University of Cyprus
Electrical and Computer Engineering Department

Fault Diagnosis and Security Monitoring in Water Distribution Systems

Demetrios G. ELIADES

PhD Thesis Proposal

Supervisor: Prof. Marios M. POLYCARPOU

Abstract: Water resources management is a key challenge that will become even more crucial in the years ahead. Water distribution systems are responsible for delivering clean water to consumers, and have an important role in sustaining certain vital societal functions. When a system fault occurs, such as water contamination or a pipe break, these societal functions may be affected negatively. In the previous years, various aspects of the security monitoring problem in water distribution systems have been examined; in addition, robust fault diagnosis algorithms have been developed within a system-theoretic framework. An open research area is the formulation of a system-theoretic framework suitable for fault diagnosis and security monitoring in water distribution systems; this is the general goal of this work. In specific, this work has four objectives. The first objective aims to formulate the monitoring and control problem of water distribution networks, in a framework suitable for sensor placement and fault diagnosis. The second objective is to find those locations in a water distribution network, where on-line quality sensors should be installed, in order to minimize the risk of a severe damage on the population; a special case is the problem of manual quality sampling scheduling, for finding where and when to take water samples to check its quality. The third objective is to design fault detection algorithms, so that a contaminant substance is detected by monitoring its reaction dynamics, by using a model-based fault diagnosis approach. An adaptive approximation model, such as a neural network, is activated after a fault has been detected, to learn the unknown fault dynamics. Furthermore, the source location of the contamination fault is estimated by considering the previous and future hydraulic dynamics. Finally, the fourth objective is to design fault accommodation algorithms which change the disinfectant concentration controller input, to accommodate the contamination fault and return the system to safe operation.

Biography: Demetrios G. Eliades received the *Diploma in Electrical and Computer Engineering* in 2004 from the Aristotle University of Thessaloniki, Greece and the *MSc in Artificial Intelligence* in 2005 from the University of Edinburgh, Scotland. He is a PhD candidate in Electrical Engineering at the University of Cyprus and a research assistant in KIOS Research Center for Intelligent Systems and Networks. He co-authored the book “Electronic Voting: Reality and Opportunities” and is the author/co-author of 4 journal articles and 11 conference papers. His background is in computational and artificial intelligence and their applications, and his research interests include water systems, fault diagnosis, optimization and evolutionary computation, embedded systems, pattern recognition and classification, e-learning and multi-agent systems. His research has been partially funded by the Cyprus Research Promotion Foundation and the University of Cyprus (15-month project *YGIESNERO* and 36-month project *ACHELOOS*). He is a student member of IEEE, IET, and is a registered engineer in ETEK; in addition, he serves as a Technology Editor in the monthly magazine *Synchrone Apopse*.

Fault Diagnosis and Security Monitoring in Water Distribution Systems

Demetrios G. Eliades

PhD Thesis Proposal

I. Introduction

THE distribution system responsible for delivering drinking water to consumers is part of the critical infrastructure and it is essential for certain vital societal functions. As a result, faults affecting the operation of a drinking water distribution system may have a negative impact on the society. Water providers are responsible to guarantee the safe and economic operation of the water distribution system, so that it complies with quality specifications and legislative requirements. Regulations require water providers to control and monitor the hydraulics and the water quality in the water distribution network which they operate, in order to deliver adequate disinfected water to all consumers. To satisfy this, water providers must collect hydraulic and quality data at various locations in the network (either manually or by using sensors) and control appropriately the system. Through this, water providers are able to detect faults related with the hydraulic dynamics (pressures, flows) or quality dynamics (such as disinfectant and contaminant concentration).

Hydraulic and quality faults may affect the water distribution's normal operation. Pipe breaks, leakages or pump/valve malfunctions relate with hydraulic faults and may cause, as a consequence, quality faults. Quality faults may be due to natural contamination (e.g. bacterial growth), or due to faults at the disinfect concentration control; these faults may pose significant threat to the consumer health. For example, in the town of Nokia, Sweden, in November 2007, due to an accident, sewage water was injected into the town's drinking water distribution, affecting thousands of people and forcing the authorities to impose a complete ban on all water usage. A special case of quality faults are due to intentional contamination; this differs from natural or accidental failure since it can be assumed that the malicious injection of a contaminant in the distribution network will occur at the most neuralgic location and at the most critical time instance, to cause the greatest damage. Water providers should be able to detect the faults fast enough and apply appropriate measures for correcting them, so that the consumers are not significantly affected.

Fault diagnosis and accommodation describes the process of detecting and isolating a fault by applying various techniques to monitor the changes in the states, learn the fault dynamics and adapt the input so that the system returns to safe operation. In addition to that, security monitoring describes the supervision of the distribution network in order to minimize the potential economic losses and the damages inflicted on the consumers as a result of a fault or a malicious attack. Security monitoring in water distribution systems is an area of increasing research interest, and requires the synergy of various fields, such as control systems, water engineering, optimization and risk management.

The fault risk in water distribution systems motivates our research. Our goal is to develop fault diagnosis and security monitoring methodologies so that water providers may provide better quality of service, while reducing the consumer exposure risk to faults. In the previous years, various aspects of the security monitoring problem in water distribution systems have been examined; in addition, robust fault diagnosis algorithms have been developed within a system-theoretic framework. An open research area is the formulation of a system-theoretic framework suitable for fault diagnosis and security monitoring in water distribution systems; this is the problem we investigate. This research can be broken down to four specific objectives, as follows:

I. To formulate the problem of monitoring and control in water distribution networks, using a mathematical framework which is suitable for the sensor placement and the fault diagnosis problem. The hydraulic and quality dynamics are coupled with the impact dynamics characterizing the "damage" of a contamination in a water distribution system, to construct a number of realistic fault scenarios.

II. To find the locations in a water distribution network where quality sensors should be installed, as well as where and when to conduct manual quality sampling, in order to monitor the water chemical characteristics. A number of risk metrics will be considered and a multi-objective optimization problem will be formulated and solved.

III. To design fault detection and isolation algorithms suitable for water systems, following a model-based analytical redundancy approach. In specific, fault detection will be achieved by comparing the magnitude of the state estimation error with an adaptive threshold, when certain detectability conditions are satisfied. Identification algorithms will be designed to find what substance type has caused the fault, and an adaptive approximation model, such as a neural network model, will be activated to approximate the unknown dynamics. Isolation algorithms will be designed to find the contamination source location within a network.

IV. To design fault accommodation algorithms which utilize the approximated fault dynamics so that the system returns to normal operation. Control laws will be developed in order to regulate the disinfectant concentration within some desired levels, and in the case of the fault, adapt the control law by using the approximated dynamics, in order to accommodate the contamination fault.

This proposal presents published results in relation to Objectives I and II, as well as some preliminary results in relation to Objective IV. The remaining parts of the research are currently under development and final results will be presented within the following months. The proposal is organized as follows. Section II provides background information on faults affecting water systems, as well as an overview of the state-of-the-art on the sensor placement and fault diagnosis research in water systems. Section III formulates the problem of monitoring and control of water distribution systems and Section IV presents the sensor placement problem for water distribution networks. Section V introduces the fault detection and isolation problem in water systems and in Section VI, the fault accommodation problem in water systems and the disinfectant concentration control problem are discussed. Finally, in Section VII, conclusions are drawn and the impact of this work is discussed.

II. State-of-the-art Overview

A. Faults in Water Distribution Systems

Modelling methodologies for hydraulic and quality dynamics and their faults, have received significant attention during the last decade [1], [2].

Hydraulic dynamics in water systems correspond to pressures at nodes, flows in pipes and water levels in storage tanks. Usually, an iterative method is applied to solve a set of mass and energy conservation equations derived from the water distribution network topology and structural characteristics [3]. Although these iterative methods are able to capture the hydraulic behaviour of a distribution system, the uncertainty is significant due to the unknown and time-varying water demands by the consumers, which cause the hydraulic dynamics to fluctuate throughout the day; these demand outflows are usually considered as periodic.

Quality dynamics in water systems correspond to the concentration of various contaminant or disinfectant substances, as well as other water chemical parameters, such as pH or turbidity. Contaminants and disinfectants travel along the water flow according to the advection and reaction dynamics. Advection is the transport mechanism of a substance in a fluid, which can be modelled as a partial differential equation. In a pipe, the advection equation is usually modelled as a first-order hyperbolic partial differential equation, which can be solved using various numerical methods [4]. This describes the change of a substance concentration in space and time. Reaction dynamics describe the change in the substance concentration due to decay or growth, as well as the reaction with other substances that may exist. These equations combined describe the propagation dynamics.

