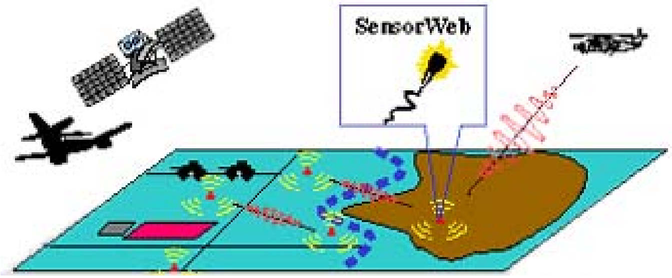


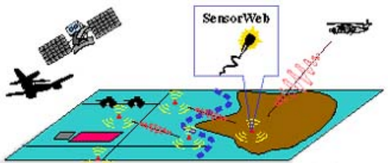
Stability and resource allocation

Tommi S. Jaakkola

MIT AI Lab

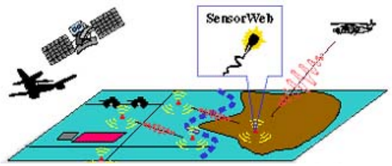
SensorWeb MURI Review Meeting

June 14, 2002



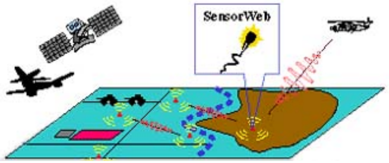
Research topics

- Inference in large sensor networks (with M. Wainwright and A. Willsky); IT 1, RCA 5
- Robust combination of information sources (with A. Corduneanu); IT 1&2, RCA 5 (& 6)
- Competitive estimation (with A. Corduneanu); IT 1&2, RCA 5
- Scalable information acquisition (with H. Siegelman); IT 2, RCA 4&5



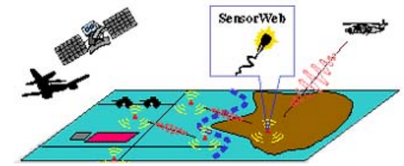
Outline of the talk

- Stability and source allocation
  - Robust combination of information from heterogeneous sources
  - Extension to competitive estimation (adversarial context)
- Resource allocation
  - Efficient acquisition of information through a limited information channel



