SAND2006-5463C

Sensor Placement to Satisfy Water Security and Operational Objectives

Bob Carr[1], *Jonathan Berry*[1], *William E. Hart*[1], *Regan Murray*[2], *Cindy Phillips*[1], *Jean-Paul Watson*[1]

[1]Discrete Algorithms and Mathematics Department, Sandia National Laboratories, Albuquerque, New Mexico

[2]NHSRC, Environmental Protection Agency, Cincinnati, Ohio

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94-AL-85000

Outline

- Modeling contaminant warning system (CWS) design as an optimization problem.

- Review of the SPOT sensor placement capabilities
  - A canonical problem formulation
  - Related problem formulations

CWS Design

Technical Goal: placement of sensors for the CWS within a budget

Developing an optimization model for sensor placement:

- How do sensors work?
- What is the design basis threat?
- How does a utility respond to detection events?
- What performance measures are used to evaluate the CWS?
- What are the potential sensor locations?

How do sensors work?

Two categories of sensors:

- Contaminant specific sensors
- Water quality sensors

Possible sensor modeling assumptions:

- Perfect sensors
  - No false positives or false negatives
  - Use a detection limit/threshold
- Simple imperfect sensors
  - Fixed false positive and false negative rates
  - These rates may depend on the detection threshold

What is the design basis threat?

Need to specify the nature of the contamination threats

- Used to compute the baseline risk of contamination
- Used to evaluate CWS sensor placement designs

Contaminant information:

- How it impacts human health and/or network infrastructure
- How contaminant flows (e.g. including decay)

Injection information:

- Time and location of contaminant events
- Duration of injection
- Injection volume and/or mass rate

How does a utility respond to detection events?

Utility response

- Timely warning of a potential water contamination incident
- Response actions to minimize public health impacts

Idea: treat utility response as a simple delay

- Sample collection time
- Laboratory analysis time
- Data transmission and analysis time
- Time for public order to be effective

Note: the public's response will likely be gradual, but we can select an acceptable threshold for compliance

What performance measures are used to evaluate the CWS?

Possible objectives:

- Minimize response time
- Minimize health impacts
- Minimize extent of contamination
- Minimize volume of water that enters the water network
- Minimize number of failed detections
- Minimize cost
- Minimize political risk…

Observations:

- Some objectives can be treated as constraints (e.g. costs)
- The EPA is most interested in health objectives

What are the potential sensor locations?

Location requirements:

- easy access
- electricity
- physical security
- data transmission capability
- sewage drains

Location categorizations

- All junctions in a water distribution model
- All junctions that correspond to utility owned nodes
- All junctions that correspond to public owned nodes

Further Modeling Considerations

What data is available for modeling?

- Many types of data can be difficult to obtain
- Balance model fidelity with required prediction accuracy
- Model fidelity includes an estimate of the data uncertainties!

There are many modeling trade-offs

- Need to make decisions about CWS given these trade-offs
- Trade-off analysis will require analysis/optimization of many related models
- There will generally be no single best sensor placement

Sensor Placement Optimization with SPOT

SPOT: Sensor Placement Optimization Toolkit

- Being developed at Sandia Labs
- Collaborators: UC, CU Denver, IBM, PNNL, UNM

Motivations:

- Scalable solvers for large-scale problems
  - 10,000s junctions and pipes
- Flexible solvers that can optimize many different objectives
- Flexible specification of performance constraints
- Fast solvers
- Methods for rigorously evaluating solver performance

A Canonical Problem Formulation

Minimize expected impact of contamination events

- Over a selected set of times and locations

Cost constraint

- Limit number of sensors
- Limit cost of installation

Note:

- This is the formulation considered by most of the literature
- This assumes the adversary has no knowledge of how attack time/location relates to impact of the attack

An Integer Programming Formulation

Variables:

- α – attack likelihood
- *w* – attack impact
- *b* – attack witness indicator
- *s* – sensor placement indicator

$$\text{minimize} \quad \sum_{a \in A} \sum_{i \in L} \alpha_a w_{ai} b_{ai}$$
$$s.t.$$
$$\sum_{i \in L} b_{ai} = 1 \qquad \forall a \in A$$
$$b_{ai} \le s_i \qquad \forall a \in A, i \in L$$
$$\sum_{i \in L} s_i \le S_{\max}$$
$$s_{ij} \in \{0,1\}$$

