Resilience of water systems in wake of disruptions

Lina Sela

Dept. of Civil, Architectural & Environmental Engineering, UT Austin

The Bad News

CNN The New York Times USA TODAY

Water crisis in Flint, Mich., federal state of emergency

January, 2016

LEAD LEVEL COMPARISONS

Water contamination in Flint, Mich., compared with that of Detroit – Flint's original source for purified water.

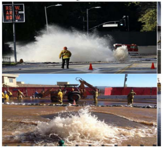
90th percentile[1] levels of lead exposure (in parts per billion):

Los Angeles Times

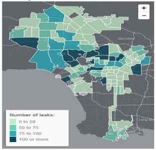
L.A.'s aging water pipes; a $1-billion dilemma

February, 2015

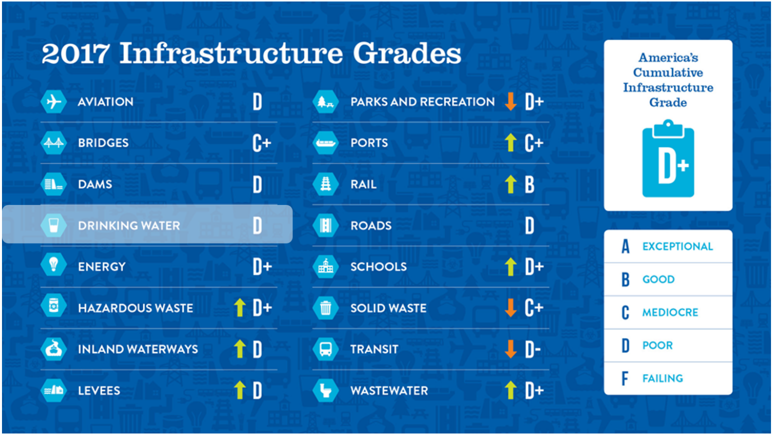
Leaks in L.A. water grid



Leaks by area, 2010 to 2014

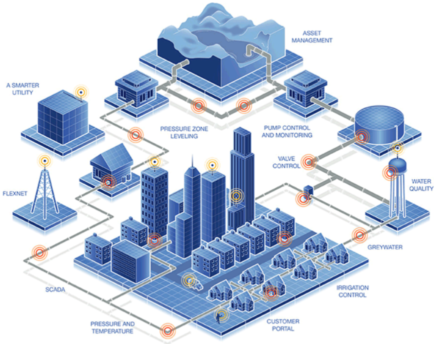


The Bad News

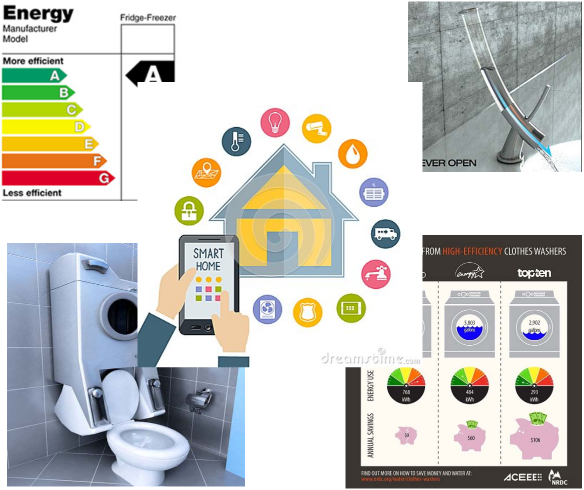


2017 Infrastructure Grades
AVIATION D
BRIDGES C+
DAMS D
DRINKING WATER D
ENERGY D+
HAZARDOUS WASTE D+
INLAND WATERWAYS D
LEVEES D
PARKS AND RECREATION D+
PORTS C+
RAIL B
ROADS D
SCHOOLS D+
SOLID WASTE C+
TRANSIT D-
WASTEWATER D+
America's Cumulative Infrastructure Grade
D+
A EXCEPTIONAL
B GOOD
C MEDIOCRE
D POOR
F FAILING

