A Novel Event-driven Hybrid Compression Technique for Grid-Edge Waveform Data

Yuru Wu1, Biao Sun1, Wei Qiu1, He Yin1, Yuqing Dong1, Yu Liu1, Yilu Liu12

1 The University of Tennessee, Knoxville
2 Oak Ridge National Laboratory

Objective
Continuous points-on-wave (POW) measurements with timestamps, also called syncwaveforms, provide valuable insights into transient and dynamic events under the high renewable penetration. However, the substantial volume of POW data engenders formidable challenges in the realms of data transmission and storage. To tackle these challenges, this paper presents an event-driven hybrid compression technique, designed to optimize the compression ratio while simultaneously ensuring the preservation of essential information.

Methodology

Data pre-processing:
Analog-to-digital conversion $x[n] = A \cos \left(2\pi n \frac{f}{f_s}\right) + N(n)$
Recoverable $L$-order Delta operation
Forward: $x'[n] = \Delta(x'(n)) = x'[n+1] - x'[n]$  
Reverse: $x'[n] = x'[0] + \sum_{k=1}^{n} x'[k]$
Recoverable periodic $L$-order Delta operation
Forward: $x'[n] = \Delta_p(x'(n)) = x'[n+m] - x'[n]$, $m = \frac{f}{f_s}$  
Reverse: $x'[n] = x'[n-m] + x'_{p}[n-m] = x'[n] + \sum_{k=1}^{n} x'[j + km]$
Event indicator: $e_i = \frac{1}{m} \sum_{n=0}^{m-1} x'[k]$  

Metrics for evaluation:
- Compression ratio (CR): $CR = \frac{\text{Bits of original data}}{\text{Bits of results}}$
- Percentage Relative Difference (PRD): $PRD = \frac{\sum_i^m (v_i - f_i)}{\sum_i^m f_i}
- Frequency mean square error (FMSE): $FMSE = \sqrt{\frac{\sum_i^m (f_i - f_0)^2}{m}}$
- Spectrum mean square error (SMSE): $SMSE = \sum_i \left( \log(P_{xx,i}) - \log(P_{xx,j}) \right)^2$

Implementation

Compression techniques:
Lossy compression: Digital wavelet transform (DWT) + periodic Delta + simple-8b encoding
Lossless compression: Direct Delta + simple-8b encoding+ LZMA compression

Conclusion:
To maintain an appropriate balance between data fidelity and compression performance, an event indicator is employed to detect the occurrence of dynamic events. This strategy effectively reduces data volume while maintaining a reliably accurate depiction of waveform characteristics, frequency, and spectrum information.