IP model:

- Objective is to minimize the expected impact of all attacks
- Very general formulation
  - Can capture different objectives/networks
- Can be solved with COTS software

Scalability Challenges

IP Size:

- *n* junctions (10,000's)
- *m* contamination times (100's)
- Up to *nm* contamination "locations"
- Up to n^2m contamination impact values

Observations:

- A 64-bit workstation is required to solve applications with 1000s of junctions and many contamination times

Need: ability to solve applications with 10,000s of junctions

Scalable Sensor Placement IPs

Observation: in practice, there are many junctions with the same nonzero impact values

- Flow simulations give the same value to junctions that are exposed at approximately the same time

Idea: reformulate the IP to only express each impact value once for each contamination event

Impact:

- 10x reduction in problem size
- Can solve problems with 10,000s junctions (on 64-bit machines)

Sensor Placement Heuristics

Observation: the IP formulation is closely related to the well-known p-median facility location problem

Idea: apply a standard p-median meta-heuristic to sensor placement

Impact:

- Can rapidly solve problems with 10,000s of junctions (minutes)
- Can solve problems 500x larger than competing methods
- Heuristic solutions are often optimal
- Can exactly bound the value of any given heuristic solution using an LP-relaxation of the IP formulation!

Summary For The Canonical Problem

Technical Capability

- Can optimize for many competing objectives
- Can quickly find near-optimal solutions for large-scale problems
- Can find exact optimal solutions for large-scale problems

Solver Technology

- Using commercial modeling tools: AMPL, CPLEX
- Using facility location heuristic methods

Impact: can solve large-scale, real-world sensor placement problems

Related Problem Formulations

Staged Placement

- Determine location for sensors given that some sensors have already been installed
- Staged installation allows for budget allocations over several years
- Can be easily accommodated within SPOT

Sensor Failures

- Can model sensor failures with false-positive and false-negative failures
- The IP model becomes very large (or nonlinear)
- Studies with heuristic solvers are encouraging

Related Problem Formulations (cont'd)

Aggregation of Sensor Detection

- For a given contamination event, can treat nearly-identical impact values as the same
- Enables significant reduction of IP model size
- Can solve exactly solve this related IP, but the final solution is provably near-optimal

Multi-objective Optimization

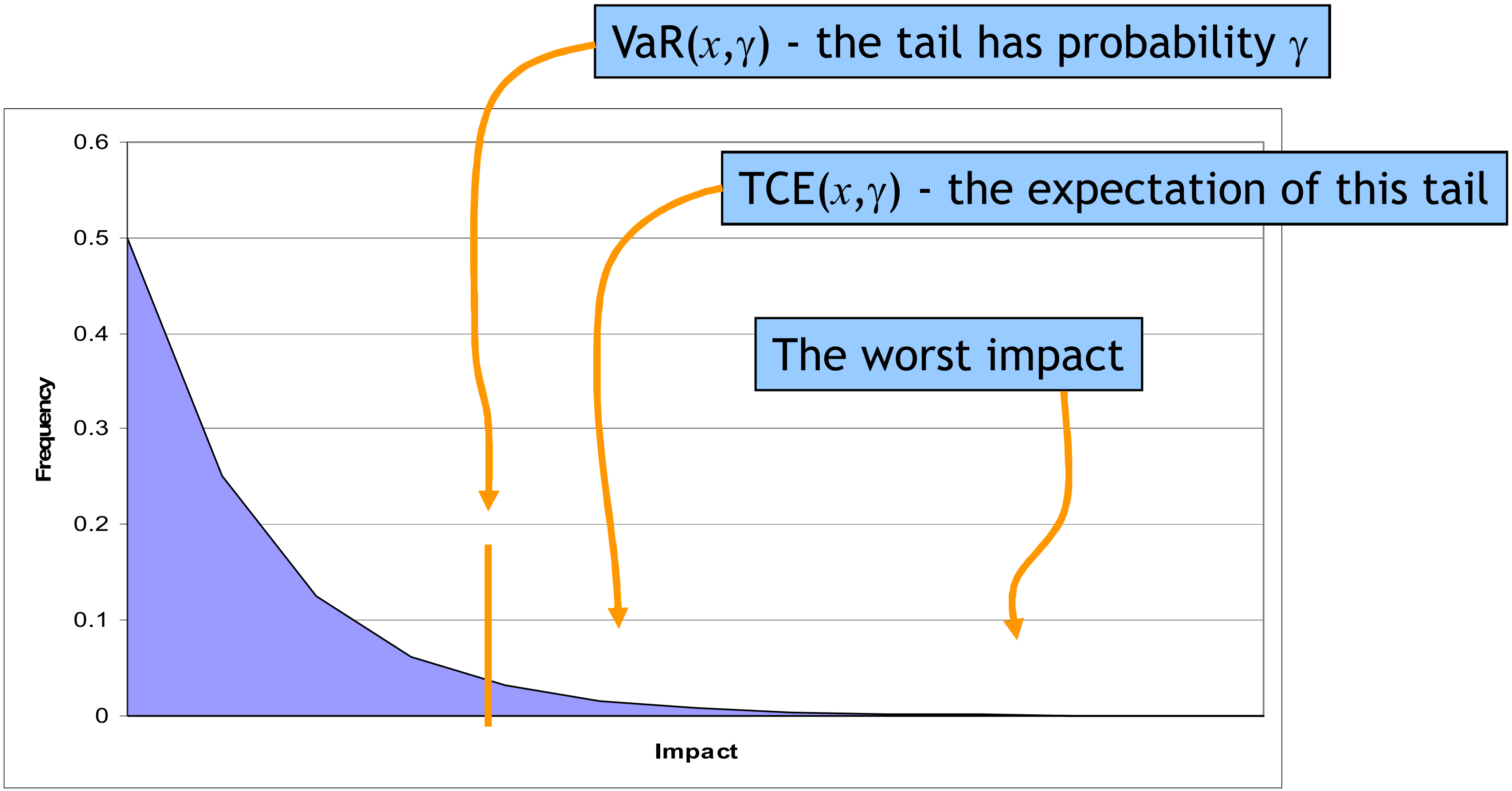
- SPOT can use side-constraints to constraint secondary objectives
- Solver performance is poor with more than one constraint
- General multi-objective methods could be applied to this type of problem