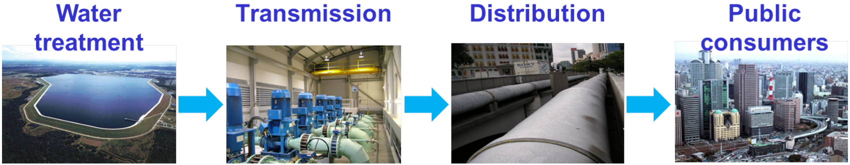
The Good News: Smart Cities



The Good News: Smart Homes

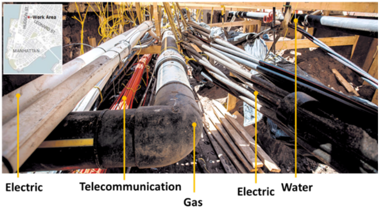


Infrastructure systems



Reduce:

- Water loss
- Water quality
- Energy requirements
- Infrastructure failures
- Supply interruptions



Sensor placement

Objective

- Sensor placement for detection and **location identification** of failures

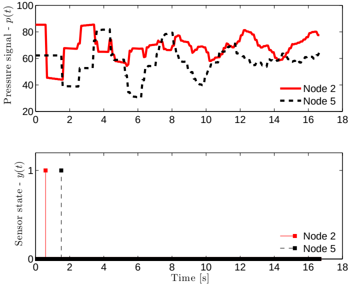
Approach

1. Influence model
   - Network and sensing models
2. Combinatorial optimization
   - The minimum test cover (MTC) problem
   - Augmented greedy solution algorithm

- L. Sela and S. Amin. ""Robust sensor placement for pipeline monitoring: Mixed integer and greedy optimization." *Advanced Engineering Informatics*, 2018.
- L. Sela, W. Abbas, X. Koutsoukos, and S. Amin. "Minimum test cover approach for fault location identification in flow networks." *Automatica*, 2016.
- W. Abbas, L. Sela, X. Koutsoukos, and S. Amin. "An efficient approach to fault identification in urban water networks using multi-level sensing." *ACM BuildSys 2015*.

Influence model

Sensing:



$\mathcal{L} = \{\ell_1, \ldots, \ell_n\}$ – set of n failure events

$\mathcal{S} = \{S_1, \ldots, S_m\}$ – set of m sensor locations

Detection:

$$y_{S_i}(t, \ell_j) = \begin{cases} 1 & \text{if } \xi\left(p_{i,t} - \hat{p}_{i,t}\right) \geq \varepsilon, \\ 0 & \text{otherwise.} \end{cases}$$

Fault signature:

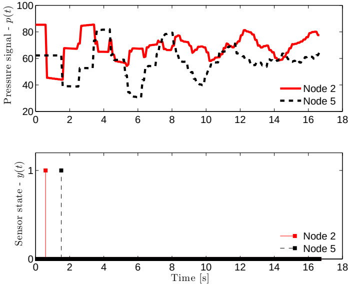
$$\mathbf{y}_{S_i}(\ell_j) = \begin{cases} 1 & \text{if } y_{S_i}(t, \ell_j) = 1, \quad \text{for any } t > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Fault matrix:

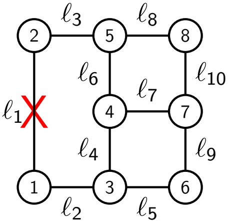
$$\mathcal{M}\left(\mathcal{L}, \mathcal{S}\right) = \begin{bmatrix} \mathbf{y}_{\mathcal{S}}(\ell_1) \\ \mathbf{y}_{\mathcal{S}}(\ell_2) \\ \vdots \\ \mathbf{y}_{\mathcal{S}}(\ell_n) \end{bmatrix}$$

Influence model

Example:



$$\mathcal{M}(\mathcal{L},\mathcal{S}) = \begin{array}{c} \\ \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \\ \ell_5 \\ \ell_6 \\ \ell_7 \\ \ell_8 \\ \ell_9 \\ \ell_{10} \end{array} \begin{array}{c} \begin{array}{cccccccc} S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 & S_8 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}$$



Detection as MSC

Detection

The *detection problem* is to select the minimum number of sensors $S \subseteq \mathcal{S}$, such that when a detectable event occurs, at least one sensor in S detects the event.

