Localization in Wireless Networks

Presented by: Rich Martin

Joint work with: David Madigan, Eiman Elnahrawy, Wen-Hua Ju, P. Krishnan, A.S. Krishnakumar

Rutgers University and Avaya Labs

WINLAB IAB, June 2005
Wireless Explosion

- Technology trends creating cheap wireless communication in every computing device

- Radio offers localization opportunity in 2D and 3D
  - New capability compared to traditional communication networks

- Challenge: Can we localize all device radios using only the communication infrastructure?
  - How much existing infrastructure can we leverage?

- 3 technology communities
  - WLAN (802.11x)
  - Sensor networks (802.15)
  - Cell carriers(3G)
• Build general purpose localization & search analogous to general purpose communication & search!

• Can we make finding objects in the physical space as easy as Google?
Vision to reality

• Getting closer …

This talk:
• Localize with only existing infrastructure

• Ad-hoc
  – No more labeled data (our contribution)
Radio-based Localization

• Signal decays linearly with log distance in a laboratory setting

\[ S_j = b_{1j} + b_{2j} \log(D_j) \]
\[ D_j = \sqrt{(x - x_j)^2 + (y - y_j)^2} \]

– Use triangulation to compute \((x, y)\) » Problem solved

• Reality is …
  – noise, multi-path, reflections, systematic errors

\([-80, -67, -50]\)
\((x?, y?)\)
Supervised Learning-based Systems

• Training

• Offline phase
  – Collect “labeled” training data \([X,Y], S1, S2, S3, ..\]}

• Online phase
  – Match “unlabeled” RSS
  – \([?, ?], S1, S2, S3, ..\] to existing “labeled” training fingerprints
Previous Work

• People have tried almost all existing supervised learning approaches
  – Well known RADAR (nearest neighbor)
  – Probabilistic, e.g., Bayes a posteriori likelihood
  – Support Vector Machines
  – Multi-layer Perceptrons
  – ...
  [Bahl00, Battiti02, Roos02, Youssef03, Krishnan04, …]

• All have a major drawback
  – Labeled training fingerprints: “profiling”
  – Labor intensive (286 points in 32 hrs => 6.7 min/point)
  – Need to be repeated over the time
Contribution

Used Bayesian Graphical Models (BGM):

- Performance-wise: comparable
- Minimum labeled fingerprints
- Adaptive
- Simultaneously locate a set of objects

• **Advantage: zero-profiling**
  - No more “labeled” training data needed
  - Unlabeled data can be obtained using existing data traffic
Outline

• Motivations and Goals

• Experimental setup

• Bayesian background

• 3 Bayesian Models: M1, M2, M3

• Comparison to previous work

• Conclusions and Future Work
Experimental Setup

- 3 Office buildings
  - BR, CA Up, CA Down
- 802.11b
- Different sessions, days
- All give similar performance
- Use BR as example

BR: 5 access points, 225 ft x 175 ft, 254 measurements
Bayesian Graphical Models

- Encode dependencies/conditional independence between variables

Vertices = random variables
Edges = relationships

Example \([(X,Y), S]\), AP at \((x_b, y_b)\)

Log-based signal strength propagation

\[
S = b_1 + b_2 \log(D)
\]

\[
D = \sqrt{(x - x_b)^2 + (y - y_b)^2}
\]
Model 1 (Simple): labeled data

Position Variables

Distances

Observed Signal Strengths

Base Station Propagation constants (unknown)

$X_i \sim \text{uniform}(0, \text{Length})$  \quad Yi \sim \text{uniform}(0, \text{Width})$

$S_i \sim N(b_{0i} + b_{1i} \log(D_i), _{-i})$, i=1,2,3,4,5

$b_{0i} \sim N(0,1000)$, i=1,2,3,4,5 \quad b_{1i} \sim N(0,1000)$, i=1,2,3,4,5
<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Labeled: training</strong></td>
<td>• Probability distributions for all the unknown variables</td>
</tr>
<tr>
<td>[(x1,y1),(-40,-55,-90,..)]</td>
<td>• Propagation constants</td>
</tr>
<tr>
<td>[(x2,y2),(-60,-56,-80,..)]</td>
<td>– b0i, b1i for each Base Station</td>
</tr>
<tr>
<td>[(x3,y3),(-80,-70,-30,..)]</td>
<td>• ((x,y)) for each mobile ((?,?))</td>
</tr>
<tr>
<td>[(x4,y4),(-64,-33,-70,..)]</td>
<td></td>
</tr>
</tbody>
</table>

**Unlabeled: mobile object(s)** |

| [(?,?),(-45,-65,-40,..)] |
| [(?,?),(-35,-45,-78,..)] |
| [(?,?),(-75,-55,-65,..)] |
Solving for the Variables

• Closed form solution doesn’t usually exist
  » simulation/analytic approx

• We used **MCMC simulation** (Markov Chain Monte Carlo) to generate predictive samples from the joint distribution for every unknown \((X,Y)\) location
Example Output
Performance results

M1 better

Comparable
Model 2 (Hierarchical): labeled data

Allowing any signal propagation constants too constrained!

Assume all base-stations parameters normally distributed around a hidden variable with a mean and variance

- Intuition:
  - Same hardware should generate same signal propagation constants
  - Systematic bias in different environments (e.g. a closet)
M1, M2, SmoothNN Comparison

M2 similar to M1, but better with very small training sets
Both comparable to SmoothNN
No Labels

- Challenge: Position estimates without labeled data

- Observe signal strengths from existing data packets (unlabeled by default)

- No more running around collecting data..
- Over and over.. and over..
Labeled: training
[(x1,y1),(-40,-55,-90,...)]
[(x2,y2),(-60,-56,-80,...)]
[(x3,y3),(-80,-70,-30,...)]
[(x4,y4),(-64,-33,-70,...)]

Unlabeled: mobile object(s)
[(?,?),(-45,-65,-40,...)]
[(?,?),(-35,-45,-78,...)]
[(?,?),(-75,-55,-65,...)]

• Probability distributions for all the unknowns

• Propagation constants
  – b0i, b1i for each Base Station

• (x,y) for each (?,?)
Model 3 (Zero Profiling)

- Same graph as M2 (Hierarchical) but with (unlabeled data)

Why this works:
- [1] Prior knowledge about distance-signal strength
- [2] Prior knowledge that access points behave similarly
Results Close to SmoothNN

Leave-one-out error (feet)

Average Error in Feet

Size of input data

Zero Profiling M3

SmoothNN

UNLABELED

Labeled
Other ideas

- **Corridor Effects**
  - What? RSS is stronger along corridors

Variable $c = 1$ if the point shares $x$ or $y$ with the AP

**No improvements**

**Informative Prior distributions**
Comparison to previous work

Error CDF Across Algorithms

- More ad-hoc
- Adaptive
- No labor investment
Conclusions and Future Work

• First to use BGM
• Considerable promise for localization
• Performance comparable to existing approaches
• Zero profiling! (can we localize anything with a radio??)

• Next Steps
  – Validate with 802.15
  – Variational approximations
  – Tracking
  – Some extra infrastructure
    • Incorporate Angle of Arrival
    • Time of Flight