Large Scale Wireless Network Systems: Experience, Lessons, and Theories

Xiang-Yang Li

www.cs.iit.edu/~xli  xli@cs.iit.edu
Acknowledgments

- The results presented here are supported by NSF, HongKong RGC, NSF China
- With students Cheng Wang, ShaoJie Tang, Xiaohua Xu, XuFei Mao, Wei Dong, Kebin Liu, Shi Li, etc
- Many collaborators: Prof. Yunhao Liu, and Prof. Zhao
Collaborators
PhD Students (graduated 10)

UNCC  GSU  financial  Google  W. Oregon

Tsinghua  financial  Motorola  Temple  Toledo
PhD and MS Students (current)
Wireless Sensor/Actuator Networks

Bridging the digital world and physical world

Information

Physical Environment

Sensed Data
Challenges

Size of now

Future

Log 16387~ 40 years

Cubic Inch

Cubic Millimeter
More Challenges

- Naming
- Localization
- Energy Supply
- Dynamic programming
- Security
- Fault detection, modeling, diagnosis
Presentation Outline

- Experience and Lessons from Large Scale WSN System Design and Deployment
  - OceanSense
  - GreenOrbs
  - CitySee
  - Waste-Water Processing

- Asymptotical Capacity of Large Scale Wireless Networks
  - Network model, and asymptotical capacity
  - Literature review
  - Our results summary
  - Our approaches
Experience and Lessons

LARGE SCALE WIRELESS SENSOR NETWORK SYSTEMS
Strategic Plan

- Environment
- Transportation
- Smart Grid
- Security
- Green Building
- Agriculture
- Logistic and Supply Chain
- Industry Monitoring
- Health Care

Example: “Sensing China” 2009--
## Some WSN Systems/ Ours

<table>
<thead>
<tr>
<th>System (Affiliation)</th>
<th>Deployment manner</th>
<th>System Scale</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great Duck (2002)</td>
<td>Outdoor, battery</td>
<td>~150</td>
<td>3 months</td>
</tr>
<tr>
<td>VigilNet (Uni. of Virginia)</td>
<td>Outdoor</td>
<td>200</td>
<td>3~6 months</td>
</tr>
<tr>
<td></td>
<td>Battery power</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motelab (Harvard Uni.)</td>
<td>Indoor</td>
<td>190</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Tethered power</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SensorScope (EPFL)</td>
<td>Outdoor</td>
<td>97</td>
<td>6 months</td>
</tr>
<tr>
<td></td>
<td>Battery power</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trio (UC Berkeley)</td>
<td>Outdoor</td>
<td>557</td>
<td>4 months</td>
</tr>
<tr>
<td></td>
<td>Solar-powered</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jindo Bridge</td>
<td>Outdoor</td>
<td>113 Nodes, 680 sensors</td>
<td>2-4 months</td>
</tr>
<tr>
<td></td>
<td>Battery power</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clemson Intelligent River</td>
<td>Outdoor</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>GreenOrbs (HKUST, IIT,...)</td>
<td>Outdoor/battery</td>
<td>1000</td>
<td>1 year</td>
</tr>
<tr>
<td>CitySee (Tsinghua, WuXi...)</td>
<td>Outdoor/battery</td>
<td>1500</td>
<td>&gt;1 year</td>
</tr>
</tbody>
</table>
1. OceanSense (QingDao)
2. GreenOrbs (HangZhou)
3. CitySee (WuXi)
OceanSense

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

ILLINOIS INSTITUTE OF TECHNOLOGY

Tsinghua University
OceanSense Project (2007-2008, video)

- The first sea environment monitoring sensor network system in China
  - More than 120 sensor nodes
  - Temperature, Light, Sea depth
  - More than one year duration
  - Deployed in the Yellow sea near Qingdao, China
- With Prof. Y. Liu, Z. Guo, etc
Experiences and Lessons

- Systems that work in labs fail horribly in practice
  - System run out of battery in a week (labs run in months)
    - Routing protocol and system design
    - Faults detection and diagnosis
  
- Nodes destroyed by water
  - Fixed deployment? Large flexibility? Tide?
  - Balance between accuracy, coverage, and sustainability

- People factors!
  - Nodes stolen by people
    - they are not interested in the node, but the sticks
GreenOrbs 绿野千传

http://www.greenorbs.org/
GreenOrbs (2009-)

About 1000 sensors deployed at multiple phases, places

Joint work with Prof. Liu from HKUST, and Prof. Dai from HDU, Prof. Zhou from ZFU, Prof. Zhao from Xi’An JTU, Prof. Gu from Tsinghua, Prof. Ma from BUPT, and several others
GreenOrbs - building blocks (1)

- **Hardware**
  - TelosB mote with MSP430 processor and CC2420 transceiver
  - Sensors

