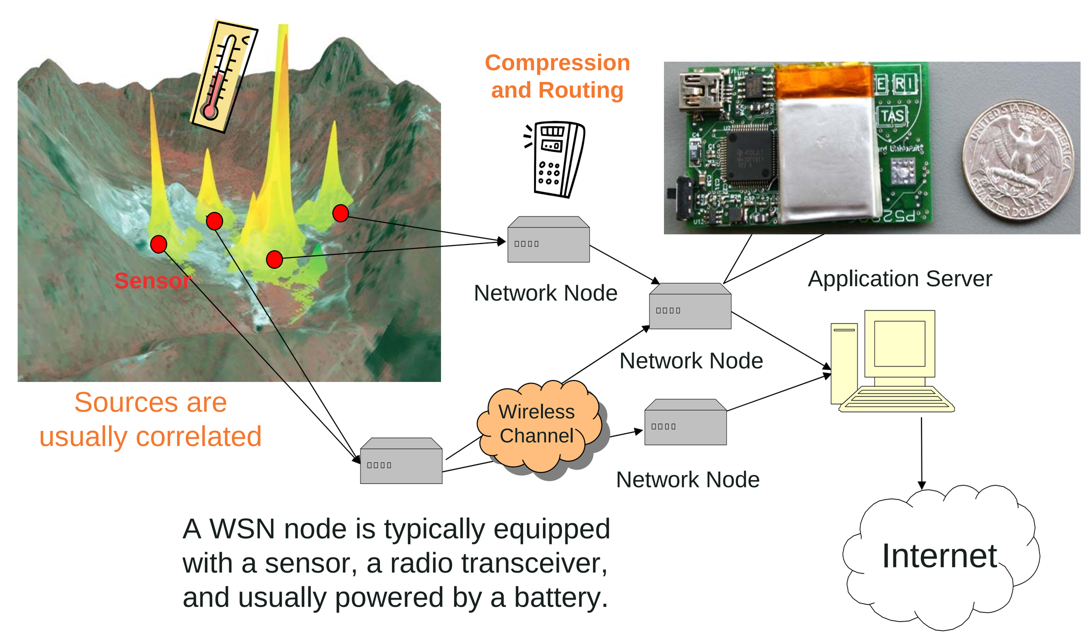
Opportunistic Source Coding for Data Gathering in Wireless Sensor Networks

Tao Cui, Lijun Chen, Tracey Ho, Steven H. Low, and Lachlan L. H. Andrew

Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125

Introduction

A wireless sensor network (WSN) is a networked sensor system that can cooperatively monitor physical or environmental conditions, such as temperature, motion or pollutants, at different locations. It has broad applications such as environment and habit monitoring, healthcare applications, and military surveillance and target tracking.

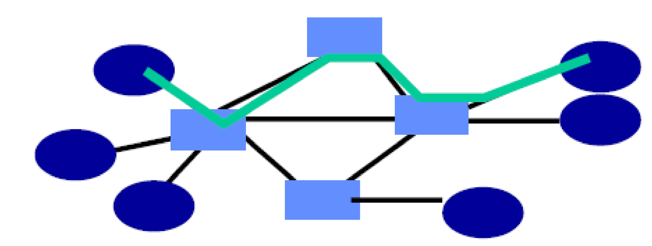


We study data gathering, where information sampled at distributed sensor nodes needs to be transported to central base stations for further processing and analysis.

Challenges: Limited battery power and bandwidth; Unreliable wireless channel.

Two Main Ingredients

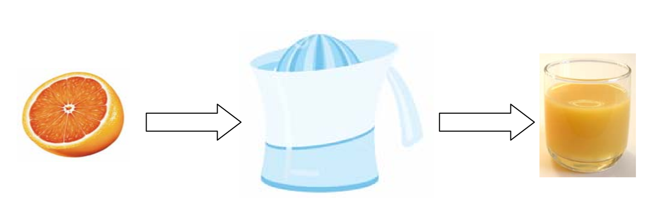
Routing: Selects paths along which to send information and data.



Existing work implicitly assumes routing techniques similar to those in wireline networks using a fixed minimum cost path.

As wireless transmissions are error-prone, sequential forwarding of packets along a fixed path may incur many retransmissions

Source Coding: Encodes observed information using fewer bits, reducing wireless transmission and energy consumption.