Minimum set cover (MSC)

Let $\mathcal{L}$ be a finite set of elements, and $\mathcal{C} = \{C_i : \ C_i \subseteq \mathcal{L}\}$ be the collection of given subsets of $\mathcal{L}$. The minimum set cover is to find $\mathcal{C}_s \subseteq \mathcal{C}$ with the minimum cardinality such that $\bigcup_{C_i \in \mathcal{C}} C_i = \bigcup_{C_j \in \mathcal{C}_s} C_j$.

Proposition

The detection problem is equivalent to the MSC problem where $f_D(\mathcal{C}_S) = \left| \bigcup_{C_i \in \mathcal{C}_S} C_i \right|$ is the detection function, $C_i \subseteq \mathcal{L}$ is the set of link failure events detected by the sensor S_i, i.e., $C_i = \{\ell_j \in \mathcal{L} | \ \mathbf{y}_{S_i}(\ell_j) = 1\}$.

Solving the MSC

The greedy approach

- In each iteration select:
  (a) Select $C_{i^*} \in \mathcal{C}$ covering the most uncovered elements in $\mathcal{L}$.
  (b) Add to current set $\mathcal{C}^* \leftarrow \mathcal{C}^* \cup \{C_{i^*}\}$.
  (c) Repeat until all elements in $\mathcal{L}$ are covered or no new element can be covered by any $C_i \in \mathcal{C}$.
- Best approximation ratio of $\mathcal{O}(\ln n)$.
- Running times $\mathcal{O}(mn)$. Can be made faster by reducing the number of function evaluations exploiting the *submodularity* property. *Lazy greedy* (Krause et al 2008).

Solving the MSC

The greedy approach

- In each iteration select:
  (a) Select $C_{i^*} \in \mathcal{C}$ covering the most uncovered elements in $\mathcal{L}$.
  (b) Add to current set $\mathcal{C}^* \leftarrow \mathcal{C}^* \cup \{C_{i^*}\}$.
  (c) Repeat until all elements in $\mathcal{L}$ are covered or no new element can be covered by any $C_i \in \mathcal{C}$.
- Best approximation ratio of $\mathcal{O}(\ln n)$.
- Running times $\mathcal{O}(mn)$. Can be made faster by reducing the number of function evaluations exploiting the *submodularity* property. *Lazy greedy* (Krause et al 2008).

Identification as MTC

Identification

The *identification* problem is to select the minimum number of sensors $S \subseteq \mathcal{S}$ that *uniquely* detect the events in $\mathcal{L}$.
Pair-wise event $\{\ell_i, \ell_j\}$ is *detectable*, if there exists a sensor that gives different outputs for ℓ_i and ℓ_j, $\exists S_p \in \mathcal{S}: \ \mathbf{y}_{S_p}(\ell_i) \neq \mathbf{y}_{S_p}(\ell_j)$.

Minimum test cover (MTC)

The MTC is to find $\mathcal{C}_t \subseteq \mathcal{C}$ with the minimum cardinality such that if for a pair of elements $\{\ell_u, \ell_v\} \in \mathcal{L}$, there exists $C_i \in \mathcal{C}$ that contains either ℓ_u or ℓ_v but not both, then there exists some $C_j \in \mathcal{C}_t$ that also contains either ℓ_u or ℓ_v, but not both.

Proposition

The problem of identification of link failures in networks is equivalent to the MTC problem.

Example cont.:

Detection: $\{S_2, S_4\}$

1	1	1	0	1	0	0	0
1	1	1	1	0	1	0	0
1	1	0	1	1	0	0	1
1	0	1	1	1	1	1	0
1	0	1	1	0	1	1	0
0	1	1	1	1	0	1	1
0	0	1	1	1	1	1	1
0	1	0	1	1	0	1	1
0	0	1	1	0	1	1	1
0	0	0	1	1	1	1	1

- All events are detected
- Only three unique sensor outputs

Identification: $\{S_1, S_2, S_3, S_5\}$

1	1	1	0	1	0	0	0
1	1	1	1	0	1	0	0
1	1	0	1	1	0	0	1
1	0	1	1	1	1	1	0
1	0	1	1	0	1	1	0
0	1	1	1	1	0	1	1
0	0	1	1	1	1	1	1
0	1	0	1	1	0	1	1
0	0	1	1	0	1	1	1
0	0	0	1	1	1	1	1