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Function</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensirion Sht11</td>
<td>Temperature &amp; Humidity</td>
<td>SensirionSht11C</td>
</tr>
<tr>
<td>Hamamatsu S1087</td>
<td>Illuminance</td>
<td>HamamatsuS1087ParC</td>
</tr>
<tr>
<td>Internal Voltage Sensor</td>
<td>MCU-Internal Voltage</td>
<td>VoltageC</td>
</tr>
<tr>
<td>GE Telaire 6004</td>
<td>Content of CO2</td>
<td>Self-developed</td>
</tr>
</tbody>
</table>
System based on TinyOS 2.x.
- Low Power Listening LPL
- Data collection: CTP
- Parameter dissemination: DRIP
# GreenOrbs Deployments

<table>
<thead>
<tr>
<th>Place</th>
<th>Area</th>
<th>Duration</th>
<th>Battery</th>
<th>Scale</th>
<th>Network Diameter</th>
<th>Duty Cycle</th>
<th>Data Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>University woodland #1</td>
<td>20,000 m²</td>
<td>1 month (2008)</td>
<td>800 mAh 1.5V</td>
<td>50</td>
<td>6 hops</td>
<td>No</td>
<td>15 Mbytes</td>
</tr>
<tr>
<td>University woodland #2</td>
<td>20,000 m²</td>
<td>10 months (2009)</td>
<td>2200 mAh 1.2V</td>
<td>120</td>
<td>10 hops</td>
<td>5%</td>
<td>272 Mbytes</td>
</tr>
<tr>
<td>University woodland #2 and #3</td>
<td>40,000 m²</td>
<td>1 year (2009.12~)</td>
<td>~8000mAh, 1.5V</td>
<td>330</td>
<td>12 hops</td>
<td>8% or No</td>
<td>140 Mbytes</td>
</tr>
<tr>
<td>Tianmu Mountain</td>
<td>200,000 m²</td>
<td>1.5 months (2009)</td>
<td>~8000mAh, 1.5V</td>
<td>50</td>
<td>10 hops</td>
<td>5%</td>
<td>3 Mbytes</td>
</tr>
<tr>
<td>Tianmu Mountain</td>
<td>200,000 m²</td>
<td>1.5 year (2009.10~)</td>
<td>~8000mAh, 1.5V</td>
<td>200</td>
<td>~ 20 hops</td>
<td>5%</td>
<td>10 Mbytes</td>
</tr>
</tbody>
</table>

![Campus Image](image-url)
Tianmu Mountain Deployment

GreenOrbs Deployment Area
TianMu Mountain Deployment
Real Deployment
Working station
Deployment in Forest
WSN Nodes
GreenOrbs Deployments

Screen capture from our video
App 1: Canopy Closure Estimates

- Using WSN for forestry measurements

App 2: Fire Risk Prediction

- Fire prediction vs. fire detection
- Microscopic vs. macroscopic prediction
- Devices designed for rangers
  - ranger’s trace, collect data
Study on the classical forestry theory of **climax community**

- Equilibrium broken after declaring Tianmu mountain as forest preserve (bamboo prevails)
Experience and Lessons

- Ensuring Performance of Large Scale Multi-Hop WSN is extremely challenging
  - The data packet loss ratio is higher initially for ~20 hops WSN
  - The network capacity is limited to support all nodes with large sampling frequency
  - Environment Factors
    - Deployment difficulty --- trees, and bamboos.
    - Flooding destroy some nodes
    - 4-season weather-proof?
CitySee
City-Wide Urban Sensing
In the process of deploying 10,000 sensors in city environment. A project that needs wide range of domain experts.
System Architecture

Data Center and Apps

MESH Networks

WSN
1500 Sensor Nodes Deployed

A snapshot of part of the deployment
Wireless Sensor, Mesh Nodes Designed
Sensor nodes, mesh routers deployed
CitySee: City-Wide Urban Sensing

- First phase deployment at Wuxi City (2011.5 to 2012.9)
  - 1100 nodes with temperature, humidity, light
  - 100 CO2 nodes
  - 4 Mesh nodes
  - 1.2 KM²

- Missions: ~2013.6
  - 4000+ sensor nodes with temperature/humidity and light sensors
  - 500+ nodes with CO₂ sensor
  - Some new nodes with other GHG measurements
  - Cover 20KM² urban area in Wuxi City, China

- Eventually: 10,000 nodes, 100 KM²
CitySee – Monitoring Areas

Applications
Environmental monitoring, Carbon sink/emission measurement, pollution detection
CitySee Monitoring/Visualization System

- Network monitoring ([video](#))
  - Network scale
  - Link scale
  - Node scale
  - Fault management
  - Localization, syn,

- Data visualization ([video](#))
  - Spatial
  - Temporal
  - Anomaly, outlier,
Experiences and Lessons

- It is extremely difficult and complicated to deploy in city environment
  - Need coordination with and approve from almost all government departments
  - Functionality is not enough
    - Need a nice-looking design to be approved for city
    - Cost reduction is a must when you need so many nodes
      - Node cost
      - Labor cost
    - Location constraints of deployment
      - physical constraints for placement and signal quality
      - quality of service constraints for quality of monitoring, and
      - Cost constraints
  - Co-existence with other wireless technologies
    - WiFi Interference
Some New Devices Designed Recently