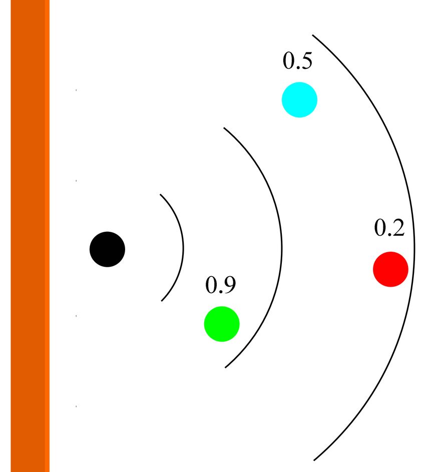
The Lempel-Ziv (LZ) algorithm is a popular algorithm for lossless source coding; variants of LZ are used in gzip and GIF images.

- LZ is vulnerable to packet loss (requires more transmissions)
- The destination has to run the LZ decoding algorithm once for each coding step in reverse order

Basic Idea of OSCOR

Opportunistic Source Coding and Opportunistic Routing (OSCOR)

★ OSCOR does opportunistic routing by exploiting Wireless Broadcast Advantage

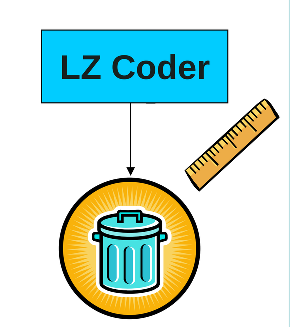


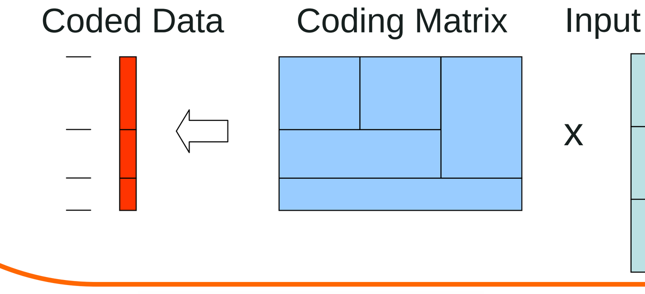
- As every packet can be received by possibly any node within the sender's communication range, the chance that no node receives the packet is small
- The packet is opportunistically forwarded after it is received by possibly multiple neighboring nodes

The probability that no node receives the packet sent by the black node is only 0.1x0.5x0.8=0.04

★ OSCOR uses Random Network Coding for data compression

- LZ encoder is used to determine only the *size* of the coded data; its output is discarded
- Coded data at each node is a random linear combination of the received data





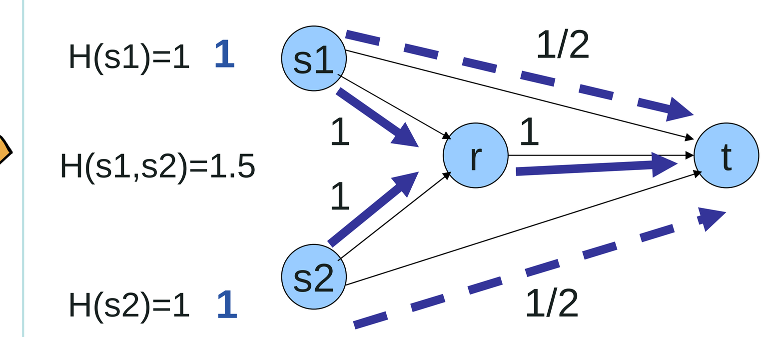
Advantages

★
- With opportunistic routing and forwarding, the number of retransmissions is reduced, and reliability and throughput are improved. Path is not pre-specified but determined online.
- With random network coding, the data received by the destination is a linear transform of the sources' data. Backtracking decoding is avoided.

Example

Two sources: s1,s2. Both s1 and s2 send 1 packet each to the destination t. If 1 packet from s1 and 1 packet from s2 are compressed at r, only 1.5 packets need to be sent.

Link delivery probabilities are shown along the edges of the following graph.



Average number of transmissions

$$\frac{1}{4}(2+3+3+3.5) = 2.875$$

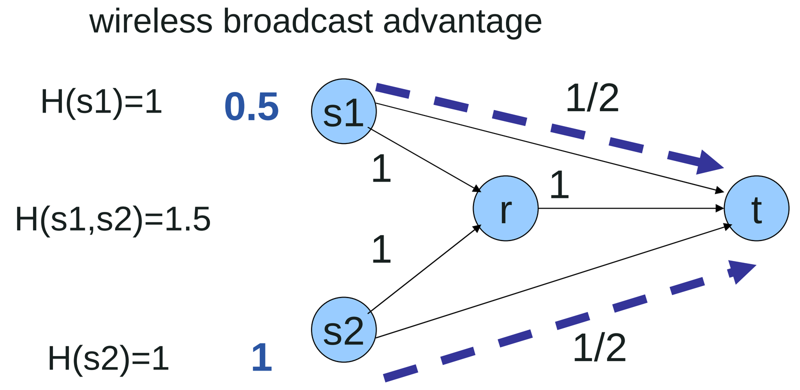
Probability	0.25	0.25	0.25	0.25
# of Transmissions	2	3	3	3.5
Event	(t, t)	(r, t)	(t, r)	(r, r)