In water research, various numerical schemes have been applied for modelling the quality dynamics in distribution systems. In specific, Eulerian and Lagrangian [5], [6] schemes have been proposed. In general, the Eulerian schemes segment the network into a number of fixed finite volumes, at which the substance is propagated and measured at each time step. On the other hand, in the Lagrangian schemes, the substance is measured in discrete parcels of water, which move along the pipes [6]. A state-space representation for the propagation dynamics was presented in [7]. An event-based Lagrangian scheme is considered for solving the propagation equations in the open-source software EPANET, provided by the Environmental Protection Agency in the USA [8]. The software has recently been extended to integrate chemical reactions with multiple chemical substances [9].

When a fault occurs in a water distribution system, it may affect the hydraulic dynamics, the quality dynamics or even both. There are two categories of faults affecting water distribution systems: hydraulic faults and quality faults.

Hydraulic faults can be due to leakages, pipe bursts, blocked pipes or malfunctioning pumps and valves [10]. A burst is an abrupt fault requiring immediate attention, and it is usually easy to isolate. On the other hand, a leakage is an incipient fault and is, in general, difficult to detect and isolate. Hydraulic faults can be detected by analyzing flow and pressure data acquired from SCADA systems. Hydraulic faults may cause quality faults, since contaminants may infiltrate the distribution system.

Quality faults may occur due to water contamination by certain substances, usually chemical, biological or radioactive, which travel along the flow of water and may exhibit decay or growth dynamics. Certain disinfectants, such as chlorine, are used in prescribed concentrations, to maintain the drinking water quality by preventing bacteria growth and neutralizing chemical agents [10]. According to the World Health Organization, a free chlorine residual concentration must exist in drinking water distribution systems, with minimum target concentration $0.2\ \frac{mg}{L}$ at the point of delivery and $0.5\ \frac{mg}{L}$ for high-risk circumstances [11]; it is common practice is to supply water with a few tenths of a milligram per litre of chlorine residual. Contaminants can also be intentionally injected in the network, during a well informed terrorist attack [12]; this is an important problem whose damage may be reduced by using security monitoring mechanisms such as specialized on-line sensors.

B. Sensor Placement for Water Quality Monitoring

The problem of where to place facilities[1] or sensors in order to keep certain objectives and constraints satisfied within a network, has been examined in various research disciplines such as operational research, combinatorial optimization, systems theory and control.

The "Set Covering" method was one of the first mathematical formulations of the problem and it has been applied in various fields, such as facility location [13]. According to this approach, an integer optimization program is formulated for determining a set of nodes from a topological graph to install facilities, so that all the remaining nodes are next to at least one facility. A related approach is the "Maximal Covering" formulation described in [14] for computing a set of nodes which maximize the population served in an area within a certain distance. A similar formulation was considered in [15] for selecting the locations to install water quality sensors in drinking water distribution systems, so that the largest volume of water consumed is examined. The authors proposed a scenario-based approach which segments a day into time-periods, corresponding to different flow patterns, and the optimization was solved for all scenarios simultaneously. Following this formulation, for solving bigger networks, other researchers examined the use of heuristics [16] or genetic algorithms in [17]. A multi-objective weighted-sum extension of [15] was presented in [18] which considered certain physical network characteristics and time-delays.

A mathematical formulation suitable for the security issues related with the location selection, is the "p-median" [19], with the objective to minimize the "maximum distance" of a facility. A similar formulation was examined in [20], [21] for water distribution systems. By considering a number of contamination

[1] In this context, *facility* can be a public service such as a police or a fire station.

scenarios and their impacts, the authors formulated a mathematical program to minimize the average "contamination impact". A multi-objective extension was further examined in [22]; however for solving the mixed-integer optimization program, significant computational power was required. In [23] the authors proposed a modification of this formulation, to determine locations for monitoring disinfection byproducts.

Within the water resources management community, the design competition of the "Battle of the Water Sensor Networks (BWSN)" in 2006, instigated significant research interest and discussion on security issues of water distribution systems [12]; the task was to find sets of locations to install quality sensors using two real benchmark networks, so that a number of objectives are optimized under various fault scenarios. Most of the participants formulated a multi-objective integer optimization program [24]. According to the BWSN instructions, only the average impact of observed faults was considered; furthermore, different fault scenarios could have different solutions.

C. Disinfectant Concentration Control in Water Distribution Systems

From a control systems perspective, both the hydraulic and water quality dynamics in a DWDN are influenced by disturbances due to unknown water demands. Large transport time-delays and storage tanks affect the system with respect to quality control. Other factors such as variations in the temperature or faults may affect the system. Controlling disinfectant residuals has gained attention in recent years [25]–[48]. A number of studies have examined the problem within an off-line framework for finding a disinfectant mass input scheduling algorithm using optimization techniques such as linear programming [26], [33], least squares [40], genetic algorithms [34], [36], [39] and goal programming [25]. Optimization of water pumps operation has also been considered in relation to the quality control problem in [37], [41], using non-linear optimization and genetic algorithms respectively.

A robust model predictive control scheme [29]–[32] and an adaptive control framework considering periodic variation of demands [38], [42] have also been examined. Both are promising approaches regarding the on-line quality control problem.

The Model Predictive Control (MPC) framework [49] is specifically suitable for systems with time-delays and hard constraints. The MPC approach is based on explicitly using a system's model to predict the outputs at a future time, and then computing the control sequence by minimizing an objective function. In [29], the uncertainties in the time-varying parameters, errors in model structure, actuator and measurements errors have been modelled within a set-bounding approach. Safety zones are applied that narrow the constraints in order to deal with uncertainties. From a different approach, the adaptive control framework presented in [38] uses Fourier series to capture the periodic nature of the system's dynamics. By using this information, the time varying coefficients of the identification model are replaced with a number of sinusoid terms with weights which are estimated on-line.

D. Fault Diagnosis and Accommodation in Water Systems

Fault diagnosis is an important part of the modern control theory and practice [50], and is usually dividend into the following processes: fault detection, isolation and identification. Fault detection traditionally is achieved by monitoring certain system parameters and if a fault caused these parameters to go above or below certain threshold values, an alarm is raised [51]. This method however is influenced by the system noise, and can trigger false alarms; moreover, a fault can affect more than one system states, thus making it difficult to isolate and identify it, i.e. to determine its type and its magnitude [50].

In the previous years, model-based analytical redundancy approaches have received significant attention in the fault diagnosis research. By using a mathematical representation of the system dynamics, residual quantities are calculated; this signal is an indication of whether a fault has occurred in the system, even when the states are subject to uncertainties [50]. Recently, the exploitation of computational intelligence techniques in combination to model-based fault diagnosis has been considered. In [52], the authors presented a robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems. By using on-line state measurements, a fault is detected by comparing the threshold of the maximum uncertainty with the magnitude of the estimation error. After a fault has been detected, an approximation estimator is activated to learn the fault dynamics, and the fault is isolated from a bank of possible faults. Fault accommodation in the system can be achieved by suitably reconfiguring the control signal, so that the fault dynamics are compensated [53].

In water systems, contamination fault detection has yet to receive significant attention, as it is commonly assumed that quality sensors are specialized and are capable of measuring a contaminant substance concentration. In practice contamination fault detection may be achieved by measuring how some standard chemical characteristics, such as chlorine concentration and pH, are affected due to a reaction with a contaminant. Some empirical studies have been performed which demonstrate how different contaminants affect these parameters [54]. Moreover, the problem of detecting the source location of a contamination has been examined, mostly from an algorithmic viewpoint, by using various quality modelling techniques [55], [56].

III. Problem Formulation for Monitoring and Control of Water Distribution Systems

In this section we formulate the problem in a framework suitable for monitoring and control, by using a state-space representation of the hydraulic and quality dynamics. The state-space equations are based on the first-order hyperbolic partial differential equations which describe the advection and reaction of a substance in a water distribution system. Furthermore, we formulate the impact dynamics in a state-space formulation which couples the quality dynamics with states corresponding to metrics measuring damages due to a fault. We construct a finite set of realistic fault scenarios parameters, depending on the fault specifications, and based on some operational constraints, we further construct the sets of locations and time instances where and when quality sampling can be conducted. Through a series of simulations of the dynamic equations, we build the overall-impact matrix, which corresponds to the damage caused by a fault in the system until it is detected at a certain location and the system has stopped.

A. Hydraulic and Quality Dynamics

Consider a water distribution network composed of pipes, junctions and water storages. The topology of this network can be represented as a graph with edges corresponding to pipes, and nodes corresponding to junctions and water storages. For modelling purposes, each pipe in the network is *a priori* virtually segmented into a number of finite volume cells. The set of all finite volume cell indices is $\mathcal{V}$ and its subset of all node indices is $\mathcal{V}_0$. Within the i-th cell or junction, let $x_i(k)$ denote the average concentration of a certain contaminant at discrete time k. The vector $x(k)$ is the state of the contaminant concentration dynamics. It is assumed that the injection of a substance can occur at any one of the nodes contained in $\mathcal{V}_0$.

