FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Resilient Sensor Placement for Faults and Attacks in Water Distribution Networks

W. Abbas[1], L. Perelman[2], S. Amin[3] X. Koutsoukos[1],

[1]Vanderbilt University
[2]UT Austin
[3]MIT

Berkeley UNIVERSITY OF CALIFORNIA
Massachusetts Institute of Technology
NSF
MICHIGAN
VANDERBILT UNIVERSITY

Motivation

Motivation:

* Leakages in urban water networks can cause huge economic losses, and health risks.
* Sensors for network monitoring are vulnerable to cyber attacks

Objective:

Sensor placement for an efficient and **resilient localization** of pipe failures in the presence of faults and attacks.

Challenges:

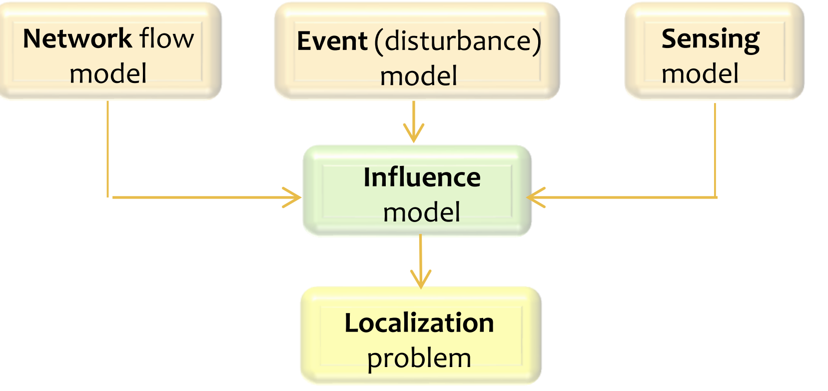
Pipe failure uncertainty, budget constraints, sensor errors, event detection and localization.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Contributions

Previous Work:

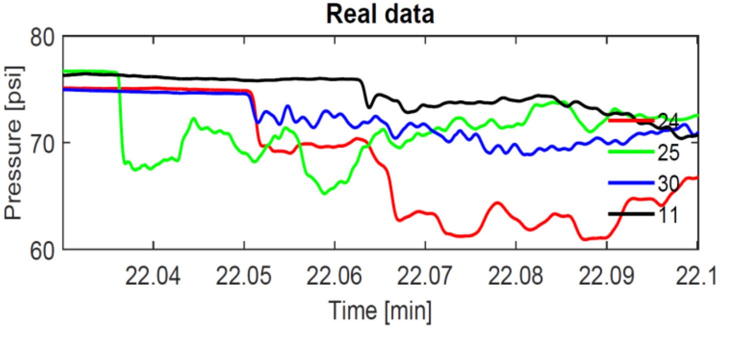
* Using combinatorial optimization, an efficient sensor placement algorithm for the localization of failure events in water networks. *(Automatica 2016)*
* Improved localization through multi-level sensing *(ACM BuildSys 2015)*



Current

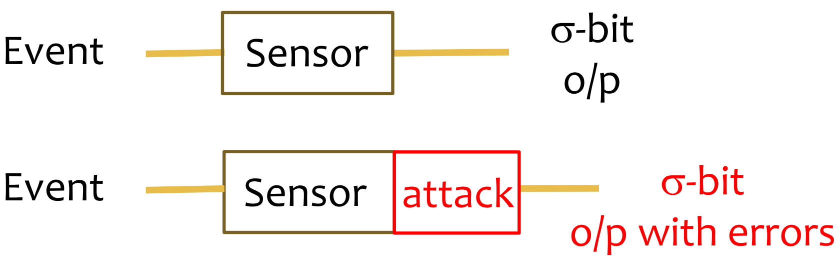
* Localization in the presence of sensor errors (faults and attacks)

W. Abbas, L. Perelman, S. Amin, and X. Koutsoukos, "Resilient Sensor Placement for Fault Localization in Water Distribution Networks" ***(ACM/IEEE ICCPS 2017)***



Attacks and Faults Resulting in Sensor Errors

Single sensor error



Error (attack): One or more of the output bits are flipped.

