

Designing Community-Based Intelligent Systems for Water Infrastructure Resilience

Nalini Venkatasubramanian
University of CA, Irvine
nalini@uci.edu

Craig A Davis
C A Davis Engineering
cadavisengr@yahoo.com

Ronald T. Eguchi
ImageCat, Inc,
rte@imagecatinc.com

ABSTRACT

In this paper, we discuss how data-driven approaches using emerging IoT and machine learning based analytics can revolutionize the resilience and efficiency of urban water systems. Key challenges in creating a next generation water infrastructure includes issues of how and where to place instruments to gather a wide variety of information useful for improving operational efficiencies and for damage detection after major disasters. We discuss how an understanding of deployed infrastructure in diverse geographies and the dynamics of interconnected systems can help design more effective placement of technology solutions. We showcase recent work illustrating how knowledge of network structures and their behavior can help to more effectively instrument and gather operational data and how AI-based approaches utilizing geospatial data more effectively can help to maintain real-time awareness of system states which allows decision makers to more effectively monitor and control their systems.

CCS CONCEPTS

• Software and its engineering → Software organization and properties → Contextual software domains → Software infrastructure → Middleware

KEYWORDS

Cyberphysical water systems. Resilience, Geospatial information, extreme events

ACM Reference format:

Nalini Venkatasubramanian, Craig Davis and Ron Eguchi. 2020. Designing Community Based Intelligent Systems for Water Infrastructure Resilience. In 3rd ACM SIGSPATIAL Workshop on Advances in Resilient and Intelligent Cities (ARIC'20), November 3–6, 2020, Seattle, WA, USA. ACM, New York, NY, USA, 4 pages, <https://doi.org/10.1145/3423455.3430303>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org ARIC'20, November 3-6, 2020, Seattle, Washington USA

© 2020 Association for Computing Machinery.
ACM ISBN 978-1-4503-8165-9/20/11...\$15.00

1. Motivation

Water is a critical resource and key lifeline for communities worldwide; the generation, treatment, distribution and maintenance of water workflows is typically managed by local governments and water districts. Many of these critical infrastructures are now reaching their useful life; this coupled with increased complexity and loads makes them more vulnerable to failures. Despite the fact that most water purveyors do implement long-term capital improvement programs, the budgets associated with these programs are generally phased in over long periods of time, often decades, and complete system replacement may take over a century. Because of this, enhancing resilience - especially against natural disasters - is slow and sometimes ineffective as parts of these systems continue to age and degrade [1]. Growth in sensing and communication technologies accompanied with big data analytics today allows us to view water workflows as a community wide cyberphysical system (CPS). The water CPS involves a range of stakeholders (suppliers, distributors, consumers) - accessibility to information about the state of water networks at finer levels of granularity can help better manage system complexities and isolate failures more accurately. To design robust water systems, we argue that one must characterize the structure, components and operation of community water systems into a CPS-enabled platform. The information-centric approach helps provide insight into dynamics (e.g. varying demands, small/large disruptions) and improve infrastructure control and lifeline service availability.

In this article, we propose an integrative approach to enabling resilience in community water systems by combining model-driven approaches with data-driven platforms (Sec 2). In particular, we study platforms that integrate heterogeneous data from GIS systems and IoT-based devices instrumented in infrastructures along with physics-based modeling and simulation tools to create an understanding of the current operational state of lifelines. Key challenges that arise in this context include addressing cost/utility tradeoffs in where and how to deploy fine-grained sensors and IoT devices (Sec 3); techniques to gather and fuse information from heterogeneous, live data sources, to bring about operational efficiencies and disaster resilience (Sec 4). We finally discuss opportunities to exploit big-data technologies and AI-based predictive analytics for structural and process improvements in lifelines (Sec 5).

While our discussions are couched in the context of potable water systems, the principles apply broadly to other water infrastructures (wastewater, stormwater) and can be used to address interdependencies with other lifeline systems such as energy and transportation. For this, multisector academia-industry-government partnerships are critical.