- Temp, Light
- Outdoor CO2
- CO2, Solar
- Dust
- Indoor CO
- Indoor SO2
- Indoor CO2
- Mesh Nodes, Solar panel
- Mobile Terminal
Some New Devices Designed Recently

- Pressure Sensor
- Magnet Sensor
- Oil Pressure Terminal
- Water Depth Sensor
- Water level
ILIGHT

CS557: Cyber-Physical
System examples (iLight)
Track Objects

Blue line: real trace
Red line: computed trace
Estimating heights

Testing environment and estimated height
Estimation error at most 2 cm w.h.p
WASTE WATER PROCESS
Managing Loosely Coupled Networked Control Systems with External Disturbances

This is a joint project with Professor ShangPing Ren (CS), Professor Fouad Teymour (ChE) and Professor Paul Anderson (CE).

Professor Xiangyang Li is the lead.
Managing Loosely Coupled Networked Control Systems with External Disturbances

- Characteristics and challenges of the targeted problem domain
  - Widely distributed physical systems
  - Has historical data, but also has large unpredictable factors
  - Wide range of timing granularities
  - Large spectrum of abstraction levels for events

- Research Focus
  Develop algorithms and timing analysis approaches to ensure that loosely coupled networked control systems satisfy timing constraints with different timing granularities; and develop event model to model and reason about events at different abstraction levels, ranging from simple sensor signals, to human control actions.
General Overview

- The goal of this research is to
  - understand the waterway as a Cyber-Physical System (CPS), and
  - to provide a set of strategies and tools for meeting new U.S. Environmental Protection Agency standards for ammonia regulation scheduled to take effect in 2014.

- Plan to update the waterway's systems by
  - extending the systems with available technologies such as wireless sensors and networks, and
  - by providing real-time, on-line monitoring and process control to minimize energy demands and carbon footprint associated with nutrient control.
Chicago Waterway System (NSF)

Joint work with Prof. Ren, Prof. Anderson and Prof. Teymour

Stickney WRP (world largest)

Ammonia sensor  Dissolved Oxygen sensor
Prof. XiangYang Li and 15 students visit Stickney WRP on March 17th, 2011.

One of the many visiting and meetings with local WRP's
Tanks
DO Sensors
Lab-used Sensors
Secondary Processing
Control Room
Control room
Key water quality indicators

- Chemical assessment
  - Oxygen saturation or dissolved oxygen (DO)
  - Chemical oxygen demand (COD)
  - Biochemical oxygen demand (BOD)
  - PH
  - Nitrate
  - Phosphate

- Physical assessment
  - Temperature
  - Total suspended solids (TSS)
  - Turbidity

- Biological assessment
Water resource systems

- Source waters
  - Lakes
  - Rivers
  - Groundwater

- Water treatment systems
  - Municipal treatment
  - Industrial treatment

- Wastewater treatment systems
  - Municipal wastewater
  - Industrial wastewater

- Stormwater
Cyber physical systems

- Intelligent sensor networks and software applied to more efficient and effective operations
- Information from the network/software system results in a change in operations
  - Range in the response times:
    - Real-time response for water/wastewater treatment
    - Long-term response times for watershed management
Optimization problem

- Minimize the cost of the system required to realize more efficient and effective operations
- Subject to:
  - Human health
  - Environmental quality
  - Commerce & industry
  - Other constraints...
Specific issues

- Network design
- Simulation and control models
- CPS operations
Illinois DO Water Quality Criteria

- **General Use Water**
  - March to July
    - >5.0 mg L⁻¹ minimum at all times
    - > 6.0 mg L⁻¹ 7-day mean
  - August to February
    - > 3.5 mg L⁻¹ minimum at all times
    - > 4.0 mg L⁻¹ 7-day mean
    - > 5.5 mg L⁻¹ 30-day mean

- **Secondary Contact and Indigenous Aquatic Life Use**
  - > 4.0 mg L⁻¹ / 3.0 mg L⁻¹ minimum at all times

**Is Continuous DO Monitoring Necessary to Determine Compliance with Standards?**
Are perpetual CDOM programs practical and reasonable?
Current MWRDGC DO Monitoring

- October 1994 to May 1996 weekly DO surveys in the Chicago River System. Water samples collected manually, chemically fixed in field, returned to laboratory for titration.

- 1998 initiated a comprehensive CDOM program to characterize Chicago River System
  - Initial focus on the Chicago River System for a two-year period
  - Program expanded to Calumet River System in 2001
  - Program further expanded to wadeable streams in 2005.