The probability that only r receives the packet from s1 and t receives the packet from s2 is 0.25. In this case, 3 transmissions are required

Previous Approaches

Distributed Source Coding (DSC)

– Assumes global statistics are known (not practical); does not exploit wireless broadcast advantage

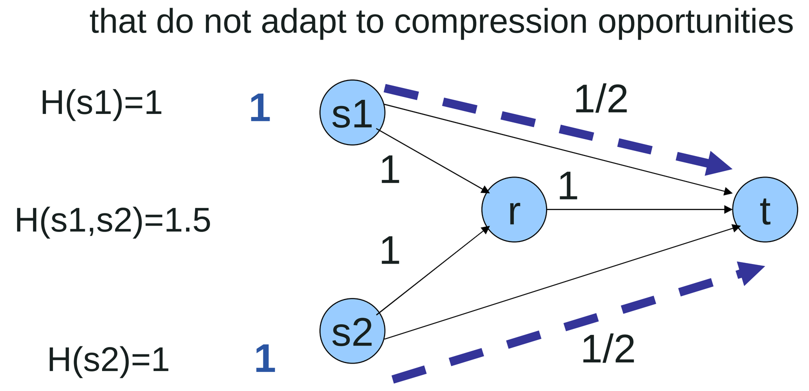


Average

$$\frac{1}{1/2}+\frac{0.5}{1/2}=3$$

Routing Driven Compression (RDC)

– Does not exploit wireless broadcast advantage; uses shortest paths that do not adapt to compression opportunities



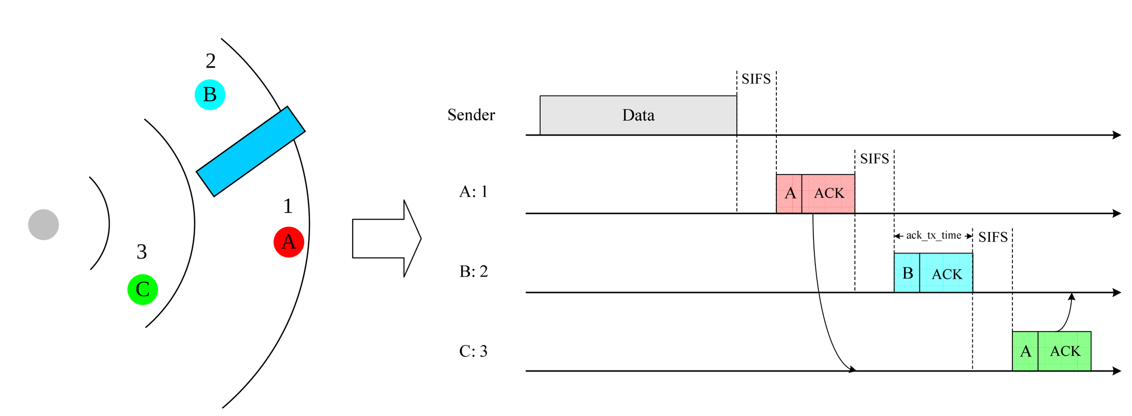
Average

$$\frac{1}{1/2}+\frac{1}{1/2}=4$$

OSCOR outperforms not only RDC but also DSC

Making it Work

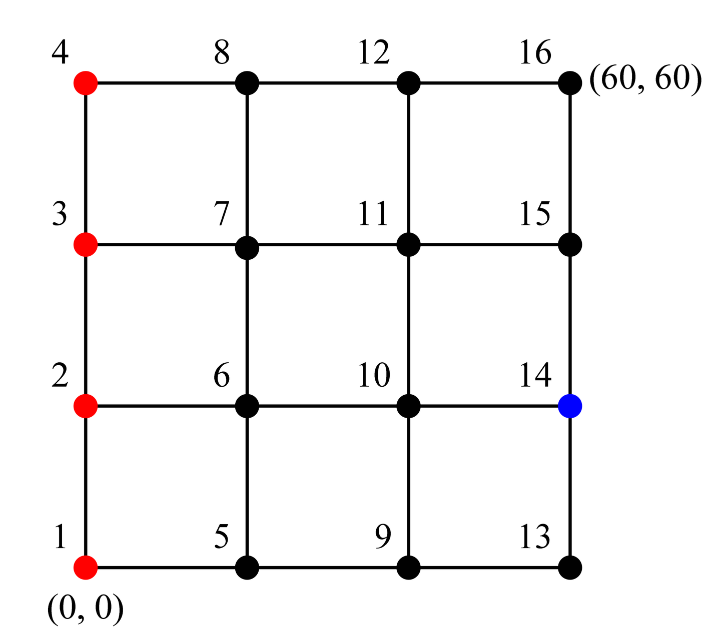
- A consensus protocol is needed to coordinate wireless transmission and packet forwarding
  - Slight modification to IEEE 802.11
    - Neighbors are prioritized according to the cost of the shortest path to the destination
    - All receivers ack packet, delayed based on priority
    - ACK indicates the highest priority ACK heard
  - The receiving node with the highest priority forwards the packet



