Formalization: Least sensitive point

Given

- a water distribution network $W = (V, E)$ with
  - a set of N nodes (junctions, tanks, reservoirs) $V = \{v_n | n \in \{1, \ldots, N\}\}$
  - a set of edges (pipes, valves, pumps) connecting the nodes: $E \subseteq V \times V$
- demand values $\mathbf{X} \in \mathbb{R}^{N \times T}$ where $x_{n,t}$ is the demand for v_n at time t
- a prediction function $f_{pred} : \mathbb{R}^N \times \mathbb{N} \to \{0, 1\}$ taking a vector of demand values $\boldsymbol{x}_t := (x_{1,t}, \ldots, x_{n,t})$ and a time $t \in \{1, \ldots, T\}$ to indicate whether the network is ok (0) or under attack (1)
- a time window $\{t + k | k \in \{0, \ldots, K\}\}$ with fixed size $K \in \mathbb{N}_0$

We are trying to introduce a maximal unnoticed change to the demand at one of the nodes.

$$\max_{\substack{n \in \{1,\ldots,N\} \\ t \in \{1,\ldots,T-K\} \\ \delta \in \mathbb{R}}} \|\delta\| \tag{1}$$

$$\text{s.t.} \quad f_{pred}(\boldsymbol{x}_{t+k} + \delta \boldsymbol{e}_n, t + k) = 0 \quad \forall k \in \{0, \ldots, K\} \tag{2}$$

Where $\boldsymbol{e}_n$ is the n-the canonical basis vector of the $\mathbb{R}^N$.

The **least sensitive point** is the node in the network, which solves the maximal unnoticed change problem, that is

$$\operatorname{lsp}(W) = v_{n^*} \tag{3}$$

where

$$n^* := \operatorname*{arg\,max}_{n \in \{1,\ldots,N\}} \max_{\substack{t \in \{1,\ldots,T\} \\ \delta \in \mathbb{R}}} \|\delta\| \tag{4}$$

$$\text{s.t.} \quad f_{pred}(\boldsymbol{x}_t + \delta \boldsymbol{e}_n, t) = 0 \tag{5}$$

Ideas for an approximation

For fixed $\boldsymbol{x}_t$ and t, one can try to approximate n^* by means of the absolute value of the gradient of f_{pred} with respect to the nodal demand.

$$\left| \frac{\partial f_{pred}(\boldsymbol{x}_t, t)}{\partial x_{n,t}} \right| \tag{6}$$

This method would require differentiability of f_{pred}. Also, a good initial guess for t or alternatively a gradient computation for several interesting timesteps is needed. The demand $\boldsymbol{x}_t$ must be known or approximated. In order to achieve a good approximation for n^*, gradients of f_{pred} should be smooth, i.e. a small gradient at $\boldsymbol{x}_t$ should imply a small gradient at $\boldsymbol{x}_t + \epsilon$.

Sensors

Usually, a few of the nodes will be equipped with pressure sensors. Given $S \leq N$ sensors in the network, the predictive function can be expressed as a composition

$$f_{pred} = f_{model} \circ f_{measure} \tag{7}$$

where

- $f_{measure} : \mathbb{R}^N \to \mathbb{R}^S$ maps the demands $\boldsymbol{x}_t$ for some fixed time t to pressure measurements $\boldsymbol{y}_t \in \mathbb{R}^S$
- $f_{model} : \mathbb{R}^S \times \mathbb{N} \to \{0, 1\}$ uses the measured presuure values $\boldsymbol{y}_t$ and the timestep t to predict whether the network is under attack

Using this composition, the derivative above can be re-written as

$$\frac{\partial f_{pred}(\boldsymbol{x}_t)}{\partial p_{n,t}} = \frac{\partial f_{model}(\boldsymbol{y}_t, t)}{\partial f_{measure}(\boldsymbol{x}_t)} \frac{\partial f_{measure}(\boldsymbol{x}_t)}{\partial x_{n,t}} \tag{8}$$

For the gradient of $f_{measure}$ one could use hydraulic simulation.