The advection-reaction equations [4] describing the propagation of a contaminant in a water distribution network can be expressed in a state-space formulation:

$$x(k+1) = A(k)x(k) + g(x(k)) + B_F\varphi(k),$$

where $A(k)$ is a matrix which characterizes the advection dynamics, and $g(\cdot)$ is a function which describes the reaction dynamics of the contaminant. Let B_F be a matrix describing the locations of the injected contaminant. The function $\varphi(\cdot)$ describes the change in the contaminant concentration due to a substance injection.

When a substance enters a pipe in which a fluid flows, the substance moves along with that flow. Inside a pipe, and by neglecting axial dispersion, a first-order hyperbolic partial differential equation describes the change in space and time of the contaminant concentration in the fluid. Numerical methods, such as the "Finite Volume Method" [4] can be employed to approximate the set of partial differential equations which describe the advection dynamics. In water storage nodes, the state is the average contaminant concentration in the finite volume. For junction nodes, the state is the concentration at the point of junction outflow.

The concentration of a substance in the water may exhibit growth or decay, or it may react with other substances flowing in the bulk water or residing at the pipe walls. For instance, if the substance is radioactive, first-order decay is considered.

Contamination faults can occur due to the intentional (malicious) or accidental injection of a contaminant at some of the nodes in the water distribution network. The function $\varphi(k)$ corresponds to the signals of the injected contaminant concentrations. These signals have a certain start time and duration, and are non-negative. Function $\varphi(k)$ can be represented through a number of linearly parameterized basis functions, such as pulses or radial-basis functions. The motivation behind the use of a linearly parameterized form of the fault function, is that it simplifies the process of computing a finite set of fault parameter matrices, either through grid sampling or otherwise. This will be useful during the solution methodology for sensor placement.

From a practical viewpoint, the contaminant injection is measured in terms of the injected contaminant mass per unit time ($\frac{mg}{min}$), while the state-space formulation is described in terms of contaminant concentration ($\frac{mg}{L}$). The fault function $\varphi_i(k)$ affecting the i-th node can be therefore expressed as the fraction of contaminant mass injection rate and the water inflow rate at that node.

B. Impact Dynamics

For each node where water is consumed, an impact value (which measures the "damage" caused by a contamination fault), can be computed at each time step. This impact value can be expressed by using epidemiological terms (e.g. how many people are affected), or by using economic terms (e.g the cost of productivity loss). Other impact measures which can be considered, is the consumed volume of contaminated water which exceeds a certain concentration threshold [12]. In general, the impact of a fault depends on the volume and contaminant concentration of the contaminated water consumed. This can be described by a dynamic equation, as we show next.

Let $\xi(k)$ be the impact state vector which describes the "damage" caused at each demand node, at discrete time k, after a contaminant has been injected somewhere in the network. For the demand node w_i, a state-space representation of the impact dynamics is given by

$$\xi_i(k+1) = \xi_i(k) + f_\Xi(x_{w_i}(k), d_{w_i}(k)),$$

where $d_{w_i}(k)$ is the outflow demand (in $\frac{m^3}{s}$) at demand node w_i, and $f_\Xi(\cdot)$ is non-negative function which characterizes the impact increase at each time step. Furthermore, let $f_\Omega(\xi(k))$ be a function which returns a number characterizing the overall impact, or total "damage", which has been caused by a certain contamination fault.

C. Contamination Scenarios

For some security-related problems in water distribution networks, it is useful to consider a number of representative fault scenarios, within a specific time-period, while the system is assumed to be operating under normal conditions. A scenario is comprised of two elements: a) the parameter matrix of the basis functions which describes the fault signal, and b) the time-delay in shutting down the system.

Let $\mathcal{Z}$ be the set of all fault parameter matrices whose elements are within certain bounds. In addition we define $\mathcal{T}_d$ as a set of various discrete time delays considered, for stopping the system after a fault has been detected. We define the set

$$\mathcal{H} \subset \{(\Theta, t_d) \mid \Theta \in \mathcal{Z}, t_d \in \mathcal{T}_d\}$$

as the finite subset of realistic fault scenarios scenarios two-tuples, with size $N_h = |\mathcal{H}|$. These scenarios can be selected through grid or random sampling.

D. *Quality Sampling Measurements*

For the security problems examined in this work, we assume that from the nodes indices set $\mathcal{V}$, a subset $\mathcal{V}_s \subseteq \mathcal{V}$ corresponds to candidate locations for installing sensors or for conducting manual sampling (at certain times). It is useful to define $\mathcal{T}_m$ as the set of discrete time instances when manual sampling can be performed (e.g. during working hours), within one period (i.e. one day). From these, we define the two-tuples set $\mathcal{Q}_m$, corresponding to candidate sampling node indices and sampling times, given by

$$\mathcal{Q}_m = \{(s, t_s) \mid s \in \mathcal{V}_s, t_s \in \mathcal{T}_m\}.$$

In addition, we define the set $\mathcal{Q}_s$, of candidate sensing node indices where sensors can be installed, such that

$$\mathcal{Q}_s = \{(s, t_s) \mid s \in \mathcal{V}_s, t_s \in \mathcal{T}_s\},$$

where $\mathcal{T}_s$ is the finite set of time instances when the on-line sensor takes measurements throughout a day; this depends on its sampling time. If the measurements are continuous, we consider the set $\mathcal{T}_s$ as empty.

Finally, we define $\mathcal{Q} = \{\mathcal{Q}_s, \mathcal{Q}_m\}$ as the set of all feasible sensing nodes and manual sampling times.

E. *Detection and System Stop Time*

For each fault scenario, in order to evaluate the damage caused, it is useful to compute the detection time when an on-line quality sensor or manual sampling is applied at a node.

In practice, if a fault is not detected using sensor technologies or manual sampling, it will propagate for some discrete time T_{other}, after its first occurrence, until it is detected through other means, such as customer complaints, hospitalizations, etc. In addition, we define a discrete maximum time T_{max}, so that for any fault scenario, its impact will not increase after that time.

For a scenario fault $h \in \mathcal{H}$, a sampling $q \in \mathcal{Q}$ and a concentration threshold $\epsilon > 0$ above which a contaminant is assumed to be detected, we construct the function $\zeta(h, q, \epsilon)$, which maps its parameters to a time step at which the fault has been detected and the system has been stopped. Depending on the scenario, the accommodation time-delay is added to the fault detection time, to compute the fault detection and system-stop time. In the case of installed sensors, the fault detection time corresponds to the minimum time when a certain fault changes the state at a certain location above the ϵ threshold; in the case of manual sampling, this corresponds to the sampling time when the concentration of a contaminant was above the ϵ threshold. Finally, when no fault was detected through these means, the fault detection time step is given by the injection time plus the minimum time T_{other}.

In practice, fault detection time may be different for an on-line sensor and manual sampling at a certain node. The difference between them is that the on-line sensor will detect a fault as soon as a certain contaminant concentration threshold is surpassed; on the other hand, manual sampling will detect a fault only if the contaminant concentration at the sampling node is above a certain concentration.

F. *Overall Impact*

The overall impact is the "damage" caused by a contamination fault, measured through certain impact metrics, such as the number of people infected or the volume of polluted water consumed. We compute the overall impact for each fault scenario with respect to each detection and system halt time. The impact of a fault with parameter matrix Θ at time k will be denoted as $\psi(k; \Theta)$.

For a certain impact metric, for $N_q = |\mathcal{Q}|$ samplings and $N_h = |\mathcal{H}|$ scenarios, we define $\Omega \in \mathcal{Z}$ as the overall impact matrix, of size $N_h \times N_q$; the (i, j)-th element of this matrix is given by

$$\Omega_{(i,j)} = \psi(\zeta(h, q, \epsilon); \Theta)$$

where $h = (\Theta, t_d) \in \mathcal{H}$, $q \in \mathcal{Q}$ and $\epsilon > 0$.

A number of different impact metrics can be considered; for each, a different overall impact matrix is constructed according to our methodology. For N_ξ overall-impact metrics, we compute N_ξ overall-impact matrices, which belong to the set $\mathcal{O} = \{\Omega_1, ..., \Omega_{N_\xi}\}$.

G. Results

The work presented in this section has been published and presented in peer-reviewed journals and conferences [57]–[60].

IV. Sensor Placement for Water Distribution Networks

In this section we discuss a sensor placement methodology in water distribution networks. First, we examine the general quality monitoring problem using a graph-theoretic approach, for determining how many and at which locations, quality sensors should installed, to maximize the measurement redundancy while minimizing the installation costs. Afterwards, we utilize the results from the previous section, in order to find the neuralgic locations in the network which must to be physically secured, as they are considered as possible locations of faults with extreme consequences. We then formulate and solve the risk-oriented, sensor placement (and manual sampling) problem, in a multi-objective optimization framework.