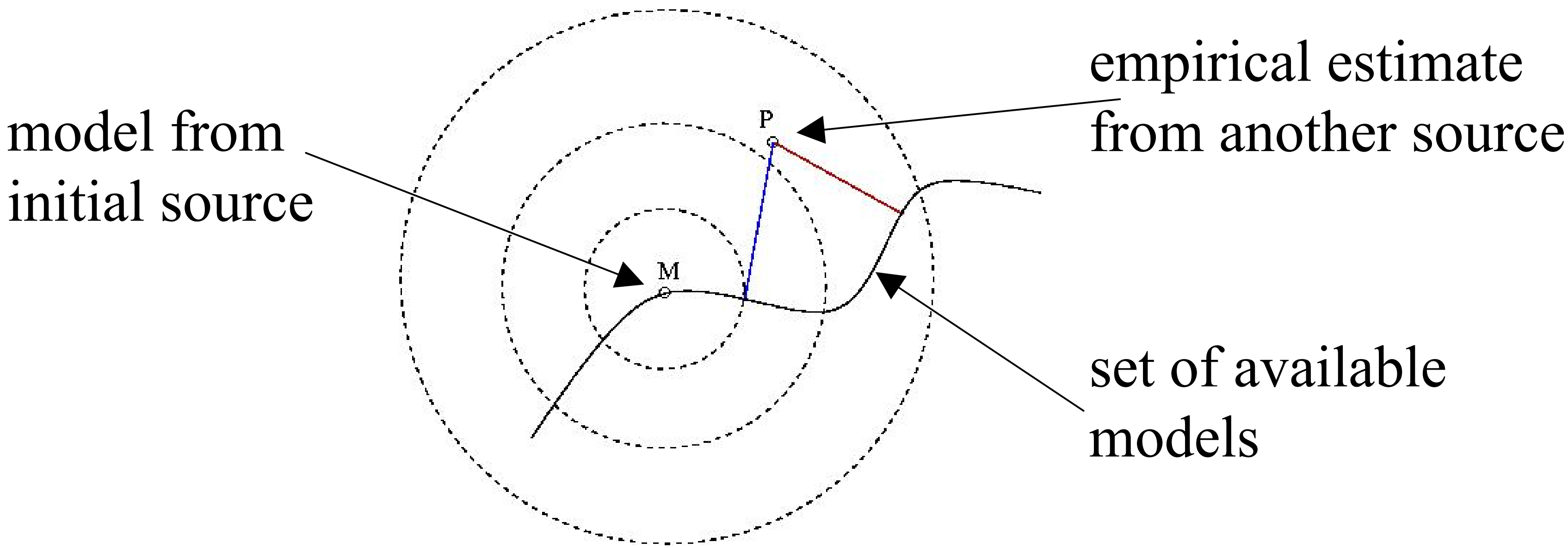
Part I: source allocation

- Heterogeneous sensors (e.g., acoustic and infrared) yield complementary views
  - How much do we rely on each source?
  - How do we resolve conflicts among the data sources?
  - How do we ensure that the estimation process remains stable?