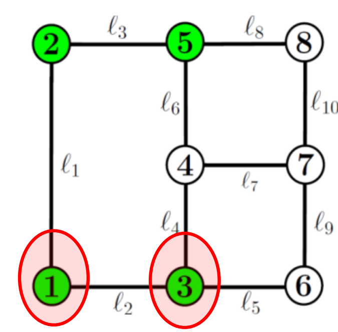
Error Sources: Sensor degradation, Cyber attacks

- At most e sensors can give incorrect outputs.
- The error model can be used to model a class of attacks in which an attacker takes control of at most e sensors and changes their output in any possible way.

Multiple sensors errors

Given a set of m sensors, at most e of them can give incorrect outputs for an event.

Example:



e = 2

Sensors	S_1	S_2	S_3	S_5
Correct o/p	0 1	1 0	0 0	0 0
Possible o/p with 2 errors	1 0	1 0	1 1	0 0
	…	…	…	…

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Problems

Problem 1: (Resilient Sensor Placement)
How to place **m** sensors, each with a σ-bit output, to maximize the number of events that can be localized accurately, even if **e** of the deployed sensors give errors? At the same time, how can we evaluate such a sensor placement in water distribution networks?

	Explanation
m	No. of sensors
σ	No. of output bits (features)
e	No. of erroneous sensors
L	Localization performance

Problem 2: (Tradeoffs Between System Variables)
What is the trade-off between **m, e, σ**, and the localization performance in the context of sensor placement for fault localization. In particular, fixing any two variables, what is the relationship between the remaining two?

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

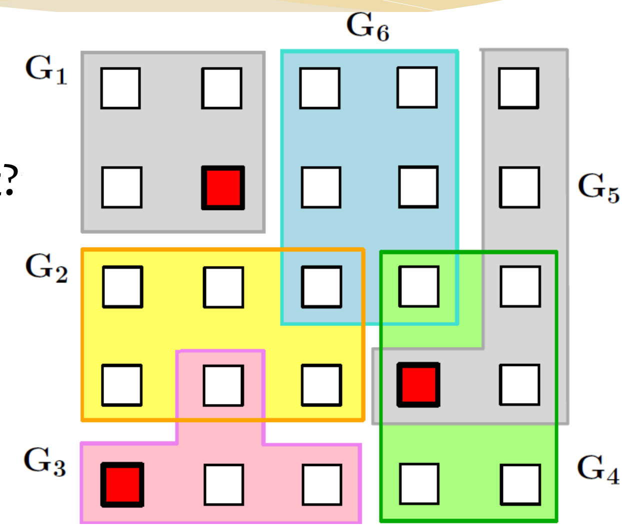
Non-adaptive Group Testing (NAGT) and Resilient Sensor Placement

Group Testing (GT):

- Set of elements with some defective ones.
- Elements are divided into groups.
- Questions are asked, "If G_i contains a defective element?
- Answers are either "yes" or "no".

Non-adaptive GT: All groups are made a priori.

Non-adaptive GT with Unreliable Tests: Some questions may be answered incorrectly.



Objective:

Design groups (queries) so that defective elements could be identified by asking the minimum number of queries.

Results:

Various results are known, e.g., at least $O((d^2/\log d)\log n)$ queries are always required, and there exist schemes that can achieve NAGT in $O(d^2 \log n)$ queries, e.g., *(Macula 1997, Porat and Rothschild 2011, Mazumdar and Mohajer 2014)*

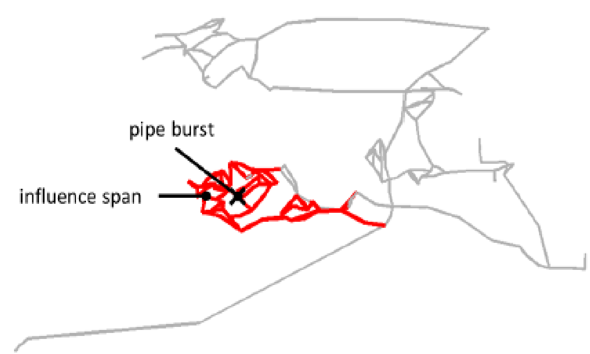
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

NAGT and Resilient Sensor Placement- Impact of physical dynamics

NAGT		Resilient Sensor Placement
Elements	→	Pipes
Defective elements		Pipes with bursts and leakages
Groups		Sensors
Tests (queries)		Sensors outputs
Unreliable tests		Sensors outputs with errors

However, there is a major difference.

