Wireless Sensor Networks for Environmental Monitoring: Energy-Efficient Inference

Paul Flikkema

Wireless Networks Research Laboratory
Department of Electrical Engineering
Northern Arizona University

October 2007
Outline of this talk

- Introduction/Motivation
  - Requirements & Applications
- WiSARDNet hardware/software
  - Transducers
- Energy efficiency
  - Hardware and software
  - MAC
  - Inference

Challenges:
- Distributed system
- Unreliable wireless channel
- Severe energy limitations
- Packaging and installation
Requirements – Environmental Sensing

- Minimal invasiveness
- Long battery life
  - Aggressive energy management
  - Target: > 12 months
- Scientific accuracy
- Support of a broad spectrum of probes
- Support transparent incremental deployment
- Scalable in network size and density
- Ease of installation and maintenance
- Support of internet connectivity via
  - Terrestrial (cellular)
  - Satellite
- Rugged, weatherproof packaging
- Low life cycle cost
CA Coastal Redwoods

Fern mat

Wired sensing infrastructure requires over 1 km of cable per tree
Wireless Ad Hoc Networking of Physically-Embedded Sensors

**Need:** Dense, minimally-invasive array of sensors to monitor microclimate variables such as temperature and light. Standalone or wired sensor arrays are invasive and difficult to deploy and operate.

**Opportunity:** Wireless networking of the sensors

- dramatically improve coverage and spatial density, and ultimately, our understanding of microclimates...
- ...while greatly reducing the total monitoring cost
2nd Generation WiSARD

Modular hardware design
- Dual-processor architecture
- Three-board stack

Gateway nodes will use memory/time/802.11 board in place of sensor data acquisition board
G2 WiSARDNet Design

Communication and Networking
- 902 – 928 MHz ISM band
- Non-Coherent Binary FSK (NC-BFSK) modulation
- Slow time/frequency hopping spread spectrum via pseudo-random number generator
- CRMA radio channel sharing algorithm
  - Distributed control
  - Local information
  - Scalable

Self Organization
- Periodic search for new nodes
- On-demand search for lost nodes (under development)
- Can add, move, or delete nodes

Power Management
- Monitor power status
- Report battery voltage
- Adjustable radio transmit power (under development)

Scheduler
- Time-triggered s/w architecture
- Dynamic scheduling of communication
- Online-configurable sample rate (coming soon)

User Interface
- Command line from PC
- User selection of ID, sample rates
- On-line diagnostics

P. Flikkema
G2 WiSARD Capabilities

**Built-in probe interfaces**
- 12-bit A/D conversion
- 4 temperature channels
  - thermocouple
- 4 light (PAR) channels
  - photodiode
- 2 general purpose probe channels, two power outputs and two CCP modules (Capture/Compare/PWM)
  - Soil moisture
    - Decagon Ech2oprobe
  - Serial communication with intelligent probes
  - Sap flow (future)

**Interface for multiple additional intelligent probes**
- One-wire bus

**Provision for external energy supplies**
- Supports autonomous switching between internal and external energy sources
- Battery-backed solar
G2 WiSARD Functional H/W Design

Transceiver
Radio Board

Micro-Controller PIC18F8720
Memory SRAM
FLASH/FRAM

Power Mgmt

System Time

Brains Board

SPI+

Micro-Controller PIC18F452

Sensor Board

Analog I/O

1-Wire
Light(4)
Temp(4)
General Purpose(2)

PWM(2)
Pwr(2)

Int Power
Ext Power

Enclosure
Temperature Probe

- **Thermocouples**
  - Based on voltage potential between two bimetallic junctions
  - Rugged; low thermal mass
  - Accuracy: absolute error < 0.5 C (tests with water bath)
  - Low cost

Diagram:

- ADC
- Copper
- Cold Junction
- Constantin
- 1-wire Digital Temperature Sensor for cold junction compensation
- Hot Junction
- Digital Output
Light Intensity Probe

- Photodiode response tailored to photosynthetic active radiation (PAR) spectrum
  - Commercial Sensor: LI-COR LI-190 Quantum Sensor
  - Low-cost home-brewed probes using Hamamatsu GaAsP photodiodes capped with acrylic diffusers
NAU PAR vs. reference standard
Soil Moisture Probe

