STEP: Semantics-Aware Sensor Placement for Monitoring Community-Scale Infrastructure

Andrew Chio[1], Jian Peng[2], Nalini Venkatasubramanian[1]

[1]University of California, Irvine, [2]Orange County Public Works

Gaining insight into Sensor Placement for Stormwater using Network Structure, Behavior, and Semantics

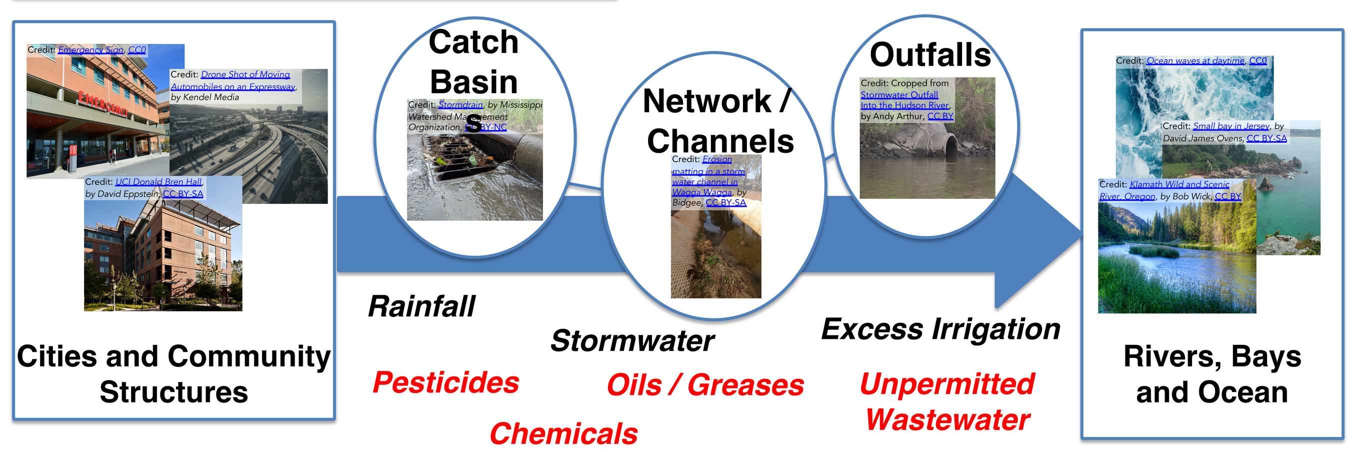
Motivation

- Urban cities and communities rely on built utility infrastructures such as water, gas and power as *critical lifelines*
- These engineered systems face issues of *resilience:* urban growth, climate change, and aging have given rise to *multiple modes of failure* which are difficult to handle due to their *continuous*, *transient*, or *sporadic* nature.
- The advent of Internet-of-Things (IoT) ecosystems and new data-driven methods show great promise for enabling next-generation smart monitoring solutions for improved operational efficiency and decision support.

How should IoT/sensor placements be designed to detect and trace anomalies to enable practical decision support for stormwater network community lifelines?

Driving Use Case: *Stormwater Networks*

Purpose and Responsibilities

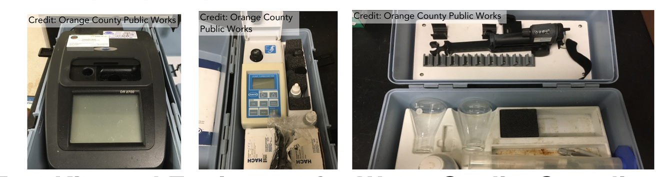


- Stormwater networks consist of catch basins, network channels, and outfalls.
- They transport rainwater and other nuisance flows from cities and communities to receiving waters, such as rivers, bays, and oceans.
- During this process, pollutants and contaminants can also be transported, which can lead to water quality impairments downstream.
- Regulations (e.g., amendment to the US Clean Water Act of 1987) prohibit pollutant discharge into MS4s

Rapid and effective monitoring of stormwater networks is essential to prevent discharges and take appropriate corrective actions