A. Quality Monitoring: A Maximum Redundancy Approach

As a first step towards security monitoring, we address the problem of selecting locations in the network for on-line quality monitoring. Although water providers are obligated to monitor the quality in the distribution network which they operate, there exist no guidelines for deciding where to measure water quality within the network. In [15], a "Maximal Covering" method was proposed. Water arriving at a node has originated from another node, unless the node is a water source; the idea is to compute "how much" each node contributes to the other nodes and to neglect those nodes with small contribution. Their work proposed a framework for water network-based approaches, which solves the problem of maximizing the monitored volume of water when the number of on-line sensors is *a priori* known.

We examine a variation of the problem posed by [15]: find the minimum number of on-line monitoring sensors required for covering the distribution network, and at which locations to install them, while maximizing the coverage redundancy. Redundancy is important in the case when a contamination fault at a certain node is not detected by the nearest quality sensor, and the contaminant propagates undetected. A binary "coverage" matrix H can be constructed by considering the average hydraulic dynamics; $H_{(i,j)} = 1$ if the water contribution of node i to node j is above a pre-specified percentage; otherwise, $H_{(i,j)} = 0$.

In addition to maximizing the coverage redundancy, the formulation takes into consideration weights corresponding to installation costs and node significance (e.g. depending on the demand volume). It is important to note that the proposed methodology can easily be extended to other relevant systems, such as power distribution networks.

We assume that it is feasible to install one or more on-line sensors at the node locations in the water distribution system. Let vector N_r be the maximum redundancy possible, i.e. the maximum number of nodes which are able to measure a sufficient water quantity from a certain node; this can be computed through $N_r = H \cdot \mathbf{1}^\top$, where $\mathbf{1} = [1, 1, ..., 1]$.

The optimization objective function $J(\chi)$ is formulated as

$$\begin{aligned} J(\chi) &= \gamma(N_r - H\chi)^\top R_s(N_r - H\chi) + \chi^\top R_c\chi && (1) \\ &= \frac{1}{2}\chi^\top(2\gamma H^\top R_s H + 2R_c)\chi + (-2\gamma N_r^\top R_s H)\chi + \gamma N_r^\top R_s N_r, && (2) \end{aligned}$$

where χ is a binary vector of size N_χ, such that $\chi_i = 1$ if a sensor is placed and $\chi_i = 0$ if not, for $i = 1, ..., N_\chi$; R_s, R_c are diagonal matrices whose non-zero elements correspond to the "significance" and "installation cost" weights of each node respectively; $\gamma = (N_r^\top R_s N_r)^{-1}$ is a normalizing factor. When all nodes are equally significant, and have the same installation costs, R_s and R_c are identity matrices.

The constant term in the optimization function can be neglected, and the optimization problem can be formulated as an integer quadratic program, such that

$$\begin{aligned} \underset{\chi \in \{0,1\}^{N_\chi}}{\operatorname{argmin}} \quad & \tfrac{1}{2}\chi^\top(2\gamma H^\top R_s H + 2R_c)\chi - 2\gamma N_r^\top R_s H\chi && (3) \\ \text{subject to} \quad & H\chi \geq b. && (4) \end{aligned}$$

The non-negative integer vector b corresponds to the minimum redundancy requirement for each node. The problem is solved using off-the-shelf Mixed Integer Quadratic Programming algorithms. Overall, this formulation will give priority to redundant coverage to nodes with the greater significance. The drawback of this algorithm is that it does not explicitly take into consideration the hydraulic and quality dynamics, nor the fault impact dynamics. In the following paragraphs, we examine the sensor placement problem in a security-oriented framework.

B. Neuralgic Locations

The first part of any security scheme in water distribution networks is to determine the locations in the water distribution network which could be considered as "high-risk" for contaminant injection, so that proper action is taken in order to secure them through physical means. For this we need to decide on a representative impact metric; a suitable choice is find the top worst-case scenarios. For a specific overall-impact matrix $\Omega \in \mathcal{O}$, we will calculate the maximum scenario impact $\overline{\omega}$, given by

$$\overline{\omega} = \max_{i=1,\ldots,N_h} \max_{j=1,\ldots,N_q} \Omega_{(i,j)},$$

where $\Omega_{(i,j)}$ is the (i,j)-th element of Ω, N_h is the number of contamination scenarios and N_q the number of feasible sensing nodes and manual sampling times.

From this, we can compute the set of the worst-case scenarios; for example, the indices set $\mathcal{Y}_0$ of the worst-5% faults is computed by

$$\mathcal{Y}_0 = \left\{ i \mid \max_{j=1,\ldots,N_q} \Omega_{(i,j)} \geq 0.95\,\overline{\omega},\ i \in \{1,2,\ldots,N_h\} \right\}.$$

Thus, from the scenario indices included in $\mathcal{Y}_0$, it is easy to construct a list of the most neuralgic locations in the network which need to be physically secured. However, due to the nature of the water distribution networks, certain contamination faults can override physical security measures; it is thus imperative to use an extra layer of protection by installing a number of on-line sensors at different locations in the network.

C. Risk-Oriented Sensor Placement and Manual Sampling

Deciding where to install quality sensors, and also where and when to perform manual sampling in a water distribution network, are problems with non-trivial solutions, which need to be addressed. By considering the mathematical formulation discussed in the previous section, the problem can be reduced to risk optimization.

In order to address the problem of security, it is important to have an understanding on what "risk" is and how it can be quantified. In general, risk is the possibility of an unpredictable future event that will result in losses, thus preventing the serving organization from meeting certain goals [61]. Risk has been examined in many fields and especially in the financial and operational research literature. In financial practice, the most commonly used risk-objective is the "Value-at-Risk" (VaR), which represents the maximum loss with a certain confidence level over a time period. This metric, however, ignores the worst scenarios, which may be crucial in the case of intentional water contamination [61]. All in all, risk management provides useful tools and insights for the problem of security in critical infrastructure systems [22].

We assume that the task is to select M_s out of $N_{qs} = |Q_s|$ locations where sensors can be installed, and M_m out of $N_{qm} = |Q_m|$ two-tuples of locations and time instances, corresponding to where and when sampling can be performed. We define $\mathcal{L}_s = \{1, 2, \ldots, N_{qs}\}$ as the set of the candidate sensing node indices, and $\mathcal{L}_m = \{N_{qs}+1, N_{qs}+2, \ldots, N_q\}$ as the set of candidate sampling location and time indices. In addition, we define $\mathbb{X}$ as the set of all solution combinations between these index sets, such that

$$\mathbb{X} = \left\{ \{X_s, X_m\} \mid X_s \in \mathcal{L}_s^{M_s}, X_m \in \mathcal{L}_m^{M_m} \right\}.$$

For N_f different objective risk functions F_i, the optimization problem is constructed as

$$\mathcal{Y} = \underset{\mathcal{X}\in\mathbb{X}}{\operatorname{argmin}} \; \{F_1(\mathcal{X};\Omega_1), ..., F_{N_f}(\mathcal{X};\Omega_{N_f})\},$$

where $\Omega_i \in \mathcal{O}$ is the overall impact matrix corresponding to each objective. For $N_f = 1$, the solution $\mathcal{Y}$ corresponds to a single set of indices, whereas for $N_f > 1$, $\mathcal{Y}$ corresponds to one or more sets of indices.

For a certain $\mathcal{X} \in \mathbb{X}$ and a certain $\Omega \in \mathcal{O}$ the i-th objective function $F_i(\mathcal{X};\Omega_i)$ is computed through a risk function. In the next paragraphs we demonstrate three risk functions, $f_{av}(\cdot), f_{max}(\cdot), f_{CVaR}(\cdot)$, corresponding to the average impact, the maximum impact and the conditional value-at-risk; these functions are suitable for the security problem in water distribution networks. As an example, an average impact objective may be $F_1(\mathcal{X};\Omega_1) = f_{av}(\mathcal{X};\Omega_1)$. In computing these risk objective functions, it is useful to define the scenario index set $\mathcal{G}^* = \{1, 2, ..., N_h\}$, corresponding to each of the N_h fault scenarios considered. In practice, a subset from the set of all scenarios could be neglected in the optimization process, depending on the risk-objective or on whether they are considered trivial with respect to their impact magnitude.

