SAND2010-6835C

Sandia National Laboratories ICS 2011 Abstracts

This document contains all the abstracts (not submitted under independent R&A) involving Sandia National Laboratories authors to be submitted to the INFORMS Computing Society 2011 conference. Each abstract is numbered with title, authors, and abstract as it will likely appear in a conference program if accepted.

Sandia is a multipurpose laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

1. Title: PySP: Modeling and Solving Stochastic Programs in Python

Authors: William Hart (01415), Jean-Paul Watson (01412), David Woodruff (U.C. Davis)

Abstract: Although stochastic programming is a powerful tool for modeling decision-making under uncertainty,various impediments have historically prevented its wide\-spread use. One key factor involves the ability of non-specialists to easily express stochastic programming problems as extensions of deterministic models, which are often formulated first. A second key factor relates to the difficulty of solving stochastic programming models, particularly the general mixed-integer, multi-stage case. Intricate, configurable, and parallel decompositionstrategies are frequently required to achieve tractable run-times. We simultaneously address both of these factors in our PySP software package, which is part of the COIN-OR Coopr open-source Python project for optimization. To formulate a stochastic program in PySP, the user specifies both the deterministic base model and the scenario tree with associated uncertain parameters in the Pyomo open-source algebraic modeling language. Given these two models, PySP provides two paths for solution of the corresponding stochastic program. The first alternative involves writing the extensive form and invoking a standard deterministic (mixed-integer) solver. For more complex stochastic programs, we provide an implementation of Rockafellar and Wets' Progressive Hedging algorithm. Our particular focus is on the use of Progressive Hedging as an effective heuristic for approximating general multi-stage, mixed-integer stochastic programs. By leveraging the combination of a high-level programming language (Python) and the embedding of the base deterministic model in that language (Pyomo), we are able to provide completely generic and highly configurable solver implementations. PySP has been used by a number of research groups, including our own, to rapidly prototype and solve difficult stochastic programming problems.

2. Confidence Intervals and Solution Quality Estimation for Sensor Network Design

Authors: Jean-Paul Watson (01412), David Woodruff (U.C. Davis), Tom Brounstein (05641), William Hart (01415) and Regan Murray (United States Environmental Protection Agency)

Abstract: The problem of placing a limited number of sensors to detect contamination incidents and maximize public health protection is well-studied in the water distribution systems analysis literature. Typically, some measure of impact is minimized over a number of hypothetical contamination events. To date, optimization formulation and algorithm design have been the primary focus of such research. A critical but ignored issue is the stochastic nature of the resulting optimization problem, and the impact of scenario uncertainty on our confidence in the quality of the final sensor configuration. A stochastic formulation and solution of the problem would provide water utilities and decision makers with confidence intervals around the predicted performance of the sensor network, and would provide more information on which to base expensive investment decisions that could affect public health. In this talk, we discuss quantitative mechanisms for computing confidence intervals on the solution quality of a sensor configuration, focusing on the multiple replication procedure introduced by Mak, Morton, and Wood. We provide experimental results for several networks, considering minimization of both the expected overall impact and the impact associated with high-impact events. Our results indicate that larger-than-expected numbers of scenario samples are required to achieve even reasonably tight (e.g., 5%) confidence intervals on solution quality in the expected impact case, and confidence intervals are significantly wider when mitigating against only high-impact events. We conclude by discussing experiments with solution stability under uncertainty, in an effort to determine if wide confidence intervals are an artifact of a conservative statistical estimation procedure, or would impact deployment.

3. Source Inversion and Optimal Sampling in Large-Scale Drinking Water Distribution Systems

Authors: Angelica Wong (Texas A&M University), Sean McKenna (06731), William Hart (01415), Carl D. Laird (Texas A&M University)

Abstract: Protecting the nations drinking water networks from accidental or intentional contamination requires effective real-time tools for detection, source determination, and response planning. Given existing technology and limited resources, real-time response systems must rely on measurements obtained by manual samples at sparse locations in time and space. Existing event detection systems, like CANARY (Sandia National Laboratories), can provide discrete positive/negative measurement data to identify the presences of contaminant. We present an integrated real-time iterative sampling strategy for finding the contamination source using discrete measurements from sparse manual grab samples taken at different sampling cycles. The inversion formulation is a mixed integer linear program formulation (MILP) that identifies a set of potential contamination sources. Additional manual grab samples are required to sufficiently narrow down the set of potential contamination sources. Optimal sampling locations are determined using an MILP formulation that maximizes the number of distinguishable pairs of events. This approach is tested on a water network model comprised of over 10,000 nodes and more than 150 time steps giving positive results.