- All events are detected
- All events are uniquely identified

Solving the MTC

Greedy solution

1. **Input:** $\mathcal{C} = \{C_1, \cdots, C_m\},\ C_i \subseteq \mathcal{L}$.
2. **Transform:** the MTC to the equivalent MSC
   - *Create a new set of events*: $\mathcal{L}^t = \{\ell^t_{12}, \cdots, \ell^t_{(n-1)n}\}$. For each unordered pair $\{\ell_i, \ell_j\}$, define a new element ℓ^t_{ij}.
   - *Create a new sets of sensors' outputs*: $\mathcal{C}^t = \{C^t_1, \cdots, C^t_m\}$, where $C^t_v = \{\ell^t_{ij} : |\{\ell_i, \ell_j\} \cap C_v| = 1\},\ \forall k \in \{1, \cdots, m\}$.
3. **Solve:** using greedy algorithm
   (a) Select $C^t_{i*} \in \mathcal{C}^t$ covering the most uncovered elements in $\mathcal{L}^t$.
   (b) Add to current set $\mathcal{C}^* \leftarrow \mathcal{C}^* \cup \{C_{i*}\}$.
   (c) Repeat until all elements in $\mathcal{L}^t$ are covered or no new element in $\mathcal{L}^t$ can be covered by any $C^t_i \in \mathcal{C}^t$.
4. **Output:** MTC, $\mathcal{C}^* \subseteq \mathcal{C}$.

Solving the MTC

Greedy solution

1. **Input:** $\mathcal{C} = \{C_1, \cdots, C_m\},\ C_i \subseteq \mathcal{L}$.
2. **Transform:** the MTC to the equivalent MSC
   - *Create a new set of events*: $\mathcal{L}^t = \{\ell^t_{12}, \cdots, \ell^t_{(n-1)n}\}$. For each unordered pair $\{\ell_i, \ell_j\}$, define a new element ℓ^t_{ij}.
   - *Create a new sets of sensors' outputs*: $\mathcal{C}^t = \{C^t_1, \cdots, C^t_m\}$, where $C^t_v = \{\ell^t_{ij} : |\{\ell_i, \ell_j\} \cap C_v| = 1\}, \forall k \in \{1, \cdots, m\}$.
3. **Solve:** using greedy algorithm
   (a) Select $C^t_{i*} \in \mathcal{C}^t$ covering the most uncovered elements in $\mathcal{L}^t$.
   (b) Add to current set $\mathcal{C}^* \leftarrow \mathcal{C}^* \cup \{C_{i*}\}$.
   (c) Repeat until all elements in $\mathcal{L}^t$ are covered or no new element in $\mathcal{L}^t$ can be covered by any $C^t_i \in \mathcal{C}^t$.
4. **Output:** MTC, $\mathcal{C}^* \subseteq \mathcal{C}$.

Solving the MTC

Greedy solution

1. **Input:** $\mathcal{C} = \{C_1, \cdots, C_m\},\ C_i \subseteq \mathcal{L}$.
2. **Transform:** the MTC to the equivalent MSC
   - *Create a new set of events*: $\mathcal{L}^t = \{\ell^t_{12}, \cdots, \ell^t_{(n-1)n}\}$. For each unordered pair $\{\ell_i, \ell_j\}$, define a new element ℓ^t_{ij}.
   - *Create a new sets of sensors' outputs*: $\mathcal{C}^t = \{C^t_1, \cdots, C^t_m\}$, where $C^t_v = \{\ell^t_{ij} : |\{\ell_i, \ell_j\} \cap C_v| = 1\},\ \forall k \in \{1, \cdots, m\}$.
3. **Solve:** using greedy algorithm
   (a) Select $C^t_{i*} \in \mathcal{C}^t$ covering the most uncovered elements in $\mathcal{L}^t$.
   (b) Add to current set $\mathcal{C}^* \leftarrow \mathcal{C}^* \cup \{C_{i*}\}$.
   (c) Repeat until all elements in $\mathcal{L}^t$ are covered or no new element in $\mathcal{L}^t$ can be covered by any $C^t_i \in \mathcal{C}^t$.
4. **Output:** MTC, $\mathcal{C}^* \subseteq \mathcal{C}$.

