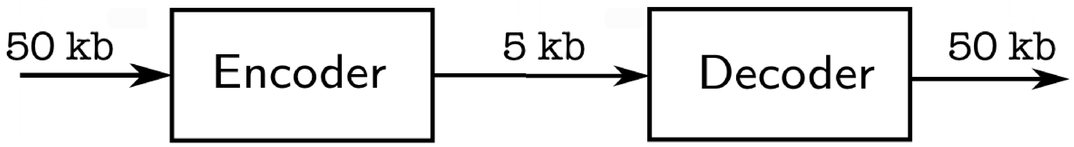
Universal Graph Compression

What is compression?



- Encode information using as few bits as possible
- Done by exploiting redundancy in data
- Requires knowing the data distribution
- Example: While encoding English text, assign "e" the fewest number of bits

Universal Compression

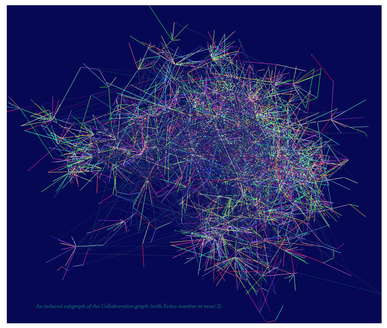
- Compress without knowing data distribution
- Well studied for 1-D data, for example time series

A Universal Algorithm for Sequential Data Compression

JACOB ZIV, FELLOW, IEEE, AND ABRAHAM LEMPEL, MEMBER, IEEE

Compression for graphs

- Compression not well-studied for graphs
- Correlations and community structure possible
- Examples: Web graph, Social networks, biological networks, Collaboration graphs[a]



[a]From `http://math.ucsd.edu/ fan/graphs/collaboration/`

My work (arXiv:2006.02643)

Method for compressing graph data with community structure that

- Achieves optimal theoretical performance (entropy)
- Achieves good empirical performance