Related Problem Formulations (cont'd)

Minimize Worst-Case Performance
- Mathematically equivalent to the p-mean problem
- Difficult to solve with IP or heuristic solvers

Robustness Formulations
- Can handle data uncertainties with min-max formulations
  - Are solvable as IPs in some cases, but they are harder
- Can handle uncertainty in contamination location/time explicitly
  - Risk measures: Value at Risk, Tail-conditioned expectation
  - Solving these formulations is still difficult
  - Can reduce worst-case impacts while maintaining good expected-case performance

Examples of Risk Measures



Final Thoughts

There are many modeling issues that have not been fully addressed
- Sensor performance/uncertainties
- Data uncertainties
- Multi-contaminant formulations

Focusing on scalability is very important for sensor placement
- Scalability has significantly influenced SPOT development
- We need a library of large-scale instances

There are “hidden” costs that need to be considered
- E.g. a large number of water scenarios may need to be analyzed
- TEVA leverages parallel computing to minimize this cost

Additional Reading

- R. Carr, H. J. Greenberg, W. E. Hart, G. Konjevod, E. Lauer, H. Lin, T. Morrison and C. A. Phillips, Robust optimization of contaminant sensor placement for community water systems, Mathematical Programming Series B, 107, pp. 337-356, 2006.
- J. Berry, W. E. Hart, C. E. Phillips, J. G. Uber and J.-P. Watson, Sensor placement in municipal water networks with temporal integer programming models, J. Water Resources Planning and Management, 132 (4), pp. 218-224, 2006.
- J. Berry, L. Fleischer, W. E. Hart, C. A. Phillips and J.-P. Watson, Sensor Placement in Municipal Water Networks, J. Water Planning and Resources Management, 131 (3), pp. 237-243, 2005.
- J. Berry, W. E. Hart, C. A. Phillips, J. G. Uber and T. M. Walski, Water quality sensor placement in water networks with budget constraints, in Proc. World Water and Environment Resources Conference, 2005.
- J.-P. Watson, W. E. Hart and R. Murray, Formulation and optimization of robust sensor placement problems for contaminant warning systems, in Proc. Water Distribution System Symposium, 2006.
- J. Berry, R. D. Carr, W. E. Hart, V. J. Leung, C. A. Phillips and J.-P. Watson, On the placement of imperfect sensors in municipal water networks, in Proc. Water Distribution System Symposium, 2006.