Water Research 139 (2018) 132–143

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Model-based approach for cyber-physical attack detection in water distribution systems

Mashor Housh*, Ziv Ohar

Faculty of Management, Department of Natural Resource and Environmental Management, University of Haifa, Haifa, Israel

ARTICLE INFO

Article history:
Received 16 October 2017
Received in revised form
13 March 2018
Accepted 14 March 2018
Available online 17 March 2018

Keywords:
Cyber-physical systems
Water distribution systems
Event detection methodology
Model-based fault detection
Cyber-attacks

ABSTRACT

Modern Water Distribution Systems (WDSs) are often controlled by Supervisory Control and Data Acquisition (SCADA) systems and Programmable Logic Controllers (PLCs) which manage their operation and maintain a reliable water supply. As such, and with the cyber layer becoming a central component of WDS operations, these systems are at a greater risk of being subjected to cyberattacks. This paper offers a model-based methodology based on a detailed hydraulic understanding of WDSs combined with an anomaly detection algorithm for the identification of complex cyberattacks that cannot be fully identified by hydraulically based rules alone. The results show that the proposed algorithm is capable of achieving the best-known performance when tested on the data published in the BATtle of the Attack Detection ALgorithms (BATADAL) competition (http://www.batadal.net).

1. Introduction

Modern critical infrastructures are often controlled by Supervisory Control And Data Acquisition (SCADA) systems and Programmable Logic Controllers (PLCs) which manage their operation. These systems, which combine physical processes with cyber networking, are defined as cyber-physical systems (Lee, 2008). The past few years have witnessed the emergence of cyber-physical systems in different components of the water supply sector, such as the water distribution system (Rasekh et al., 2016), water supply reservoirs (Bobat et al., 2015) and treatment plants (Spellman, 2013).

While cyber-physical systems make water supply management efficient and reliable, they also expose the system to cyber-physical attacks. Indeed, the threat of cyberattacks on critical infrastructure systems (e.g. electricity, communication, transportation and water networks) is becoming a major concern (Rasekh et al., 2016). In Water Distribution Systems (WDSs) in particular, these can range from stealing consumer data to damaging equipment, causing water shortages or degrading water quality (Slay and Miller, 2008). As such, cyber physical attacks on WDSs can exert economic, public health and environmental impacts, which makes them attractive targets for terrorism and cyberwarfare (Lewis, 2002; Horta, 2007; Dakin et al., 2009). According to the United States Industrial Control Systems Cyber Emergency Response Team, several cyber physical attacks have already occurred against the United States' water sector (ICS-CERT, 2016). It is thus important to develop cyberattack detection methodologies for WDSs if we are to minimize the systems' vulnerability and limit the attacks' potential damages.

Unlike WDSs, several studies have considered the issue of cyber security in power grids. Early studies of cyber security in these systems have focused on communication standards (Ericsson, 2010) in different components such as the metering infrastructure (Cleveland, 2008) and the SCADA system (Yang et al., 2006). More recently, new approaches have been developed which also consider the interaction between the physical power system and the cyber infrastructure (Xie et al., 2010; Kosut et al., 2010a,b,c; Kim and Poor, 2011; Liu et al., 2011; Hug and Giampapa, 2012; Giani et al., 2013; Liu et al., 2015). More specifically, these studies focused on investigating the attack mechanisms of false data injection in the context of power system state estimation.

The state estimation problem is a critical stage for ensuring optimal and reliable real-time operation of the power grid (Liu et al., 2015). State estimation is the process of estimating unknown state variables in a power grid based on the real-time measurements of system variables such as flow measurements and bus voltage measurements. False data injection attacks, involve situations where an attacker can inject false data into the

* Corresponding author.
E-mail address: mhoush@univ.haifa.ac.il (M. Housh).

https://doi.org/10.1016/j.watres.2018.03.039
0043-1354/

measurements such as to lead to an incorrect system state estimation, which in turn leads to false operational decisions and system failures (Yuan et al., 2011). Yuan et al. (2011) have shown that operational decisions that were based on incorrect power flow data as a result of false data injection attack can lead to an uneconomic operation and manipulation of the electricity market (Xie et al., 2010; Negrete-Pincetic et al., 2009; Esmalifalak et al., 2013) as well as to system failures (Liu et al., 2015). In this respect, Liu et al. (2015) designed a framework for identifying a set of lines that could be overloaded due to false data injections, and which should be strictly monitored to ensure the reliability and safety of power systems. Along the same lines, Bobba et al. (2010) derived a lower bound for the number of meters which should be ensured as attack free in order to prevent the manipulation of the state estimation procedure.

The most relevant studies to our work are Kosut et al. (2010a,b, c) which focused on the detection problem associated with false data injection. In this respect, Kosut et al. (2010a) developed computationally efficient heuristics for detecting false data attacks against state estimators, while Kosut et al. (2010b) developed a Bayesian framework for the same purpose. Kosut et al. (2010c) suggested a graph theoretic approach for detecting false data injections which targets the state estimation procedure. Still, and unlike our study, these previous studies only consider false data injection attacks in the context of the state estimation procedure which is a critical stage in modern power grids.

As opposed to the rich literature on power grids' cyber security, only a limited number of studies have considered the cyberattack problem in WDSs (Amin et al., 2013a; b; Perelman and Amin, 2014; Taormina et al., 2017). Amin et al. (2013a, b) consider automated open canal networks, that are characterized by unpressurized and tree like layout. For this reason, the approach suggested by Amin et al. (2013a, b) cannot be applied to pressurized and looped WDSs. Perelman and Amin (2014) introduced a network interdiction model for simulating attacks on WDSs which considers both energy and water balance, making it potentially applicable to a pressurized looped network. However, their work did not consider the attack detection problem; their objective was merely to assess the network's vulnerability to attacks. More specifically, the problem is formulated as an attacker-defender model, in which the attackers attempt to minimize the satisfied demand by removing one pipe at a time from the network (this was the only type of attack scenario considered) while the operator attempts to maximize the satisfied demand. Taormina et al. (2017) developed the epanetCPA toolbox that allows the simulation of a wide range of attack scenarios. More specifically, this toolbox utilizes the EPANET hydraulic engine (USEPA, 2013) to simulate the physical system as well as a cyber-module that simulates the system's cyber components (e.g., sensors, PLCs, and SCADAs). The tool was used for simulating six different attack scenarios on different WDS components. As such, and like Perelman and Amin (2014), Taormina et al. (2017) focus on simulating cyber-attacks rather than on developing detection systems for the identification of attack scenarios. In light of the above, there exists a lack of analytical tools designed for addressing the cyber-attack detection problem with respect to WDSs.

At this point, it is important to note the differences between the cyberattack detection problem and the physical attack detection problem in WDSs. Unlike the former, the latter has been widely studied in the context of contamination Event Detection Systems (EDS) (Arad et al., 2013; Housh and Ostfeld, 2015). More specifically, the difference between the two problems can be seen as being associated with three issues: (1) temporal resolution: the contamination event durations and detection time considered in the EDS literature are usually longer than those considered in the cyberattack problem, for example, because the number of water quality monitoring stations is limited, the time between the contamination injection until it reaches the monitoring station is considered the minimum (i.e. best) detection time; (2) spatial resolution: the algorithms in the EDS literature, often monitors a limited number of water quality monitoring stations, while in the cyberattack problem the algorithm monitors the signals from all the cyber components that are spatially distributed throughout the network; (3) attack concealment (i.e. alteration of several indicators in order to hide the attack): cyberattack problems consider the possibility of attack concealment while EDS literature does not consider the possibility of concealment.