- **Average Impact:** The average impact metric is suitable for optimizing reliability, in the case where contaminant injection can occur at any node with equal probability. This metric, however, has limitations when considering the security framework, since it does not take into sufficient consideration rare faults with extreme consequences. For a certain overall-impact matrix $\Omega \in \mathcal{O}$, a scenario index set $\mathcal{G} = \mathcal{G}^*$, and for a specific set of solution indices $\mathcal{X} \in \mathbb{X}$, the average impact across all faults is given by

$$f_{av}(\mathcal{X};\Omega) = \frac{1}{|\mathcal{G}|} \sum_{i\in\mathcal{G}} \min_{j\in\mathcal{X}} \Omega_{(i,j)}.$$

- **Maximum Impact:** The maximum impact metric is used to reduce the effect of the most extreme fault, in terms of causing the most damage. This metric is useful from a security perspective; on the other hand, it does not take into consideration the fault frequency distribution, and in specific, the frequency of extreme faults. For a certain overall-impact matrix $\Omega \in \mathcal{O}$, a scenario index set $\mathcal{G} = \mathcal{G}^*$, and for a specific set of solution indices $\mathcal{X} \in \mathbb{X}$, the maximum impact across all faults is given by

$$f_{max}(\mathcal{X};\Omega) = \max_{i\in\mathcal{G}} \min_{j\in\mathcal{X}} \Omega_{(i,j)}.$$

- **Conditional Value-at-Risk** The Conditional Value-at-Risk (CVaR) metric, which is frequently used in finance optimization applications, is defined in [62] as the average "loss" for the worst $\alpha\%$ scenarios. In the present work, "loss" corresponds to the overall impact. This metric is quite suitable for the water security problem, since it can be used to minimize extreme contamination faults while at the same time taking into account the frequency of extreme faults. For this metric, a decision maker needs to specify the parameter $\alpha \in (0, 100)$, so that only fault impacts above $(1 - \frac{\alpha}{100}) f_{max}(\mathcal{X};\Omega)$ are considered. Let $\mathcal{G} \subset \mathcal{G}^*$ be the set of extreme-fault indices, given by

$$\mathcal{G} = \{i \mid \min_{j\in\mathcal{X}} \Omega_{(i,j)} \geq (1 - \frac{\alpha}{100}) f_{max}(\mathcal{X};\Omega), i \in \mathcal{G}^*\}.$$

  Therefore, the average tail-impact metric is given by

$$f_{CVaR}(\mathcal{X};\Omega) = \frac{1}{|\mathcal{G}|} \sum_{i\in\mathcal{G}} \min_{j\in\mathcal{X}} \Omega_{(i,j)}.$$

D. Optimization Problem Solution

The different objectives presented in the previous subsection, in general will yield different solutions. Often, it is desirable to compute a set of "good" solutions which satisfy an N_f number of objectives instead of a single one. Minimizing one objective function may result in maximizing others; it is thus not possible to find one optimal solution which satisfies all objectives at the same time. It is possible, however, to find a set of solutions, laying on a Pareto front, where each solution is no worse than the other.

A feasible solution $\mathcal{X}$ is called Pareto optimal if for a set of N_f objectives, there exists no other feasible solution $\mathcal{X}^*$ such that $F_i(\mathcal{X}^*;\Omega_i) \leq F_i(\mathcal{X};\Omega_i)$ with $F_j(\mathcal{X}^*;\Omega_j) < F_j(\mathcal{X};\Omega_j)$ for at least one j. Therefore, a solution is Pareto optimal if there is no other feasible solution which would reduce some objective function without simultaneously causing an increase in at least one other objective function [63, p.779].

One of the most popular solution methodologies for multi-objective optimization problems is to assign a scalar weight for each cost function and calculate their weighted sum, so that the problem is reduced into a single-objective optimization. However, the computed solution might not belong to the set of Pareto front solutions; in addition, weight assignment is susceptible to biases by the decision maker. Instead, we consider utilizing a multi-objective evolutionary algorithm in which each objective is optimized in such a way that the algorithm computes solutions which are non-dominant to each other [64]. Various optimization techniques have been used in such problems, such as evolutionary algorithms, heuristic searching and others. In the following we present two methods for solving the multi-objective optimization problem.

The Iterative Deepening of Pareto Solutions Search approach is a heuristic methodology which searches in parallel across Pareto solutions [65]. The underlying idea is that searching should not be performed blindly within the solution space; instead, it should be guided, iteratively, towards sets of good-enough solutions, pruning those combinations that are likely of no interest. The heuristic behind this algorithm is that "good enough" solutions may appear in the various Pareto fronts. The first few iterations of the algorithm are as follows: for each feasible node in the network, the objective functions are calculated and the first Pareto front is computed. When considering a search tree, the first branches are the network nodes; next we prune all branches with solutions which do not belong in the Pareto front set. Each of the remaining branches is then expanded by adding leafs corresponding to the remaining feasible nodes (in each branch a node is allowed to appear only once). The objective functions are calculated for the new branches, a new Pareto front is computed and the branches which are not in the new Pareto front are pruned; the procedure is iterated until the tree reaches a depth equal to the number of sensors to be installed in the system. The output of the algorithm is a set of Pareto front points, which may not necessarily be on the global Pareto front.

A different procedure for solving the multi-objective placement of M_s sensors is by using the evolutionary algorithm described in [64]. The outline of the algorithm is: Randomly select $N_x \ll |\mathbb{X}|$ solutions from the set of all feasible solutions $\mathbb{X}$, and construct a "parent" set $\mathcal{P}_0 \subset \mathbb{X}$. The values of the objective functions for each solution in the parent set $\mathcal{P}_0$ are then computed. Next, the solutions are sorted according to their non-dominance and are assigned in Pareto ranks, based on which Pareto front they reside. A subset of the parent set is selected and an "offspring" population set is computed, by using the genetic operators of mutation and crossover, suitably modified so that only feasible solutions are generated. From the combined set of parent and offspring solutions, the elements are sorted according to non-dominance. A "crowding distance" metric is computed for each solution, expressing how close it is with its neighboring solutions. An N_x number of solutions is selected from the parent-offspring set, which will comprise the new parent set; this is achieved by selecting solutions with the highest Pareto rank, as well as by considering the crowding distance for better dispersion of the solutions on the Pareto front. The algorithm iterates for a certain number of epochs. The set of sensor placement solutions on the Pareto front computed within the last epoch, are the solutions of the problem. Finally, a decision maker can use higher-level reasoning in choosing the most suitable sensor placement solution from among the computed solutions.

E. Results

The work presented in this section has been published and presented in peer-reviewed journals and conferences [12], [57]–[60], [65]–[67]. In addition, the maximum redundancy formulation has been applied to power distribution networks for Phasor Measurement Units location selection; part of the results have been published in a peer-reviewed journal and a conference [68], [69] by the collaborative team.

V. Fault Detection and Isolation in Water Systems

A. Fault Detection

Currently, most of the water security research assumes that on-line quality sensors explicitly measure contaminant concentration, and that a contamination fault is detected when its concentration exceeds a certain threshold. In practice, this assumption is not realistic; specialized quality sensors may be quite expensive for the water providers. Chlorine sensors are cheaper and are already used by some water providers to monitor water quality; we will investigate how the information given by these sensors could be used to detect a contamination fault. When a contaminant is injected into chlorinated water, a reaction occurs which changes the nominal chlorine decay dynamics. The new reaction dynamics depend on the contaminant's chemical characteristics and on the initial water quality.

For fault detection we will consider a model-based approach, which assumes that the nominal dynamics are known, and the uncertainty is bounded [52]. We assume that the nominal dynamics of chlorine reacting with a contaminant substance, without uncertainty, can be described using the following dynamic equations

$$\begin{aligned} \dot{c} &= -k_b c - k_{cz} c z \\ \dot{z} &= -k_{zc} c z, \end{aligned}$$

where c is the chlorine concentration and z the contaminant concentration with initial conditions c_0, z_0 respectively; k_b is the chlorine decay rate (in this case it is assumed to be linear), and $k_{cz}, k_{zc} > 0$ are the reaction velocity which depend on the water chemical characteristics and the chemical properties of the contaminant substance. For instance, Sodium Arsenite and Organophosphate are two contaminants which react with chlorine according to these dynamics, but with different decay and reaction rates; the first reacts much faster than the second [70]–[72].

The general reaction model can be extended in water systems, and in specific for water storages. According to [8], at least three types of models are considered for describing the chemical reactions within a water storage tank: a) the continuous stirred-tank reactor (CSTR) model, at which the chemicals are perfectly mixed and spread uniformly, b) the plug-flow reactor model, at which there is no mixing of water between the different water parcels assumed to travel along the flow in the tank; c) the two-compartments mixing model, at which the tank is segmented into two perfectly mixed compartments. We will consider the chlorine reaction dynamics in a continuous stirred-tank reactor, which is in general a common assumption in water systems modelling [8]. The hydraulic and quality dynamics can be formulated, without uncertainties, as

$$\begin{aligned} \dot{c} &= \frac{q_{in}}{v} c_{in} - \frac{q_{out}}{v} c + u_c - k_b c - k_{xz} c z \\ \dot{v} &= q_{in} - q_{out} \\ \dot{z} &= -k^i_{zx} c z + u_z \end{aligned}$$

where c is the measurable chlorine concentration in the tank, q_{in}, q_{out} is the inflow and outflow, v is the tank volume, c_{in} is the disturbance concentration of the inflow water, u_c is the control signal, k_b is the chlorine decay rate, z is the contaminant concentration. The reaction rates $k_{xz}, k_{zx} > 0$ depend on the chemical characteristics.