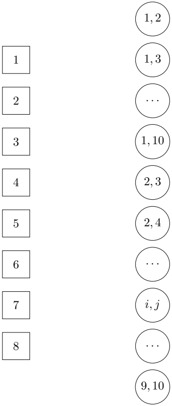
Solving the MTC

Greedy solution

1. **Input:** $\mathcal{C} = \{C_1, \cdots, C_m\},\ C_i \subseteq \mathcal{L}$.
2. **Transform:** the MTC to the equivalent MSC
   - *Create a new set of events*: $\mathcal{L}^t = \{\ell^t_{12}, \cdots, \ell^t_{(n-1)n}\}$. For each unordered pair $\{\ell_i, \ell_j\}$, define a new element ℓ^t_{ij}.
   - *Create a new sets of sensors' outputs*: $\mathcal{C}^t = \{C^t_1, \cdots, C^t_m\}$, where $C^t_v = \{\ell^t_{ij} : |\{\ell_i, \ell_j\} \cap C_v| = 1\}, \forall k \in \{1, \cdots, m\}$.
3. **Solve:** using greedy algorithm
   (a) Select $C^t_{i*} \in \mathcal{C}^t$ covering the most uncovered elements in $\mathcal{L}^t$.
   (b) Add to current set $\mathcal{C}^* \leftarrow \mathcal{C}^* \cup \{C_{i*}\}$.
   (c) Repeat until all elements in $\mathcal{L}^t$ are covered or no new element in $\mathcal{L}^t$ can be covered by any $C^t_i \in \mathcal{C}^t$.
4. **Output:** MTC, $\mathcal{C}^* \subseteq \mathcal{C}$.

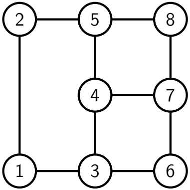
Example cont.

MTC to MSC



Sensors   *Pair-wise* events

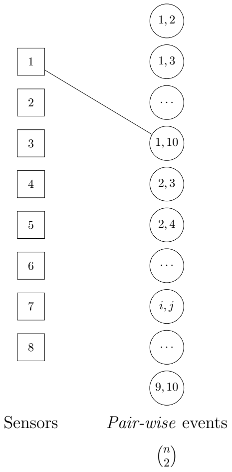
$\binom{n}{2}$

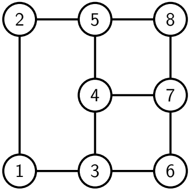


	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
ℓ_1	1	1	1	0	1	0	0	0
ℓ_2	1	1	1	1	0	1	0	0
ℓ_3	1	1	0	1	1	0	0	1
ℓ_4	1	0	1	1	1	1	1	0
ℓ_5	1	0	1	1	0	1	1	0
ℓ_6	0	1	1	1	1	0	1	1
ℓ_7	0	0	1	1	1	1	1	1
ℓ_8	0	1	0	1	1	0	1	1
ℓ_9	0	0	1	1	0	1	1	1
ℓ_{10}	0	0	0	1	1	1	1	1

Example cont.

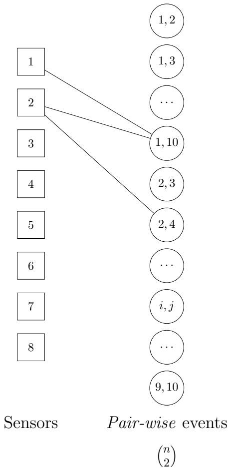
MTC to MSC

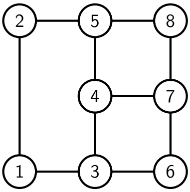




	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
ℓ_1	1	1	1	0	1	0	0	0
ℓ_2	1	1	1	1	0	1	0	0
ℓ_3	1	1	0	1	1	0	0	1
ℓ_4	1	0	1	1	1	1	1	0
ℓ_5	1	0	1	1	0	1	1	0
ℓ_6	0	1	1	1	1	0	1	1
ℓ_7	0	0	1	1	1	1	1	1
ℓ_8	0	1	0	1	1	0	1	1
ℓ_9	0	0	1	1	0	1	1	1
ℓ_{10}	0	0	0	1	1	1	1	1