The two problems are nonetheless similar in the sense that both contamination event detection systems and the cyberattack detection problem can be formulated as a Fault Detection (FD) problem where a fault is defined as abnormal system behavior. A FD problem involves the monitoring of a system with a view to identifying when a fault has occurred. Generally speaking, there are two approaches that can be taken to resolve it, namely Signal processing based FD and Model-based FD. The former infers faults directly from sensor readings, while the latter employs a simulation model as a reference model in order to analyze the discrepancy between sensor readings and reference values (Gertler, 1998). Most contamination event detection systems follow the signal processing based FD approach but the results of our recent work (Housh and Ohar, 2017a) suggest that the model-based approach outperformed classical signal-based approaches when applied to contamination event detection systems in WDSs.

The present study offers a model-based FD methodology that utilizes a physically based water hydraulics simulation model (EPANET) for detecting cyberattacks on WDSs. The model-based approach is based on the generation of residuals from a reference model, which are then evaluated by a threshold-based classification method in order to distinguish between faulty and normal behaviors. While the process of the model-based approach is similar across different disciplines (e.g. power grids (Kosek and Gehrke, 2016), smart buildings (Weimer et al., 2013), communication network faults (Cheung et al., 2007), and automotive systems (Freeman et al., 2013)), the main focus in the development of such fault detection approaches lies in the construction of the reference model and in the design of the residual classification algorithm. Amin et al. (2013a, b), for example, considered a model-based cyberattack detection methodology for a relatively simple system of cascaded gravity-flow canals, while the study offered hereunder considers general pressurized and looped WDSs. Indeed, the development of a model-based detection for such WDSs introduces challenges which were not addressed in Amin et al. (2013a, b) - as we will demonstrate in the next section. Furthermore, certain detection problems in looped WDSs require the design of different model-based approaches. Perez et al. (2014) and Meseguer et al. (2014), for example, developed a model-based approach for leakage detection and localization that cannot be applied to our problem since the network demand data associated with the present study are unobservable.

It is worth noting that an earlier version of the proposed Cyber-attack Detection System (Housh and Ohar, 2017b) competed in the Battle of cyber-attack detection algorithms (Taormina et al., 2016) organized during the 2017 Annual Congress of the Environmental Water Resources Institute (EWRI) of the American Society of Civil Engineers (ASCE). The proposed algorithm was announced as the first-place winner. It is also important to mention that, unlike our model-based algorithm, the other algorithms in the competition were signal-based algorithms in which sensor readings were analyzed by statistical and/or machine learning methods with a view to inferring dataset anomalies. The reader is referred to BATADAL (2016) for further information about the battle itself.

2. Methodology

2.1. Problem statement

The problem statement is as follows: given an attack-free dataset of hourly SCADA readings (Dataset 1) and a dataset of hourly SCADA readings with labeled cyberattacks (Dataset 2), develop a detection mechanism which maximizes detection reliability while minimizing detection time and the amount of false alarms. The SCADA readings include records for the flow rate, the inlet and outlet pressure, and the status of every pump in every pumping station, as well as the system tanks' water levels.

2.2. Model-based detection method

Our Cyber-attack Detection System (CDS) relies on a physically-based simulation model for obtaining reference values pertaining to normal system operation conditions. We used the EPANET hydraulic simulator to simulate the WDS hydraulics. The idea is to use the error existing between the EPANET-simulated values and the SCADA readings to infer abnormal system behavior. As such, it is important to rely on a well-calibrated simulation model or it will be difficult to distinguish between normal and abnormal errors.

First, the attack-free dataset (Dataset 1) is used in order to produce “normal” errors which are expected in the system due to the differences between the model and the actual WDS. Next, the simulated-attack dataset (Dataset 2) is used in order to produce errors between the model and the SCADA readings. This comparison between “normal” errors and the errors produced in the presence of attacks allows the inference of outliers that are used as indicators of a potential cyberattack.

While the framework described above is straightforward, it nonetheless introduces several challenges. The major challenge lies in the fact that the EPANET simulator is a demand driven simulator, meaning that it requires the demands for all system nodes in order to simulate the system's hydraulics. Given that demand is not an observable variable in the problem (i.e. it is not measured), the system demands must be estimated before using EPANET. With this in mind, we suggest a three-phase approach: (1) Estimating the demand on the basis of partial SCADA records of water quantity measurements (i.e. flow and tank levels); (2) Simulating the hydraulics on the basis of the estimated demand; (3) calculating and classifying the errors between the SCADA readings the simulated values.

Let us define the SCADA records at time t as the set $\mathcal{H}$, and the subset of the records which correspond to water quantity variables (i.e. flows and tanks levels) at time t as the set $\mathcal{D}$. Then, given an operator f to estimate demand, $\mathcal{D} = f(\mathcal{Q})$, and using EPANET as an operator, g, which converts the demand to system hydraulic variables, $\tilde{\mathcal{H}} = g(\mathcal{D})$, we obtain Eq. (1).

$$\mathcal{H} = \tilde{\mathcal{H}} = g(\mathcal{D}) = g(f(\mathcal{Q})) \tag{1}$$

where $\mathcal{H}$ is the actual SCADA readings and $\tilde{\mathcal{H}}$ are simulated values for the SCADA readings. Eq. (1) holds under normal operation conditions. However, the equality is not maintained when there is a “manipulation” in any element of the RHS, the LHS or both sides of the equation. In such cases, the error between the two sides is used as an indicator for a cyberattack event. Three possible cyberattack scenarios can be detected in this context:

Case 1. Only the SCADA measurements used for the demand estimation (i.e. quantity variables) are altered by an attack. In this case, the EPANET model will produce wrong pressure values and yield high errors when compared to pressure SCADA readings. A wrong ‘Off’ pump status signal in one demand zone, for example, will cause an underestimation of demand and - as a result - an overestimation of pressure in the RHS of Eq. (1).

Case 2. Only the pressure measurements are altered by the attacker. In this case the demand estimation will be accurate and EPANET will produce pressures which are close to the real pressures (RHS of Eq. (1)) but which still differ from the wrong SCADA readings (LHS of Eq. (1)).

Case 3. All SCADA readings are altered. In such a case, the demand estimation will be wrong. However, the EPANET pressure values will still accord with the system's physics/hydraulics. Major errors will thus be observed between the EPANET pressures and the SCADA pressures unless the attacker changes the pressure according to the network's physical properties.

The implementation of the process described in Eq. (1) is summarized in Fig. 1. As can be seen, the SCADA quantity variables (flows, tank levels and pump statuses) are used for estimating the demands in the network which are then simulated by EPANET in order to formulate the simulated reference values of the system hydraulics (i.e. flows, tank levels, pump statuses, and pressures). The error vectors are then calculated based on these reference values and classified into normal and abnormal operation conditions.