Based on the proposed model, we will formulate the fault detection problem and in addition, compute the detectability conditions which relate the model uncertainty with the fault magnitude, so that detection is possible. A possible extension is to consider the problem of fault detection when the input/output dynamics within certain nodes in a water distribution network have already been approximated [38], [42].

Once a fault has been detected, an on-line approximator will be activated to learn the unknown dynamics. By using the Lyapunov synthesis method, we will construct the adaptive law with which the unknown parameters of a linearized approximation function will update at each time step, so that the estimation error is reduced. In addition, by considering the characteristics of the approximated dynamics, we will examine whether a certain contaminant substance may be identified, as well as its initial concentration.

In extent, although the chlorine decay dynamics are usually assumed to be linear, in practice they may exhibit non-linear dynamics [73], [74] or time-varying decay rate. As these dynamics may change in time due to changing environmental parameters, it is useful to learn the changes by utilizing an adaptive approximation-based methodology.

B. *Fault Source Isolation*

Part of the identification and isolation problem is to determine, after a contamination has been detected, the location where and when the contaminant substance had been injected into the distribution network, given some measurements. This problem has received some attention in the recent years, mostly by considering the case where the system is monitored with on-line sensors [55], [56]. This is a non-trivial problem, especially due to the large scale nature of water distribution systems and the lack of information regarding the states in all but a few nodes; finding a unique solution for this problem, i.e. a source node, is in general extremely difficult.

Previous research has concentrated on the use of measurements from installed quality sensors; however, until a fault has been detected (or not) at more than one nodes, it is difficult to reduce the possible fault source space. In addition, the time-differences between the first detection and the followings detections at a sensor location my be large. One more disadvantage is that sensors may be decoupled, and thus making it difficult in isolating the fault's source. By utilizing manual sampling, however, these drawbacks can be alleviated. Water providers are able to dispatch their personnel at a certain time, to manually take water quality measurements at various locations.

By utilizing a backtracking algorithm (such as in [75]) along with the formulation developed in Section III, we will design a methodology for combining manual quality sampling scheduling with on-line measurements, in order to enhance fault isolation.

C. *Results*

This work is part of an ongoing research and no results have been published.

VI. Fault Accommodation in Water Systems

After a fault has been detected and its dynamics have been approximated by a model (such as a neural network), the controllers which regulate the operation of the system may be modified in order to return the system to its safe operation. Before we address the problem of changing the input signal, we discuss the challenges in relation to disinfectant concentration control.

A. *Disinfectant Concentration Control Challenges*

Controlling a large-scale system such as a water distribution system is a challenging problem. Hydraulic controllers are usually related with pumps and valves, which regulate the flows and pressures in the system. Quality controllers are usually related with the injection of some disinfectant mass (e.g. chlorine) at some locations, so that its concentration reaches a desired level. Most of the related research has examined the use of quality controller for regulating the disinfectant concentrations in water distribution systems, whereas hydraulic controllers act independently to satisfy their own objectives.

Controlling the disinfectant concentration somewhere in the network is not trivial, due to a number of reasons. The most important is that the actual dynamics at a certain moment in the system are not known, due to the fact that consumer demands are not measurable. The control problem is significantly different from the off-line sensor placement problem we examined in the previous section, as it utilized a mathematical model describing the behaviour of the system, which was based on average demand flows, and not on real-time dynamics. In practice, flows in the network are measured at water source locations (e.g. tanks), at District Metered Areas and for some large consumers. The total water demand varies continuously throughout the day, usually in a similar manner (with uncertainty) within the work days, differently in the weekends, and may be affected by seasonal and long-term changes (e.g. tourism and population increase), as well as temperature, precipitation and other parameters. There is not a common demand flow signal for all nodes in the network, since demands are based on social characteristics and the types of business which operate within a certain region. Overall, variations in the hydraulic dynamics influence the propagation time of a disinfectant substance, making control a difficult task. A related problem is that hydraulic controls may change independently, due to various reasons, thus affecting the previous and future quality control actions. Another challenge is that the decay dynamics of chlorine, which is the most widely used disinfectant, are considered in most of the research as linear, whereas in practice these dynamics may be non-linear.

In the next paragraphs we examine a controller design which tries to take into consideration some of the challenges discussed.

B. Disinfectant Concentration Control: A Soft Computing Approach

We consider a real water distribution network, with a number of chlorine booster actuators and a number of chlorine concentration sensors installed at selected locations so that the chlorine concentration at those locations is controllable. In addition to that, we consider that the flows at some "representative" locations in the network are monitored. Moreover, we consider that there exists a software model whose structural parameters have been calibrated to the actual system. The reason why to use a software model instead of a mathematical one, is that a water distribution system can be seen as a hybrid system with various discrete and continuous parameters which make it difficult to take them all into consideration by using a nominal mathematical model. What's more, software models are commonly utilized and fine-tuned by water providers in order to be able to monitor their system's behaviour in changes. We utilize the Model Predictive Control methodology [49], the underlying idea of which is to compute a sequence of future control inputs, so that an objective function is minimized over a predicted horizon. In addition, we use an adaptive demand forecasting methodology for learning the time-varying demand dynamics by using and approximation model (such as radial basis functions or Fourier series) as they change in time.

The controller description is as follows. The water distribution system has two inputs, the quality control signal $u_q(k)$ and the hydraulic controls signal $u_h(k)$, and it is affected by uncontrolled consumer demand disturbances $d(k)$. Some representative flows $q(k)$ are measured from within the system. By using an adaptive forecasting algorithm, a prediction of a certain number of future time-flows is computed. In this work we consider the use of Fourier series or Radial Basis Functions as the approximation structures in an adaptive framework. Following, we use the software model which takes as arguments the hydraulic and quality inputs, the measured flows as well as the predicted flows. An iterative algorithm is applied on the software model, in order to construct the time-varying *dynamic matrix* $\hat{H}(k)$ which will affect the future inputs, and the vector $\hat{F}(k)$ of future responses which corresponds to the system output due to past inputs. These two are used in a optimization algorithm which minimizes function $J(\cdot)$; this function corresponds to the square error between the prediction of chlorine concentration with a reference signal r_c, at some locations over a time frame. Constraints can be included in the optimization program, so that the input and output signal are bounded within some safety regions (input in $[U_{min}, U_{max}]$ and output in $[Y_{min}, Y_{max}]$). The output of the optimization is the future inputs vector U. The quality control inputs for the next time step are applied to the system, and the algorithm iterates. The optimization formulation is

$$\underset{U}{\operatorname{argmin}}\; J(U; \hat{H}(k), \hat{F}(k), r_c)$$

subject to

$$\begin{bmatrix} I \\ -I \\ \hat{H}(k) \\ -\hat{H}(k) \end{bmatrix} U \leq \begin{bmatrix} \mathbf{1}U_{max} \\ -\mathbf{1}U_{min} \\ \mathbf{1}Y_{max} - \hat{F}(k) \\ -\mathbf{1}Y_{min} + \hat{F}(k) \end{bmatrix},$$

where $\mathbf{1} = [1, 1, ..., 1]^\top$.

What follows is to research on how to extent water quality control algorithms to take into consideration the approximated fault dynamics, so that the fault is accommodated.

C. Results

This work is part of an ongoing research and some results in relation to chlorine concentration control have been presented in [76].

VII. Conclusions and Impact

Water resources management is a key challenge that will become even more crucial in the years ahead. Water distribution systems in specific play an essential part in sustaining certain vital societal functions and when faults occur, such as water contamination or pipe breaks, these functions may be affected negatively. From a system-theoretic viewpoint, monitoring and control of water distribution networks present important new challenges due to their large-scale interconnected dynamics, their structural uncertainties, the complex propagation and reaction dynamics, as well as the hydraulics uncertainties. Analytical results and algorithmic tools developed in the automatic control and control systems technology community can be employed to advance current knowledge on the issues of fault detection, isolation, identification and accommodation in water distribution systems.

Quality faults such as water contamination may affect severely the population health, disrupt the economic processes and affect the societal functions. This motivates the general goal of this work, which is to develop fault diagnosis and security monitoring methodologies for water distribution systems. In specific, this work has four objectives: a) to formulate the problem for monitoring and control of water distribution networks in a framework suitable for the sensor placement and fault diagnosis; b) to find the locations in a water distribution networks where quality sensors should be installed; c) to design fault detection and isolation algorithms suitable for water systems and d) design fault accommodation algorithms in order to compensate for the contamination fault dynamics.

The described objectives address a number of important problems, whose solution will have a significant impact on the quality of service of water providers, as well as to the society in general. Through the developed methodologies, the security weaknesses of the distribution system are revealed, and water providers are able to monitor and control their water distribution system more efficiently and with more security. Another added benefit of this work, is that the developed methodologies can be extended to other types of critical infrastructure, such as power distribution networks, where its operation is of critical importance, both for high quality-of-service, as well as for increasing human safety. This work can be considered as the first step towards a wider scope, to design and develop intelligent monitoring systems and control methodologies which will increase fault tolerance in water distribution systems.

