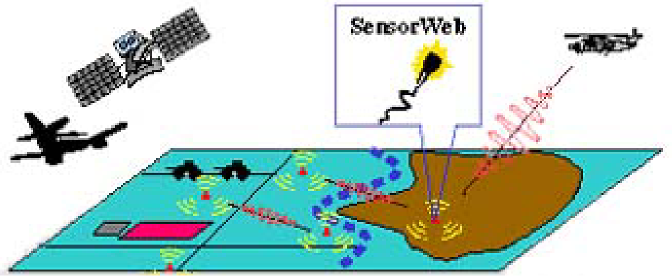
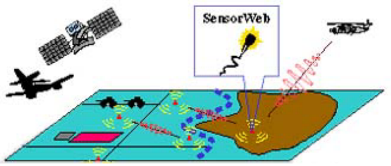


Network-constrained Estimation

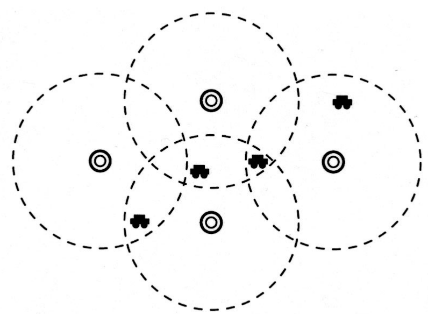
Alan S. Willsky

SensorWeb MURI Review Meeting
June 14, 2002

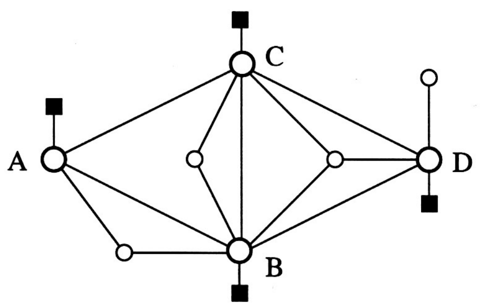


A Notional Example

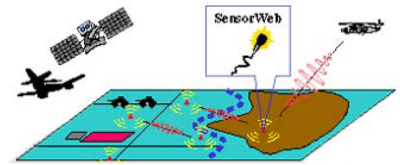
Multiple sensors with one or more bearing or location measurements



Possibly additional signal features

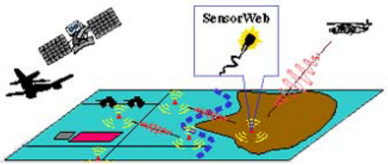


Challenge: *Scalable algorithms* for data association and estimation under network constraints



The Estimation/Association Problem-I

- Objects: $\{1, \ldots, N\}$ Sensors: $\{1, \ldots, M\}$
- O_i = objects seen by ith sensor $= \{n_{i1}, \ldots, n_{im_i}\} \subset \{1, \ldots, N\}$
- S_k = sensors seeing kth object $= \{r_{k1}, \ldots, r_{kn_k}\} \subset \{1, \ldots, M\}$
- Desired quantities
  - x_k = Object "state" (location, velocity, type, …)
  - $p(x_k)$ = "Prior" distribution



The Estimation/Association Problem-II

- Assignment and measurement permutations

  - Sensor i measurements $\{1, \ldots, m_i\}$

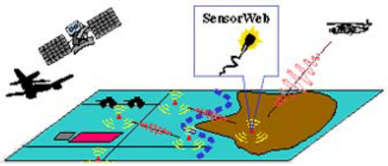
  - Permutation $\pi_{\mathrm{i}} : \{n_{i1}, \ldots, n_{im_i}\} \longrightarrow \{1, \ldots, m_i\}$

    $\pi_i(n_{ij}) =$ Sensor i measurement index for object n_{ij}

  - Assignment vector for Object $k : a_k = \{j_{k1}, \ldots, j_{kn_k}\}$

    $j_{ki} =$ Measurement index for Sensor r_{ki} observation of object k

- The data association constraint : $j_{ki} = \pi_{r_{ki}}(k)$

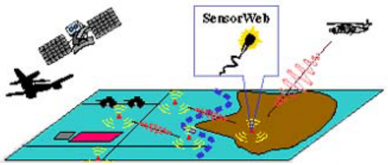


The Estimation/Association Problem-III

- Measured quantities
  - $\{y_{i1}, \ldots, y_{im_i}\}$ – measurements from Sensor i
- If $\{a_k\}$ or equivalently $\{\pi_i\}$ are known