State-of-the-Art Approaches

- Mainly consist of inspections, citizen reports, and manual site visits. Test kits and lab analysis are utilized to understand water quality.
- Is costly and ineffective for understanding discharges into the network



Test Kits and Equipment for Water Quality Sampling

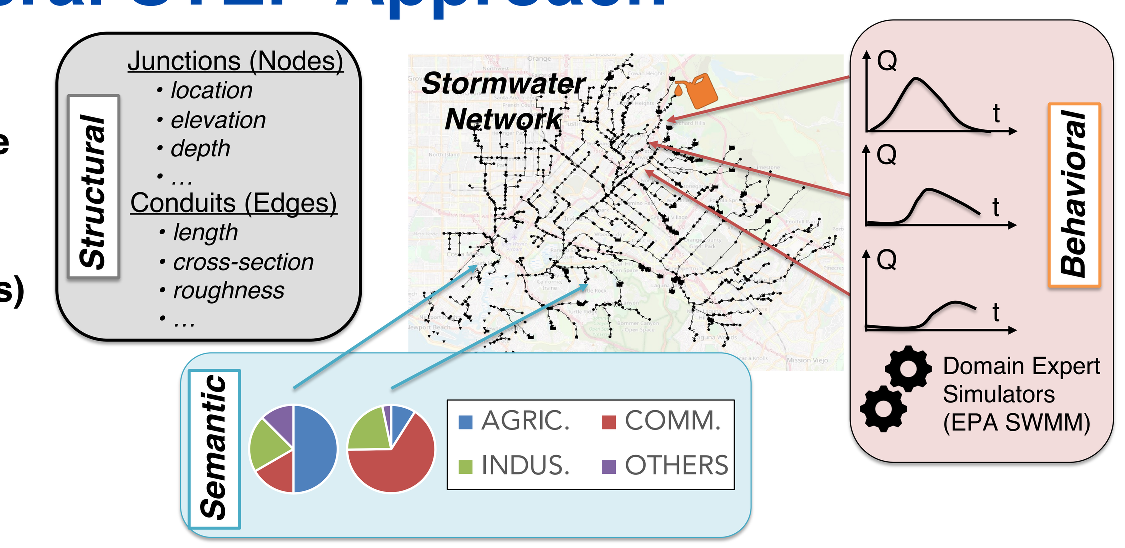
Other Challenges

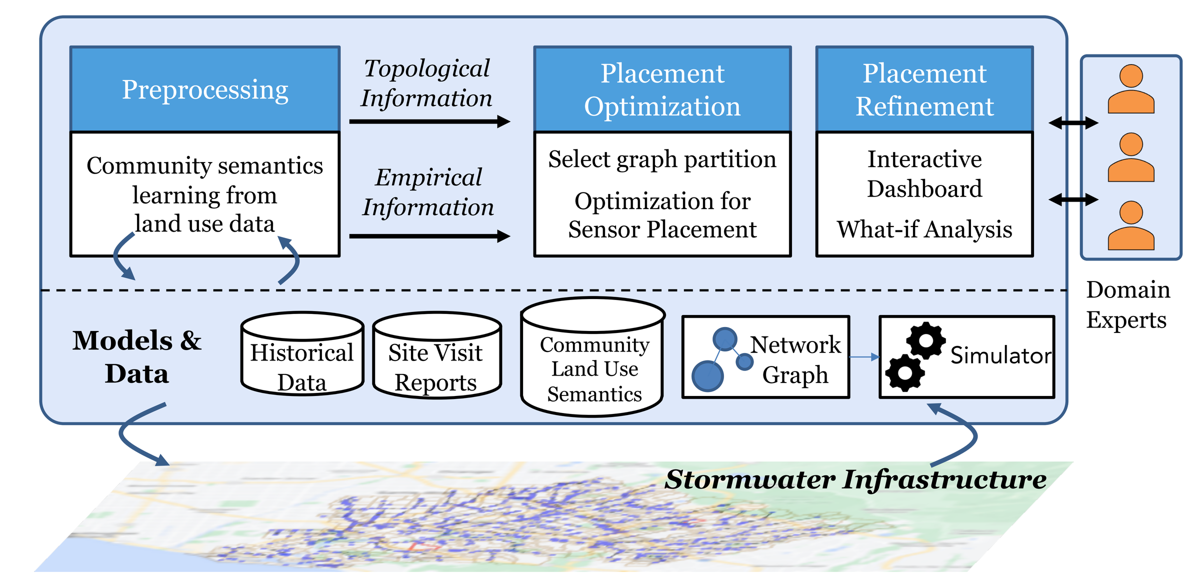
- Stormwater networks are large and geo-distributed, with regional catchment areas, and have thousands of potential entry points
- Pollutants can be *transient, spontaneous, and heterogeneous* in nature, making it difficult to detect and trace in the network

