Data Fusion Improves the Coverage of Sensor Networks

Guoliang Xing
Assistant Professor
Department of Computer Science and Engineering
Michigan State University
http://www.cse.msu.edu/~glxing/
Outline

• Background
• Problem definition
  – Coverage of large-scale sensor networks
• Scaling laws of coverage
• Other projects
  – Model-driven concurrent medium access control
  – Integrated coverage and connectivity configuration
• Personal perspectives on research
Mission-critical Sensing Applications

• Large-scale network deployments
  – OSU ExScal project: 1450 nodes deployed in a 1260X288 m² region

• Resource-constrained sensor nodes
  – Limited sensing performance

• Stringent performance requirements
  – High sensing probability, e.g., 90%, low false alarm rate, e.g., 5%, bounded delay, e.g., 20s
Sensing Coverage

• Fundamental requirement of critical apps
  – How well is a region monitored by sensors?

• Coverage of static targets
  – How likely is a target detected?
Network Density for Achieving Coverage

• How many sensors are needed to achieve **full** or **instant** coverage of a geographic region?
  – **Any** static target can be detected at a high prob.

• Significance of reducing network density
  – Reduce deployment cost
  – Prolong lifetime by putting redundant sensors to sleep
State of the Art

• **K-coverage**
  – Any physical point in a large region must be detected by at least \( K \) sensors

• **Coverage of mobile targets**
  – Any target must be detected within certain delay

• **Barrier coverage**
  – All crossing paths through a belt region must be \( k \)-covered

• **Most previous results are based on simplistic models**
  – All 5 papers on the coverage problem published at MobiCom since 2004 assumed the disc model
Single-Coverage under Disc Model

- Deterministic deployment
  - Optimal pattern is hexagon
- Random deployment
  - Sensors deployed by a Poisson point process of density $\rho$
  - The coverage (fraction of points covered by at least one sensor):
    $$c = 1 - e^{-\rho \pi r^2}$$ [Liu 2004]
Sensing Model

The (in)famous disc model

Sensor can detect any target within range $r$

Real-world sensor detection

- There is no cookie-cutter “sensing range”!

Acoustic Vehicle Tracking Data in DARPA SensIT Experiments  [Duarte 04]
Contributions

• Introduce probabilistic and collaborative sensing models in the analysis of coverage
  – *Data fusion*: sensors combine data for better inferences

• Derive scaling laws of network density vs. coverage
  – Coverage of both static and moving targets

• Compare the performance of disc and fusion models
  – Data fusion can significantly improve coverage!
Outline

• Background
• Problem definition
  – Coverage of large-scale sensor networks
• Scaling laws of coverage
• Other projects
  – Model-driven concurrent medium access control
  – Integrated coverage and connectivity configuration
• Personal perspectives on research
Sensor Measurement Model

- Sensor reading $y_i = s_i + n_i$
- Decayed target energy $s_i = S \cdot w(x_i)$
- Noise energy follows normal distribution $n_i \sim N(\mu, \sigma^2)$

$$w(x) = \frac{1}{1 + x^k}, \quad 2 \leq k \leq 5$$
Single-sensor Detection Model

- Sensor reading $y_i$
- $H_0$ – target is absent
- $H_1$ – target is present

false alarm rate:
$$P_F = P(y_i \geq t \mid H_0) = Q\left(\frac{t - \mu}{\sigma}\right)$$

detection probability:
$$P_D = P(y_i \geq t \mid H_1) = Q\left(\frac{t - \mu - s_i}{\sigma}\right)$$

$Q(\cdot)$ – complementary CDF of the std normal distribution
Data Fusion Model

- Sensors within distance $R$ from target fuse their readings
  - $R$ is the fusion range
- The sum of readings is compared again a threshold $\eta$
- False alarm rate
  $$P_F = 1 - \chi_n(n \cdot \eta)$$
- Detection probability
  $$P_D = 1 - \chi_n(n \cdot \eta - \sum w(x_i))$$

$\chi_n$ – CDF of Chi-square distribution
$w(x_i)$ – Energy reading of sensor $x_i$ from target

