# Hierarchical Residual Encoding for Multiresolution Time Series Compression

**Bruno Barbarioli, Gabriel Mersy, Stavros Sintos, Sanjay Krishnan**

University of Chicago, Department of Computer Science, Chicago, IL, USA

Full paper

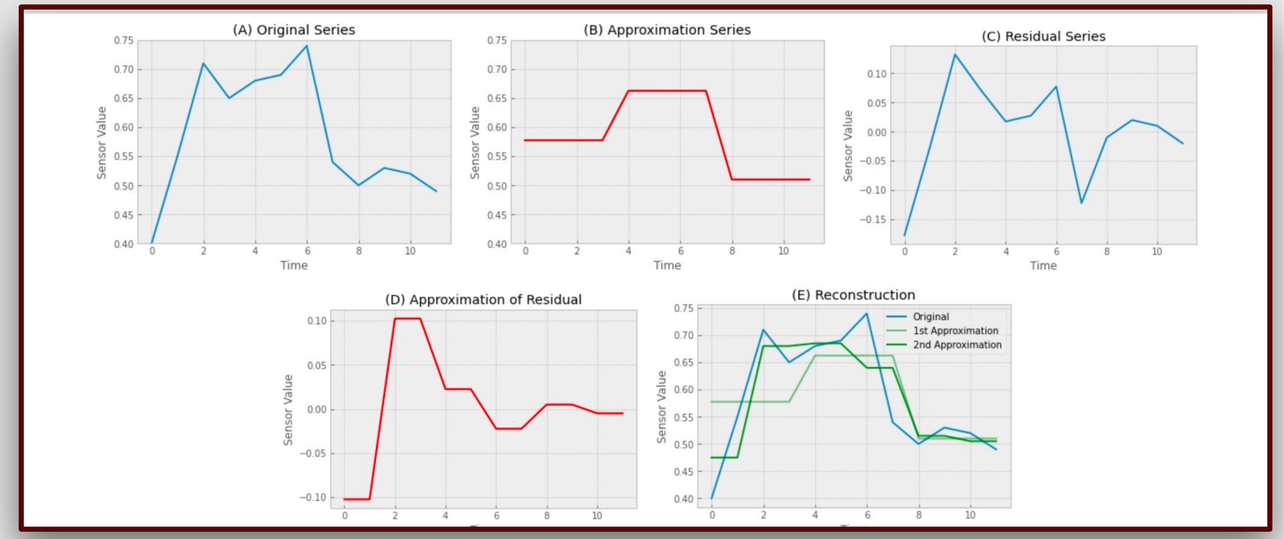
## Introduction

Different applications can have varying requirements for error bounds on time series data. In this setting, existing lossy compression approaches would have to do one of three things, all of which suffer from limitations:

1. **(Strict)** Encode at the strictest requirement
2. **(All)** Encode multiple times to meet all of the requirements
3. **(Lazy)** Re-encode the data after first encoding at the strictest requirement

**Our Contribution (HIRE):**

A time series compression algorithm that can efficiently construct multiple sub-encodings each with an individual $L^\infty$ error bound.