MTC to MSC

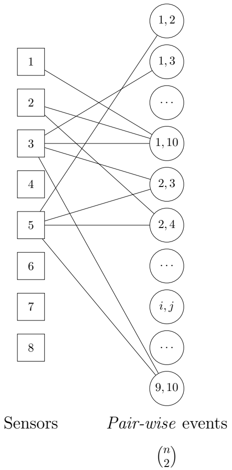




	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
ℓ_1	1	1	1	0	1	0	0	0
ℓ_2	1	1	1	1	0	1	0	0
ℓ_3	1	1	0	1	1	0	0	1
ℓ_4	1	0	1	1	1	1	1	0
ℓ_5	1	0	1	1	0	1	1	0
ℓ_6	0	1	1	1	1	0	1	1
ℓ_7	0	0	1	1	1	1	1	1
ℓ_8	0	1	0	1	1	0	1	1
ℓ_9	0	0	1	1	0	1	1	1
ℓ_{10}	0	0	0	1	1	1	1	1

Example cont.

MTC to MSC



- Equivalent MSC

$$
\begin{array}{c} \\ \ell_1,\ell_2 \\ \ell_1,\ell_3 \\ \ell_1,\ell_4 \\ \vdots \\ \ell_1,\ell_{10} \\ \ell_2,\ell_3 \\ \ell_2,\ell_4 \\ \vdots \\ \ell_9,\ell_{10} \end{array}
\begin{array}{c}
\begin{array}{cccccccc} S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 & S_8 \end{array} \\
\begin{pmatrix}
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0
\end{pmatrix}
\end{array}
$$

- Solve using the greedy algorithm:

$$f_I(\mathcal{C}_S) = f_D(\mathcal{C}_S^t)$$

Augmented greedy MTC solution

Transformed greedy solution

- Memory needed to transform MTC to the MSC in *GB*: $\binom{n}{2} \times m \times 10^{-9}$
  - $m = 1000$; $n = 1000$; $\sim 0.5GB$
  - $m = 2000$; $n = 2000$; $\sim 4GB$
  - $m = 10000$; $n = 10000$; $\sim 500GB$

Augmented greedy solution

- Avoid the complete transformation of the MTC to the MSC.

Augmented greedy MTC solution

Transformed greedy solution

- Memory needed to transform MTC to the MSC in GB: $\binom{n}{2} \times m \times 10^{-9}$
  - $m = 1000; n = 1000; \sim 0.5GB$
  - $m = 2000; n = 2000; \sim 4GB$
  - $m = 10000; n = 10000; \sim 500GB$

Augmented greedy solution

- Avoid the complete transformation of the MTC to the MSC.

Augmented greedy MTC solution

Main idea

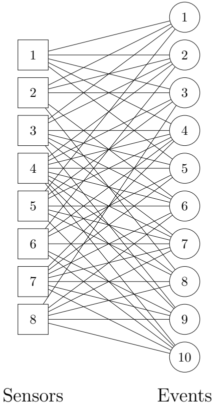
- A sensor i that detects k events (i.e., $|C_i| = k$) can distinguish between k detected events and $(n-k)$ undetected events, i.e. it detects $k(n-k)$ *pair-wise* events (i.e., $|C_i^t| = k(n-k)$).
- Let $C^* \subseteq \mathcal{C}$ be the (test) cover until the current iteration, and $\mathcal{C}_{cov}$ be the set of link failures detected by the sensors that are included in the (test) cover, i.e., $\mathcal{C}_{cov} = \bigcup_{C_u \in C^*} C_u$.
- The utility of adding C_i to C^* in each iteration is based on:
  (i) x_i – how many pair-wise events corresponding to **undetected events**, i.e., not in $\mathcal{C}_{cov}$ can be detected by C_i?
  (ii) y_i – how many undetected pair-wise events corresponding to **detected events**, i.e, in $\mathcal{C}_{cov}$ can be detected by C_i?