2. AquaSCALE: An Integrated Middleware Approach for Resilient Water Systems

A recent effort by the authors, AquaSCALE [2], focuses on the design of a computational framework to model and explore the resilience and improved performance of community water infrastructures. In AquaSCALE, we leverage dynamic data from multiple information sources including IoT sensing data, geophysical data, human input, and simulation and modeling engines to accurately and quickly identify vulnerable spots in water networks. Such sensor-simulation-data integration platforms can assist in design time tasks (e.g. optimizing IoT device placement) and run-time tasks such as fault detection (e.g. isolation of leaky pipes). AquaSCALE is implemented as a cyber-physical-human middleware that transforms input raw sensor data streams to higher-level semantic streams to capture application concepts (e.g. locations of broken pipes).

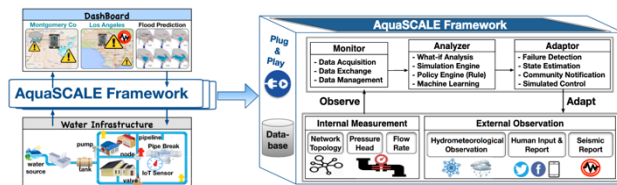


Figure 1: The AquaSCALE Framework

Different classes of stakeholders and decision makers can use information to better manage water networks. City planners can determine how to instrument water networks with monitoring devices to identify the vulnerable spots in community water infrastructures. The design also allows citizens to participate (provide and obtain) meaningful information, such as leak and flood awareness. AquaSCALE enables water agency operators to quickly locate problems and explore solutions to problems in cyberspace before instantiating them into a physical infrastructure. For example, large sections of water systems (usually entire pressure zones) may be shut down by pipe bursts, which cuts off access to water needed for dependent critical operations such as firefighting. An alternative solution to a complete shutdown might be to reroute the water to the impacted regions using alternate paths. Proactive planning and instantiation of alternate routes (i.e. through *seismic resilient networks* [3] and *resilient overlays*) during damage is an area that needs further exploration.

The AquaSCALE prototype system integrates modules for scenario generation, sensor data acquisition, model-driven simulations, ML-based analytics and dashboards to support decision-making. Given knowledge of the network structure

(from water agencies), robust simulation methods using commercial grade hydraulic simulators enhanced appropriately with IoT sensors and failure modeling was used to generate profiles of anomalous events. These profiles are then trained by diverse plug-and-play machine learning strategies to rapidly isolate anomalies and explore adaptation of network flows to mitigate impacts. The prototype builds upon multiple tools developed by domain experts - these include the industrial strength hydraulic simulation platform EPANET (from the Environmental Protection Agency, EPA) for simulating the behavior of water networks, G-WADI, a hydrological remote sensing platform and a novel adaptive Tweet Acquisition System (TAS) enables adaptive acquisition of tweets based on the situation. Using the flood model BREZO, we incorporate the hydrodynamics of flood propagation to depict the cascading impact of leak events. The integration of disparate data sources with new ML and AI techniques raises new interoperability challenges. Water infrastructure data utilizes multiple representation formats and interfaces; variations in network and data granularity and access restrictions pose further challenges to data sharing.

3. Design-Time Challenges for Deployment

Today, detection of anomalous events in civil infrastructures (e.g. pipe breaks and leaks) is time consuming and often takes hours or days [4]. Some leaks can go undetected for years resulting in large water losses and serious environmental impacts when the line eventually ruptures [5]. The goal of instrumentation is to reduce the number of lost services, the outage time, and cascading property losses resulting from damages to the water system. Since placing IoT devices in underground water networks has a high cost, it is important to develop intelligent sensor placement strategies that can reduce cost of deployment and maintenance while enabling maximum value, e.g. by reducing impact to communities under various failure events. Early efforts in this direction aim to illustrate the value of sensors [6] and explore optimal (but expensive) approaches to ensure network coverage [7].