- Subsequently, resulting data have been used for calibration and verification of a water quality model for the CAWS used in IEPA’s UAA study.
Successful CDOM Program Requires Intensive QA

- Currently, Thirty-Two Locations Monitored
- DO, Water Temperature, and Specific Conductivity Measured Hourly at All Locations, pH and Turbidity at Selected Locations
- Monitors Deployed for Seven Continuous Days in Protective Enclosures
- Monitors Exchanged over Period of 3 Days /Week (Tuesday To Thursday)
- Calibrated and Serviced Monitors Deployed Weekly to Replace Monitors Retrieved from Field
- Winkler DO Check Sample Taken During Monitor Exchanges
- Housings Cleaned Weekly April Through November, Monthly December Through March
- Cross-Sectional DO Measured at Each Monitoring Location During April, August, and November
CDOM Program QA (Cont’d)

- Eighty Monitors Purchased
- Retrieved Monitors Cleaned Weekly With Laboratory Detergent Solution
- Battery Compartment, Cable Connector Inspected for Water Leakage
- Batteries Checked with Voltmeter
- O-Rings Cleaned, Inspected and Lubricated
- DO, Specific Conductivity, Temperature Calibration Checked Daily in Holding Tanks
- Polarographic DO Sensor Membrane and Electrodes Observed Daily Under Microscope
SOME KEY CHALLENGE ISSUES
Network design

- Spatial resolution for the sensor network
- Required network communication distance
- Required network communication capacity
- What sensor array provides the most useful integrated system information?
CPS operations

- Temporal resolution for the sensor network
- Temporal resolution for process control
- Network maintenance requirements
- Balancing network costs and process control confidence
Process and control models

- Two very different model requirements
  - Simulation model to assess network design
  - Control model links sensor input to process control
Water resource examples

- Watershed
- Municipal water treatment
- Wastewater treatment
Watershed management

- Federal and state programs
- Data to calibrate watershed model
- Asynchronous process control
- Link development to water quality
Municipal & industrial process water treatment

- Security

- Groundwater versus surface water

- Changes in water quality
  - Seasonal (algal blooms)
  - Storm events

- Monitor
  - Source water
  - Processes
  - Distribution system

- Real-time process control
Wastewater treatment

- Effluent quality is regulated
- Influent quality varies
  - Seasonal (temperature)
  - Storm events (flow, loading)
- Monitor
  - Influent
  - Processes
- Real-time control
Protecting legacy critical Infrastructures against coordinated attacks

Raw water
Raw water
Raw water
Pressure Vessel

V-1
V-2 Filter
V-3
P-5 Product Booster Pump
V-4
V-5 For Consumption
V-6 Purified Water Container
V-7 Pressure Release Container
V-8

S1
S2
S3

Pressure Vessel

Raw water
Raw water
Raw water
Behavior-Based Coordination for Open Distributed Real-Time and Embedded Computing

- Characteristics and challenges of the targeted problem domain
  - Open
  - Dynamic
  - Large scale of autonomous and concurrent entities
  - Hard QoS requirements
  - Running in a real world environment which is sometimes unpredictable

- Research Focus

  Develop models, software architecture and programming language support to facilitate the design and development of such systems and further be able to verify the correctness of the systems.
HARDWARES
WSN Controlled Mobile Car

ShaoJie Tang
XuFei Mao
XiaoHua Xu
Sensor nodes

- **General**

- **Co2, GPRS, GPS**

- TI MSP430 MCU 低功耗
- IEEE 802.15.4 兼容
- Tinyos 2.x 操作系统
More sensor nodes
Sensor nodes
Sensor nodes
Sensor nodes
CO2 Sensor node
Encapsulated Node
Another one
Many nodes
Wireless Mesh Nodes

- IEEE 802.11g/a
- 30Mbps, 1km
- Distance 7km,

- MultiHop
- Auto-Routing
- Security
Mesh Nodes

- Solar Panel and Wind Energy
Sensor nodes, mesh routers deployed
Some New Devices Designed Recently

- Temp, Light
- Outdoor CO2
- CO2, Solar
- Dust

Indoor CO
Indoor SO2
Indoor CO2
Mesh Nodes, Solar panel
Mobile Terminal
Some New Devices Designed Recently

Pressure Sensor

Magnet Sensor

Oil Pressure Terminal

Water Depth Sensor

Water level
OBSERVATIONS,
EXPERIENCES, LESSONS
Approaches

deployment

measurement

Improvement
Design and protocol

analysis
Measurement Study

multiple testing deployments at a campus forest

First: Network “topology” varies over time and space
## Multiple Network Scenarios

December 2009, 29 consecutive days, 2,540,000 data packets

<table>
<thead>
<tr>
<th>Trace No.</th>
<th>Network Scale</th>
<th>Power level</th>
<th>Data Rate (pkts/hour)</th>
<th>Duration (hour)</th>
<th>Duty cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>15</td>
<td>3</td>
<td>60</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>15</td>
<td>3</td>
<td>25</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>330</td>
<td>15</td>
<td>3</td>
<td>300</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>330</td>
<td>15</td>
<td>12</td>
<td>24</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>330</td>
<td>15</td>
<td>18</td>
<td>100</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>330</td>
<td>15</td>
<td>27</td>
<td>30</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>330</td>
<td>15</td>
<td>54</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>330</td>
<td>15</td>
<td>108</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>330</td>
<td>31</td>
<td>12</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>330</td>
<td>21</td>
<td>12</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>11</td>
<td>330</td>
<td>15</td>
<td>12</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>12</td>
<td>330</td>
<td>8</td>
<td>12</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>13</td>
<td>330</td>
<td>15</td>
<td>3</td>
<td>150</td>
<td>8%</td>
</tr>
<tr>
<td>14</td>
<td>330</td>
<td>15</td>
<td>60</td>
<td>12</td>
<td>8%</td>
</tr>
</tbody>
</table>
**Routing trace**
- Routing path
- Sensor reading