Main algorithm

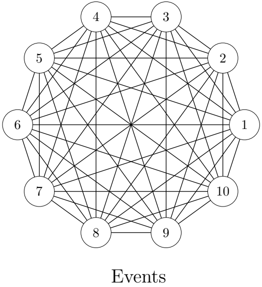
Input: $\mathcal{C} = \{C_1, \cdots, C_m\}, \ C_i \subseteq \mathcal{L}$
Output: MTC: $\mathcal{C}^* \subseteq \mathcal{C}$
Initialization: $\mathcal{C}_{cov} = \emptyset; \ \mathcal{C}^* = \emptyset; \ G_0 = \emptyset; \ j = 1; \ n = |\mathcal{L}|; \ \ w_{i^*} = 1;$
while $w_{i^*} > 0$ **do**
    $n_j \leftarrow n - |\mathcal{C}_{cov}|$
    for all i **do**
        $X_i \leftarrow (C_i \setminus \mathcal{C}_{cov})\, ; k_{i,j} \leftarrow |X_i|$
        $x_i \leftarrow k_{i,j}(n_j - k_{i,j})$
        $Y_i \leftarrow C_i \cap \mathcal{C}_{cov}$
        $y_i \leftarrow \sum_{t=0}^{j-1} |\alpha(Y_i, G_t)|$
        $w_i = x_i + y_i$
    end for
    $w_{i^*} \leftarrow \max w_i$
    if $w_{i^*} > 0$ **then**
        $\mathcal{C}^* \leftarrow \mathcal{C}^* \cup \{C_{i^*}\}$
        $\mathcal{C}_{cov} \leftarrow \mathcal{C}_{cov} \cup C_{i^*}$
        $G_j \leftarrow \beta(X_{i^*})$
        for $t = 0$ to $j - 1$ **do**
            $G_t \leftarrow G_t \setminus \alpha(Y_{i^*}, G_t)$
        end for
        $j \leftarrow j + 1$
    end if
end while

Example cont.

Initialization:

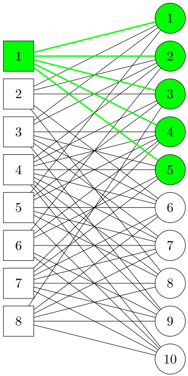


$\mathcal{C}_{cov} = \emptyset;\ \mathcal{C}^{*} = \emptyset;\ G_0 = \emptyset;\ n = 10;$

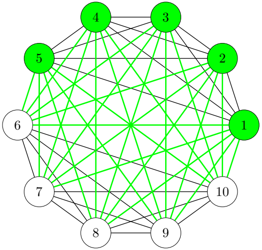


Example cont.

Iteration 1:

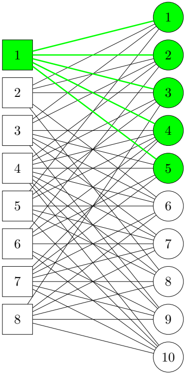


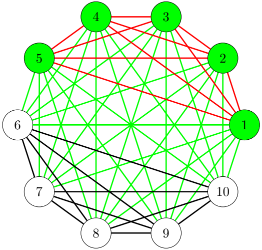
$$x_i = k_{i,1}(n - k_{i,1});$$
$$x_1 = 5(10 - 5) = 25;$$
$$y_i = 0;\ w_i = x_i + y_i;$$



Example cont.

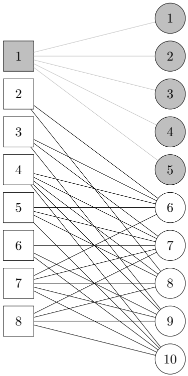
Iteration 1:



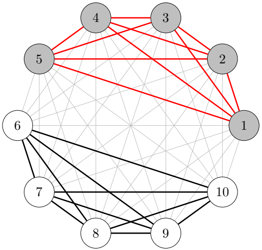


Example cont.