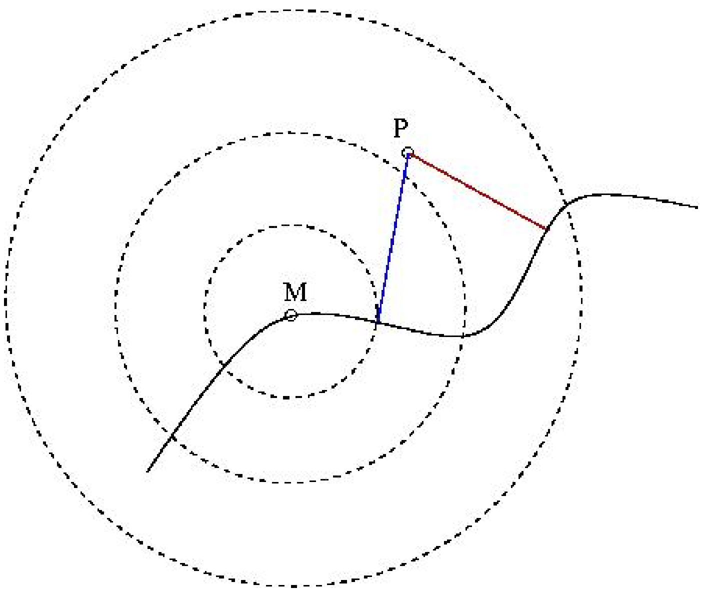


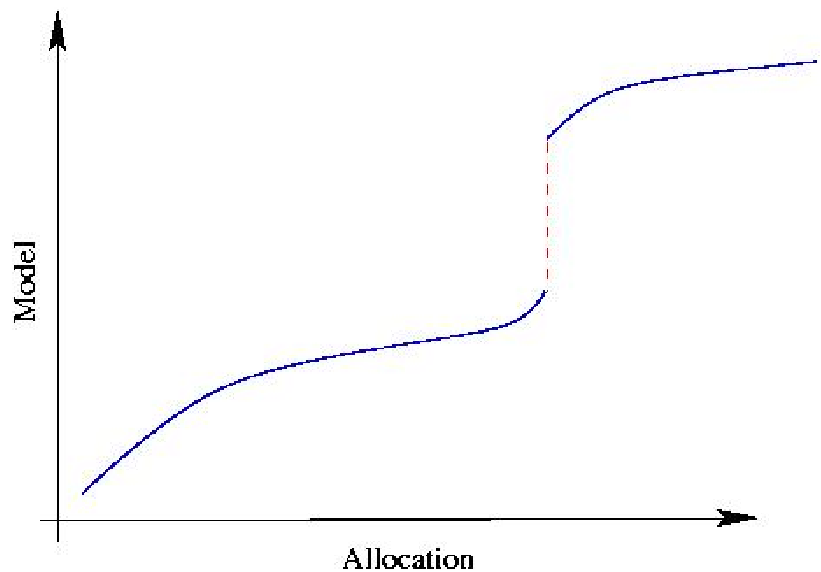
The problem

- Estimation with heterogeneous sources is inherently unstable

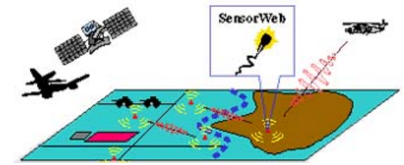


Critical points





The critical points appear almost surely as jumps, not as bifurcations



Example settings

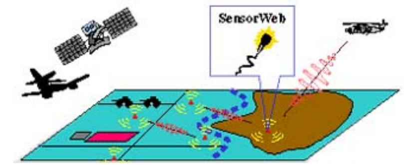
- The stability issue affects all estimation methods reducible to fixed point computations
- Example: estimation with incomplete data
  - Model $Q(x,y)$
  - Complete data log-likelihood: $D(\hat{P}_c(x,y)||Q(x,y))$
  - Incomplete data log-likelihood: $D(\hat{P}_I(x)||Q(x))$
  - Estimation criterion:

empirical estimates

$$J(Q,\lambda) = (1-\lambda)D(\hat{P}_c(x,y)||Q(x,y)) + \lambda D(\hat{P}_I(x)||Q(x))$$

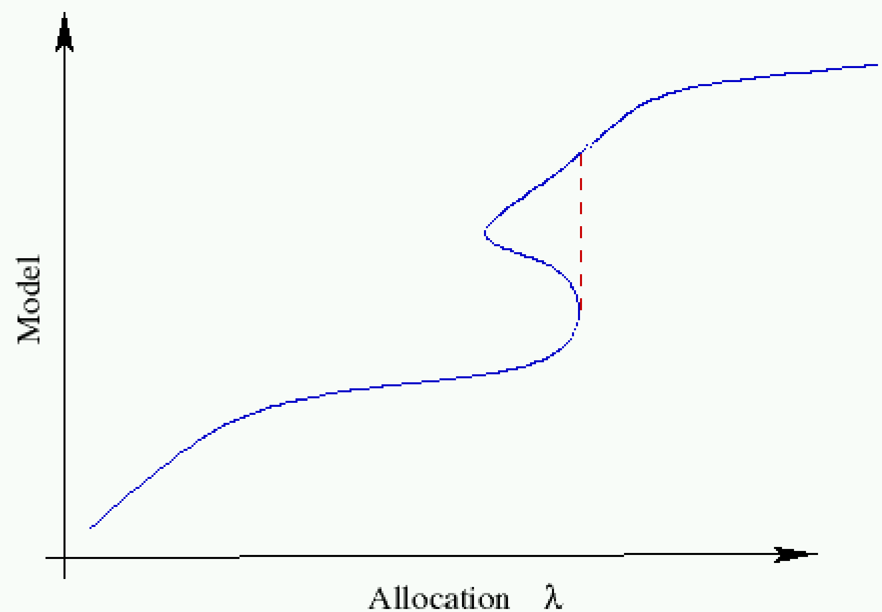
Fixed point equation: $\nabla_Q J(Q,\lambda) = 0$

allocation parameter

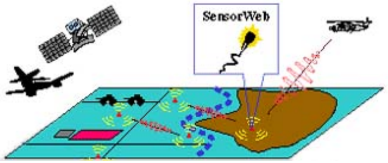


Stable identification of critical points

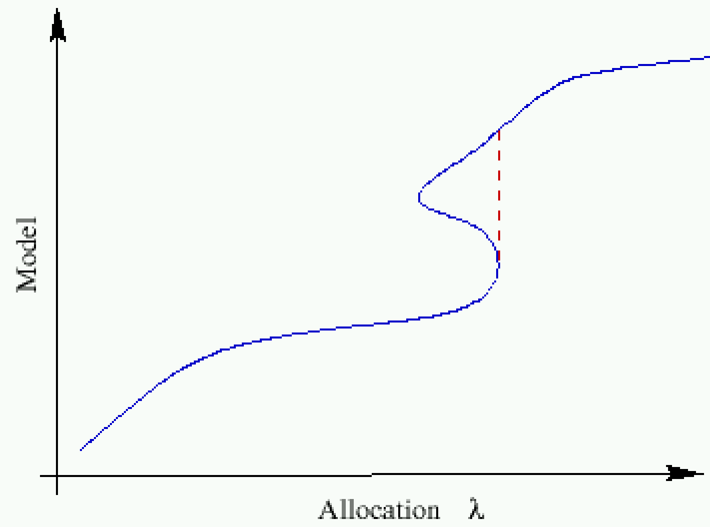
We can find a smooth curve in the joint space (Q, λ)