REFERENCES

[1] W. M. Grayman, "A quarter of a century of water quality modeling in distribution systems," in *Proc. of Water Distribution Systems Analysis Symposium*, S. G. Buchberger, Ed., vol. 247, no. 40941, 2006, pp. 1–12.

[2] D. Savic, Z. Kapelan, and P. Jonkergouw, "Quo vadis water distribution model calibration?" *Urban Water Journal*, vol. 6, no. 1, pp. 3–22, 2009.

[3] E. Todini and S. Pilati, "A gradient method for the analysis of pipe networks," in *Proc.of Computer Applications for Water Supply and Distribution*, 1987.

[4] R. LeVeque, "Nonlinear conservation laws and finite volume methods," in *Computational methods for astrophysical fluid flow*, O. Steiner and A. Gautschy, Eds. Berlin, Germany: Springer, 1998.

[5] L. A. Rossman, P. F. Boulos, and T. Altman, "Discrete volume-element method for network water-quality models," *ASCE Journal of Water Resources Planning and Management*, vol. 119, no. 5, pp. 505–517, 1993.

[6] L. A. Rossman and P. F. Boulos, "Numerical methods for modeling water quality in distribution systems: A comparison," *ASCE Journal of Water Resources Planning and Management*, vol. 122, no. 2, pp. 137–146, 1996.

[7] R. Łangowski and M. Brdys, "Monitoring of chlorine concentration in drinking water distribution systems using an interval estimator," *International Journal of Applied Mathematics and Computer Science*, vol. 17, no. 2, pp. 199–216, 2007.

[8] L. Rossman, *EPANET 2 Users Manual*, National Risk Management Research Laboratory, Office of Research and Development, U.S. Enviromental Protection Agency, Cincinnati, OH 45268, 2000.

[9] F. Shang, J. G. Uber, and L. A. Rossman, *EPANET Multi-Species Extension User's Manual*, National Risk Management Research Laboratory, Office of Research and Development, U.S. Enviromental Protection Agency, Cincinnati, OH 45268, August 2007.

[10] T. M. Walski and J. W. Male, *Water distribution systems handbook*. McGraw-Hill New York, 2000, ch. Maintenance and Rehabilitation/Replacement, pp. 17.1–17.28.

[11] WHO, *Guidelines for drinking-water quality, Vol. 1, 3rd edition incorporating 1st and 2nd addenda*. World Health Organization, 2008.

[12] A. Ostfeld, J. G. Uber, E. Salomons, J. W. Berry, W. E. Hart, C. A. Phillips, J.-P. Watson, G. Dorini, P. Jonkergouw, Z. Kapelan, F. di Pierro, S.-T. Khu, D. Savic, D. Eliades, M. Polycarpou, S. R. Ghimire, B. D. Barkdoll, R. Gueli, J. J. Huang, E. A. McBean, W. James, A. Krause, J. Leskovec, S. Isovitsch, J. Xu, C. Guestrin, J. VanBriesen, M. Small, P. Fischbeck, A. Preis, M. Propato, O. Piller, G. B. Trachtman, Z. Y. Wu, and T. Walski, "The battle of the water sensor networks (BWSN): A design challenge for engineers and algorithms," *ASCE Journal of Water Resources Planning and Management*, vol. 134, no. 6, pp. 556–568, Nov./Dec. 2008.

[13] C. Toregas and C. ReVelle, "Optimal location under time or distance constraints," *Papers in Regional Science*, vol. 28, no. 1, pp. 131–143, Dec. 1972.

[14] R. Church and C. ReVelle, "The maximal covering location problem," *Papers in Regional Science*, vol. 32, no. 1, pp. 101–118, Dec. 1974.

[15] B. Lee and R. Deininger, "Optimal locations of monitoring stations in water distribution system," *ASCE Journal of Environmental Engineering*, vol. 118, no. 1, pp. 4–16, Jan./Feb. 1992.

[16] A. Kumar, M. L. Kansal, and G. Arora, "Identification of monitoring stations in water distribution system," *ASCE Journal of Environmental Engineering*, vol. 123, no. 8, pp. 746–752, 1997.

[17] M. A. Al-Zahrani and K. Moeid, "Locating optimum water quality monitoring stations in water distribution system," in *ASCE Proc. of World Water and Environmental Resources Congress*, 2001, pp. 393–402.

[18] P. Harmant, A. Nace, L. Kiene, and F. Fotoohi, "Optimal supervision of drinking water distribution network," in *ASCE Proc. of Water Resources Planning and Management Conference*, 1999, pp. 52–60.

[19] S. Hakimi, "Optimum locations of switching centers and the absolute centers and medians of a graph," *Operations Research*, vol. 12, pp. 450–459, May/Jun. 1964.

[20] J. Berry, L. Fleischer, W. Hart, and C. Phillips, "Sensor placement in municipal water networks," in *ASCE Proc. of World Water and Environmental Resources Congress*, 2003, pp. 40–49.

[21] J. Berry, W. Hart, C. Phillips, J. Uber, and J. Watson, "Sensor placement in municipal water networks with temporal integer programming models," *ASCE Journal of Water Resources Planning and Management*, vol. 132, pp. 218–224, Jul./Aug. 2006.

[22] J.-P. Watson, H. J. Greenberg, and W. E. Hart, "A multiple-objective analysis of sensor placement optimization in water networks," in *ASCE Proc. of World Water and Environmental Resources Congress*, 2004, pp. 456–465.

[23] D. L. Boccelli and W. E. Hart, "Optimal monitoring location selection for water quality issues," in *ASCE Proc. of World Environmental and Water Resources Congress*, K. C. Kabbes, Ed., vol. 243, no. 40927. ASCE, 2007, pp. 522–527.

[24] A. Krause, J. Leskovec, C. Guestrin, J. VanBriesen, and C. Faloutsos, "Efficient sensor placement optimization for securing large water distribution networks," *ASCE Journal of Water Resources Planning and Management*, vol. 134, no. 6, pp. 516–526, Nov./Dec. 2008.

[25] S. R. Agha, "Use of goal programming and integer programming for water quality management–A case study of Gaza Strip," *European Journal of Operational Research*, vol. 174, no. 3, pp. 1991–1998, Nov. 2006.

[26] D. L. Boccelli, M. E. Tryby, J. G. Uber, L. A. Rossman, M. L. Zierolf, and M. M. Polycarpou, "Optimal scheduling of booster disinfection in water distribution systems," *ASCE Journal of Water Resources Planning and Management*, vol. 124, no. 2, pp. 99–111, 1998.

[27] M. A. Brdys, T. Chang, and K. Duzinkiewicz, "Intelligent model predictive control of chlorine residuals in water distribution systems," in *ASCE Proc. of Water Resource Engineering and Water Resources Planning and Management*, D. Phelps and G. Sehlke, Eds., vol. 111, no. 40569. ASCE, 2001, pp. 391–401.

[28] M. A. Brdys, T. Chang, K. Duzinkiewicz, and W. Chotkowski, "Hierarchical control of integrated quality and quantity in water distribution systems," in *ASCE Proc. of Water Resource Engineering and Water Resources Planning and Management*, R. H. Hotchkiss and M. Glade, Eds., vol. 104, no. 40517. ASCE, 2000, pp. 193–202.

[29] M. Brdys and T. Chang, "Robust model predictive control under output constraints," in *Proc. of IFAC World Congress*, 2002, pp. 21–26.

[30] T. Chang, "Robust model predictive control of water quality in drinking water distribution systems," Ph.D. dissertation, The University of Birmingham, 2003.

[31] T. Chang, M. A. Brdys, and K. Duzinkiewicz, "Quantifying uncertainties for chlorine residual control in drinking water distribution systems," in *ASCE Proc. of World Water and Environmental Resources Congress*, P. Bizier and P. DeBarry, Eds., vol. 118, no. 40685. ASCE, 2003, pp. 28–37.

[32] ——, "Decentralized robust model predictive control of chlorine residuals in drinking water distribution systems," in *ASCE World Water Congress 2003*, P. Bizier and P. DeBarry, Eds., vol. 118, no. 40685. ASCE, 2003, pp. 255–264.

[33] S. Constans, B. Bremond, and P. Morel, "Simulation and control of chlorine levels in water distribution networks," *ASCE Journal of Water Resources Planning and Management*, vol. 129, no. 2, pp. 135–145, 2003.

[34] G. Dandy and C. Hewitson, "Optimizing hydraulics and water quality in water distribution networks using genetic algorithms," in *ASCE Proc. of Water Resource Engineering and Water Resources Planning and Managemen*, R. H. Hotchkiss and M. Glade, Eds., vol. 104, no. 40517. ASCE, 2000, pp. 211–220.

