

# The Impact of Spatial Correlation on Routing with Compression in Sensor Networks

Sundeep Patten<sup>1</sup>, Bhaskar Krishnamachari<sup>1</sup>, Ramesh Govindan<sup>2</sup>

1. ANRG, EE-Systems, USC: <http://ceng.usc.edu/~bkrishna/anr>, 2. ENL, CS, USC: <http://enl.usc.edu>

## Data Gathering and Aggregation: Schemes for Routing with Compression

### Data-Centric Routing

#### Data Aggregation

This is widely accepted as *an essential paradigm for energy-efficient operation* of sensor networks. For data-gathering applications in which data originates at *multiple correlated sources* and is routed to a single sink, aggregation would primarily involve in-network compression of the data.

### Designing routing schemes

#### What is lacking in earlier work?

- Several data aggregation schemes have been proposed
- we seek to understand the *performance of various data aggregation schemes across the range of spatial correlations*
- Intuitively, *optimal routing* would depend on the *level of correlation* in the data originated by sources, requiring an *adaptive element* in the routing scheme.

## Optimal Routing with Compression:

Are there efficient schemes that are independent of the level of correlation?

- Our result shows that while the nature of optimal routing with compression does depend on the correlation level, surprisingly, there exists a practical static clustering scheme which can provide near-optimal performance for a wide range of spatial correlations.
- This result has important consequences—it obviates the need for sophisticated routing and compression schemes, or adaptive schemes that require the underlying correlations in the data to be “learned”.

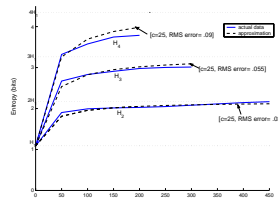
## Methodology: Study generic routing+compression schemes to obtain deeper insights for design

### Model and Metrics

#### Metrics

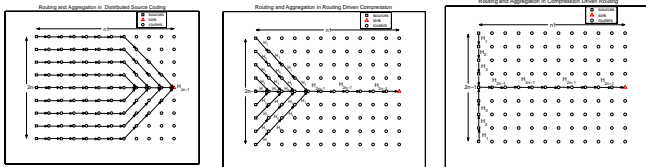
- Energy expenditure : bits x hops metric
- Uncorrelated data gathered : joint entropy of sources
  - Empirically obtained approximation
  - For n sources in a line, separated by d, having correlation level c,

$$H_n(d) = H_1 + (n-1) \left[ 1 - \frac{1}{\left(\frac{d}{c} + 1\right)} \right] H_1$$



#### Generic Routing Schemes

- Distributed Source Coding (DSC) - Idealized, Optimal
- Routing Driven Compression (RDC) - Shortest paths, compress en-route
- Compression Driven Routing (CDR) - Compress close to sources

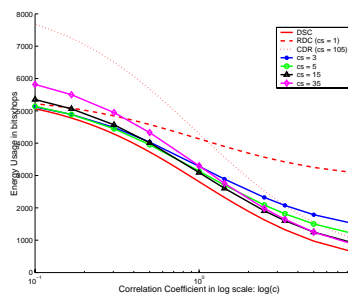


- RDC and CDR are extremes of a clustering scheme
  - In RDC, each source is a cluster in itself, optimal for c=0
  - In CDR, all sources are part of a single cluster, optimal for c=∞

### Results

#### Hybrid Scheme: Intermediate Cluster size

- Optimal cluster size depends on correlation level c
- *A particular cluster size is close to optimal over the range of c*



- We prove the existence of such a cluster size analytically using an appropriate definition for the near-optimal cluster size

#### Extension to a 2-D deployment of sources

- A simple scheme based on static, localized clustering
- Similar results as in 1-D case