2.3. Demand estimation model

The demand estimation model is primarily based on solving the network flow problem without accounting for system pressure. In other words, it is meant for solving the network's water balance problem. However, unlike the classical network flow problem where demands are given and link flows are unknown, the demands here are unknown while the flows are given for a subset of links. The system should thus satisfy the water balance as described in Eq. (2) at each time step.

$$A_t \cdot Q_t = q_t + \Delta V_t \tag{2}$$

where A_t is the network incidence matrix, Q_t is a vector of the flow in the network's links, q_t is a vector of the demand in the network's nodes, and ΔV_t is a vector of the water volume change in the network's tanks. Given the District Metering Areas (DMAs), it is possible to parametrize the demand in the network's nodes according to DMA demands. This parametrization yields a mass balance model as given in Eq. (3). The parametrization matrix Δ, is a mapping matrix between the DMAs' demands and the network's nodal demands. As such, it has a row for each node and a column for each DMA. Each row has one non-zero element which is equal to the fraction of the node's demand from its corresponding DMA demand.

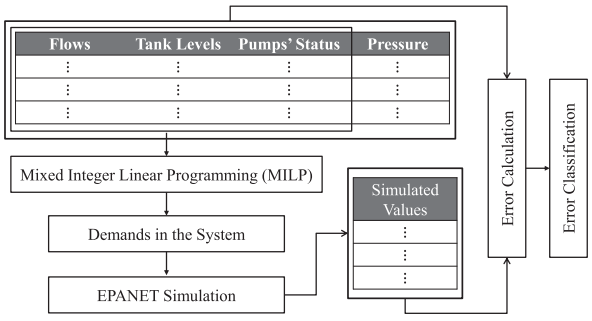

Fig. 1. Model-based detection framework.

$$A_t \cdot Q_t = \Delta \cdot d_t + \Delta V_t \quad (3)$$

where Δ is the parametrization matrix with given coefficients and d_t is a vector of the DMAs' demand.

In hours, $t \notin \tau_{on/off}$ (where $\tau_{on/off}$ is the subset of hours with pump status switches), Eq. (3) yields a unique solution for demand vector d_t. That is, given SCADA values for a subset of Q_t(i.e. flow in pumps) and a storage change from the tanks' SCADA readings, it is possible to obtain unique values for d_t when $t \notin \tau_{on/off}$. The obtained $d_t$is an estimation of DMA demands under the assumption that the pump flow at the beginning of the time step (i.e. 1 hour) is constant along this time step.

For hours $t \in \tau_{on/off}$ the flow in the system links is not given in the SCADA records. This is because the actual times where the pump switch status is not reported. For example, if a pump begins the hour with an "On" status and is turned off after 5 min, then assuming a positive flow along the entire hour will lead to an overestimation of the DMAs' demand. Nevertheless, it is still possible to use the pumps' control rules in order to predict the switch points by linking the pump flow with the tank water levels according to the specified control rules.

We formulate a Mixed Integer Linear Program (MILP) model to estimate DMA demands when $t \in \tau_{on/off}$. The MILP formulation is necessary because of the discontinuous pump control rules. The MILP model assumes a constant hourly demand and a 15-min integration time step where the pumps' statuses are determined on the basis of the control rules at the end of each integration time step. In this respect, two distinct cases must be singled out: the first is when a pump turns on during the hour, and the second is when a pump turns off during the hour. The MILP formulation given in Eq. (4) covers these two cases as well as the "no switches" case where $t \notin \tau_{on/off}$.

$$\begin{aligned}
&\text{Min} \quad \omega \cdot 1^T \cdot \left| V_t^4 - V_t^{end} \right| + (1-\omega) \cdot 1^T \cdot \left| y_t^5 - S_t^{end} \right| \\
&\text{Subject to} \\
C1: \quad & V_t^i = V_t^{i-1} + A_t \cdot Q_t^i - \Delta \cdot d_t \cdot 0.25 && \forall i = 1 \ldots 4 \\
C2: \quad & P_t^j = P_t^{0,j} && \forall j \in J_t^{off} \\
C3: \quad & P_t^j = P_t^{end,j} && \forall j \in J_t^{on} \\
C4: \quad & P_t^j = 0.5 \cdot \left(P_t^{0,j} + P_t^{end,j} \right) && \forall j \in \widehat{J}_t^{off} \cap \widehat{J}_t^{on} \\
C5: \quad & Q_t^{i,l} = y_t^{i,j} \cdot P_t^j && j = J^l \;\; \forall l = 1 \ldots dim\left(Q_t^i\right) \;\; \forall i = 1 \ldots 4 \\
C6: \quad & y_t^{i+1,j} \le y_t^{i,j} && \forall j \in J_t^{off} \quad \forall i = 1 \ldots 4 \\
C7: \quad & y_t^{i+1,j} \ge y_t^{i,j} && \forall j \in J_t^{on} \;\; \forall i = 1 \ldots 4 \\
C8: \quad & y_t^{i,j} = S_t^{0,j} && \forall j \in \widehat{J}_t^{off} \cap \widehat{J}_t^{on} \;\; \forall i = 1 \ldots 4 \\
C9: \quad & \underline{V}^j - M \cdot y_t^{i+1,j} \le V_t^{i,k} \le \overline{V}^j + M \cdot \left(1 - y_t^{i+1,j} \right) && j = J^k \;\; \forall k = 1 \ldots dim\left(V_t^i\right) \;\; \forall i = 1 \ldots 4 \\
C10: \quad & d_t \ge 0 \\
C11: \quad & y_t^{i,j} = \{0, 1\}
\end{aligned} \quad (4)$$

where V_t^i is a vector of tank volume at quarter i of hour t with the elements $V_t^{i,k}$for tank k; $V_t^0, V_t^{end}$are SCADA values for the volumes at the beginning and end of the hour respectively; S_t^0, S_t^{end} are vectors of SCADA values for the pump statuses at the beginning and end of the hour, respectively with the elements $S_t^{0,j}$ and $S_t^{end,j}$; $P_t^{0,j}, P_t^{end,j}$ are SCADA values for pump j's flow rate at the beginning and end of the hour respectively; P_t^j is pump j's flow rate at time t; $A_t$is the network incidence matrix; $Q_t^i$is a vector of link flows with the elements $Q_t^{i,l}$for link l; $\Delta$is a given transformation matrix from DMA demands to nodal demands; $d_t$is a vector of DMA demands; $y_t^{i,j}$are binary variables which take a value of 1 if pump j is On during quarter i; M is a large scalar value; $\underline{V}^j, \overline{V}^j$are tank volumes which trigger pump j On and Off respectively; ω is a weighting factor; $J^l$is the pump connected to link l; $J_t^{off}$is the set of pumps switched off during time t;J_t^{on} is the set of pumps switched on during time t; $\widehat{J}_t^{off} \cap \widehat{J}_t^{on}$ is the set of pumps which did not switch status during time t; and $J^k$is the pump controlled by tank k.