**Link trace**
- List of neighbor nodes
- RSSI, LQI, and ETX

**Node statistic trace**
- A large set of statistical information on each node
Out-band Measurement

**Overhearing**
- Multiple sniffers in the network to overhear the network traffic

**Beaconing**
- Each node actively broadcast beacons periodically

**Local logging**
- The fine-grained local events on the nodes are recorded as a backup data set for diagnosis
Measurements

Yield

• Measure the quantity of the collected data

Packet Reception Ratio / Loss Ratio

• Measure the quality of a link

Packet Delivery Ratio (PDR)

• The ratio of the amount of packets received by the destination to those sent by the source
Measures and Derivations

End-to-end delay

- The time difference between the sending time at the source node and the reception time at the sink

Correlation Coefficient

- A statistical measure of association between two variables
- This is often used for fault diagnosis of sensor networks
Traffic distribution: balanced in CTP?

5% nodes account 80% traffic.

90% nodes have very low traffic.

The traffic distribution is relatively stable over time.
Causes of Packet Losses

- PDR about 85%
- Link loss (61%) vs. Node drops (39%)
- Faulty behavior on forwarding nodes

Cumulative distribution of packet loss

Causes of packet drops on sensor nodes
Packet Loss Diagnosis

December 10, 2010; 400 nodes, 60,000m²

Data of 10 days:
1,137,430 packets received
181,862 packets lost

– The **green nodes** with PRR > 90%.
– The **red nodes** with PRR < 90%,
– The radius indicates the number of lost packets
Packet Losses: Non-ACK

- 84,030 packet loss due to non-ack
  - 46.2% of total losses
  - 68,444 caused by physical environment (bad links)
Packet Losses: Non-ACK

- 84,030 packet loss due to non-ack
  - 46.2% of total losses
  - 4,361 caused by interferences (contention <--reboot, loop)
Packet Losses: Corrupted Packets

- 9,511 corrupted packets
  - 9037 real losses (after consider retransmission)
  - ~5% of total loss
Packet Losses: Routing Loop

- 5,178 packet loss due to overflow from routing loop
  - 2.9% of total losses
  - 93% of overflow events did not result in packet loss
### Packets Loss Summary

<table>
<thead>
<tr>
<th>Root cause</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. sink-side failure</td>
<td>12.5%</td>
</tr>
<tr>
<td>1.1 vertical banding</td>
<td>12.45%</td>
</tr>
<tr>
<td>2. corruption</td>
<td>5%</td>
</tr>
<tr>
<td>3. overflow drops</td>
<td>2.87%</td>
</tr>
<tr>
<td>3.1 loop overflow drops</td>
<td>2.85%</td>
</tr>
<tr>
<td>3.2 non-loop overflow drops</td>
<td>0.02%</td>
</tr>
<tr>
<td>4. no-ack drops</td>
<td>46.2%</td>
</tr>
<tr>
<td>4.1 env-no-ack drops</td>
<td>37.6%</td>
</tr>
<tr>
<td>4.2 interference-no-ack drops</td>
<td>2.4%</td>
</tr>
<tr>
<td>5. reboot (direct impact on loss)</td>
<td>~0</td>
</tr>
</tbody>
</table>

About **35%** packet losses are **unidentified** now.
Summary of Some Observations

1. A small portion of nodes bottleneck the entire network, most of the existing network indicators may not accurately capture them.

2. The environment, although the dynamics are not as significant as we assumed, has an unpredictable impact on the sensor network performances.

3. By adjusting the operation parameters of various protocols (e.g., MAC), performance greatly improved.

4. Many challenges to make it
   1. **Sustainable** --- energy and fault diagnosis?
   2. **Scalable** --- performance bottleneck?
   3. **Robust** --- co-existence?
   4. **Predictable** ----- system stable points?
What limits the system scale?

- What is the dominant resource, first depleted when the network workload scales?

- Is such resource appropriately used?

- Where and when does resource depletion happen?

- How should existing protocols be improved to adapt to large-scale sensor network characteristics?

- How much information a network can support? How do networks scale?
CAPACITY OF LARGE SCALE WIRELESS NETWORKS
Large Scale WSN: $n$ nodes randomly placed in a square or nodes follow Poisson distribution with density $\zeta$. 
Asymptotical Capacity

- How much information a large WSN can support
  - Impact of network size $n$, and deployment size $a$
  - Impact of network model and interference model
  - Impact of different sessions, number of sessions, size of a session,
Network Model

- \( N_p(\zeta, n) \): place nodes in 2-D plane according to a Poisson point process of density \( \zeta \)
  - focus on a square \([0, (n/\zeta)^{1/2}]^2\)

- \( n_s \) sources \( S \) for \( n_s \) multicast flows, each with \( n_d \) nodes
  - Each source node \textit{randomly} selects \( n_d - 1 \) points and closest \( n_d - 1 \) nodes to these points as receivers

- Each source \( v_i \) \textit{sends} \( \lambda_i \) bits/second to all receivers.
General Node Density \( \zeta \in [1, n] \)
- random dense network (RDN, \( \zeta = n \)).
- random extended network (REN, \( \zeta = 1 \)).