[35] K. Duzinkiewicz, M. Brdys, and T. Chang, "Hierarchical model predictive control of integrated quality and quantity in drinking water distribution systems," *Urban Water Journal*, vol. 2, no. 2, pp. 125–137, 2005.

[36] G. R. Munavalli and M. S. M. Kumar, "Optimal scheduling of multiple chlorine sources in water distribution systems," *ASCE Journal of Water Resources Planning and Management*, vol. 129, no. 6, pp. 493–504, 2003.

[37] A. Ostfeld and E. Salomons, "Conjunctive optimal scheduling of pumping and booster chlorine injections in water distribution systems," *Engineering Optimization*, vol. 38, no. 3, pp. 337–352, Apr. 2006.

[38] M. Polycarpou, J. Uber, Z. Wang, F. Shang, and M. Brdys, "Feedback control of water quality," *IEEE Control Systems Magazine*, vol. 22, no. 3, pp. 68–87, June 2002.

[39] T. D. Prasad, G. A. Walters, and D. A. Savic, "Booster disinfection of water supply networks: Multiobjective approach," *ASCE Journal of Water Resources Planning and Management*, vol. 130, no. 5, pp. 367–376, 2004.

[40] M. Propato and J. G. Uber, "Linear least-squares formulation for operation of booster disinfection systems," *ASCE Journal of Water Resources Planning and Management*, vol. 130, no. 1, pp. 53–62, 2004.

[41] A. B. A. Sakarya and L. W. Mays, "Optimal operation of water distribution pumps considering water quality," *ASCE Journal of Water Resources Planning and Management*, vol. 126, no. 4, pp. 210–220, 2000.

[42] Z. Wang, M. Polycarpou, J. Uber, and F. Shang, "Adaptive control of water quality in water distribution networks," *IEEE Transactions on Control Systems Technology*, vol. 14, no. 1, pp. 149–156, 2006.

[43] M. M. Polycarpou, J. G. Uber, and U. Desai, "Design of a feedback controller for water distribution systems residual maintenance," in *Proc. of Water Resources Planning and Management Conference*, E. M. Wilson, Ed., vol. 102, no. 40430. ASCE, 1999, pp. 61–67.

[44] Z. Wang, M. Polycarpou, and J. Uber, "Decentralized model reference adaptive control of water quality inwater distribution networks," in *IEEE Proc. International Symposium on Intelligent Control*, 2000, pp. 127–132.

[45] Z. Wang, M. Polycarpou, J. Uber, and F. Shang, "Adaptive periodic control for chlorine residual maintenance indrinking water distribution networks," in *IEEE Proc. Conference on Decision and Control, 2001.*, vol. 5, 2001.

[46] M. Propato, J. G. Uber, F. Shang, and M. M. Polycarpou, "Integrated control and booster system design for residual maintenance in water distribution systems," in *ASCE Proc. of World Water and Environmental Resources Congress*, D. Phelps and G. Sehlke, Eds., vol. 111, no. 40569. ASCE, 2001, pp. 390–399.

[47] Z. Wang, M. M. Polycarpou, F. Shang, and J. G. Uber, "Design of feedback control algorithm for chlorine residual maintenance in water distribution systems," in *ASCE Proc. of World Water and Environmental Resources Congress*, D. Phelps and G. Sehlke, Eds., vol. 111, no. 40569. ASCE, 2001, pp. 377–386.

[48] ——, "Adaptive control formulation for chlorine residual maintenance in water distribution systems," in *Proc. of Water Resource Engineering and Water Resources Planning and Management*, R. H. Hotchkiss and M. Glade, Eds., vol. 104, no. 40517. ASCE, 2000, pp. 192–201.

[49] F. Camacho and C. Bordons, *Model Predictive Control*, 2nd ed. Springer, 2004.

[50] J. Chen and R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*. Norwell, MA, USA: Kluwer Academic Publishers, 1999.

[51] R. Isermann, *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Springer Verlag, 2006.

[52] X. Zhang, M. Polycarpou, and T. Parisini, "A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 4, pp. 576–593, Apr. 2002.

[53] ——, "Adaptive fault-tolerant control of a class of nonlinear MIMO systems," in *IEEE Proc. Conference on Decision and Control*, Dec. 2008, pp. 398–403.

[54] Y. Jeffrey Yang, R. Haught, and J. Goodrich, "Real-time contaminant detection and classification in a drinking water pipe using conventional water quality sensors: Techniques and experimental results," *Journal of Environmental Management*, vol. 90, no. 8, pp. 2494–2506, 2009.

[55] A. Preis and A. Ostfeld, "A contamination source identification model for water distribution system security," *Engineering Optimization*, vol. 38, no. 8, pp. 941 – 947, Dec. 2007.

[56] A. E. De Sanctis, F. Shang, and J. G. Uber, "Real-time identification of possible contamination sources using network backtracking methods," *Journal of Water Resources Planning and Management*, 2009.

[57] D. G. Eliades and M. M. Polycarpou, "Security issues in drinking water distribution networks," in *Proc. of the International Workshop on Computational Intelligence in Security for Information Systems*, 2009, pp. 69–76.

[58] ——, "Security of water infrastructure systems," in *Proc. of Critical Information Infrastructure Security*, ser. Lecture Notes in Computer Science, R. Setola and S. Geretshuber, Eds., no. 5508, 2009, pp. 360–367.

[59] D. G. Eliades, M. M. Polycarpou, and B. Charalambous, "Security issues in drinking water distribution networks," *Journal of Information Assurance and Security*, vol. 4, pp. 500–508, 2009.

[60] D. G. Eliades and M. M. Polycarpou, "A fault diagnosis and security framework for water systems," *IEEE Trans. Control Systems Technology*, 2009, (accepted for publication).
[61] D. Rosen and S. A. Zenios, *Handbook of Asset and Liability Management*. Elsevier, 2006, ch. Enterprise-wide Asset and Liability Management: Issues, Institutions, and Models, pp. 1–23.
[62] S. Uryasev, *Probabilistic Constrained Optimization: Methodology and Applications*. Springer, 2000.
[63] S. Rao, *Engineering Optimization: Theory and Practice*. Wiley-Interscience, 1996.
[64] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
[65] D. Eliades and M. Polycarpou, "Iterative deepening of pareto solutions in water sensor networks," in *ASCE Proc. of Water Distibution System Analysis Symposium*, 2006, pp. 114–132.
[66] D. G. Eliades and M. M. Polycarpou, "Multi-objective optimization of water quality sensor placement in drinking water distribution networks," in *Proc. European Control Conference*, 2007, pp. 1626–1633.
[67] D. G. Eliades, B. Charalampous, and M. M. Polycarpou, "A security-oriented manual quality sampling methodology for water systems," in *Proc. of Electrical and Computer Engineering Student Conference*, 2009.
[68] S. Chakrabarti, D. Eliades, E. Kyriakides, and M. Albu, "Measurement uncertainty considerations in optimal sensor deployment for state estimation," in *IEEE Proc. of International Symposium on Intelligent Signal Processing*, 2007, pp. 1–6.
[69] S. Chakrabarti, E. Kyriakides, and D. Eliades, "Placement of synchronized measurements for power system observability," *IEEE Transactions on Power Delivery*, vol. 24, no. 1, pp. 12–19, Jan. 2009.
[70] K. A. Umberg, "Performance evaluation of real-time event detection algorithms," Master's thesis, University of Cincinnati, 2006.
[71] M. Dodd, N. Vu, A. Ammann, R. Kissner, H. Pham, M. Berg, and U. von Gunten, "Kinetics and mechanistic aspects of As (III) oxidation by aqueous chlorine, chloramines, and ozone: Relevance to drinking water treatment," *Environ. Sci. Technol*, vol. 40, no. 10, pp. 3285–3292, 2006.
[72] Q. Zhang and S. Pehkonen, "Oxidation of diazinon by aqueous chlorine: kinetics, mechanisms, and product studies," *J. Agric. Food Chem*, vol. 47, no. 4, pp. 1760–1766, 1999.
[73] F. Hua, J. R. West, R. A. Barker, and C. F. Forster, "Modelling of chlorine decay in municipal water supplies," *Water Research*, vol. 33, no. 12, pp. 2735 – 2746, 1999.
[74] J. C. Powell, J. R. West, N. B. Hallam, C. F. Forster, and J. Simms, "Performance of various kinetic models for chlorine decay," *ASCE Journal of Water Resources Planning and Management*, vol. 126, no. 1, pp. 13–20, 2000.
[75] F. Shang, J. Uber, and M. Polycarpou, "Particle Backtracking Algorithm for Water Distribution System Analysis," *ASCE Journal of Environmental Engineering*, vol. 128, no. 5, pp. 441–450, 2002.
[76] D. G. Eliades, E. C. Kyriakides, and M. M. Polycarpou, "Enhanced robustness in model predictive control of water quality using adaptive estimation methods," in *Proc. of Computing and Control for the Water Industry*, 2007.