The optimization problem in Eq. (4) is a MILP problem which results in the demand values for all DMAs. Note that the absolute values in the objective function can be easily transformed into linear constraints. The first constraint represents the water balance in the tanks over the course of 1 h at 0.25-h time steps. Constraints 2–4 consider the non-zero flow values during any hour, while Constraint 5 represents all the pump operation combinations (assuming a status change can only occur at the end of each quarter-hour). Constraints 6–7 guarantee that any pump will only be triggered On or Off once during the hour. Constraint 9 represents the control rules and governing volumes for each pump-tank pair. For example, when $V_t^{i,k} > \overline{V}^{j^k}$, the value will be $y_t^{i+1,j^k} = 0$ in order to satisfy the constraint, and thus pump J^k will be turned off during the next quarter of the hour. Constraint 6, in turn, guarantees that it will be turned off during the remainder of the hour.

It is worth noting that the estimation method described above is based solely on the network water balance at each time step and thus does not require the demand's statistical properties. Moreover, the demand estimation methodology is scalable to large networks (i.e. the size of the MILP does not increase significantly), because: 1) The MILP is formulated for each time step independently meaning that extended simulation has no impact on its size; 2) The MILP is formulated on the DMA scale (i.e. an aggregated scale), meaning that the size of the problem is a function of the number of DMAs, not a function of the number of pipes and nodes. The MILP will thus be tractable in large networks since the number of DMAs is significantly smaller than number of nodes.

2.4. Error classification

The demands derived from the MILP model are used as an input for an EPANET simulation such as to produce reference values for

the system's hydraulics (i.e. pump flows, status, pressures and tank levels). These reference values are then used for generating errors between the model and SCADA readings. We first generate the errors in the training dataset (i.e. Dataset 1), which represents the no-attack normal operation of the system. The different components' errors are then analyzed in order to calculate thresholds for "normal" over- and under-estimation errors (i.e. positive and negative errors). These thresholds are then used for defining outliers in the error vectors. We follow a multilevel approach in which we record the thresholds for different moving average values of training dataset error. That is to say, given the error vector e_i for hydraulic components (e.g. flow valve, pressure sensor or tank), it becomes possible to calculate moving averages with $0-K$ lags for each hydraulic component i , thus producing $K+1$ error vectors, $e_i^k \ \forall k = 0...K$. We then define two error thresholds as the α and $1-\alpha$ percentiles for each of these error vectors, which are in turn defined as $\underline{e}_i^k$and $\overline{e}_i^k$ respectively. These thresholds will then be used for identifying outliers in datasets which include cyberattacks (i.e. Dataset 2). To declare an outlier at time t with moving average lag-k we require that at least δ_1 hydraulic components experience an outlier. However, the declaration of an event from the system requires that at least δ_2 out of the $K+1$ moving averages outliers are flagged.

An outlier in hydraulic component i and moving average lag-k is declared if the error at time t, $e_{i,t}^k$, is below $\underline{e}_i^k$or above $\overline{e}_i^k$. As such, this outlier's identification process produces $(K+1)\cdot I$ outlier vectors O_i^k, with the element $O_{i,t}^k$being 1 if time t of hydraulic component i and moving average lag-k is an outlier, with $O_{i,t}^k$ being zero otherwise (Eq. (5)).

$$O_{i,t}^k = \begin{cases} 1 & if \ \ e_{i,t}^k < \underline{e}_i^k \ \ or \ \ e_{i,t}^k > \overline{e}_i^k \\ 0 & else \end{cases} \tag{5}$$

Given the outliers' values at time t, $O_{i,t}^k \ \forall i \forall k$, it is necessary to define a decision rule that would declare an alarm based on the obtained outlier values. For the purposes of the present study, this decision rule is defined in Eqs. (6) and (7).

$$b_t^k = \begin{cases} 1 & if \quad \sum_{i=1}^{I} O_{i,t}^k \geq \delta_1 \\ 0 & else \end{cases} \tag{6}$$

$$a_t = \begin{cases} 1 & if \quad \sum_{k=0}^{K} b_t^k \geq \delta_2 \\ 0 & else \end{cases} \tag{7}$$

where a_t is the value of the alarm at time t, which takes a value of one if an alarm is declared and zero otherwise; $b_t^k$is an auxiliary variable used for defining a_t; and δ_1, δ_2 are parameters which should be calibrated in order to optimize system performance.

The decision rule in Eq. (6) indicates that we require at least δ_1 hydraulic components to experience an outlier error (i.e. the definition of b_t^k) if we are to declare an alarm at time t with moving average lag-k. On the other hand, at least δ_2 out of the $K+1$ moving average alarms need to be flagged (i.e. the definition of a_t) if we are to declare a system alarm.

Fig. 2 demonstrates the alarm identification process for an illustrative example with four hydraulic components while setting $K = 2$,$\delta_1 = 2$, and $\delta_2 = 3$. Given the error vectors $e_i \ \forall i = 1..4$ and the thresholds $\underline{e}_i^k$, $\overline{e}_i^k \ \forall i \forall k$ (which are obtained from the $\alpha = 0.01$ and $1-\alpha = 0.99$ percentiles of the attack-free dataset) the alarms identification process is summarized in Fig. 2. Since $K = 2$, the four given error vectors are used for building the moving average errors for lags 0–2 and thus creating 12 error vectors (i.e. four for each moving average), as can also be seen in Fig. 2. The outlier identification process which defines the three binary vectors, $b^k \ \forall k = 0..2$, is based on $\delta_1 = 2$. In other words, we require two hydraulic component outliers in order to define an outlier from moving average lag-k. In moving average lag-1 at time 1278, for example, the errors of component 2 and component 4 exceed the error thresholds and thus b^1 it has a value of 1 at time 1278. The alarm identification process is based on $\delta_2 = 3$, which implies that at least three moving average outliers should be declared in order to raise an alarm. This condition is satisfied at time 1278, where the three vectors, $b^k \ \forall k = 0..2$, have a value of one.

2.5. Pre-solve procedure

In addition to the model-based analysis, a pre-solve procedure is performed in order to check the SCADA input data for times in which there is a contradiction in the physical rules/parameters, between the SCADA data and specific elements in the analyzed network. This procedure does not depend on the demand derivation and on the EPANET simulation and thus, in a sense, is "deterministic." The occurrence of an event can thus be declared with a high degree of reliability in cases of physical rule violations. The pre-solve procedure includes three simple checks:

a. Do the reported pump or valve statuses fit the reported flows? If the pump status is reported as Off, for example, then the flow should be zero. An alarm is declared if this is not the case.
b. Does the reported tank level meet the tank's physical boundaries?
c. Does the pump's operation point (the pump head and flow) fall on the pump's characteristic curve? More specifically, since it is possible to calculate the pump's head given the inlet and the outlet pressure in each pump station, this head should fall on the pump's characteristic curve for the reported flow rate. An alarm is declared if this is not the case.

The alarms declared during the pre-solve procedure are added to the alarms indicator a_t. In other words, $a_t = 1$if at time t the answer to one or more of the above three questions is no.