Provided that the Jacobian of $T(Q, \lambda) = \nabla_Q J(Q, \lambda)$ has full rank, $T(Q, \lambda) = 0$ defines a smooth 1-dim manifold in the joint space (Q, λ).



Homotopy continuation

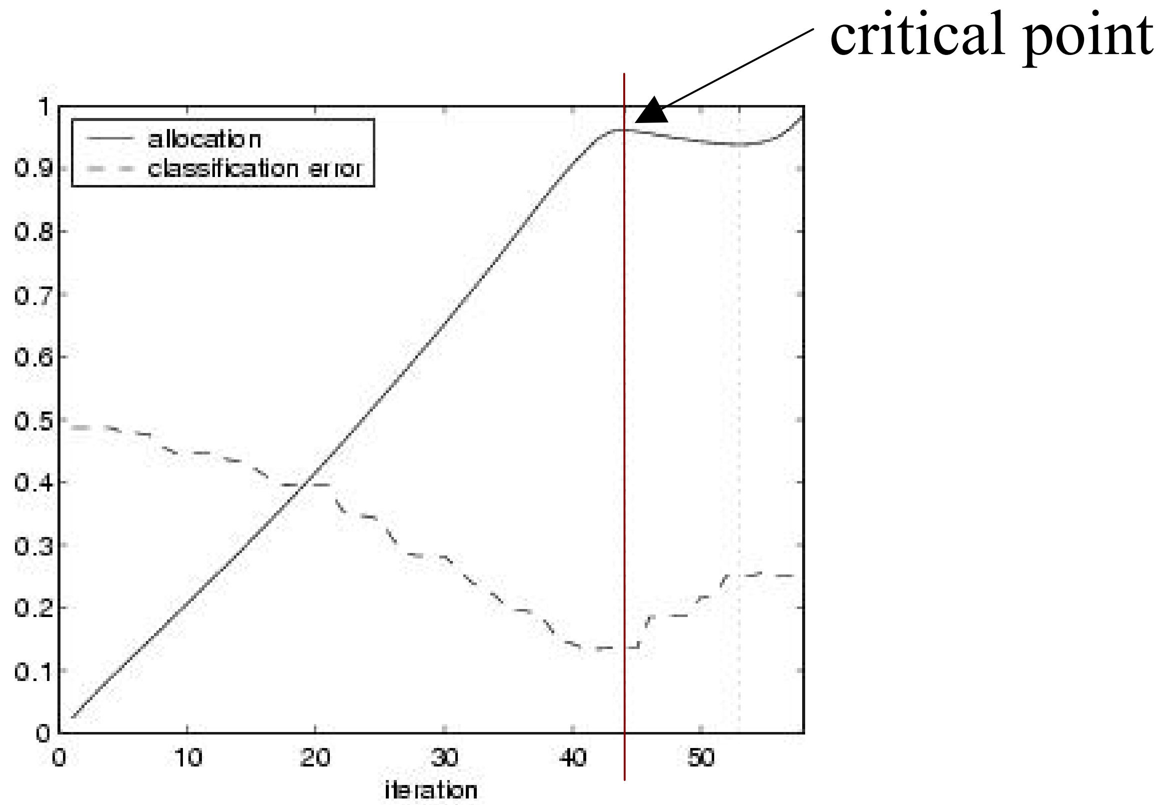


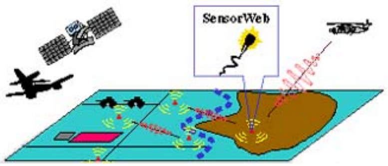
s parameterizes the curve

$$\left[\nabla_Q^2 J(Q,\lambda) \quad \frac{\partial}{\partial\lambda}\nabla_Q J(Q,\lambda)\right]\left[\begin{array}{c} dQ/ds \\ d\lambda/ds \end{array}\right] = 0$$