- Typically in NAGT, any set of elements can be grouped together to make a test.
- In sensor placement, groups (tests) are coming from the physical system, i.e., any set of pipes cannot be grouped together.



FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Localization with Sensor Errors

Events	S_1	S_2	S_3	S_4	…	…	S_{total}
Pipe 1	0 0 0 1	1 0 0 1	1 0 1 0	0 1 0 1	…	…	1 1 1 0

Events	S_1	S_4	…	S_m
Pipe 1	0 0 0 1	0 1 0 1	…	0 1 0 1

← Signature of pipe 1 failure.

At most e sensors can be erroneous.

Two events can be distinguished in the presence of at most **e erroneous sensors** as long as the hamming distance between their signatures is at least **(2e + 1).**

Therefore, select sensors such that pairs of signatures that are at least (2e+1) hamming distance away from each other is maximized.

Set Multi-Cover for Localization with Errors

Pair-wise influence matrix:

Pair-wise events	S_1	S_2	S_3	S_4	…	S_{total}
Pipes (1,2)						

Set Multi-Cover Problem

Given a set of elements L, and a collection C of subsets of L. Select the minimum number of subsets in C such that each element in L is contained in at least k of selected subsets. For our problem, k = 2e+1.

* Simple greedy gives (1+ln a)-approximation algorithm *(Vazirani 2001)*
* A randomized algorithm with an approximation ratio *(Berman et al. 2007)*

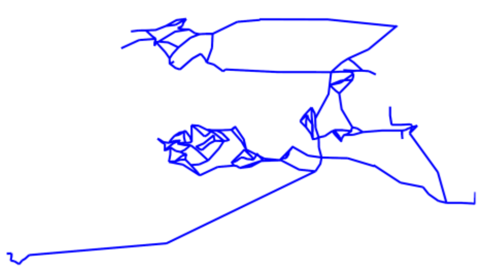
$$(1 + o(1)) \ln (a / k) \quad \text{if } (a / k) \geq 7.39$$

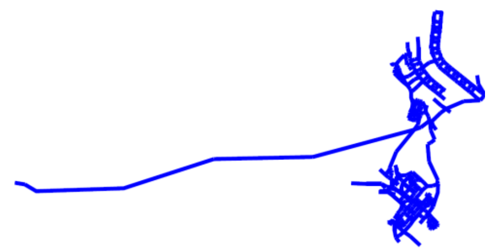
$$1 + 2(a / k)^{1/2} \quad \text{if } (a / k) < 7.39$$

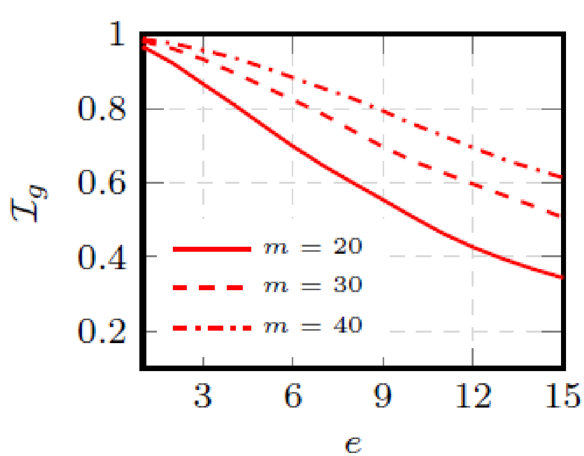
We solve the set multi-cover problem on the pair-wise influence matrix.

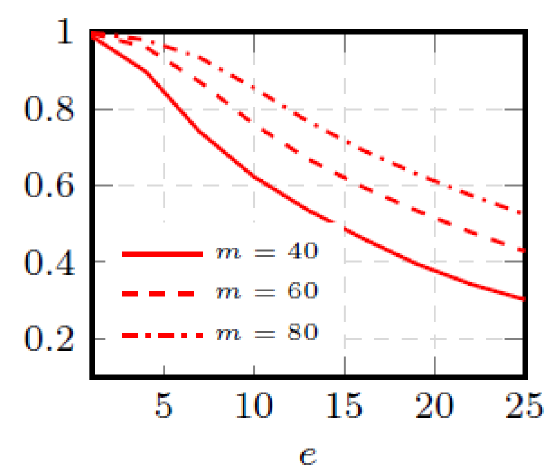
Numerical Evaluation

Generalized Identification Score:





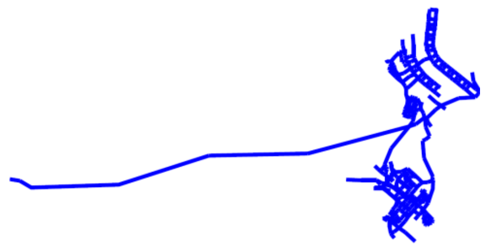


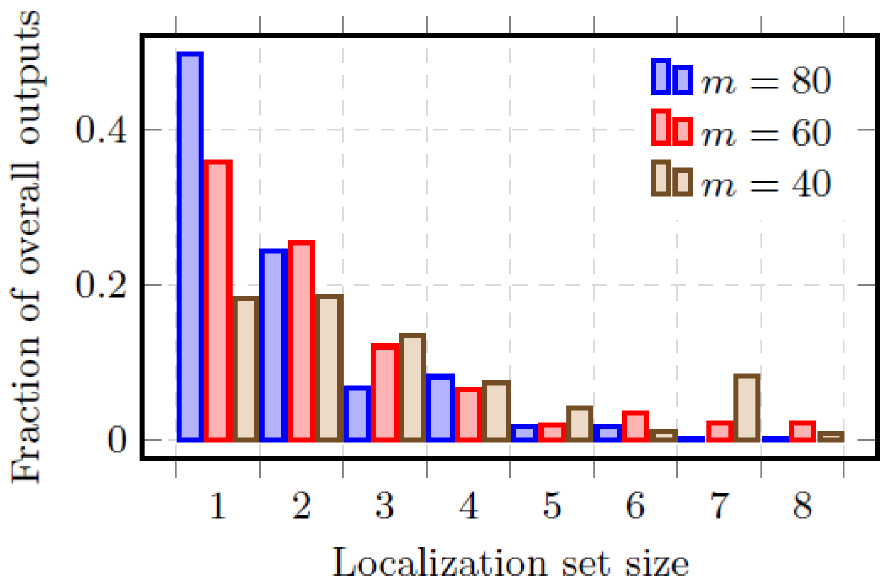


FORCES
FOUNDATIONS OF RESILIENT
CYBER-PHYSICAL SYSTEMS

Numerical Evaluation

Localization Set Size:



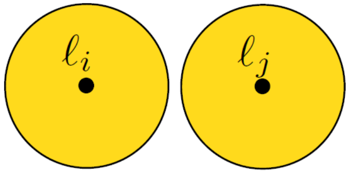


The percentage of outputs that have localization sets of size at most 5 is about 90%, 80%, and 58% for m = 80, 60, and 40 respectively.

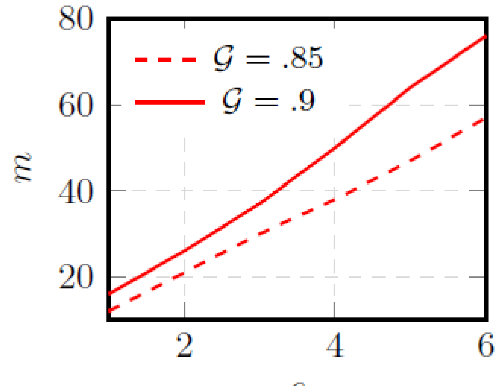
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Numerical Evaluation

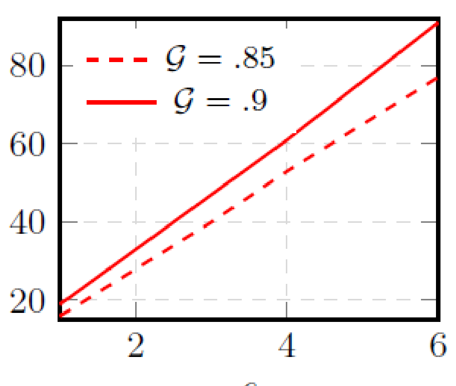
G = fraction of pair-wise events whose signatures are at least (2e+1) hamming distance apart.



For a fixed G, the ratio e/m (approximately) remains the same, i.e., m increases almost linearly with e.



(a) Water network 1



(b) Water network 2

Example:
In water network 2,
For G = .85, the ratio e/m = .08,
and for G = 0.9, e/m = .065

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Thank you