2.6. Calibration

We used enumeration in the calibration phase in order to determine the optimal values of α, δ_1, δ_2 which maximize a performance measure that accounts for the True Positive Rate (TPR), True Negative Rate (TNR) and Average Detection Time (ADT) as defined in Eq. (8). In this enumeration procedure, we first discretize the three parameters within a predetermined range to create a 3D grid, and then evaluate the classification performance in each grid point. The grid point with the best performance is selected as possessing the best values for the three parameters of the classification methodology. It should be noted that this calibration process is performed only after the errors are generated. As such, the calibration problem will always include three parameters, regardless of WDS network size, meaning that the calibration problem is solvable by the enumeration procedure even when a large-scale network is analyzed.

$$\eta = \omega_1 \cdot TPR + \omega_2 \cdot TNR + \omega_3 \cdot \eta_{ADT} \tag{8}$$

where η is the overall CDS performance measure; $\omega_i \ \forall i = 1...3$ are weighting factors, and $\eta_{ADT}$is the ADT performance measure which scales the ADT to a 0–1 range similar to the TPR and TNR measures.

The TPR and the TNR are defined in Eqs. (9) and (10).

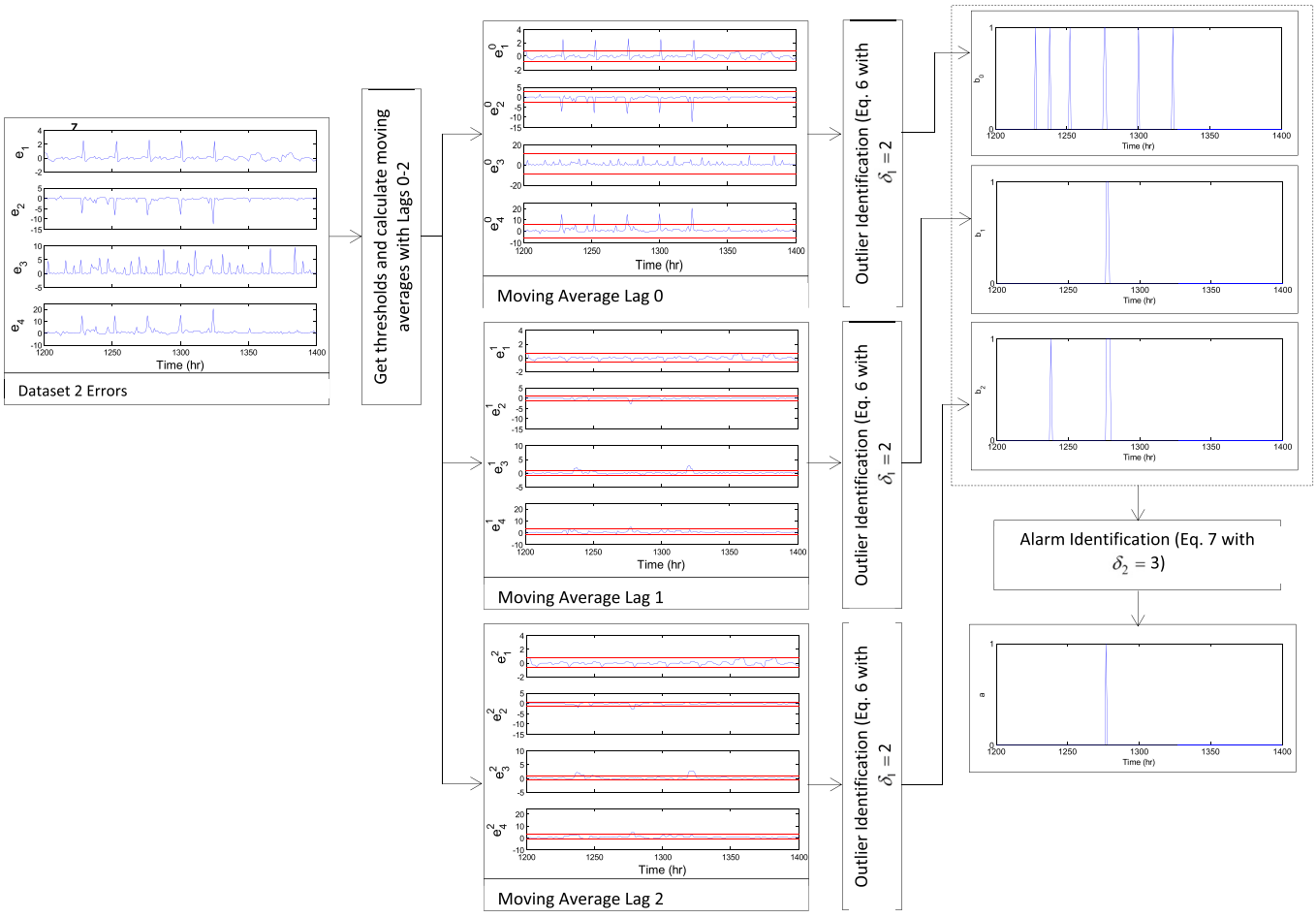

Fig. 2. Illustrative example of the alarm identification process with four hydraulic components and where $K = 2$, $\delta_1 = 2$, and $\delta_2 = 3$.

$$TPR = \frac{TP}{TP + FN} \tag{9}$$

$$TNR = \frac{TN}{TN + FP} \tag{10}$$

where TP is true positive alarms, that is, the number of times the system was under an attack that was recognized by the algorithm; FN is false negative alarms, that is, the number of times where the algorithm failed to detect that the system was under attack; TN is true negative alarms, that is, the number of times the algorithm identified situations in which the system was not under attack correctly; and FP is false positive alarms, that is, the number of times in which the algorithm generated an alarm incorrectly because the system was not actually under attack.

The detection time of the j-th event, DT_j, is defined as in BATADAL (2016) by using the difference between the time, $t^s_{a,j}$, at which an alarm is raised and the time, $t^s_{e,j}$, at which the event started. If the event was not detected, we set DT_j as the event duration, Δt_j, as given in Eq. (11).

$$DT_j = \min\left(t^s_{a,j} - t^s_{e,j}, \Delta t_j\right) \tag{11}$$

The ADT performance measure, η_{ADT}, is thus as defined in Eq. (12).

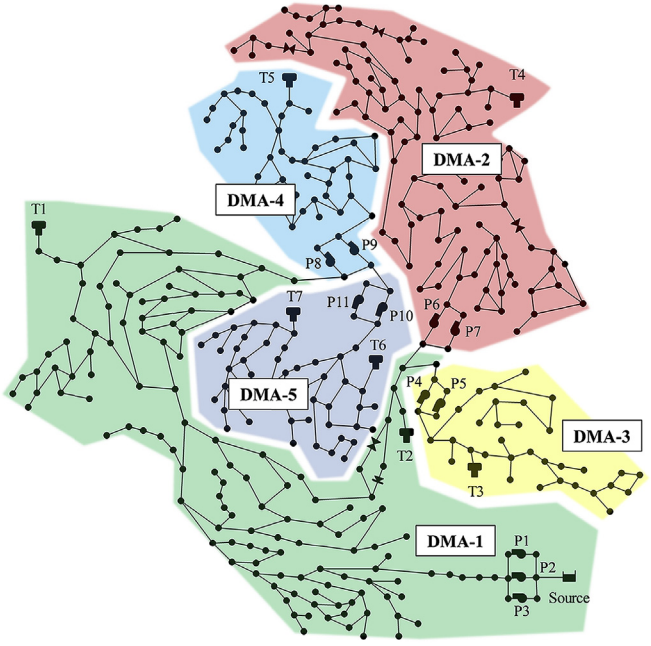

Fig. 3. The C-TOWN network; the colors represent the defined District Metering Areas (DMAs). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

$$\eta_{ADT} = 1 - \frac{1}{n_e}\sum_{j=1}^{n_e}\frac{DT_j}{\Delta t_j} \tag{12}$$

where n_e is the total number of events. η_{ADT} ranges between 0 and 1 with $\eta_{ADT} = 1$ representing an ideal case in which all events are immediately detected, and $\eta_{ADT} = 0$ representing the worst possible performance where none of the events are detected.