$$y_{i\pi_i(n_{ij})} \text{ measures Object } \pi_{\mathrm{i}}(n_{ij})$$

$$(\text{e.g. } y_{i\pi_i(n_{ij})} = f(x_{\pi_i(n_{ij})}) + \text{noise}$$

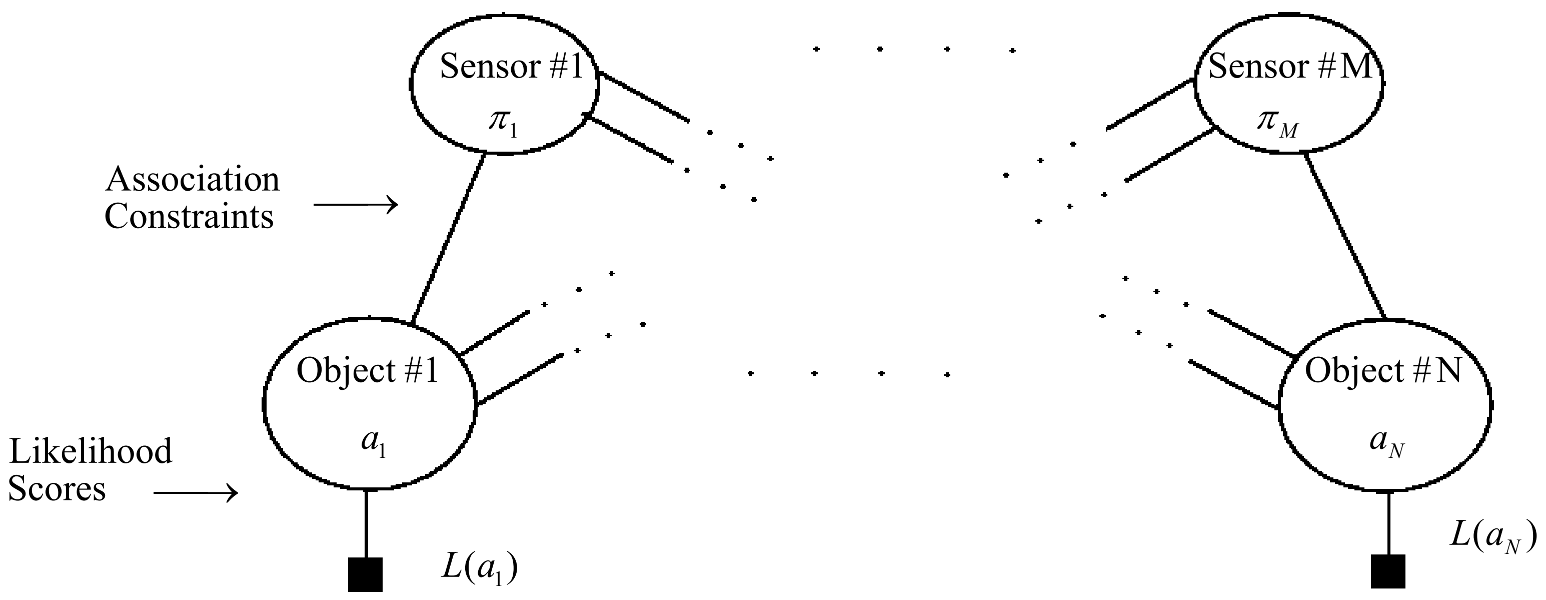


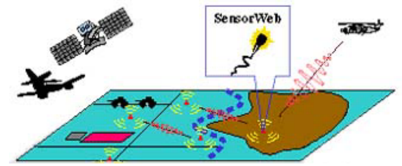
The Estimation Problem

- *Given* the assignments/permutations, compute the optimal estimates for each object as well as the *likelihoods* for each set of assignments to each individual object
  - A graphical model estimation problem
  - The likelihoods for each set of assignments to each object act as "scores" for optimal data association

The Association Problem

- Given the "scores", determine the optimal (or nearly optimal) set of assignments
  - This is a graphical model optimization problem





Fusion and Inference on Graphical Models

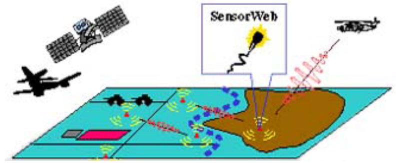
- $\mathcal{G} = (\mathcal{V}, \mathcal{E}), \mathcal{V} =$ nodes, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V} =$ edges
- $\mathrm{C} =$ set of cliques $C \subset \mathcal{V}$
- $x_s, s \in \mathcal{V}$ - random variables / vectors at nodes of the graph, forming a Markov random field
- Given label "compatibility functions" $\psi_C(x_C)$

$$P(\{x_S \mid s \in \mathcal{V}\}) \propto \prod_{C \subset \mathrm{C}} \psi_C(x_C)$$

- Objective

Estimation : Compute $P_s(x_s)$

Optimization : arg max $P(\{x_S \mid s \in \mathcal{V}\})$



Trees are Nice

- If the graph is acyclic, the distribution factorizes:

For Estimation

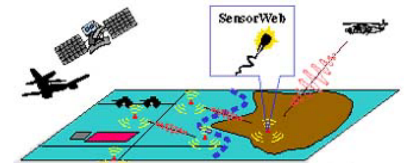
$$P(\{x_s \mid s \in \mathcal{V}\}) = \prod_{s \in \mathcal{V}} P_s(x_s) \prod_{(s,t) \in \mathcal{E}} \frac{P_{st}(x_s, x_t)}{P_s(x_s) P_t(x_t)}$$

For Optimization

$$P(\{x_s \mid s \in \mathcal{V}\}) \propto \prod_{s \in \mathcal{V}} \overline{P}_s(x_s) \prod_{(s,t) \in \mathcal{E}} \frac{\overline{P}_{st}(x_s, x_t)}{\overline{P}_s(x_s) \overline{P}_t(x_t)}$$

$$\overline{P}_s(x_s) = \max_{\{x_t \mid t \neq s\}} P(\{x_s \mid s \in \mathcal{V}\})$$

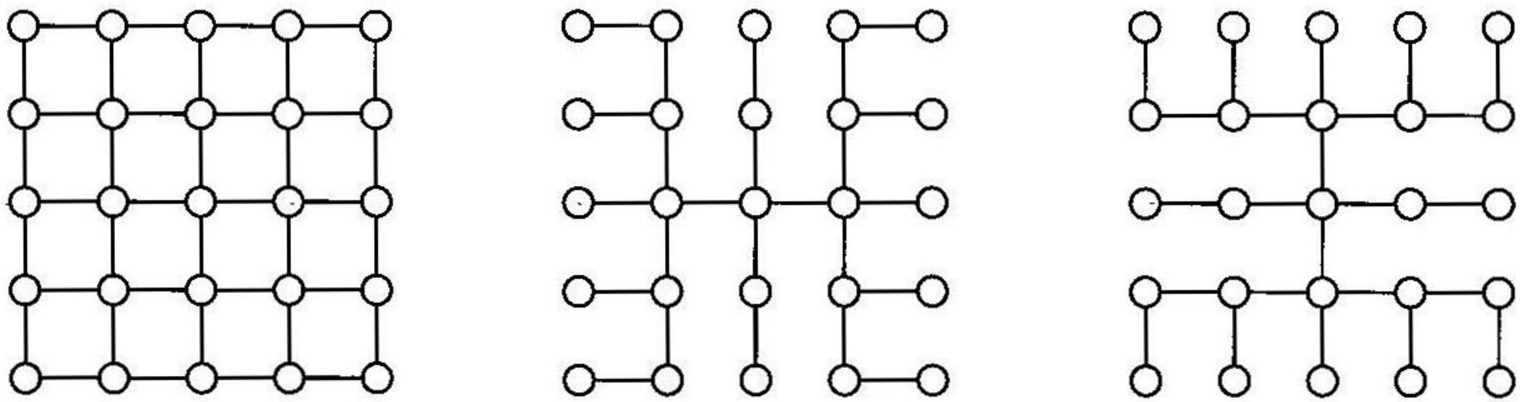
- Furthermore, these factorizations can be computed by sequences of local passing of messages

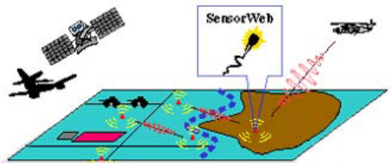


Exploiting acyclic structure

- Last time, introduced three classes of algorithms:
  - **Embedded Tree Estimation Algorithms**
  - Recursive Cavity Models (for linear and nonlinear estimation)
  - **Tree Reparameterization Algorithms (for discrete, continuous, hybrid estimation and graphical optimization)**