End of Iteration 1:



$\mathcal{C}_{cov} = \{1, 2, 3, 4, 5\};\ n = 5;$

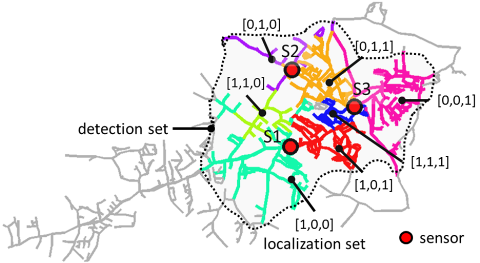


$G_1 = \{\{1, 2\}, \{1, 3\}, \cdots, \{4, 5\}\};$

Application example:

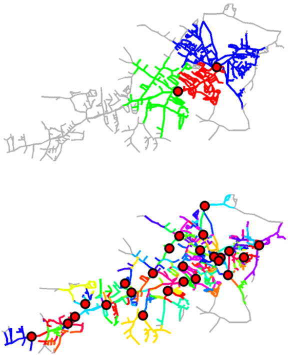
Net9@KY

- Daily supply $\sim 1.5M[\frac{gal}{day}]$; $260[km]$ pipe length;
- > 950 junctions; > 1100 pipes;



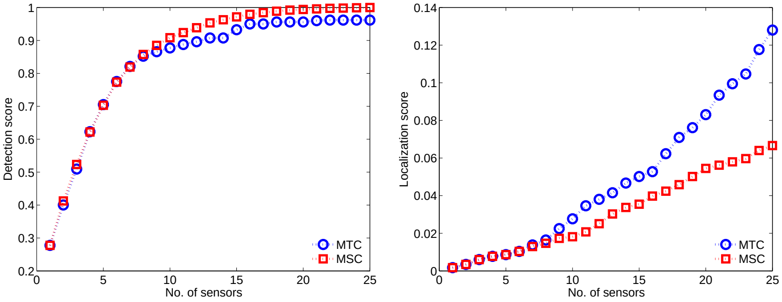
Adopted from Jolly et al 2014

Net9@KY cont.

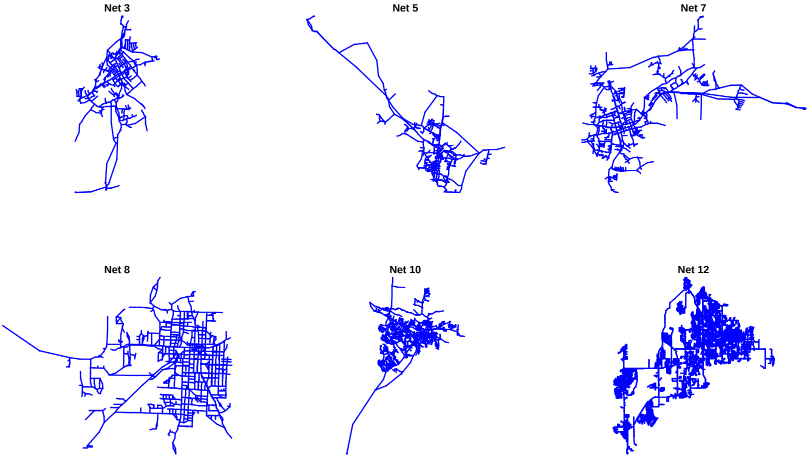


MTC vs. MSC

Net9@KY cont.



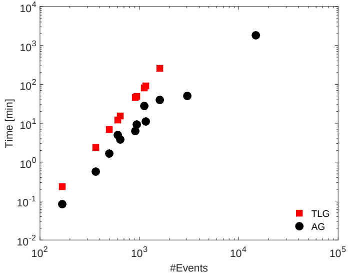
Simulations



Computations

Network	No. of sensors	No. of pipes	TLG [min]	AG [min]
Net1	48	168	0.23	0.08
Net2	98	366	2.39	0.58
Net3	134	496	6.93	1.65
Net4	138	603	11.98	4.93
Net5	164	644	15.58	3.85
Net6	258	907	45.46	6.31
Net7	139	940	49.12	9.31
Net8	195	1124	80.55	28.07
Net9	359	1156	91.57	11.06
Net10	408	1614	257.41	39.48
Net11	712	3032	–	50.53
Net12	1001	14822	–	1800.08

TLG - transformed lazy greedy; AG - augmented greedy;



- TLG – $\mathcal{O}\left(\binom{n}{2}\right)$
- AG – $\mathcal{O}\left(\sum_i^{m_j}\binom{k_i}{2}\right)$
- $\sum_i \binom{k_i}{2} \leq \frac{k}{n}\binom{n}{2}$