3. Application

The CDS proposed above has been implemented on the data published in the BATtle of the Attack Detection ALgorithms (BATADAL) competition (BATADAL, 2016). The competition considers the C-Town WDS layout and hourly SCADA readings from each of the five pumping stations and seven tanks as well as one flow control valve (Fig. 3). The SCADA readings include records for the flow rate, the inlet and outlet pressure, and the status of each pump in every pumping station, as well as the system tanks' water levels. The BATADAL data included three datasets: (a) Dataset 1, which includes readings from 12 months before the installation of cyber-physical devices (e.g. devices with telemetric capabilities, such as pressure readings, tank level sensors, etc.). As such, it is guaranteed to be attack-free and can be used for studying normal system operations; (b) Dataset 2, which includes readings from approximately 6 months after the installation of the cyber-physical devices. This dataset includes attacks which were discovered by the operators and thus these attacks are labeled attacks; (c) Dataset 3, which includes readings from 3 months of labeled attacks. Datasets 1, 2 were used for the system calibration, while Dataset 3 was used to test its performance. The attacks in Datasets 2, 3 were simulated using the epanetCPA toolbox (Taormina et al., 2017).

3.1. Results

The developed CDS was coded using MATLAB and begins with a pre-solve procedure which includes three simple checks for the violation of trivial physical conditions as detailed in the previous section. In order to demonstrate the pre-solve procedure, let us examine the operation points of pump 10 as shown in Fig. 4. The figure shows the pump head, calculated as the difference between the outlet and the inlet pressure, versus the pump flow rate. Under normal conditions, any point should fall on the pump's characteristic curve. Fig. 4 shows that the points for times 2028–2051 do not correspond to the characteristic curve and thus indicate abnormal SCADA readings which violate the laws of physics applicable to the pump. And, indeed, the system was in fact under attack as described by the second attack in Table 1 during the aforementioned time points (2028–2051).

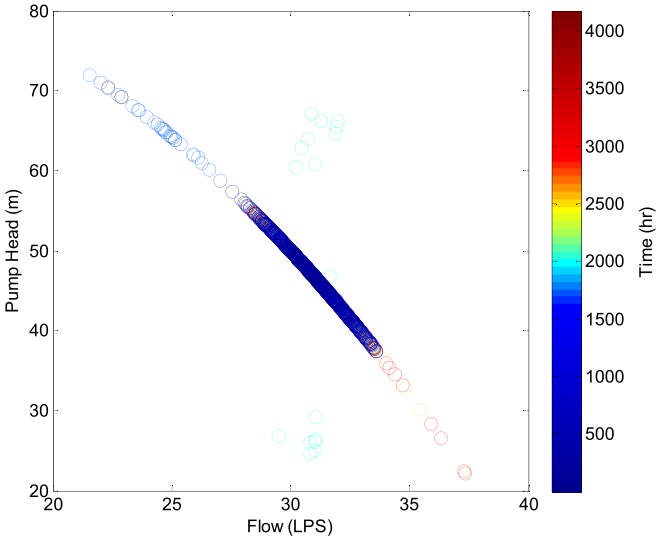

Fig. 4. Pump 10 operation points as derived from Dataset 2 SCADA readings.

The CDS begins the actual model-based detection after the pre-solve procedure. As shown in Fig. 1, the model-based methodology relies on estimated demands derived from the MILP model in order to produce reference values for the system's normal behavior. As such, the quality of these reference values depends on the accuracy of the estimated demands. The MILP program was formulated using Yalmip toolbox (Lofberg, 2004) where CPLEX 12.6.1 (International Buisness Machines, 2014) was used as the optimization problem solver. As explained previously, the MILP is formulated for each time step independently, as such it is a small size MILP (43 continuous variables and 60 binary variables) which is solved efficiently (average CPU time 0.25 s, on i7-4700MQ CPU @ 2.4 GHz machine).

Using the original demand patterns (which are not observable to our detection methodology), Fig. 5 presents the performance of the demand estimation methodology with respect to the five DMAs and with respect to the total demand of the C-Town network (Fig. 3). The results show the histograms of the relative errors for the six estimated DMA demands when using Dataset 1. All estimation errors have near zero mean and a standard deviation of up to 0.3%, demonstrating the accuracy of the demand estimation methodology.

By relying on Dataset 1 and Dataset 2, and after applying the demand estimation methodology and calculating the error vectors, it becomes possible to calibrate CDS parameters α, δ_1, and δ_2 against a dataset with labeled attacks. With this in mind, our calibration phase employed BATADAL competition Dataset 2, which includes seven attacks as detailed in Table 1. We then determined the optimal values for the three parameters which maximize the CDS's overall performance, η, where $K = 9$. These optimal values are 0, 2, and 2 for α, δ_1, and δ_2 respectively, and using them produces a system performance score of $\eta = 0.97$ on Dataset 2.

As can be seen in the calibration result illustrated in Fig. 6, the CDS captured all seven of the attacks detailed in Table 1 during the calibration phase, but at the cost of three false alarms. Still, the value of η, which is very close to 1, indicates that near-perfect performance is obtained in the calibration phase because these false alarms are very short and do not exert a significant impact on the performance measure.

We further tested the methodology by applying the CDS on Dataset 3, which also includes seven attacks as detailed in Table 2. This dataset is provided in BATADAL (2016) for testing purposes, and thus was not used in the CDS's development or calibration.

According to Table 2, there is an attack during time 298–367 which impacts Tank T3 and Pump 4's operation in DMA 3. Fig. 7 presents the data readings for Tank 3 and Pump 4 with the readings during event time in red. However, Fig. 7 alone is insufficient for distinguishing any abnormal behavior during the attack through the mere examination of these readings. Nevertheless, and since the proposed CDS examines all readings simultaneously for violations of physical behavior and/or operation rules, the CDS is still capable of detecting that there is something wrong with the readings at these points in time.

Table 1
Description of attacks included in the training Dataset (i.e. Dataset 2).