General Session
- multicast capacity \( n_d \in [1, n] \)
- unicast capacity \( n_d = 1 \).
- broadcast capacity \( n_d = n \).

General Number of Sessions \( n_s \in (1, n] \)
- \( n_s = \Theta(n) \)
Various network capacities for multicast

- **Total Capacity**: $\sum_{v_i \in S} \lambda_i$

- **Minimum Capacity**: $\phi_{n_d} (n) = \min_{v_i \in S} \lambda_i$

- **Average Capacity**: $\sum_{v_i \in S} \lambda_i / n_s$
Interference and Link Models

- Fixed Range Protocol Interference Model (PrIM)
  - Link rate $w$ bps
  - Transmission range $r$
  - Interference range $R$
  - Receiver $v$ should not be interfered by other senders

Idealistic, but give us a reasonable scenario to study
Physical Interference Model

- Node $u$ can send to $v$ successfully at a given data rate only if

$$SINR = \frac{P_u \ell(u, v)}{\sigma + \sum_{w \in \text{sending}} P_w \ell(w, v)}$$

at node $v$ is at least a threshold value

- Node $u$ can send to $v$ successfully at a given data rate only if

$$SINR = \frac{P_u \ell(u, v)}{\sigma + \sum_{w \in \text{sending}} P_w \ell(w, v)}$$

at node $v$ is at least a threshold value
The capacity of link $u \rightarrow v$ is

$$B \log(1 + S\text{INR}), \text{ where}$$

$$S\text{INR} = \frac{P_u \ell(u, v)}{\sigma + \sum_{w \in \text{sending}} P_w \ell(w, v)}$$

Assume that all nodes have same power $P_u$, and

$$\ell(u, v) = \min(1, \|u - v\|^{-\beta}),$$

for a constant $\beta > 2$
Feasible Rate Vector

A data rate vector

\[ \lambda = (\lambda_1, \lambda_2, \cdots, \lambda_{n-1}, \lambda_n) \]

is feasible if there is a spatial and temporal scheme for scheduling transmissions such that by operating the network in a multi-hop fashion and buffering at intermediate nodes when awaiting transmission, every node \( v_i \) can send \( \lambda_i \) bits/sec average to its chosen destination nodes.
The per flow multicast capacity of a class of random networks is of order $\Theta(f(n))$ bits/sec

- if there are deterministic constants $c > 0$ and $c < c' < +\infty$ such that

$$\lim_{n \to \infty} \Pr(\min \lambda_i(n) = c \cdot f(n) \text{ is feasible}) = 1$$

$$\lim_{n \to \infty} \inf \Pr(\min \lambda_i(n) = c' \cdot f(n) \text{ is feasible}) < 1$$
Factors affecting capacity, but not studied here

- Noise Strength and Distribution,
- Dynamic Power Adjustment,
- Traffic Profile, Topology,
- Latency, Heterogeneity,
- Mobility,

- Channel diversity, multi-user, MIMO,
- Network Coding (application layer, physical layer)
- Successive Interference Cancelation
- Cognitive radio
- .....
RESULTS REVIEW AND
OUR RESULTS SUMMARY
Some Milestone Results for Unicast

<table>
<thead>
<tr>
<th>Unicast Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ random $\Rightarrow \Theta(W)$ $\Rightarrow \Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ NC $\Rightarrow \Theta\left(\frac{W}{\sqrt{n}}\right)$ GC</td>
</tr>
</tbody>
</table>

- **A Scare, 00-01** Gupta, Kumar: Per-flow unicast throughput under PrIM is $\Theta(W/\sqrt{n \log n})$

- **Mobility Matters, 2002** Grossglauser and Tse, $\Theta(W)$ via mobility and power-adjustment, large delay.

<table>
<thead>
<tr>
<th>Unicast Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ random $\Rightarrow \Theta(W)$ $\Rightarrow \Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ NC $\Rightarrow \Theta\left(\frac{W}{\sqrt{n}}\right)$ GC</td>
</tr>
</tbody>
</table>

- **Network Coding Does Not Matter, 2006** Li, Goeckel and Towsley: $\Theta(1/\sqrt{n \log n})$ with NC & PrIM.