$\eta$ – threshold

$R$ – fusion range
Outline

• Background
• Problem definition
  – Coverage of large-scale sensor networks
• Scaling laws of coverage
  – Coverage of static targets
• Other projects
  – Model-driven concurrent medium access control
  – Integrated coverage and connectivity configuration
• Personal perspectives on research
\((\alpha, \beta)\)-coverage

- A physical point \(p\) is \((\alpha, \beta)\)-covered if:
  - The system false alarm rate \(P_F \leq \alpha\)
  - For target at \(p\), the detection prob. \(P_D \geq \beta\)

- \((\alpha, \beta)\)-coverage is the fraction of points in a region that is \((\alpha, \beta)\)-covered:
  - Full (0.01, 0.95)-coverage: system false alarm rate is no greater than 1%, and the prob. of detecting any target in the region is no lower than 95%
Extending the Disc Model

• Classical disc model is deterministic
• Extends disc model to stochastic detection
  – Choose sensing range $r$ such that if any point is covered by at least one sensor, the region is $(\alpha, \beta)$-covered

$$ r = w^{-1} \left( \frac{Q^{-1}(\alpha) - Q^{-1}(\beta)}{\delta} \right) $$

$\delta$-- signal to noise ratio $S/\sigma$

• Previous results based on disc model can be extended to $(\alpha, \beta)$-coverage
Disc and Fusion Coverage

• Coverage under the disc model
  – Sensors independently detect targets within *sensing range* $r$

• Coverage under the fusion model
  – Sensors collaborate to detect targets within *fusion range* $R$
(α, β)-coverage under Fusion Model

- The (α, β)-coverage of a random network is given by

\[ c = \mathbb{P} \left( \frac{\sum_{i \in F(p)} S_i}{\sqrt{N(p)}} \geq \sigma \left( Q^{-1}(\alpha) - Q^{-1}(\beta) \right) \right) \]

- \( F(p) \) – set of sensors within fusion range of point p
- \( N(p) \) – # of sensors in \( F(p) \)

![Diagram showing coverage and optimal fusion range](image-url)
Network Density for Full Coverage

• $\rho_f$ and $\rho_d$ are densities of random networks under fusion and disc models

$$\rho_f = \mathcal{O}\left(\frac{2r^2}{R^2} \cdot \rho_d\right), \quad c \to 1$$

• Sensing range is a constant

$$r = w^{-1} \left( \frac{Q^{-1}(\alpha) - Q^{-1}(\beta)}{\delta} \right)$$

• Opt fusion range grows with network density

$\Rightarrow \rho_f < \rho_d$ when high coverage is required
Network Density w Opt Fusion Range

• When fusion range is optimized with respect to network density

\[ \rho_f = O \left( \rho_d^{1 - 1/k} \right), \quad c \to 1 \quad , \ 2 \leq k \leq 5 \]

• When k=2 (acoustic signals)

\[ \rho_f = O \left( \sqrt{\rho_d} \right) \]

• Data fusion significantly reduces network density
Network Density vs. SNR

• For any fixed fusion range

\[
\frac{\rho_f}{\rho_d} = \mathcal{O}\left(\delta^{2/k}\right), \quad c \to 1
\]

• The advantage of fusion decreases with SNR
Trace-driven Simulations

- Data traces collected from 75 acoustic nodes in vehicle detection experiments from DARPA SensIT project
  - $\alpha = 0.5, \beta = 0.95$, deployment region: $1000\text{m} \times 1000\text{m}$
Simulation on Synthetic Data

- $k=2$, target position is localized as the geometric center of fusing nodes
Conclusions

• Bridge the gap between data fusion theories and performance analysis of sensor networks
• Derive scaling laws of coverage vs. network density
  – Data fusion can significantly improve coverage!
• Help to understand the limitation of current analytical results based on ideal sensing models
• Provide guidelines for the design of data fusion algorithms for large-scale sensor networks
Outline

• Background
• Problem definition
  – Coverage of large-scale sensor networks
• Scaling laws of coverage
  – Coverage of static targets
• Other projects
  – Model-driven concurrent medium access control
  – Integrated coverage and connectivity configuration
• Personal perspectives on research
Improve Throughput by Concurrency