Each point along the curve necessarily satisfies the fixed point condition $\nabla_Q J(Q,\lambda) = 0$

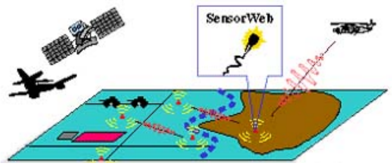
Typical results





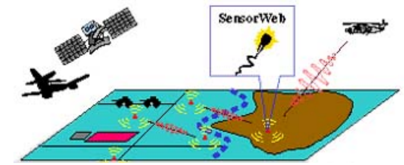
Summary of part I

- Data fusion is often unstable
- We can restore stability by identifying and avoiding critical points
  - homotopy continuation provides an efficient way of identifying stable data allocations
  - the methodology is applicable for most estimation settings



Extension: competitive estimation

- Estimation/decisions often have to be made in an adversarial context
- Robust decisions can be found with competitive (game theoretic) estimation



Competition, solution

- Two interpretations, two criteria

$$\max_R \min_Q \Big[\text{Loss}(Q,R) - \text{Loss}(Q) \Big]$$

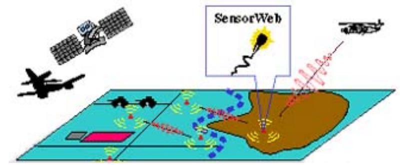
adversary

decision maker

loss resulting from interaction

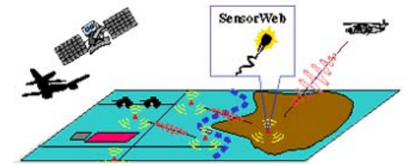
$$\min_Q \max_R \Big[\text{Loss}(Q,R) - \text{Loss}(R) \Big]$$

  - homotopy continuation applies as before
  - critical points arise as before (but can desirable in this context)

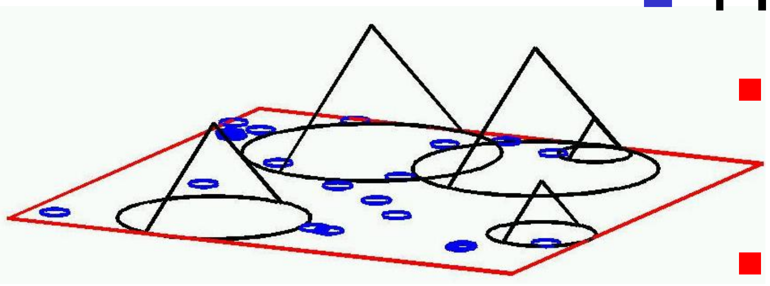


Part II: resource allocation

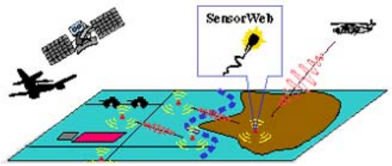
- The problem here is information acquisition (e.g., locating assets) with minimal resources
- The key question is how the available resources should be used/allocated
- Technical components:
  - sensor models
  - information channel
  - beliefs and inference (scalability)



The framework

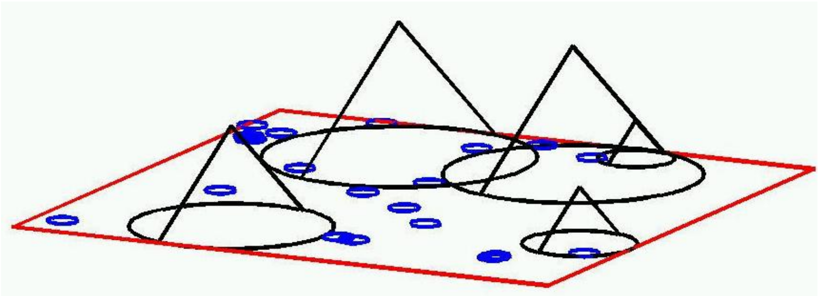


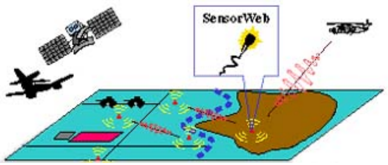
- Model:
  - Multi-resolution sensors
    - response model
  - Limited information channel
    - number of sensors that can be queried in parallel
  - Processing
    - maintaining beliefs
    - query optimization
- Key requirement: scalability