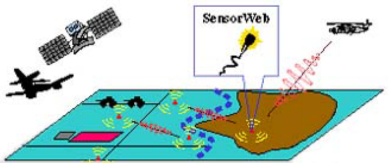
Embedded Trees



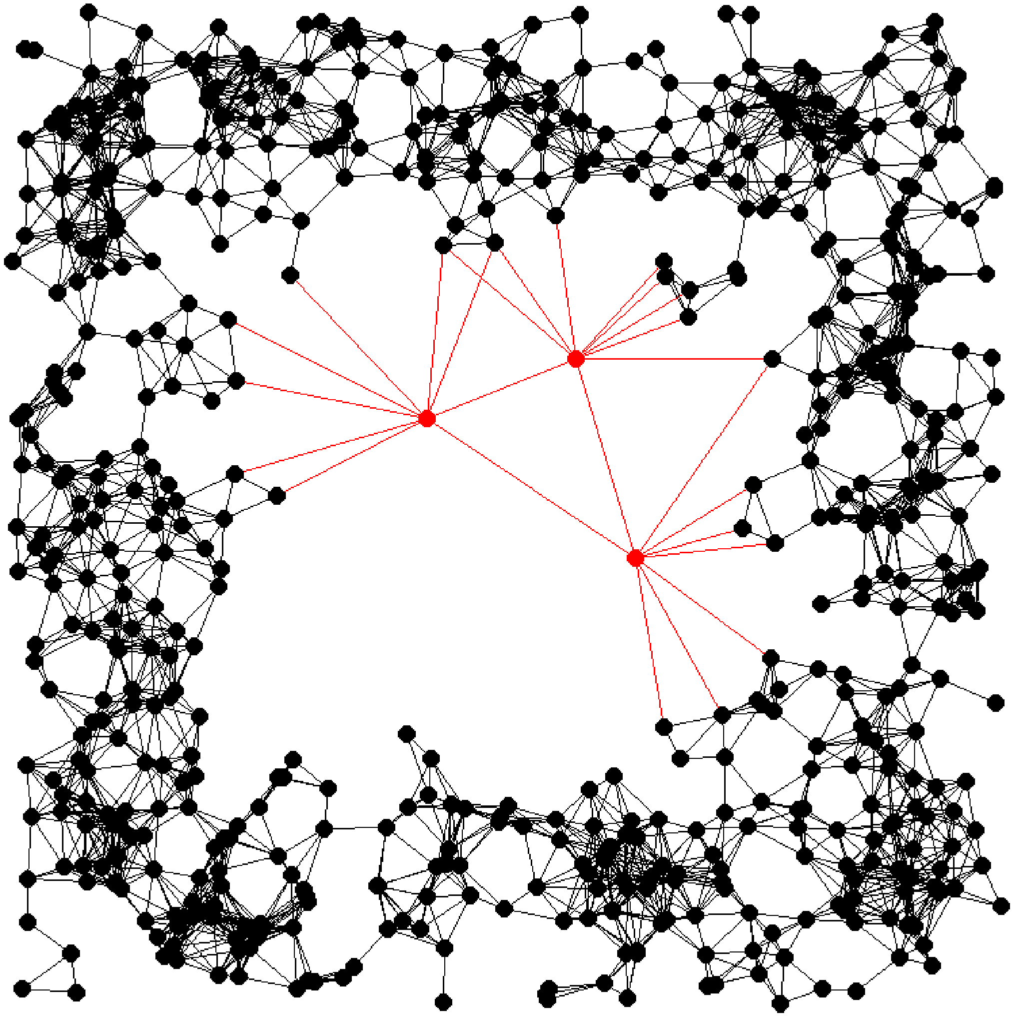


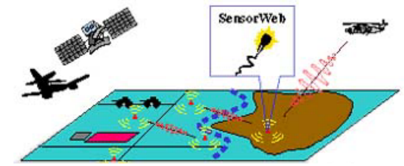
ET (continued)

- Previous results
  - Algorithms that (if they converge) yield not only optimal estimates but also correct error statistics
- Recent progress
  - Demonstration of excellent convergence properties
    - Using multiple trees
    - Using "preconditioner" concepts (tree computation followed by "local" relaxation steps
    - Sufficient conditions for convergence