The General STEP Approach

Network Perspectives

- Several aspects of the network can provide insight into effective sensor placements
- ***Structural Aspect***: Physical properties and characteristics of network junctions (nodes) and conduits (edges)
- ***Behavioral Aspect***: Responses to various stimuli in the network, and their impact
- ***Semantic Aspect***: Influences from specific land uses of a community on anomalies





The STEP Workflow

- Use historical data and network semantics to construct a set of new realistic anomalies
- Leverage key topological and empirical properties to enable graph partitioning and MILP placement optimization
- Refine computed placement using the STEP toolkit and interactive dashboard for practical deployments

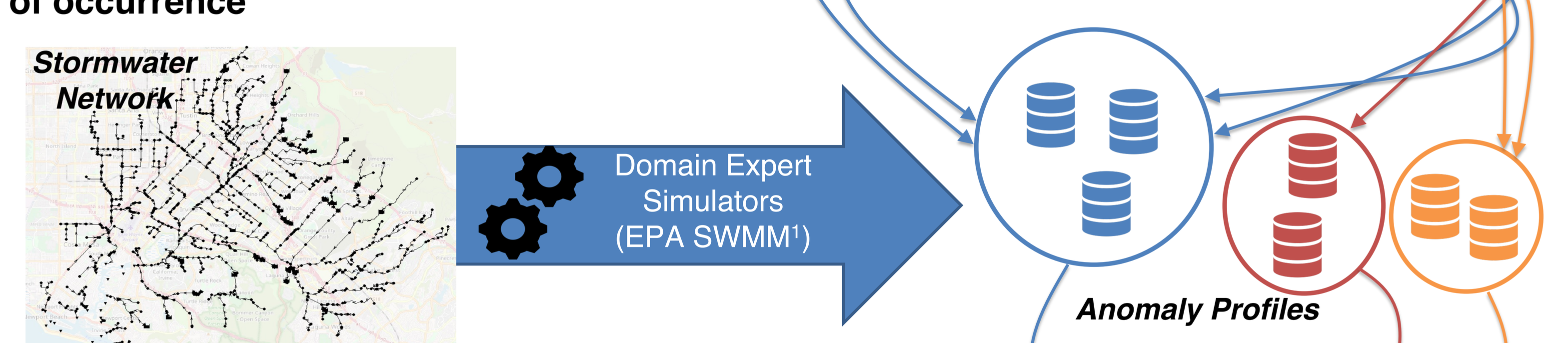
Generating Realistic Anomalies

Extracting anomalies from historical data

- Simulate anomalies uniformly in network and cache into database
- Cluster anomalies into *profiles* based on the similarity of their impact in the network
- Map historical instances of anomalies to constructed profiles to estimate likelihood of occurrence