We argue that a more holistic approach which integrates geospatial and community-level information in the infrastructure design process is required for effective technology deployment. In initial work [8], community structures (e.g. locations of critical infrastructures, public organizations, schools) and demographics are used to characterize “impact” - the degree to which failure events can impact societal processes. Such geosocial correlations are used to design an impact-driven algorithm to inform and guide sensor placement strategies. For this, we begin with utilizing a hydraulic model of a water distribution network (e.g. EPANET), to simulate potential failure events and develop flood maps resulting from each of these failure events. Using open-source services such as OpenStreetMaps, we map the water network to the surrounding community and obtain census demographic/population data as well as locations of critical

infrastructure – e.g. hospitals, senior homes, in the community and compute an impact score for each failure. The algorithm assigns a utility score to each potential sensor location based on the amount of impact that it would be mitigating as a function of the time taken to detect the failure as well as its impact score. These utility scores are used to drive sensor placement decisions. We evaluated the impact-driven sensor placement approach on multiple real-world water networks - the WSSC network (Montgomery County, USA), Richmond network (Yorkshire, UK), and the Wolf Cordera Ranch network (Colorado Springs, USA). Experimental results indicate an 80% reduction in impact while using a comparable number of sensors (to graph-theoretic approaches) thus showcasing the effectiveness of community impact in sensor placement.

In recent years, mobile water sensing technologies (e.g. SmartBall from Pure Technologies) have been used to augment in-situ sensing; these units can be inserted into and extracted from the water pipes through existing infrastructure like manhole covers, fire hydrants, etc. [9]. These mobile sensors flow along with the water while collecting information about the surrounding pipeline as well as the chemical composition of the water. The mobile sensors are dynamic and flexible since they can be inserted at any location at any time, and hence incur low operation costs. However, they lack large sensing ranges, and their movement is entirely dependent on the flow of water thus resulting in the need for probabilistic approaches to model movement. We developed and validated a hybrid insitu-mobile monitoring framework [9] for water distribution networks that leverages the impact model described earlier. Results demonstrated the effectiveness of our hybrid monitoring approach as compared to strategies that relied on just one type of sensor. We observed that the hybrid monitoring framework was able to achieve 79% reduction in impact with up to 68% greater cost efficiency compared to approaches using traditional coverage heuristics, and 30% reduction in impact while being 52% more cost efficient compared to approaches that attempt to minimize impact.

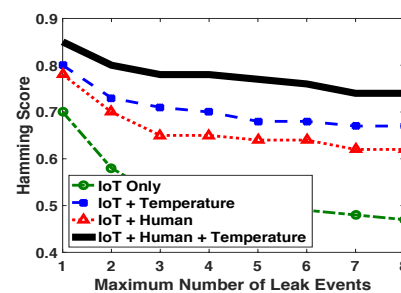
4. Techniques for Resilient Operation

In this section, we discuss how AquaSCALE's data-driven approaches can assist water agencies dealing with day-to-day operational issues (leak detection in pipelines) and support resilience to extreme events (e.g. earthquakes).

Modeling and detecting operational failures in water systems: Two types of operational failures are pertinent in potable water distribution networks - component failures resulting in water loss [2] and contaminant injection resulting in poor water quality [10]. Here, we focus on the case of pipe failures. The AquaSCALE approach integrates hydraulic models of water distribution networks, observations from external sources (remote sensing GIS), and human input using social media sources to model leak events and IoT measurements in the water infrastructure under failure scenarios. Our experience indicates that while IoT measurements by

themselves may work well to identify a single leak event, it may not be sufficiently accurate to isolate multiple concurrent failures; a combination of diverse information sources is useful here. For example, a significant temperature changes and pressure cycles can cause pipe breaks [5]. Coarse grained environmental information can be used to obtain ambient temperatures and fused with IoT data and human information to improve leak localization [2].