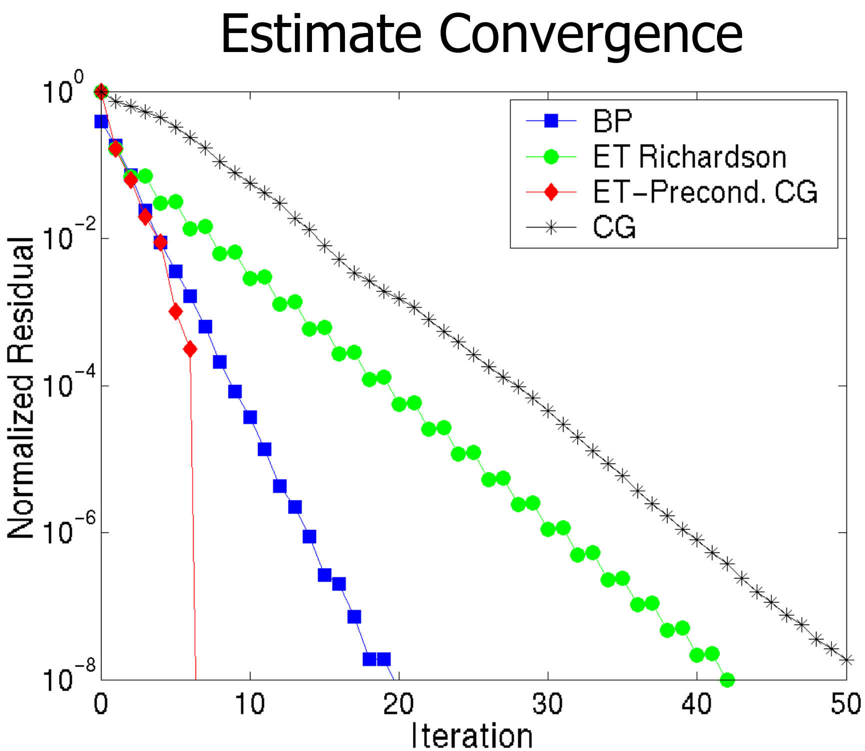


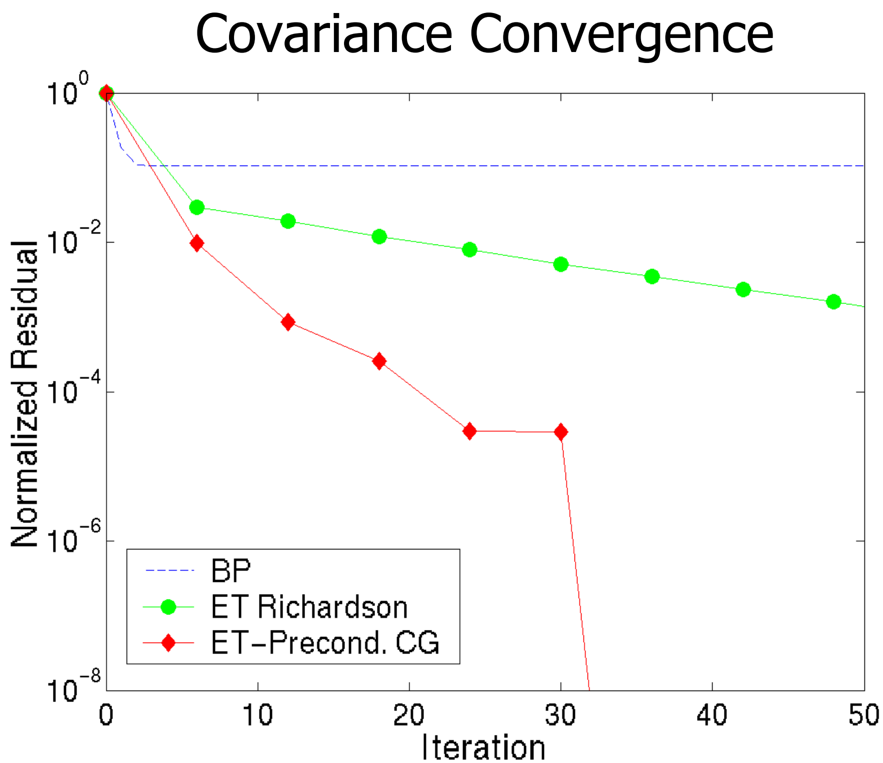
Network of 600 sensor nodes

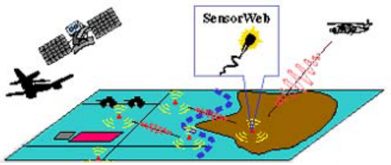




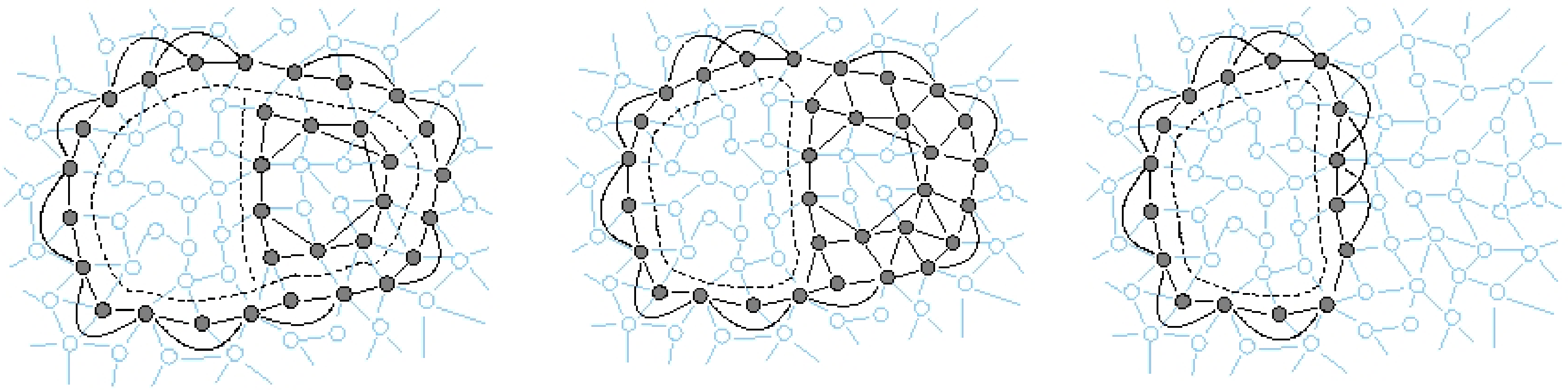
Estimate and covariance convergence results