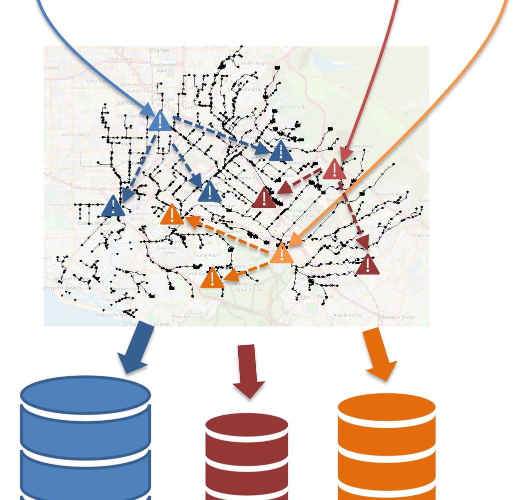
Historical Dataset D

Time t	Node v	Flow f	Obs. o
2017-05-23, 10:02	FVES@D05	0.03	Turbidity: 1.2
2017-06-30, 08:22	FVES@D05	0.14	Elec. Cond.: 906
2019-05-15, 08:00	COF07S01	0.28	Turbidity: 4.03
2019-07-24, 09:10	COF07S01	0.21	Temp. : 26.24
2019-08-07,08:28	GC02S0172DS	0.09	Elec. Cond.: 6113
2019-07-17,08:25	IRVF06P06	0.11	Turbidity: 3.5
...	...	...	...



Generating new anomalies through semantics

- Select an anomaly profile from which to generate a new anomaly
- Pick semantic land use "cause" from anomaly profile
- Pick origin node based on nearby area of selected semantic land use
- Sample all other properties of the new anomaly based on average / standard deviation of values in profile (duration, amount, phenomenon produced)



[1]EPA. 2023. EPA Stormwater Management Model (SWMM). https://www.epa.gov/water- research/storm- water- management- model- swmm

Placement Optimization and Refinement

Placement Optimization

Objective: Coverage $\mathcal{COV}$

Ability to capture and observe anomalies in the network

$$COV(\mathcal{X},\mathcal{A},\mathcal{G}) = \frac{1}{|\mathcal{V}|}\sum_{v_i\in\mathcal{V}} covered(v_i,\mathcal{X},\mathcal{A}(v_i))$$

$$covered(v_i,\mathcal{X},\mathcal{A}(v_i)) = \mathbb{1}\left[\sum_{\alpha_k\in\mathcal{A}(v_i)}\sum_{v_j\in\mathcal{V}}\sum_{s_l\in\mathcal{S}} x_{lj}OB(l,k)PT(k,j) \geq \rho\,|\mathcal{A}(v_i)|\right]$$

Objective: Traceability $\mathcal{TR}$

Ability to use observations to track the origin of an anomaly

$$TR(\mathcal{X},\mathcal{A},\mathcal{G}) = \frac{1}{|\mathcal{A}|}\sum_{\alpha_k\in\mathcal{A}}\sum_{s_l\in\mathcal{S}(\mathcal{P}_k)}\sum_{v_j\in\mathcal{V}}\left|\mathcal{V}^{up}_{v_j,\alpha_k,\mathcal{X}}\right| \Big/ |\mathcal{V}|$$

Empirical

Topological

Topological

Betweenness Centrality $\mathcal{BTN}$

# of anomalies passing through nodes

$$\mathcal{BTN}(v_j) = \sum_{\alpha_k\in\mathcal{A}} \mathbb{1}\left[time(\alpha_k, v_k^*, v_j) \leq \tau\right]$$

Branching Complexity $\mathcal{BC}$

Degree of merging/splitting at nodes

$$\mathcal{BC}(v_j) = \begin{cases} 1 & \text{if IsRoot}(v_j) \\ \mathcal{BC}^{max}_{pa(v_j)} + \sum_{v_i\in pa(v_j)} \frac{\mathcal{BC}(v_i)}{\mathcal{BC}^{max}_{pa(v_j)}} - 1 & \text{else} \\ \text{where: } \mathcal{BC}^{max}_{pa(v_j)} = \max_{v_i\in pa(v_j)} \mathcal{BC}(v_i) & \end{cases}$$