We explored a multiphase approach where robust simulation using an enhanced version of hydraulic simulators such as EPANET or WNTR (from Sandia National Labs and the EPA) were used offline to generate profiles of anomalous events. The offline robust simulations and fault profiles are used for rapid coarse fault isolation that in turn inform statistical approaches for fine-grained localization of failures. To perform fine-grained localization, AquaSCALE uses plug-and-play analytics to drive multiple machine learning based techniques for fault isolation at both offline and online phases. Offline, the objective is to train a set of classifiers to generate a robust profile model using measurements collected



from the water infrastructure. Given the number of available devices and network structure, AquaSCALE uses a robust clustering

technique to select a group of locations based on pressure head and flow rate readings. Multiple simulations are then executed under different leak conditions and a classifier profile model is learned. The classifier in AquaSCALE uses a hybrid approach named HybridRSL and is a combination of Random Forests (RF) and Support Vector Machines (SVM) via Logistic Regression, because RF and SVM remain robust with decreasing number of IoT sensors, and Logistic Regression has low variances and is less prone to overfitting. We conducted extensive evaluation of the proposed approaches under diverse failure scenarios using real water networks in MD and CA. Weather data from weather service datasites during the collection period were obtained - the east coast of the US experienced extremely cold temperatures, while the west coast experienced high precipitation due to El Nino effects. To validate the role of human reports, we collected/gathered 30 million “leak-related” tweets posted in the US during this period using a tweet acquisition system with data cleaning capabilities. The integrated approach allowed us to identify multiple leaks with high levels of accuracy (Hamming score of 87%) in reasonable time limits (minutes). The HybridRSL ensemble approach was able to handle varying amounts of data and multiple failures.

Water Infrastructure under Extreme

Events: Natural disasters and other types of hazards have resulted in different types of water service disruptions, and caused financial, social, environmental and human health consequences. The ability to maintain delivery of water supplies during and after catastrophic events is critical to ensure public safety and welfare.

Preliminary work has shown the value of AI-based techniques such as graphical models to perform state estimation, i.e. capture interdependencies in a complex water network topology as a structured probabilistic framework. Here, a factor graph representation is used to characterize the non-linearity of fluid flow in a water network that is dynamically altered by leaks, breaks and damage during extreme events such as earthquakes [11]. Scalability is a concern with large networks; techniques for network decomposition using articulation points followed by the distributed Gauss-Newton Belief Propagation (GN-BP) based inference, can deliver optimal estimates of the system state in near real-time. Using case studies from actual failure patterns documented during several California earthquakes, including the 1994 Northridge earthquake, we validated the robustness of our estimation techniques to identify faulty zones, infer service losses and disruptions accurately (MSE of 0.02) in a relatively fast manner (within 1s). Such analytical methods are applicable to extreme events including severe-cold and freeze-thaw conditions that result in multiple simultaneous breaks.

5. Emerging Directions

The resilience methods discussed above can be generalized to other lifelines – in particular, any system built underground, such as wastewater, storm drainage, natural gas, and liquid fuel systems. These spatially distributed networks can be monitored at different points within the system. As in water, existing sensor networks (SCADA), when stressed by extreme events (earthquakes) they fail to provide the necessary readings to accurately pinpoint where damage has occurred and to what extent. Adaptation of the technologies described in this paper can significantly improve the resilience of these systems during extreme events and in turn, improve community resilience of interdependent systems. A newly initiated effort, SWADE (Smartwater Data Exchange) aims to capture and manage interdependencies across multiple water systems (potable water, wastewater and stormwater networks) through flexible data representation and data exchange. SWADE allows sharing of raw data from agencies as well as the results of analytics, predictions and impacts.