Sensors

- The sensors are assumed to appropriately cover the domain (identifiability)
- Characteristics of sensors (detectors)
  - static or dynamic definition
  - resolution, sensitivity
  - cumulative detection
- Sensor responses are captured by the detection probabilities $P(y = 1|r)$





Beliefs, expected response

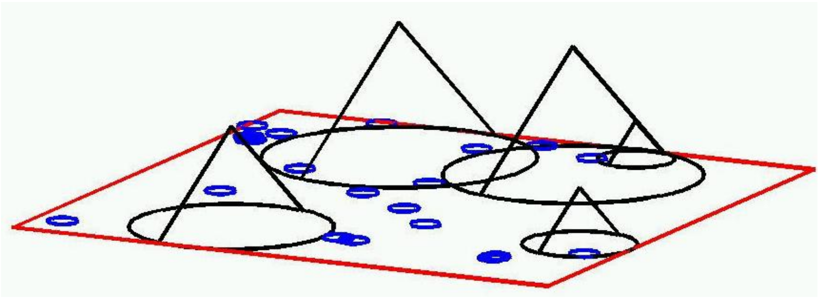
- We maintain a factored belief over elements/locations

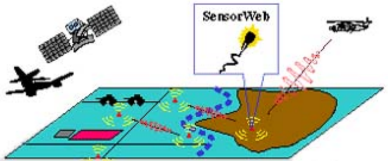
$$P(r|\theta) = \prod_{x \in \mathcal{X}} \theta_x^{r_x} (1 - \theta_x)^{1-r_x}$$

- The expected response from a sensor is given by

$$P(y_c = 1|\theta) = \sum_r P(y_c = 1|r) P(r|\theta)$$

where $y_c = 1$ signifies "detection"



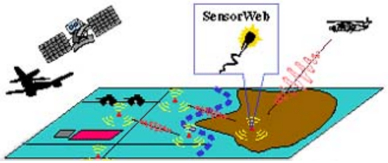


Maintaining beliefs

- We have to revise our beliefs (e.g., about asset locations) after each sensor response

  We project the posterior back into the factored beliefs

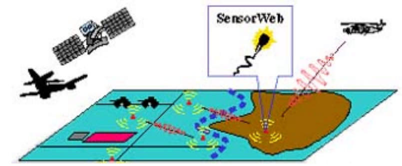
$$P(r;\theta') = \arg\min_{Q\in\mathcal{P}} D(P(r|\hat{y}_c,\theta)||Q(r))$$



Query optimization

- The expected information rate from a sensor often cannot be evaluated efficiently.
- We instead optimize a **lower bound**

$$I(y_c; r|\theta) \geq E_{y_c}\left\{\sum_{x \in c} D\left(\theta_{x;y_c} \,\|\, \theta_x\right)\right\} \stackrel{def}{=} I_p(y_c; r|\theta)$$