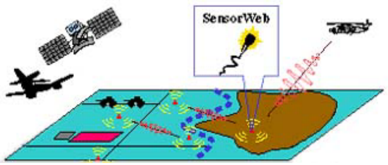






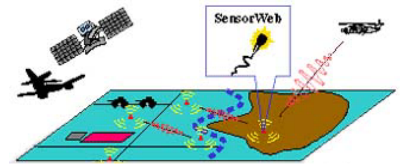
Recursive Cavity Models





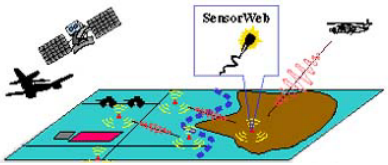
RCM (continued)

- Previous results
  - Last year we introduced the RCM concept
- Recent progress
  - Demonstration of efficiency and accuracy of RCM procedures with "boundary thinning"
  - Extension from linear models to general discrete and hybrid models
  - Theoretical framework for establishing stability and performance bounds from boundary thinning



Tree-Reparameterization Algorithms

- Previous results (for estimation *only*)
  - Introduction of the framework
  - Characterization of fixed points of iterations
  - Some convergence results
  - Initial work on characterizing errors in resulting estimates



The TRP Concept

- For *any* embedded acyclic structure:

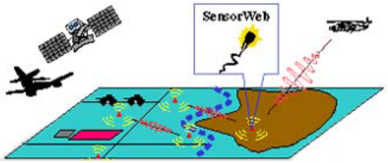
For Estimation

$$P(\{x_s \mid s \in \mathcal{V}\}) = \prod_{s \in \mathcal{V}} T_s(x_s) \prod_{(s,t)\in\mathcal{E}} \frac{T_{st}(x_s, x_t)}{T_s(x_s) T_t(x_t)} \times \text{Remainder}$$

For Optimization

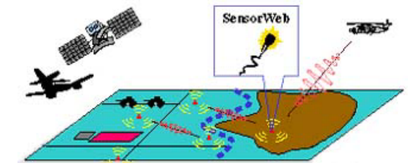
$$P(\{x_s \mid s \in \mathcal{V}\}) \propto \prod_{s \in \mathcal{V}} \overline{T}_s(x_s) \prod_{(s,t)\in\mathcal{E}} \frac{\overline{T}_{st}(x_s, x_t)}{\overline{T}_s(x_s) \overline{T}_t(x_t)} \times \text{Remainder}$$

$$\overline{T}_s(x_s) = \max_{\{x_t \mid t \neq s\}} T(\{x_s \mid s \in \mathcal{V}\})$$