- Time domain reflectometer (TDR)
  - measures the dielectric constant of a medium via traversal time of electromagnetic pulse in transmission line buried in the medium
  - Expensive; relatively complex
- Low-cost probe (Decagon ECH2O)
  - ca. $80/probe
  - measures the soil dielectric constant (proportional to volumetric water content) from rate of change of voltage on a sensor embedded in the medium
  - can calibrate to soil type in lab using gravimetric water content
Weather Station Peripheral

- Vaisala WXT-510 smart probe for above-canopy data
  - Wind Speed and Direction
    - ultrasonic vector anemometer
  - Liquid/Hail Precipitation
    - piezoelectric impact detector
  - Relative Humidity/VPD
  - Temperature
  - Barometric Pressure
Trial Deployment, April 2005: Grasslands site, C. Hart Merriam Elevational Gradient

Data from All Sites

Data from One Site

Microburst and brief cloud cover

P. Flikkema
A few lessons learned so far...

- Packaging is a challenge
  - Weatherproofing vs. probe interfaces
- Probe costs are significant
  - Probeset costs will exceed sensor node costs
- Deployment is time-consuming
  - 3D space
  - Truly “embedded” probes
- Requirements for deployment will vary by site
- Correlation radii vs. transmission range
- Maintenance and QA
  - Probe models in non-stationary environments

These problems are harder and more expensive than any of the research challenges
Energy-Efficient Inference: Strategy

1. Minimize useless radio operation
   - transmitting when there is no relevant node to receive
   - listening when no relevant node is transmitting

2. Transmit only what is necessary to solve the problem of model/data inference
   - exploit spatio-temporal redundancy of the data
   - use coding to protect data
CRMA: Overview

- MAC algorithms for this application
  - Requirements and characteristics
  - Related work
- Energy vs. bandwidth efficiency
- Why (pseudo) random access?
- CRMA: features, the algorithm, and PCR extension
- Overview of analysis
Introduction

*Wireless ad-hoc networks are unique:*

- Nodes have limited energy supplies
- Global information is expensive
  - Algorithms should be distributed
- Unreliable links & shared, distributed channels
- Topologies may be dynamic (e.g., mobile nodes)
Go after the biggest problem first

- Energy is very limited
  - Solar too expensive and unreliable
- Where is the most energy consumed?
  Transmission and reception of one bit consumes approx. $6 \times 10^4$ times the energy required for execution of one microcontroller instruction
- Goal: minimize useless radio operation
MAC Protocols for Wireless Ad-Hoc Nets: Requirements

- Energy and bandwidth efficiency (trade-off)
- Latency and throughput
- Scalability: with
  - network size
  - node resources (e.g., multiple radios)
- Robustness to network dynamics (topology changes, node failures)
- Admit energy-aware routing
- Exploit diversity (time, frequency, angle, space)
- Robustness to interference
MAC Layer Characteristics

- Handling contention/collisions:
  - Avoidance: deterministic algorithms, assign resources
  - Resolution: random algorithms, recovery strategies

- **Key tradeoff:** *proactive* vs. *reactive* coordination
  - Proactive coordination implies less contention, better energy efficiency (e.g. reservation TDMA)
    - But efficiency drops when topology is dynamic
    - May not scale well
  - Reactive mechanisms offer simplicity and poor energy efficiency (e.g., ALOHA)
Some Related Work

- Link-based – nodes negotiate for contention-free slots
  - Sohrabi and Pottie (99)
- Some local contention: nodes advertise when they will be listening
  - Rozovsky and Kumar [SEEDEX] (01)
  - Heidemann and Estrin (02)
  - Mergen and Tong [RASC] (02)
    - Inspired by SEEDEX
    - Link-based assignment of resources
Energy vs. Bandwidth Efficiency

- Each application has unique set of characteristics that drive the MAC design
- In general, in wireless sensor networks with low information rates, we aggressively trade bandwidth for energy efficiency
- Large amount of excess bandwidth allows randomized access
- Low topology dynamics allows for energy efficiency via proactive coordination
Why Use a (Pseudo) Randomized Access Algorithm?