4. Title: FORMULATING AND ANALYZING MULTI-STAGE SENSOR PLACEMENT PROBLEMS

Authors: William Hart (01415), Regan Murray (United States Environmental Protection Agency), Jean-Paul Watson (01412) and David Woodruff (U.C. Davis)

Abstract: The optimization of sensor placements is a key aspect of the design of contaminant warning systems for automatically detecting contaminants in water distribution systems. Although researchers have generally assumed that all sensors are placed at the same time, in practice sensor networks will likely grow and evolve over time. For example, limitations for a water utility's budget may dictate an

staged, incremental deployment of sensors over may years. We describe optimization formulations of multi-stage sensor placement problems. The objective of these formulations includes an explicit trade-off between the value of the initially deployed and final sensor networks. This trade-off motivates the deployment of sensors in initial stages of the deployment schedule, even though these choices typically lead to a solution that is suboptimal when compared to placing all sensors at once. These multi-stage sensor placement problems can be represented as mixed-integer programs, and we illustrate the impact of this trade-off using standard commercial solvers. We also describe a multi-stage formulation that models budget uncertainty, expressed as a tree of potential budget scenarios through time. Budget uncertainty is used to assess and hedge against risks due to a potentially incomplete deployment of a planned sensor network. This formulation is a multi-stage stochastic mixed-integer program, which are notoriously difficult to solve. We apply standard commercial solvers to small-scale test problems, enabling us to effectively analyze multi-stage sensor placement problems subject to budget uncertainties, and assess the impact of accounting for such uncertainty relative to a deterministic multi-stage model

5. Title: Recent Developments in Coopr: A COmmon Optimization Python Repository

Authors: William Hart (01415), Carl Laird (Texas A&M University), Jean-Paul Watson (01412), David Woodruff (U.C. Davis)

Abstract: Coopr is a collection of Python optimization-related packages that supports a diverse set of optimization capabilities for formulating and analyzing optimization models. This talk will describe recent updates to Coopr. This includes new modeling capabilities for generalized disjunctive programming and other nonlinear models, new solver capabilities like Benders decomposition, and a simple GUI for creating and solving Pyomo models.

6. Methodology to Optimally Select Flushing Locations in a Water Distribution System

Authors: Terra Haxton (US Environmental Protection Agency), Regan Murray (US Environmental Protection Agency), David Hart (06731), Robert Janke (US Environmental Protection Agency)

Abstract:

Drinking water security is an important area of research for the U. S. EPA, and work is ongoing in the areas of prevention, detection, mitigation, and decontamination. Upon determination of a possible contamination threat in a drinking water distribution network, a variety of response actions (e.g., public notification and operational changes) can be pursued in order to minimize public health and economic impacts and ultimately return the utility to normal operations. Flushing is a relatively common operational response option employed by utilities to address water quality concerns. The effectiveness of flushing alternatives can be evaluated by modeling water distribution system dynamics. This paper will present a modeling and simulation approach in conjunction with optimization software to identify beneficial flushing response options that a utility could pursue to help reduce the spread of contamination following an incident. Assuming a network of online, continuous, spatially-distributed sensors, an optimal hydraulic response tool supplied with utility-specific flushing characteristics can be utilized to identify the best hydrant locations to flush. Different impact measures will be employed to evaluate the most effective flushing strategies.

7. Sensor Placement in Water Distribution Systems with Multi-Species Water Quality Models

Authors: Regan Murray (US Environmental Protection Agency), Robert Janke (US Environmental Protection Agency), Terra Haxton (US Environmental Protection Agency), William Hart (01415), Tom Taxon (Argonne National Laboratory)

Abstract: Sensor placements for municipal water networks are evaluated based upon how well they might detect a given suite of potential contamination events. The impact of each event depends upon the concentration of contaminants through the network over time. This determines both where and when contaminant might cause damage to the network and to the population, and when and where the contamination can be detected. Initial studies relied on EPANET, the industry-standard hydraulic and water quality modeling software package. EPANET simulates only a single reactive constituent such as chlorine. Thus most studies assume contaminants act like a "tracer" passing through the distribution system without interacting with other chemicals or with the biofilms on pipes.