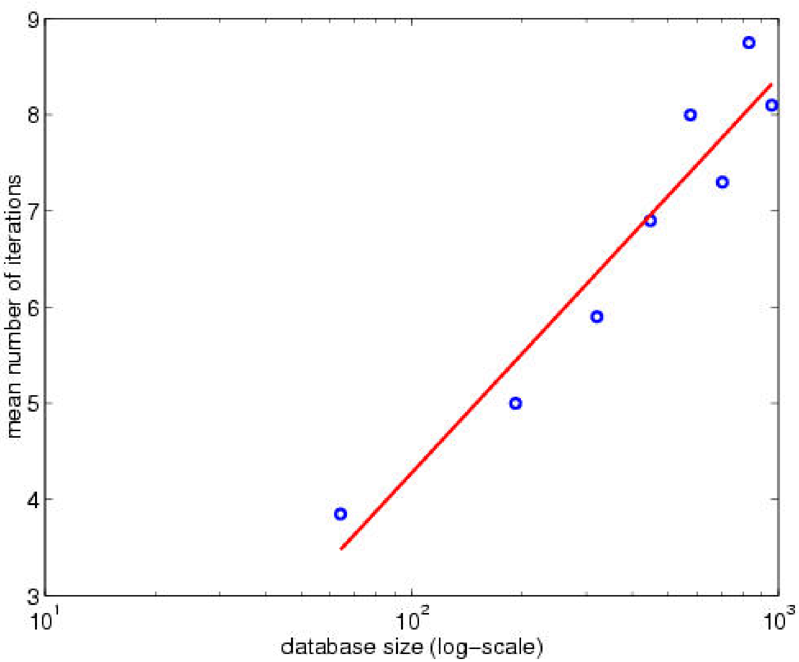
Query optimization cont'd

- To select k sensors for a query, we combine the lower bound with a series of conditional projections

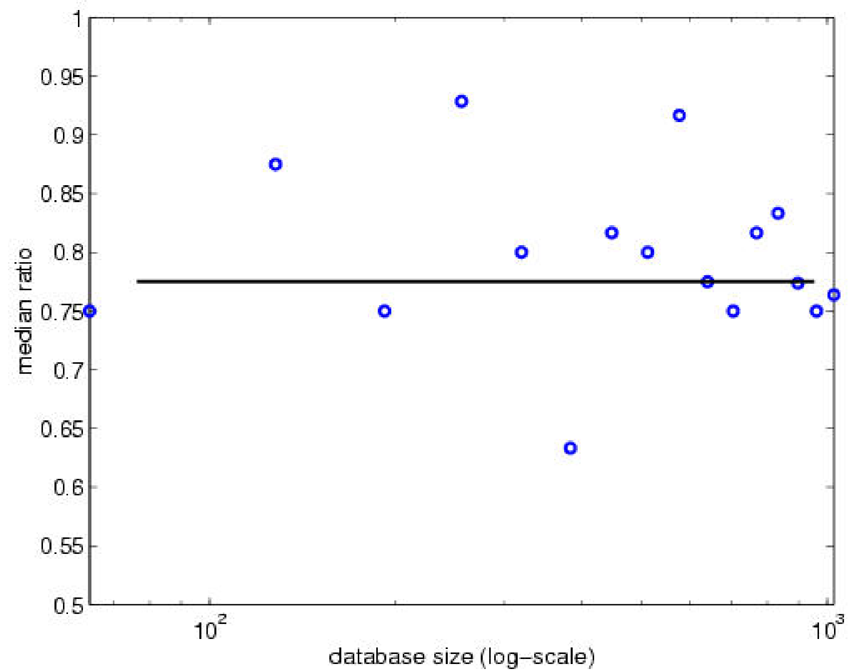
$$
\begin{array}{rcll}
 & & \textit{Optimization} & \textit{Conditional projection} \\
c_1 & = & \mathrm{argmax}_c\, I_p(y_c; r|\theta), & \\
c_2 & = & \mathrm{argmax}_c\, I_p(y_c, y_{c_1}; r|\theta), & \theta'_{x:y_{c_2}} = E\left\{ \theta_{x:y_{c_1}, y_{c_2}} \,\middle|\, y_2 \right\} \\
c_3 & = & \mathrm{argmax}_c\, I_p(y_c, y_{c_2}; r|\theta'), & \theta''_{x:y_{c_3}} = E'\left\{ \theta'_{x:y_{c_2}, y_{c_3}} \,\middle|\, y_{c_3} \right\} \\
 & \cdots & &
\end{array}
$$

$\Rightarrow$ k selections in time $\mathcal{O}(kn)$ (with cached reconstruction of non-additive components)

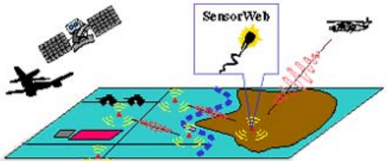
Example results



overall scaling of search



scaling with query size



Summary

- Information queries from a collection of sensors can be performed in a scalable manner
  - The algorithms scale linearly with domain/channel size
  - The sensors/detectors limited by “cumulative” detection
- Extensions:
  - Incorporation of specific sensors characteristics
  - Analysis and coordination of heterogeneous sensors