗虚假数据注入攻击策略[J]. 高电压技术, 2021, 47(7): 2342-2349.
HE Xi, TU Chunming, YU Li. Measurements processing method of distribution network state estimation and anti false data injection attack strategy[J]. High Voltage Engineering, 2021, 47(7): 2342-2349.
- [10] GADDE P H, BISWAL M, BRAHMA S, et al. Efficient compression of PMU data in WAMS[J]. IEEE Transactions on Smart Grid, 2016, 7(5): 2406-2413.
- [11] SAYOOD K. Introduction to data compression[M]. USA: ELSEVIER, 2012.
- [12] WEN Lulu, ZHOU Kaile, YANG Shanlin, et al. Compression of smart meter big data: A survey[J]. Renewable and Sustainable Energy Reviews, 2018, 91: 59-69.
- [13] ZHANG Fang, CHENG Lin, LI Xiong, et al. Application of a real-time data compression and adapted protocol technique for WAMS[J]. IEEE Transactions on Power Systems, 2015, 30(2): 653-662.
- [14] Souza J D, Assis T, Pal B C. Data compression in smart distribution systems via singular value decomposition[J]. IEEE Transactions on Smart Grid, 2017, 8(1): 275-284.
- [15] 吴毓峰, 李富盛, 余涛, 等. 基于残差双重注意机制网络的电力数据压缩与高精度重建[J/OL]. 电网技术, 2021: 1-15[2022-05-01]. DOI:10.13335/j.1000-3673.pst.2021.1419.
WU Yufeng, LI Fusheng, YU Tao, et al. Power data compression and high-precision reconstruction based on residual dual attention mechanism network[J/OL]. Power System Technology, 2021: 1-15[2022-05-01]. DOI:10.13335/j.1000-3673.pst.2021.1419.
- [16] 赵洪山, 马利波. 基于张量 Tucker 分解的智能配电网大数据压缩[J]. 中国电机工程学报, 2019, 39(16): 4744-4753.
ZHAO Hongshan, MA Libo. Big data compression of smart distribution systems based on tensor tucker decomposition[J]. Proceedings of the CSEE, 2019, 39(16): 4744-4753.
- [17] 王鹤, 李石强, 于华楠, 等. 基于分布式压缩感知和边缘计算的配电网电能质量数据压缩存储方法[J]. 电工技术学报, 2020, 35(21): 4553-4564.
WANG He, LI Shiqiang, YU Huanan, et al. Compression acquisition method for power quality data of distribution network based on distributed compressed sensing and edge computing[J]. Transactions Of China Electrotechnical Society, 2020, 35(21): 4553-4564.
- [18] KIM S, CHO N I. Hierarchical Prediction and Context Adaptive Coding for Lossless Color Image Compression[J]. IEEE Transactions on Image Processing, 2014, 23(1): 445-449.
- [19] 邢龙平, 李东晖, 胡川川. 基于模糊逻辑的无损视频压缩的算法[J]. 计算机应用, 2012, 32(10): 2859-2862.
XING Longping, LI Longhui, HU Chuanchuan. Lossless video compression algorithm based on fuzzy logic[J]. Journal of Computer Applications, 2012, 32(10): 2859-2862.
- [20] LOUIE H, MIGUEL A. Lossless compression of wind plant data[J]. IEEE Transactions on Sustainable Energy, 2012, 3(3): 598-606.
- [21] YAN Lei, HAN Jiayu, XU Runyan, et al. Model-free lossless data compression for real-time low-latency transmission in smart grids[J]. IEEE Transactions on Smart Grid, 2021, 12(3): 2601-2610.
- [22] CUI Yang, HE Yingjie, XIONG Xiong, et al. Algorithm for identifying wind power ramp events via novel improved dynamic swinging door[J]. Renewable Energy, 2021, 171: 542-556.
- [23] UNTERWEGER A, ENGEL D. Resumable load data compression in smart grids[J]. IEEE Transactions on Smart Grid, 2015, 6(2): 919-929.
- [24] Gallager R, Voorhis D V. Optimal source codes for geometrically distributed integer alphabets[Corresp.][J]. IEEE Transactions on Information Theory, 1975, 21(2): 228-230.
- [25] MOON Y H. A new coeff-token decoding method with efficient memory access in H.264/AVC video coding standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(16): 729-736.

- [26] ZHAO Jinli, YE Yuzhuan, YU Hao, et al. Improved SVD-based data compression method for synchronous phasor measurement in distribution networks[J]. International Journal of Electrical Power and Energy Systems, 2021, 129: 1-11.

收稿日期: 2022.04.02。

作者简介:



习伟

习伟(1980), 男, 博士研究生, 教授级高工, 研究方向为电网数字化及芯片化, E-mail: xiwci@csg.cn;

李鹏(1973), 男, 博士, 教授级高工, 研究方向为电网数字化、智能化, E-mail: lipeng@csg.cn;

李鹏(1981), 男, 博士, 教授, 研究方向为分布式发电与智能配电网的运行、仿真与分析, E-mail: lip@tju.edu.cn。