• Enable concurrency by controlling senders' power
Received Signal Strength

• 18 Tmotes with Chipcon 2420 radio
• Near-linear $\text{RSS}_{\text{dBm}}$ vs. transmission power level
• Non-linear $\text{RSS}_{\text{dBm}}$ vs. $\log(\text{dist})$, different from the classical model!
Packet Reception Ratio vs. SINR

0~3 dB is "gray region"

- Classical model doesn't capture the gray region

parking lot, no interferer  office, no interferer  office, 1 interferer
C-MAC Components

- Implemented in TinyOS 1.x, evaluated on a 18-mote test-bed
- Performance gain over TinyOS default MAC is >2X
Performance Evaluation

- Implemented in TinyOS 1.x
- 16 Tmotes deployed in a 25x24 ft office
- 8 senders and 8 receivers
Experimental Results

- System throughput (Kbps) vs. Time (second)
- System throughput (Kbps) vs. Number of Senders

- Improve throughput linearly with num of senders.
Deterministic Coverage + Connectivity

- Select a set of nodes to achieve
  - K-coverage: every point is monitored by at least K sensors
  - N-connectivity: network is still connected if N-1 nodes fail

A network with 1-coverage and 1-connectivity
Connectivity vs. Coverage: Analytical Results

• Network connectivity does **not** guarantee coverage
  – Connectivity only concerns with node locations
  – Coverage concerns with **all** locations in a region

• If \( \frac{R_c}{R_s} \geq 2 \)
  – K-coverage \( \Rightarrow \) K-connectivity
  – Implication: given requirements of K-coverage and N-connectivity, only needs to satisfy \( \max(K, N) \)-coverage
  – Solution: Coverage Configuration Protocol (CCP)

• If \( \frac{R_c}{R_s} < 2 \)
  – CCP + connectivity mountainous protocols
Research Summary

• Data fusion in sensor networks
  – Coverage [MobiCom 09]; deployment [RTSS 08]; mobility [ICDCS 08]
• MAC protocol design and architecture
  – C-MAC: concurrent model-driven MAC [Infocom08]
  – UPMA: unified power management architecture [IPSN 07]
• Sensornet/real-time middleware
  – MobiQuery: spatiotemporal query service for mobile users [ICDCS 05, IPSN 05]
  – nORB: light-weight real-time middleware for networked embedded systems [RTAS 04]
• Controlled mobility
  – Mobility-assisted spatiotemporal detection [ICDCS 08, IWQoS 08]
  – Rendezvous-based data transport [MobiHoc 08, RTSS 07]
• Power management
  – Minimum power configuration [MSWiM 07, MobiHoc 05, TOSN 3(2)]
  – Integrated coverage and connectivity configuration [TOSN 1(1), SenSys 03]
  – Impact of sensing coverage on geographic routing [TPDS 17(4), MobiHoc 04]
  – Real-time power-aware routing in sensor networks [IWQoS 06]
  – Data fusion for target detection [IPSN 04]
"Measure" of good research

• Good research is hard to plan or measure!
• Papers
  – Reputation of conferences/journals
  – Citations
• Demos/prototypes/systems
  – Economical/social impact
Find a research problem -- Dos

• Getting familiar with "real stuff"
  – What're real apps of sensor nets?
  – What're the fundamental limitations of practical sensor nets?
  – What're the unique properties of sensor nets?

• Paper reading
  – General vs. special topics
  – Are the fundamental assumptions challenged?
    • Circular sensing/comm. models
    • Random mobility

• System or theory?
  – System work: focuses more on impls & evaluation, simple (but non-trivial) ideas
  – Theoretical work: deep solution and analysis of interesting problems
  – Problem driven, works for me, but not the only way
Find a research problem — Don'ts

– This problem hasn't been studied, let's publish a trivial solution before others find it
– I used xxx theory in this paper, let's apply it to next paper
– This idea works for Internet, let's apply it to sensor nets
– I have improved this famous algorithm's performance by 20%, let's publish it
– This algorithm can do 3 different things by combining existing ideas
Problem solving