- Allows local (one-hop) proactive coordination where contention/interference are critical
- Interference from the rest of the net is made noise-like; implies good scalability
- Admits exploitation of node’s communication resources
- Provides robustness to external interference and channel fading
CRMA Features

- **Clique**s – a node’s cliques are subsets of its sets of one-hop neighbors
  - Includes unicast and multicast subsets
  - Generalizes notion of links

- Cleanly exploits **multiple and multi-channel radios**
  - *Locally deterministic/globally random hopping* among a set of orthogonal time/frequency/code channels

- **Predictive conflict resolution**
  - Nodes can predict and resolve collisions between cliques they belong to
CRMA Protocol

- Time slotted into frames; frames divided into slots
- **Proactive coordination:**
  - Node cycles through cliques
  - Node contacts members of the clique; they cooperate to share a common package of information:
    - A pseudo-random number generator (PRNG)
    - A seed
    - A start-of-first-frame time
- **Operation:** clique members update their common PRNG’s each contact frame/slot, agreeing on future frame slot allocations
Mitigating Collisions

- In CRMA, two types of conflicts:
  - **Soft conflicts** occur when proximate cliques allocate the same time-frequency-code slot
    - Causes multiple-access interference that is mitigated by excess bandwidth and randomized access
  - A **hard conflict** can occur when the number of cliques intersecting at a node exceeds the number of radios
    - Nodes can use **predictive conflict resolution (PCR)** to look to the future, identify hard conflicts, and resolve them
CRMA - Summary

- Balances proactive and reactive coordination mechanisms
- Couples local deterministic coordination of communication with global averaging of interference
- Scalable
- Highly energy efficient
- Robust to interference and unreliable wireless medium
Trial Deployment, April 2005: Grasslands site, C. Hart Merriam Elevational Gradient

Data from All Sites

Data from One Site

Microburst and brief cloud cover

P. Flikkema
Light Variation in One Tree

Light Intensities at Different Levels in Site #133

- Light Intensity (lux/50cm²) at Tree Top
- Light Intensity (lux/50cm²) at Mid Tree
- Light Intensity (lux/50cm²) at Ground

P. Flikkema
Sensor Readings at One Site

Environmental Parameters Measured at Site #156

- Light Intensity (W/m²)
- Soil Moisture (%) at Uphill
- Soil Moisture (%) at Downhill
- Air Temperature (°C)
- Soil Temperature (°C)
Dynamic Sensor Networks: Context

- WSN technologies maturing
- Network designs are becoming application-specific
- Driving applications differ dramatically
  - Target characteristics
  - Physical node size
  - Availability of internal/ambient energy
  - Mobility
- Range of network densities and node capabilities
  - size and expense of physical transducers
  - deployment cost
Application: Dynamic Inference of Ecosystem Data and Processes

- Motivating application: Revolutionize understanding of environmental change
  - forecast how altered climate and CO2 will impact biodiversity and carbon storage in the biosphere
- Challenge: endow network with sufficient explanatory power under significant energy consumption constraints
- Tight coupling between the sensed and the sensors
  - sensors are even more deeply embedded in their environments
Properties

- Data is highly heterogeneous: natural scales range from
  - Meters to landscapes
  - Seconds to years
  - and is strongly non-stationary
- Dynamic model/data-driven control of
  - Sampling
  - Communication
  - Estimation and prediction
  - Model inference
  according to their relative values and costs:

Integrate spatio-temporal sensing with modeling and prediction in an adaptive framework
Architecture

Accommodate limited in-network resources
Complement them with out-of-network capability
- Relaxed energy constraints
- Massive processing power available
- Latency

- Both should support multiple concurrent models

Approach
- Adaptive in-network joint estimation and coding (inner, fast control loop)
- Supervisory out-of-network processing (outer, slow control loop) – model inference
Dynamic Sensor Networks: Components of Control

- **SLIP**: scalable landscape inference and prediction
- **DONC**: dynamic out-of-network control
- **NIP**: network inference and prediction
- **DINC**: dynamic in-network control
- **Dynamic sampling and reporting**

Flow of information:
- SLIP to DONC via predictions
- DONC to NIP via models and decisions
- NIP to DINC via data, real-time estimates, uncertainty
- DINC to SLIP via scheduling
- SLIP to data
- DONC to data

Data flow:
- SLIP to DONC
- DONC to NIP
- NIP to DINC
- DINC to SLIP

P. Flikkema
Goal: maximum predictive power at ecological model level for a given energy cost

Where to be

Model-Driven Dynamic Optimization: In- and Out-of-Network Control of Sampling and Reporting

Easy

Model-driven measures
Sensor Networks: Design Space

How do realities of WSN implementations affect energy-efficient inference?