Event No.	Start time step	End time step	Event description
1	1728	1777	Changing the pump 10 operation rules (which are based on tank T7) and altering T7's water level readings.
2	2028	2051	Similar attack to 1 with additional concealment in the upstream pumping station's flow and status readings
3	2338	2397	False low levels in T1 which alters the pumps' operation and causes a concealed overflow
4	2828	2921	Similar to 3 with additional concealment in the upstream pumping station's flow, pressure and status readings
5	3498	3557	Pump 7 specific speed reduced to 0.9 of the nominal speed
6	3728	3821	Pump 7 specific speed reduced to 0.7 of the nominal speed with additional concealment in the downstream tank's (T4) level
7	3928	4037	Similar to 6 with additional concealment in pump 6 &7's flow and status readings

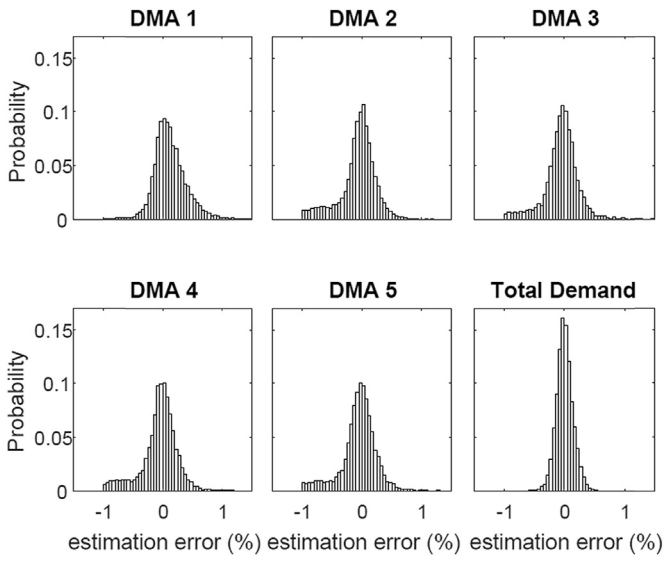

Fig. 5. Performance of the demand estimation methodology.

Fig. 8 presents the outliers from the tanks and pressure sensors for different moving average lags (note: we only show 4 out of the 10 moving average lags 0–9). Each of the subfigures in Fig. 8 shows the different hydraulic components on the y-axis, that is Tanks 1–7, Pumps 1–11, and pressure sensors 1–12. The blue stamps are added when an outlier is obtained in one of the components' signals.

The results show that several components in different lags were able to detect the attacks. For example, the third event was detected by many components in moving average lags 0–9. Nevertheless, some attacks were better captured in the large moving average lags. For example, the first event is detected by two components in moving average lags 0 and 1, but by five components in moving average lags 8 and 9.

Our error classification methodology requires one outlier vector from each moving average lag (Eq. (6)). This outlier's vector is obtained by requiring $\delta_1 = 2$ components to violate the thresholds. For example, the outliers in moving average lag 0 between times 0–1500 will not give rise to an outlier in moving average lag 0 indicator b_t^0 because only one component flagged an outlier at each time instance. The events identification process uses the moving average outliers,b_t^k, to raise an alarm for a cyber-attack event. More

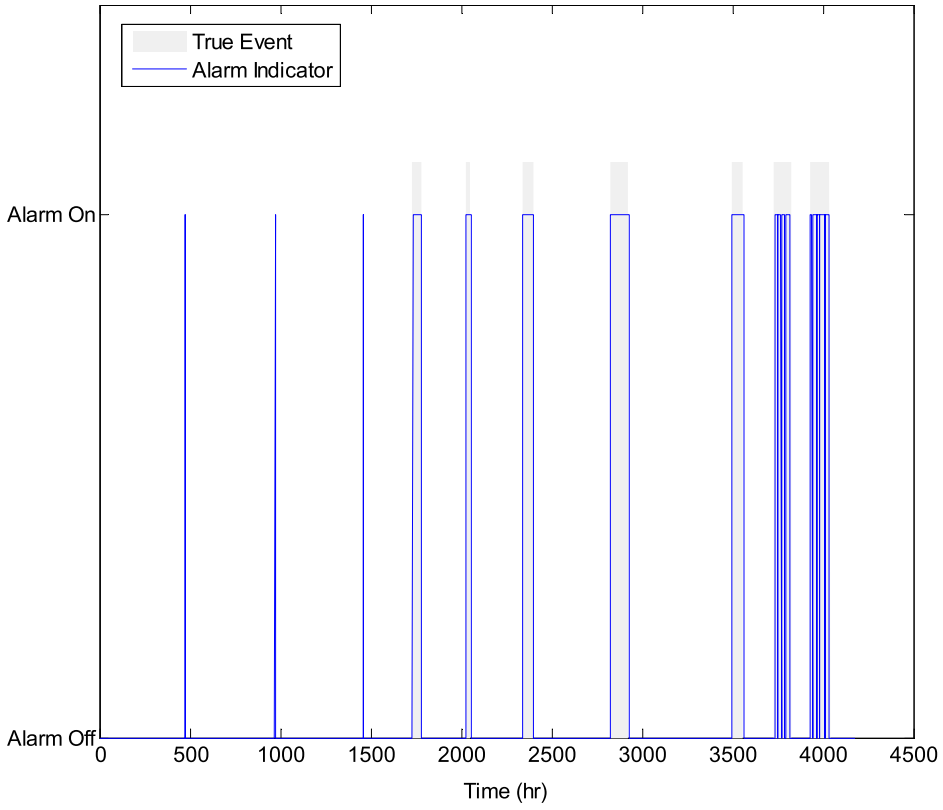

Fig. 6. Calibration results of the Cyber-attack Detection System on Dataset 2.

Table 2
Description of the attacks in the testing Dataset (i.e. Dataset 3).

Event No.	Start time step	End time step	Event description
1	298	367	Changes to the operation rules of pumps 4 and 5 (which are based on tank T3) and alterations of T3's water level readings, pump flows and pump statuses.
2	633	697	Alterations of water level readings in T2 which are sent to the valve operation rules while concealing the low water level changes from the SCADA.
3	868	898	Pump 3 is switched on
4	938	968	Pump 3 is switched on
5	1230	1329	Like attack 2 but with additional concealment of the status, flow and the inlet/outlet pressure readings in the valve
6	1575	1654	Changes to the operation rules of pumps 10 and 11 (which are based on tank T7) with concealments of T7's water level readings, pump flows, statuses and inlet/outlet pressures.
7	1941	1970	False level readings in T4