• Model with solid theoretical foundation
  – Never be afraid of trying new approaches or theoretical tools
  – Geometry, graph theory, probability,…

• Break a big problem into smaller problems
  – Solve special cases first, add more twists

• Don't have to be good at theory
  – Best results are often obvious in retrospect: “anyone could have thought of that”

• Identify your weakness
  – Ask for collaborations, change approaches

• Read, think, discuss
What to do when gets stuck

• Find causes
  – Need new theory tools – look for collaborators
  – Need working systems – build or borrow
  – Problem is not as interesting as expected
• Be persistent
  – Don't give it up too soon
• Be flexible
  – Change directions before too late
• Deal with paper rejections
  – Opportunities to improve
  – Focus on the big goal
What makes a good paper I

• Significant and interesting problem
  – Clear statement/formulation of the problem
  – Why it is interesting/important, always back up with real apps/systems
  – Why existing solutions can't work
What makes a good paper II

• Interesting solutions
  – Simple but elegant
  – In-depth analysis, good theoretical results
  – New system design/implementation methodology

• Solid performance evaluation
  – Impl and evaluation on real testbeds
  – Thorough realistic simulations
What makes a good paper III

• Good writing
  – Exciting abstract/introduction
  – Complete and concise related work
  – Clean problem description and assumptions with justifications
  – Good balance of analysis, proofs and/or protocol/system description
  – Smooth flow (forward/backward references), good illustrations
  – More than 5 polishing rounds before submission
What makes a bad/average paper

- Old problems
  - "yet another routing protocol"
  - "we improve an existing solution by 20%"
- Shaky assumptions
  - Made in favor of your solutions
  - Complicated your problems
- Average solutions
  - Centralized solutions without good bounds
  - Heuristics without insights
- Questionable performance evaluation
  - No comparisons with existing solutions
  - No reasonable explanation to unexpected results
    - Never throw away bad data and only show good ones!
  - Ignorance of important issues
    - packet loss, interference, perfect sensing, change 10 parameters in a run
  - Lack of enough details: "implemented it in NS2, and here are results".....
    - Which version of NS2? What wireless prop. model? Energy model?
Advices on presentation

- Allocate 1.5-2 minutes per slide
  - When running out time, skip some slides rather than rush through all slides
- Do >3 dry runs with friends/critics
- Tape a practice talk (audio/video tape)
- Other tips
  - Don't use fonts smaller than 18
  - Don't use too fancy templates
  - Don't put too much text on one slide
  - Face audience, don't read slides
  - Explain diagrams clearly
What makes a successful PhD student

• Motivation
• Motivation
• Motivation
Acknowledgement

• Students
  – Rui Tan, Mo Sha

• Collaborators
  – Benyuan Liu, Jianping Wang.....
Michigan State University

• First land-grant institution
  – Founded in 1855, prototype for 69 land-grant institutions established under the Morrill Act of 1862

• One of America's Public Ivy universities

• Big ten conferences
  – University of Illinois, Indiana University, University of Iowa, University of Michigan, University of Minnesota, Northwestern University, Ohio State University, Pennsylvania State University, Purdue University, University of Wisconsin

• Single largest campus, 8th largest university in the US with 46,648 students and 2,954 faculty members

• Rankings of 2008
  – 80th worldwide, Shanghai Jiao Tong University’s Institute of Higher Education
CSE@MSU

• People
  – 27 tenure-stream faculty
  – Each year awards approximately 100 BS, 40 MS, and 10 PhD degrees in Computer Science

• Research
  – 9 research laboratories, with annual research expenditures exceeding $3.5 million

• Rankings
  – 15th graduate program in US, a recent article of Comm. of ACM
  – Top 100, Shanghai Jiao Tong University’s Institute of Higher Education
My Group

• Research
  – Sensor networks
    • Data fusion, power management, voice streaming, controlled mobility
  – Low-power wireless networks
    • MAC, Interference management
  – Cyber-physical systems

• Students
  – Supervise 4 PhDs (CityU and MSU), 2 MS
  – Co-supervise 4 PhDs (CAS, CWM, UTK, MSU)