EPANET-MSX is a new extension to the EPANET simulation tool recently released to the public. It can handle multiple types of reacting material by solving a set of differential algebraic equations. EPANET-MSX requires mathematical models and parameter values for simulating specific reactions. Researchers have initial models for contaminant adsorption, inactivation, and attachment to biofilms. This talk discusses the impact of multi-species simulation on sensor placement. The simlator affects impact measures, thus the objective function parameters, and ultimately sensor placement. Using two real-world networks, we estimate the impact of multi-species simulations.

8. Title: The Max-Flow Network Inhibition Problem
Author: Cynthia Phillips (01412)
Abstract: In the max-flow network inhibition problem, we wish to expend a limited budget maximally reducing the transporation capacity between two specific nodes in a network. That is, each edge has a capacity and a destriction cost. We wish to (partially) remove a set of edges to maximially reduce the resulting maximum s-t flow. This talk will review some of the classic complexity results for this problem, and describe the algebraic and geometric intuition behind a bicriteria approximation algorithm for this problem (also somewhat classic). The approximation follows directly and simply from a linear programming relaxation. The strategy returned by this algorithm will either be superoptimal but somewhat overbudget, or within budget, but somewhat supoptimal. We will summarize more recent experience and computational experiments.

9. Computational Challenges for Designing Contamination Warning Systems: Case Studies and Options for Implementation

Authors: Robert Janke (US Environmental Protection Agency), Katherine Klise (06731), Cynthia Phillips (01412)

Abstract: Sensors in water distribution systems (WDS) can detect accidental or intentional contamination. Evaluating the consequence of contamination to WDS networks has been a research focus area within the federal government, namely EPA, in recent years. The Environmental Protection Agency, Sandia National Laboratories, and Argonne National Laboratory have developed the Threat Ensemble Vulnerability Analysis and Sensor Placement Optimization Tool (TEVA-SPOT) (U.S. EPA, 2010) to evaluate the vulnerabilities of distribution systems to contamination events and determine where best

to place sensor monitoring stations. Using an ensemble of contamination events, TEVA-SPOT evaluates the contaminant consequence to the network by calculating an impact measure for each event. These impacts contribute to the objective function for optimizing sensor placement. A WDS model for a large municipality can contain many thousands of nodes and interconnecting pipes. As a result, evaluating large numbers of contamination events in such a WDS and then designing a sensor network can be computationally challenging, requiring a powerful workstation. This presentation will provide the latest research findings on ways to minimize the computational difficulties while still fully evaluating vulnerabilities and ensuring the best sensor network design. Case study examples will be provided along with options for easier implementation.

10. Online Event Detection from Water Quality Data

Authors: David B. Hart (06731), Sean A. McKenna (06731), Terra Haxton (US Environmental Protection Agency), Regan Murray (US Environmental Protection Agency)

Abstract: Rapid and accurate detection of contamination incidents in drinking water is critical for notifying consumers of threats and risks to public health and for making remediation and recovery decisions. CANARY is a software package that performs on-line, multivariate, event detection from networked sensor data. Employing statistical forecasting and classification algorithms, CANARY continuously analyzes time series signals for anomalous conditions and reports the probability of an event. Noisy data are adaptively filtered to accurately identify anomalous events while minimizing false positive detections. Data aggregation over multiple time steps is accomplished through a binomial event discriminator (BED) algorithm that processes the number of outlier time steps against the expected number under background conditions. Parameterization of the BED and other parameters controlling the history window and the threshold to identify outliers at each time step allows for the tradeoff between sensitivity and specificity to be examined. Simulated water quality events are superimposed on measured background water quality signals to examine this tradeoff. Additionally, the delay to detection from the onset of an event is examined. Results of this tradeoff and of the delay calculations are presented.

11. Sensor Placement with Sensitivity Tuning for Municipal Water Networks

Authors: David B. Hart (06731), William E. Hart (0415), Sean McKenna (06731), Regan Murray (US Environmental Protection Agency), Cynthia Phillips (01412)

Abstract: We consider the problem of placing sensors in a municipal water network when we can choose both the location of sensors and the sensitivity and specificity of the contamination warning system. Sensor stations in a municipal water distribution network continuously send sensor output information to a centralized computing facility, and event detection systems at the control center determine when to signal an anomaly worthy of response. Although most sensor placement research has assumed perfect anomaly detection, signal analysis software has parameters that control the tradeoff between false alarms and false negatives. We describe a nonlinear sensor placement formulation, which we heuristically optimize with a linear approximation that can be solved as a mixed-integer linear program. We report the results of initial experiments on a real network and discuss tradeoffs between early detection of contamination incidents, and control of false alarms.