We next discuss the role of robust simulations to project damage/failure impacts and introduce interventions and alerts. Using robust simulations, we have created large stochastic catalogs of earthquake simulations, or “event sets” that allow water system operators to evaluate system-level consequences for the entire network. A recent study for City of LA [12] indicates that knowledge of system risk from each

plausible event has allowed the City to focus on those earthquakes that are likely to cause the most serious disruptions. The advances above, along with cloud computing, has empowered water companies to take a more realistic and meaningful look at potential mitigation measures that will effectively reduce future risks from damaging earthquakes.

Robust simulations are also a critical tool to capture multihazard and interdependent impacts of climate change and address solutions. During earthquakes, a serious concern for communities is the occurrence of urban fires, often caused by broken gas lines (San Francisco Fire, 1906). A functioning water system is key to extinguishing these fires, and a major priority for water system operators is to quickly identify where system damage limits the water flow and pressure and isolate these damages to prevent total water drawdown. Intense wildfires also draw significant water flow, potentially exceeding system capacity, and must be dealt with through system monitoring and requiring non-essential customers to ration usage and reserve flow for firefighting. The integrative middleware approach, presented here, can further be extended to monitor multiple lifelines - water, natural gas, and transportation systems to ensure a safe and resilient infrastructure.

ACKNOWLEDGMENTS

This work was supported by the NSF CNS1528995 and CNS1952247). We acknowledge the AquaSCALE research team and partners, WSSC and LADWP for input and feedback.

REFERENCES

- [1] C.A. Davis. 2018. Assuring Seismic Resilience for the Los Angeles Water System. In Proceedings 11th National Conference on Earthquake Engineering, Earthquake Engineering Research Institute. June 25-29, Los Angeles, CA,
- [2] Q. Han, et al. "Toward An Integrated Approach to Localizing Failures in Community Water Networks". IEEE ICDCS, 2017.
- [3] Davis, C. A., 2018, "Creating a Seismic Resilient Pipe Network for Los Angeles," ASCE Pipelines 2018, UESI, Toronto, CN, ASCE.
- [4] Bardet, Jean-Pierre, and Richard Little. "Epidemiology of urban water distribution systems." *Water Resources Research* 50.8 (2014): 6447-6465.
- [5] Bardet, J.P., et al. 2010. Expert Review of Water System Pipeline Breaks in the City of Los Angeles during Summer 2009. Report to the Steering Committee on Water Pipeline Breaks of the City of Los Angeles. April 9, 2010. http://clkrep.lacity.org/onlinedocs/2009/09-2232-s1_rpt_cao_04-13-10.pdf.
- [6] Kartakis, Sokratis et al. "Waterbox: A testbed for monitoring and controlling smart water networks." Intl Wkshp on CPS for Smart Water Networks. 2015.
- [7] Perelman, Lina Sela, et al. "Sensor placement for fault location identification in water networks: A minimum test cover approach." *Automatica* 72 (2016).
- [8] P. Venkateswaran, Q. Han, R. Eguchi, and N. Venkatasubramanian. "Impact Driven Sensor Placement for Leak Detection in Community Water Networks". In IEEE ICCPS 2018.
- [9] P. Venkateswaran et al , Augmenting In-situ with Mobile Sensing for Adaptive Monitoring of Water Distribution Networks, in ICCPS, 2019.
- [10] Q. Han et al. . AquaEIS: Middleware Support for Event Identification in Community Water Infrastructures. In ACM Middleware 2019.
- [11] Q. Han et al. "Enabling State Estimation for Fault Identification in Water Distribution Systems under Large Disasters". in 37th IEEE Intl Symp on Reliable Distributed Systems (SRDS), 2018.
- [12] Y.Lee et al. , "Earthquake Scenario Development for Seismic Performance and Risk Evaluation of the City of Los Angeles Water System Pipeline Network," 11th Intl Exchange Forum: Seismic and Multi-Hazard Preparedness and Response Practices for Drinking Water Utilities, 2019, Los Angeles, CA.