Semantic Entropy $\mathcal{SE}$

Skewness of distribution of upstream land uses

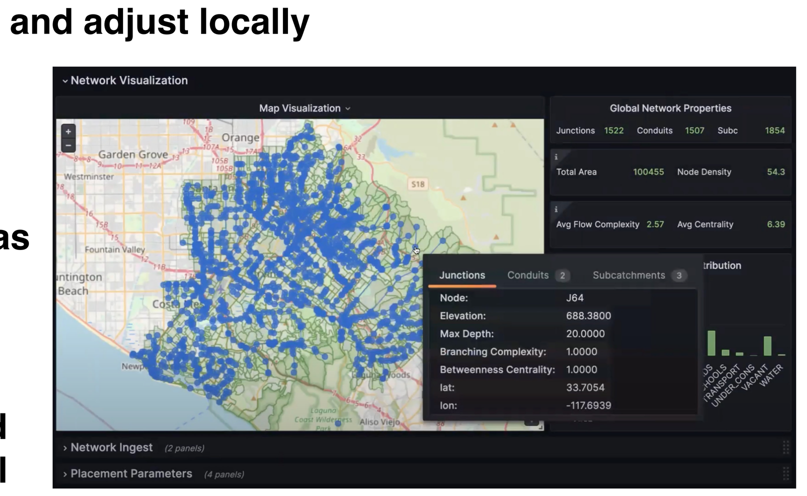
$$\mathcal{SE}(\mathcal{U},\mathcal{G}^{up}_{v_j}) = \sum_{u_m\in\mathcal{U}(v_j)} \lambda_m \cdot \left(-P(u_m)\log P(u_m)\right)$$

$$\text{where: } P(u_m) = \sum_{v_i\in\mathcal{V}^{up}_{v_j}} \left(\frac{Area(v_i,u_m)}{\sum_{u_m\in\mathcal{U}} Area(v_i,u_m)}\right)$$

- Partition stormwater graph based on key network properties; select nodes that maximize $\mathcal{BTN}$, and minimize $\mathcal{BC}$ and $\mathcal{SE}$
- Find ideal placements on partitioned subgraphs using *MILP optimization*
- Merge subgraph placement solutions, and adjust locally

Placement Refinement

- Ideal placement generated algorithmically may be infeasible in practice due to external factors, such as potential vandalism, location-specific communication issues, and physical barriers preventing human access
- The STEP toolkit includes a dashboard for domain experts to refine a potential placement as needed

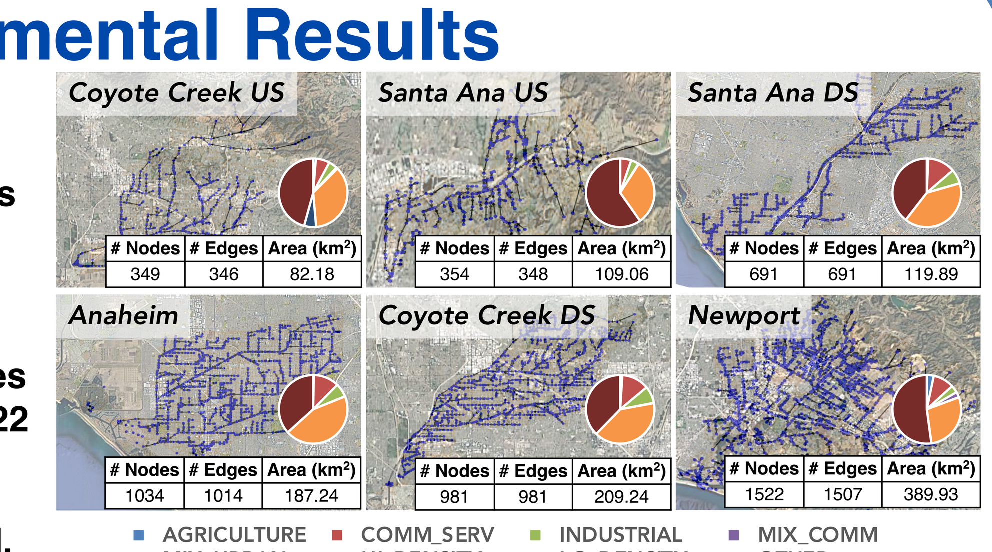


https://github.com/andrewgchio/STEP

Experimental Results

Experimental Setup

- 6 EPA SWMM networks of varying sizes provided by Orange County Public Works
- 7 primary semantic land uses
- 5 real types of sensors considered[1,2,3]
- 1292 historical grab samples of anomalies from 30 different locations from 2006-2022
- 6 baseline comparison algorithms
- Measured number of anomalies detected, traceability, and node coverage