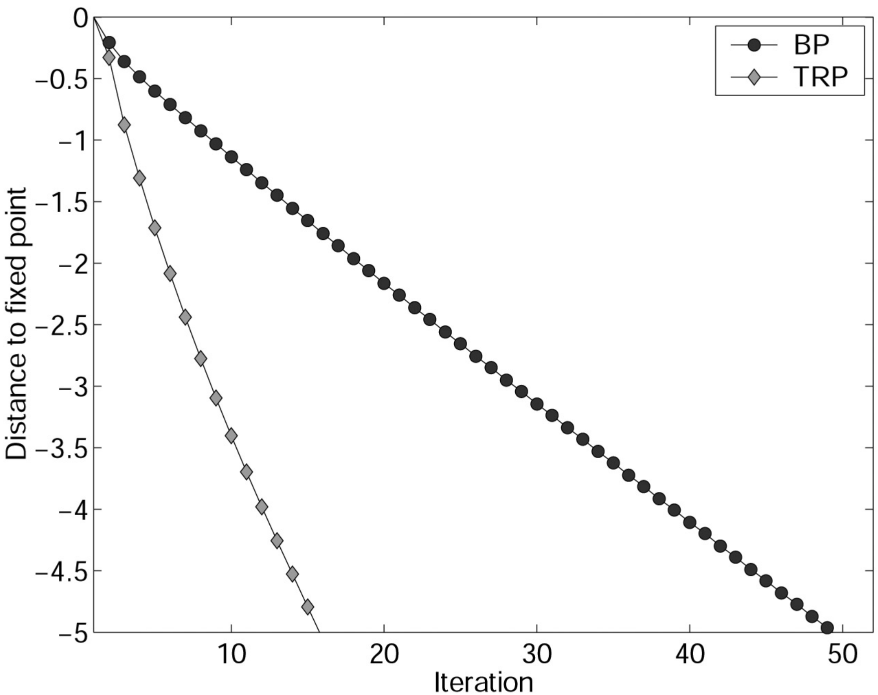


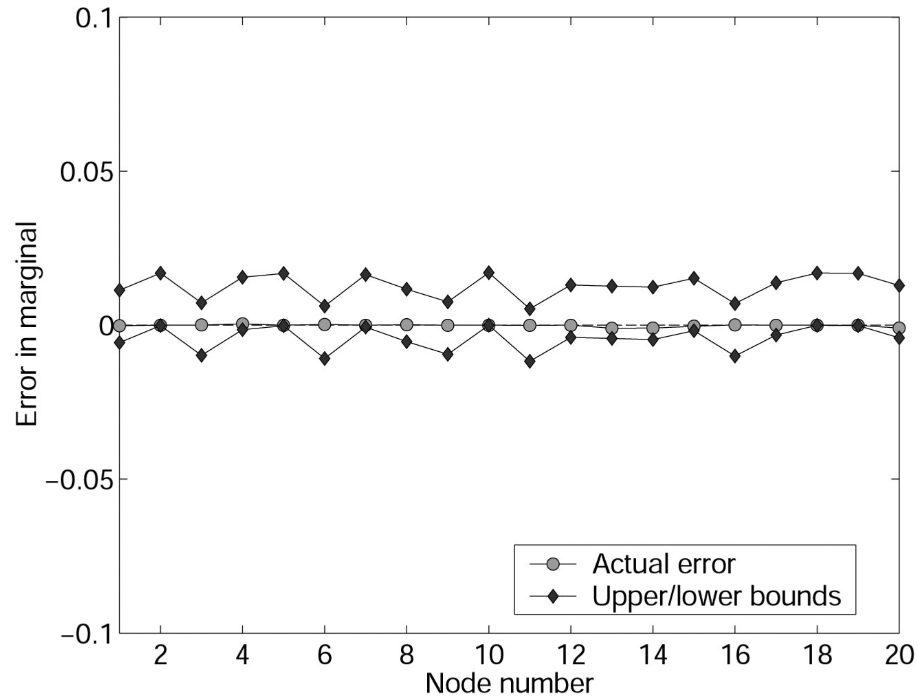
TRP: Recent Progress, Part I

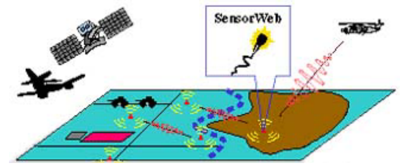
- Demonstration of superior performance in many cases (without optimizing choices of trees)
- Error characterization and bounds
  - The key is the TRP representation which allows error representation in terms of expectations over tree-distributions
  - Optimal Bounds: Weighting over all trees
    - There are *lots* of trees!
    - Convex analysis comes to the rescue



Sample TRP Estimation Results

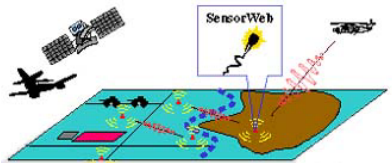






TRP: Recent Progress, Part II

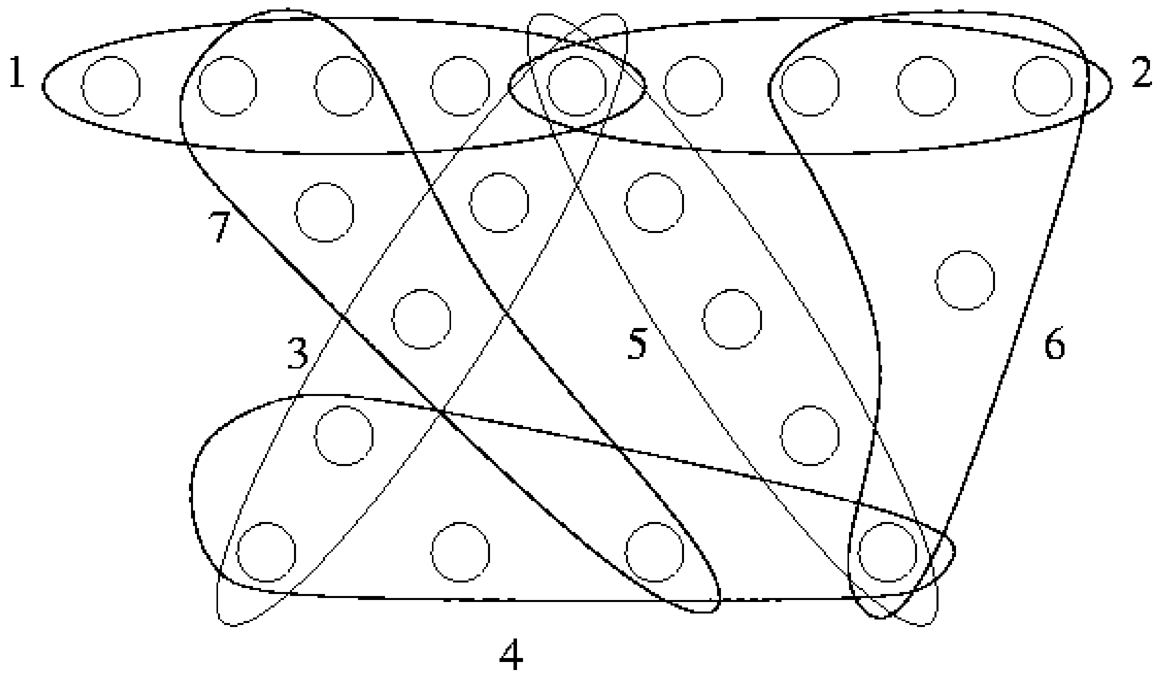
- TRP for optimization (rather than estimation)
  - Characterization of large classes of distributed algorithms: Rewriting global "cost" in terms of locally computable costs through message passing
    - Fixed point characterization
    - Clarifying when this works even in the acyclic case
    - Bounds
  - Use of "reweighting" concept to obtain algorithms that yield *optimal* solutions
    - Yields distributed optimal solution to the data association problem