## Methodology

**Methodology Summary:**

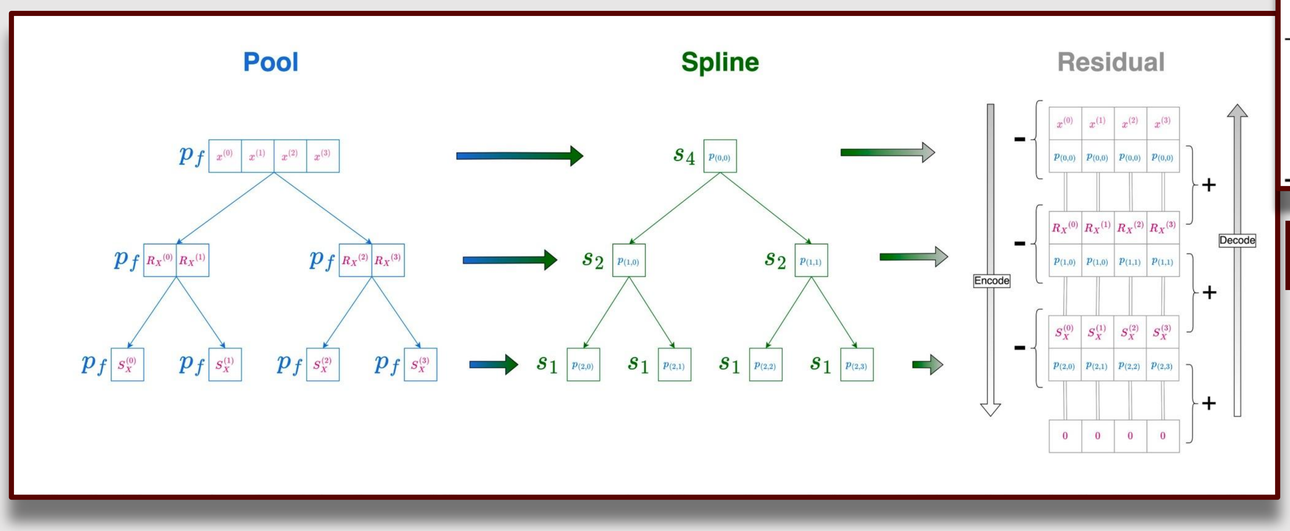
1. **(Hierarchical)** Apply encoding operations at incrementally finer granularities
2. **(Residual)** Connect the encodings in an additive decomposition by recursively applying approximation to the **residual** signal, i.e., the difference between the approximation and original
3. **(Encoding)** Two operations: **pool** as a low-dimensional summary of a time series window and **spline** as a lossy approximation of the window. Uniform quantization is applied to the encoded pooled values according to the strictest threshold $\varepsilon^*$

**The Multiresolution Problem**

$Cx = C_1 \oplus C_2 \oplus \ldots \oplus C_l$ *(combined sub-encodings)*

$\varepsilon_1 \quad \varepsilon_2 \quad \ldots \quad \varepsilon_l$ *(multiple errors)*

Optimizations improve codec time complexity from *O(Tk)* to *O(T)* where *k* is a log factor.




## Results Overview

We conducted an experiment on an edge device to demonstrate the concept. We can see from the results below in *Table 4* that our method outperforms the selected baselines overall.

| Algorithm | Scheme | IL | TS | RO | LS |
|---|---|---|---|---|---|
| QTRC | Strictest | 4.8s | 3.8MB | 0.3s | 3.8MB |
| QTRC | All | 44.5s | 1.3MB | 0.2s | 12.9MB |
| QTRC | Lazy | 4.8s | 1.3MB | 5.1s | 3.8MB |
| GZip | Lossless | 0.9s | 28MB | 0.17s | 28MB |
| HIRE | Multiresolution | 1.2s | 2.0MB | 0.18s | 5.8MB |

*Table 5* below breaks down the compression ratio (lower is better) for each individual $L^\infty$ error threshold. Our method only needs to store the strictest one.

| Thresholds | 0.15 | 0.10 | 0.075 | 0.050 | 0.025 | 0.010 | 0.0075 | 0.005 | 0.0025 | 0.001 |
|---|---|---|---|---|---|---|---|---|---|---|
| HIRE | 0.008 | 0.009 | 0.011 | 0.014 | 0.023 | 0.038 | 0.045 | 0.057 | 0.082 | 0.116 |
| IdGZ | 0.555 | 0.555 | 0.555 | 0.555 | 0.555 | 0.555 | 0.555 | 0.555 | 0.555 | 0.555 |
| QTRC | 0.005 | 0.006 | 0.007 | 0.009 | 0.014 | 0.025 | 0.029 | 0.036 | 0.051 | 0.076 |

## Conclusion

- We presented HIRE, a novel system for **multiresolution compression** that uses hierarchical residual encoding to store time series data at multiple error thresholds.
- Our experiments validate that HIRE performs better than the baselines at multiresolution compression for edge computing applications.
- Theoretical details and the full experimental evaluation can be found in our paper.