- Star or multi-hop: logical or physical
- Is phase coherence possible?
- Most WSN radios are simple
  - Binary symmetric channel
- Is time synchronization maintained?
- Source and channel coding
Example: multi-sensor time series of temperature. Can view as

- **streams of numbers**
  - use general-purpose source codes (e.g., delta modulation)

- **correlated spatio-temporal process**
  - use a parameterized statistical model to drive adaptive space-time sampling and reporting

- **high-level model input**
  - sampling driven by the needs of ecological models
  - e.g., leaf efficiency/tree growth model: sampling rate may be high because of sensitivity of high-level model, even when dynamics appear slow
Energy-Efficient Inference: Strategy

1. Minimize useless radio operation assuming something useful will be sent
   - transmitting when there is no relevant node to receive
   - listening when no relevant node is transmitting

2. Transmit only what is necessary to solve the problem of model/data inference
   - exploit spatio-temporal redundancy of the data
   - use coding to protect data
Energy-Efficient Inference: Strategy (2)

- Strategy should be PHY-layer aware
  - radio warm-up time & synchronization preamble are significant overhead, so should look at censoring to save packets, not bits
- Don’t send differences
- Don’t send every measurement
  - Will have multiple correlated measurements from which to infer the measurement/model
  - Take advantage of prior information when available
- Can consider energy/latency tradeoff
- Keep in mind other layers
  - Does LLC layer include ARQ?
A First Step

- Dynamic Reporting

- +/- $\varepsilon$ threshold used to determine whether measurement should be reported

  $$|x - x^*| > \varepsilon$$

- Implies simple
  - prior distribution
  - error criterion based on that distribution

- Is there a better, yet simple, way?

  $$E(x - x^*)^2$$
Example: Precision vs. # of reporting sensors: linear estimate of a sampled random field

Projection onto random spatial basis

Minimal prior knowledge (a priori mean and priori variance)

Flikkema ISCC 2006
Compressive Sensing

- Uber-goal: use basis that gives sparse (efficient) representation (encoding) of information
  - Project data onto basis
- Compressive sensing theory shows that randomized basis can be efficient for data-gathering sensor nets
- Excellent potential when no priors are available
energy-constrained inference in sensor nets

slope = pwr density

efficiency limit

spatial granularity limit

fidelity

power consumption density (W/m²)
Managing Uncertainty and Errors

- Network as instrument
- Measurement errors
  - Transducer noise
  - Channel errors $\Rightarrow$ failures
    - bit errors and lost packets
    - A bit error can be worse than a lost packet
    - E.g., getting the direction bit correct!

How can we manage them?
Across the network layers

- **Old way:**
  Separate source and channel coding
  - Fails for short block lengths found in sensor networks

- **Joint source-channel coding**
  - Optimize compression and channel error correction
  - How to do this in a sensor network…
    - Slepian-Wolf +?
FEC coding provides redundancy that protects information bits from communication errors.

Every communication link requires the use of coding to achieve capacity.

A (good) code’s *rate* is a rough indicator of its power.

Coding can also improve energy efficiency.

Optimum per-link coding requires long delays.

Simple codes:
- Repetition
- Single parity-check
Bayesian decoder/estimator

Integrates
- Prior probabilistic spatial-temporal data model
- Network characteristics
  - Can incorporate knowledge of
    - Symbol/packet failures
    - Selective reporting (suppression/censoring)
- Measurements (data)
- Properties of FEC code

Source coding: exploit redundancy in data as encoded in model
- At destination, symbols are exploited as joint information/check symbols

At symbol/bit level, received coded measurements must satisfy
**Parity constraints** from FEC code
**Equality constraints** from model
At the destination

- Received coded measurements form *source-channel product codeword*

\[
SC \text{ code rate} = f(\text{FEC code, transmission policy})
\]
Global source model + FEC improves inference

Channel bit errors

Source Model
MV Gaussian
3 measurements
\( s^2_{info} = 1 \)
\( \sigma^2 = 0.4 \)

FEC channel code
(7,4,3) Hamming

BER: bit error rate

PCMP Iterative decoder

Bayesian (MAP) limit

WSN Regime

\( \rho: \text{BSC probability of error} \)

\[ s^2_{info} = 1 \]
\[ \sigma^2 = 0.4 \]
Take Home Messages

- **Energy efficiency**
  - Use radio only when communicating
  - Communicate only what is necessary

- **Management of uncertainty with an energy-constrained distributed instrument**
  - Coding
  - Use of information about data and the network

- **Joint management of data & models**

- **Inference in data collection, not just data analysis**