Small Example

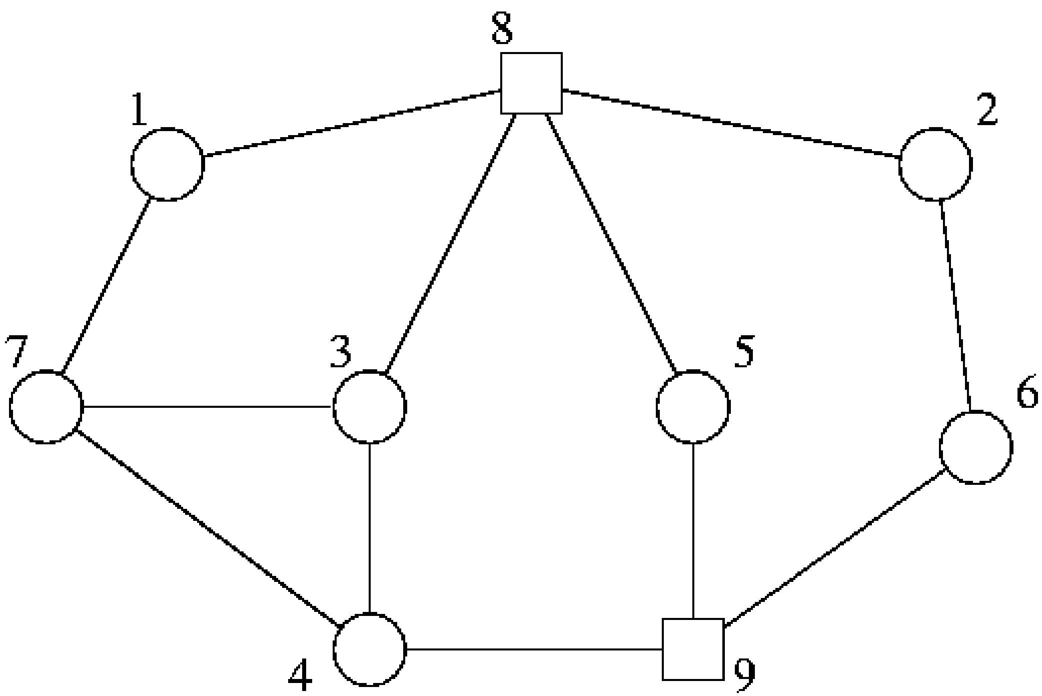
- 7 "sensors" (either all different sensors at same point in time or fewer sensors with measurements at multiple times)
- 21 targets
- Each "sensor" sees 5 targets
- Key issue: How organize hypotheses?
  - Target-centric? (best for centralized fusion)
  - Sensor-centric? (distributed)
  - Hybrid, driven by dynamic structure

Example structure

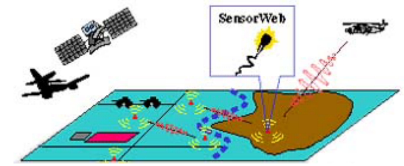


- Sensor-centric global hypothesis space is huge even for this problem

Hybrid Sensor-Target Representation



- Message passing algorithm yields distributed association solution *very* quickly and efficiently



Where to from here?

- Exploitation of framework for target tracking
  - Explicit (rather than implicit) representation of time, combining RCM and TRP
  - Incorporation of false, missed alarms, new objects
- Extension of optimization results to include both *querying and stopping* (as in sequential tests)
- Expanding the tradeoff space
  - Effect of local memory
  - Effect of nonlocal (or multi-hop) communications