<table>
<thead>
<tr>
<th>Unicast Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ random $\Rightarrow \Theta(W)$ $\Rightarrow \Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ NC $\Rightarrow \Theta\left(\frac{W}{\sqrt{n}}\right)$ GC</td>
</tr>
</tbody>
</table>

- **Channel Model Does Matter, 2007** Franceschetti et al., $\Omega(W/\sqrt{n})$ when using Gaussian Channel.
Milestone Results: Unicast

Total Unicast Capacity for RDN under Protocol Model, 

\[ n_s = \Theta(n), \text{ Gupta and Kumar } [IEEE TIT 2000]. \]
Milestone Results: Unicast

Unicast Capacity for RDN under Physical Model, $n_s = \Theta(n)$, Gupta and Kumar [IEEE TIT 2000].
Milestone Results: Unicast

Total Unicast Capacity for RDN, REN under Gaussian Model, \( n_s = \Theta(n) \), Franceschetti et al. [IEEE TIT 2007].
Milestone Results: Unicast

Total Unicast Capacity for RDN under Gaussian Model, $n_s = \Theta(n)$, Keshavarz-Haddad and Riedi [WiOPT2007].
Our Results: Unicast

Total Unicast Capacity for REN under Gaussian Model,
\( n_s = \Theta(n) \), Li et al [MobiCom2008].
Milestone Results: Unicast

Total Unicast Capacity for mobile RDN under Physical Model (i.i.d. mobility model), $n_s = \Theta(n)$, Grossglauser and Tse [INFO2002],
Milestone Results: Broadcast

Broadcast Capacity for RDN under Protocol Model, $n_s = \Theta(n)$, Keshavarz-Haddad et al. [MobiCom 2006].
Milestone Results: Broadcast

Broadcast Capacity for REN under Gaussian Model, \( n_s = \Theta(n) \), Zheng et al. [INFOCOM 2006].
Our Results: Multicast

Multicast Capacity for REN under Protocol Model, $n_s = \Theta(n)$, Li et al. [MobiCom 2007].
Summary of Our Results

- The aggregate multicast capacity of \( n \) sessions is

\[
\Lambda_{n_d}(n) = \begin{cases} 
\Theta(\sqrt{\frac{n}{\log n}} \cdot \frac{W}{\sqrt{n_d}}) & \text{when } n_d = \mathcal{O}\left(\frac{n}{\log n}\right), \\
\Theta(W) & \text{when } n_d = \Omega\left(\frac{n}{\log n}\right)
\end{cases}
\]

- Our results unify previous results

1. **Unicast** (when \( n_d = 2 \)): \( \Theta(\sqrt{\frac{n}{\log n}} \cdot W) \) by Gupta and Kumar

2. **Broadcast** (when \( n_d = n \)): \( \Theta(W) \) by Keshavarz-Haddad et al., MobiCom’06.

3. **Multicast** \((n_s = n^\epsilon \text{ and } n_d = n^{1-\epsilon})\), by Shakkottai, Liu, Srikant, Mobihoc’07.
Our Results: Multicast

Multicast Capacity for REN under Gaussian Model, $n_s = \Theta(n)$, Li et al. [MobiCom 2008].
Our Results: Multicast

Multicast Capacity for REN under Gaussian Model, $n_s = \Theta(n)$,

Wang, Li et al. [INFOCOM 2009].
Our Results: Multicast

Multicast Capacity for REN under Gaussian Model, \( n_s = \Theta(n) \), Wang, Li et al. [INFOCOM 2010].
Our Results: Multicast

Multicast Capacity for REN under Gaussian Model, $n_s = \Theta(n)$, Wang, Li et al. [INFOCOM 2011].
Our Results: Multicast

\[ \Lambda(n, n_d) \]

Multicast Capacity for RDN under Gaussian Model, \( n_s = \Theta(n) \), Wang, Li et al. [INFOCOM 2011].
Observations:

- There are two typical models in terms of scaling patterns of the network:

Criteria of Scaling Patterns:

- Dense Scaling Model: $\zeta = \Omega(\log n)$
- Extended Scaling Model: $\zeta = o(\log n)$
GENERAL APPROACHES
Multicast under Protocol Model

- **Data Copies Argument** (upper bound)
  - Estimate the expected (or asymptotic lower bound) number of nodes $N(b)$ that received (or listened) a bit $b$.
  - Capacity at most $n \cdot W / N(b)$ since all nodes receive at rate at most $n \cdot W$. 
Upper-bound Proof Flow

Capacity Upper-Bound \( \frac{nW}{N(b)} \)

Data Copies Lower-Bound \( N(b) \geq \frac{\tau \sqrt{n_d r \cdot n}}{2c_0 a} \)

Area \( D(T) \) of active nodes \( \geq \frac{\tau \sqrt{n_d a r}}{c_0} \) w.h.p.

Density \( \frac{n}{a^2} \) w.h.p.

Edge Length of \( T \) \( \geq \rho T \sqrt{n_d a} \) w.h.p.

Area \( D(T) = \Theta(\|T\| \cdot r) \)

Length of EMST \( \geq \tau \sqrt{n_d a} \) w.h.p.