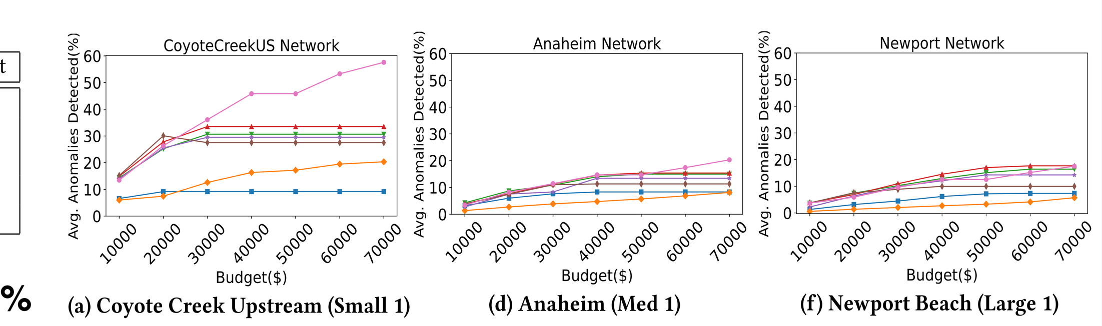


EPA SWMM Networks used for evaluation

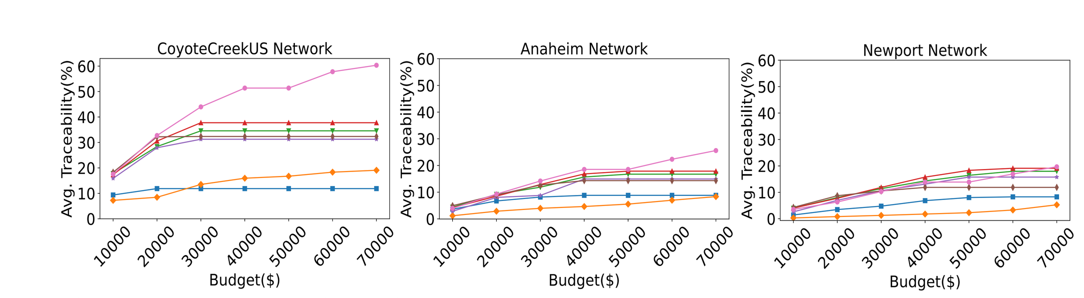
Table 1: Sensors considered in placement

Phenomenon	Accuracy	Hardware & Depl. Cost	Op. Cost
Turbidity	11.6%	\$100	\$300
Depth	1 mm	\$150	\$350
Temperature	$0.5^{\circ}C$	\$200	\$300
Electric Cond.	10%	\$150	\$300
Velocity	5 mm/s	\$150	\$350

- STEP detected ~35%, ~32% and ~1% more anomalies for the small, medium, and large networks than the best baseline
- The traceability provided by the STEP placement was ~30%, ~43% and ~3% better than the best baseline for the small, medium and large sized networks
- Additional results provided in paper



Number of anomalies detected for sampled small, medium and large network



Traceability for sampled small, medium and large network

[1]S Catsamas et al., *Characterisation and development of a novel low-cost radar velocity and depth sensor*. In SPN 2022.
[2] B. Shi et al. *A low-cost water depth and electrical conductivity sensor for detecting inputs into urban stormwater networks*. In Sensors 2021.
[3] M Wang et al., *An Innovative Low-cost Turbidity Sensor for Long-term Turbidity Monitoring in the Urban Water System*. In ICUD 2021.

Reference: Andrew Chio, Jian Peng, and Nalini Venkatasubramanian. 2023. STEP: Semantics-Aware Sensor Placement for Monitoring Community- Scale Infrastructure. In The 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '23), November 15–16, 2023, Istanbul, Turkey. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3600100.3623752

Acknowledgements: This work is supported by the UC National Laboratory Fees Research Program Grant No. L22GF4561, and National Science Foundation NSF Grants No. 1952247 and 2008993.