- Routing depends not only on link quality but also on compression opportunities
  - Link quality and data compression ratio of each node are learnt online
  - The effect of data compression is not additive
  - **Solution:** Modified Dijkstra's Algorithm

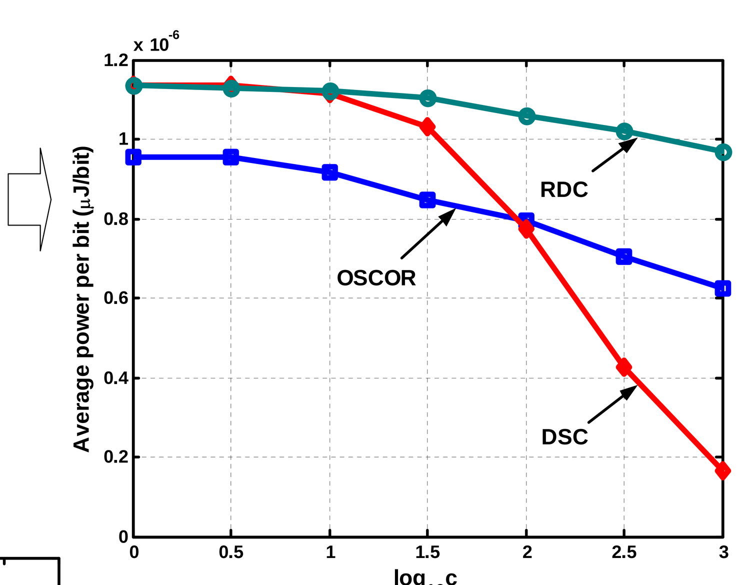
Experimental Results

A 4x4 grid network (coordinates in meters)

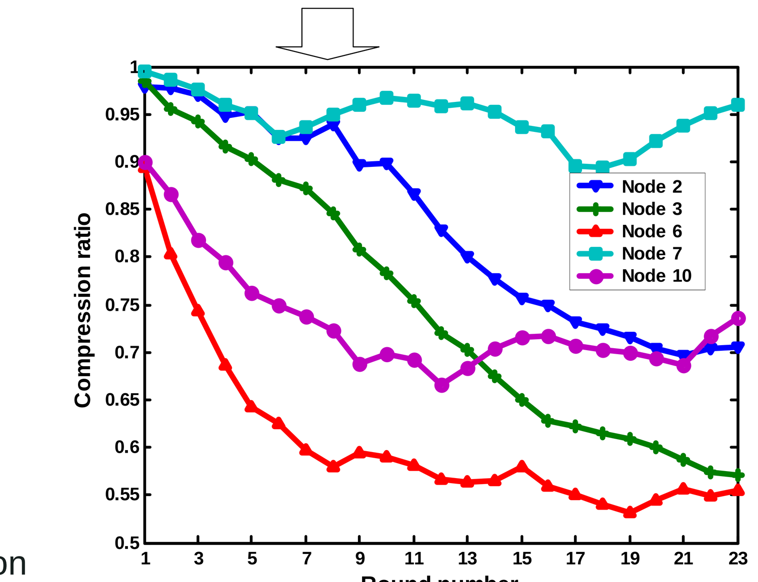


- Sources: 1, 2, 3, 4; Destination: 14
- Sources' correlation follows the jointly Gaussian data model.
  - $\sigma_{i,j} = \exp(-d_{i,j}/c)$, where $d_{i,j}$ is the distance between nodes i and j and c is a correlation parameter
- Simulations are based on IEEE 802.11b standard

- The average power consumption per bit versus the correlation parameter c.
- The sources have perfect knowledge of global statistics in DSC (this is not assumed for OSCOR and RDC).



The evolution of compression ratio as a function of rounds (or time) with OSCOR.



- OSCOR adaptively learns a path online with the highest possible compression and the best possible link quality.
- OSCOR saves power for moderate correlation over both DSC and RDC without requiring the global statistics at the sources.

Supported by

DARPA, Caltech's Lee Center for Advanced Networking, and Microsoft Research