\( \|T\| \geq \rho \|\text{EMST}\| \)
Lower-Bound: Routing and Scheduling

- Build EMST
  - Routing structure using EMST as backbone
  - Need to bound the conflict and total data copies
    - The lower-bound of multicast tree length w.h.p. is $\text{EMST}$
    - Maximum number conflicting flows in the network w.h.p
      - Using VC dimension (proved to be $O(\log n_d)$), and VC theorem
Lower-bound Proof Flow

Capacity Lower-Bound $\frac{nW}{\Delta \cdot N(b)}$

Data Copies Upper-Bound $N(b)$

Area $D(T)$ of active nodes $\leq$

Density $\leq 2 \frac{n}{a^2}$ w.h.p.

Edge Length of $T$ $\leq$

Area $D(T) = \Theta(||T|| \cdot r)$

Length of EMST $\leq$

$||T|| \leq \varphi_2||EMST||$
Multicast under Gaussian Model

- Divide the links in multicast into links among giant components (formed by short links), and other links.
Relationship between links

- Consider giant component with link length at most $l_c$

  Define the max distance between any node not in GC and the giant cluster by $\bar{l}_c$

\[
\text{If } l_c = o(\sqrt{\log n / \zeta}) \text{ then } \zeta \cdot l_c \cdot \bar{l}_c = \Omega(\log n)
\]
There is a link $uv$, *that will be used by many flows (say $f$)* ⇒ *the minimum data rate*

$- \min \lambda_i \leq \text{rate supported by } uv / f$
There is an isolated cluster \( C \) of nodes, and \( f \) flows will have links going inside this cluster \( \Rightarrow \) the minimum data rate

\[
\min \lambda_i \leq \text{total rate supported by links reaching } C / f
\]
Lower Bounds Techniques

- Highway systems
  - Cell is of $O(1)$ nodes inside
  - from *percolation* theory
  - First used by Tse et al
Our New Techniques

- Parallel Arterial Road Systems
  - longer links to connect isolated nodes to highway

- Parallel Scheduling
Other Capacity Results

- Hybrid Wireless Networks
  - Backbone networks + ad hoc networks
  - Asymptotic capacity for multicast

- Cognitive Radio Networks
  - Primary Networks + Secondary Networks
  - Asymptotic capacity for multicast

- Mobile Wireless Social Networks
  - Social Networking + Mobile Networks

- Capacity for other operations
  - Data collection and Data aggregation
  - SelectCast, AnyCast,
  - Capacity and Delay Tradeoffs
Summary

- Experience and Lessons from Large Scale WSN System Design and Deployment
  - OceanSense
  - GreenOrbs
  - CitySee

- Asymptotical Capacity of Large Scale Wireless Networks
  - Network model, and asymptotical capacity
  - Literature review
  - Our results summary
  - Our approaches
Other Research Concentrations

- Application of Sensor Networks
  - Wastewater processing (CPS medium project)
  - Mobile health
- Algorithms for wireless networks
  - Offline scheduling,
  - Online scheduling and optimization
  - Game theory and economics
  - Cognitive radio networks
- Social Networks
  - Information propagation
  - Team formation/link predication
Cyber Physical Systems
Cognitive Radio Networks
Mobile, Social Networks

- Privacy and security
- Energy saving
- Location, navigation

- Influence computation
  - Churn prediction
- Relationship learning
- Sentiment Analysis
- Spam detection
OUR GROUP

CS595: Foundations of Cyber-
Theoretical Studies

- Algorithm Design and Analysis of Practical Questions
  - Wireless ad hoc networks
  - Wireless sensor networks
  - RFID
  - Cognitive networks
  - **Online optimization (little regret)**
  - Computational geometry
  - Game theory and its applications
  - Information theory (such asymptotical behavior of large scale networks)
Where do we publish?

- Journals
  - IEEE/ACM Transactions on Networking, TPDS, Computers, JSAC, and so on
  - ACM Transactions, and so on

- Conferences
  - ACM MobiCom, ACM Mobihoc, ACM STOC, ACM SODA, ACM EC
  - IEEE INFOCOM, ICNP, ICDCS, and so on

- Well recognized and accepted in the community
Where do our students go?

- Graduated students (7 PhDs)
  - (4) Faculty at North Carolina Charlotte, Washington State University, Minnesota State University, BUPT (China)
  - Researcher at Google,
  - Game designer and truck industry
  - Financial industry
Professors, and students
Collaborators
Students and Collaborators
PhD Students (graduated, current)
PhD and MS Students (current)
More PhD students

- TaeHo Jung
- Cheng Bo
- JunZe Han
MS students

- Yue Tao (EE, IIT)
- Eric Sze Ching Duan (CS, IIT)
- SuFeng Niu (EE, IIT)
- PengQian Hu (CS, IIT)
- GuoBiao Yang (CS, IIT)
- YiTian Pan (CS, IIT)
- Chan Guo (CS, IIT)
- YanJie Wang (CS, IIT)
- Hao Bian (CS, IIT)
- Unsuk Heo (CS, IIT, undergraduate)
- Juan Garcia (CS, IIT)
- Siddharth Shankar (CS, IIT)
- Wei Wang
- YiFan Zhu
- Shufan Wang
Thank you!

Xiang-Yang Li
Professor, IIT, USA
www.cs.iit.edu/~xli
xli@cs.iit.edu