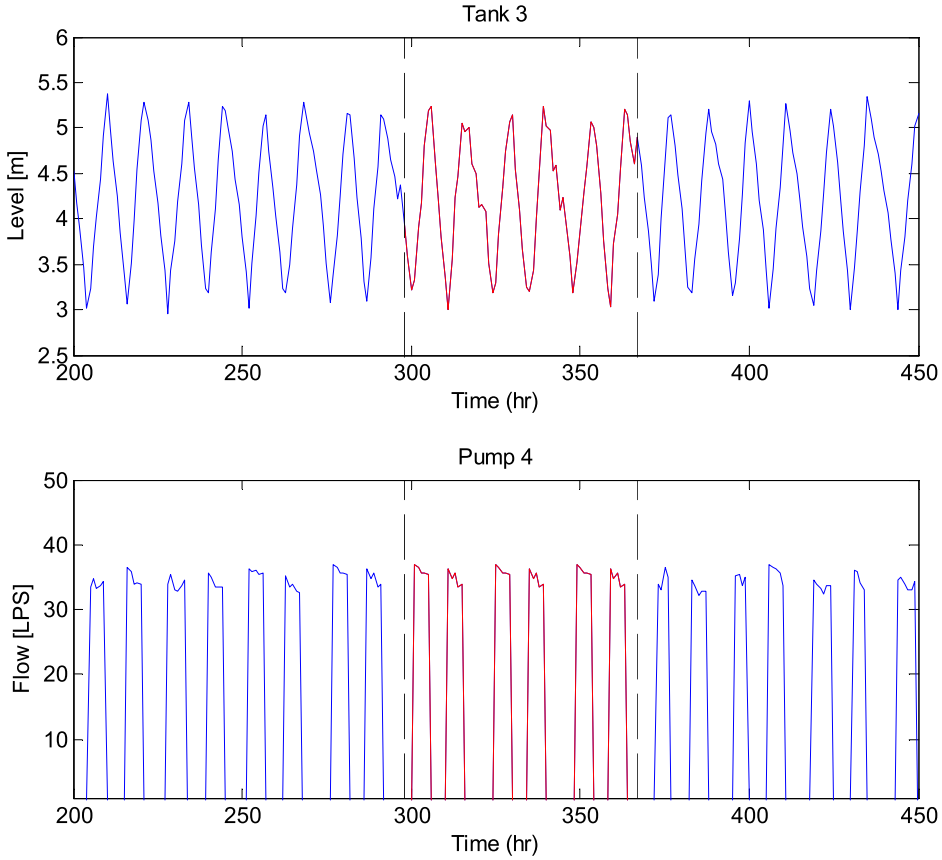

Fig. 7. SCADA readings for Tank 3 and Pump 4 in Dataset 3 in adjacent and during event No. 1.

specifically, an alarm is raised if at least $\delta_2 = 2$ moving average outliers are found. As such, and since none of $b_t^k \ \forall k = 0..9 \ \forall t = 0..1500$ is flagged, no alarm is raised in the course of times 0–1500.

Fig. 9 presents the algorithm's performance for the calibrated values $\delta_1 = 2$ and $\delta_2 = 2$. The results show that the alarms (i.e. aggregation of the pre-solve and the model-based procedure) were capable of detecting all the labeled events with one false positive and an overall performance of $\eta = 0.99$. The pre-solve procedure was able to capture some of the events too, as also shown in Fig. 9. The third event, for example, was detected by the pre-solve procedure because the SCADA reading during this period did not pass the third check, that is to say, the pumps' pressure difference and flow rate did not fall on the pump's characteristic curve (in a manner akin to the case presented in Fig. 4). Nevertheless, since the pre-solve procedure cannot indicate events such as tank levels altering or events with good concealment, it is not recommended as a standalone procedure. In such case, it will yield poor performance of TPR = 0.25, TNR = 1 (since the pre-solve has no FP), and overall score of $\eta = 0.81$.

3.2. Sensitivity analysis

In order to examine the sensitivity of the algorithm to errors in the demand estimation, we have added a normally distributed noise to the estimated DMAs' demand which are obtained from the MILP model. The sensitivity of the CDS is examined in Fig. 10 which presents the performance of the model-based model for different

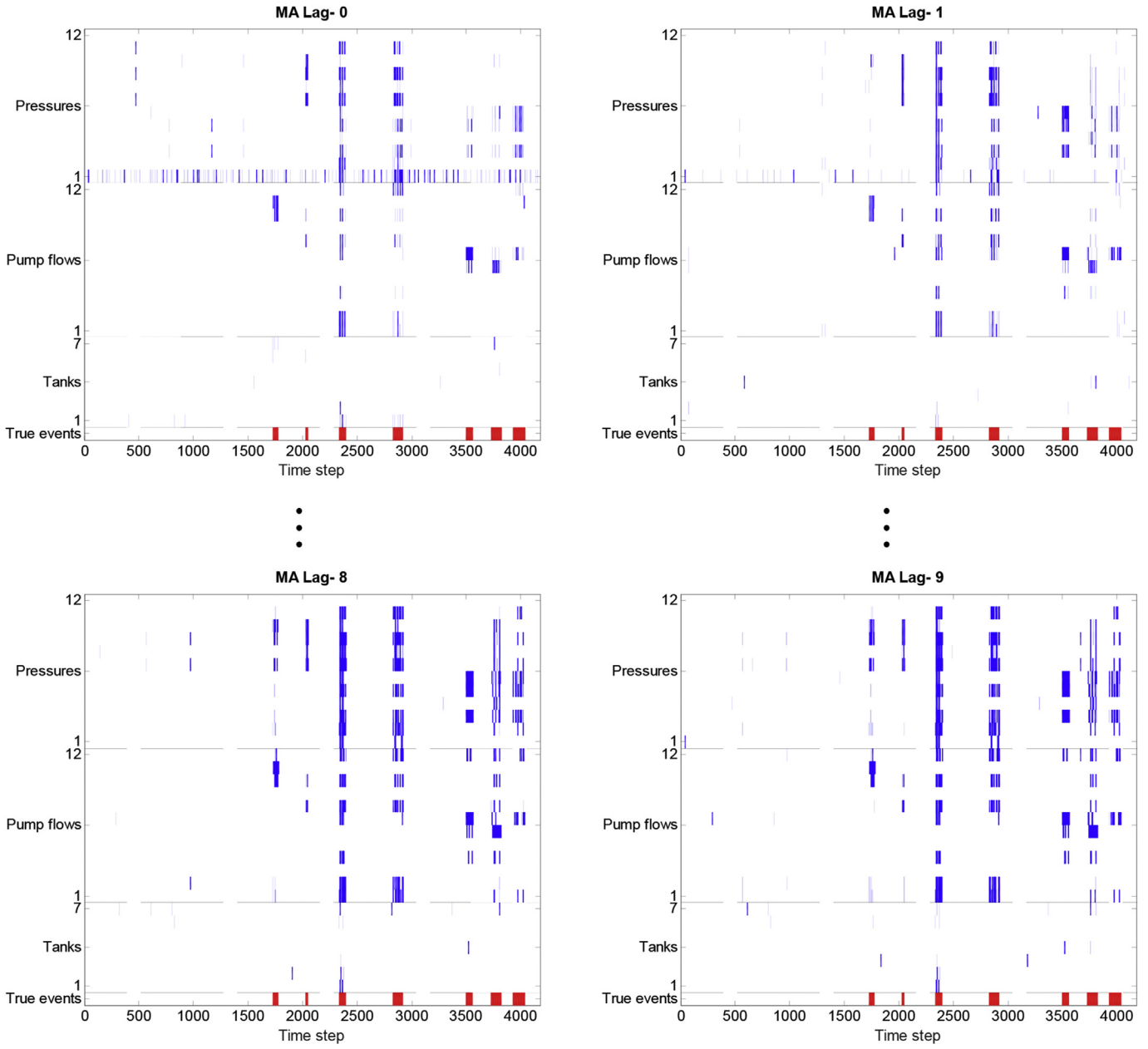

Fig. 8. Outlier identifications for different hydraulic components of the system in different Moving Average (MA) Lags.

levels of noise. We have considered three level of noise which are normally distributed around zero with different level of coefficient of variation (CV). The results show only minor deterioration even in relatively high CV values. This robust performance is attributed to the fact that the calibration procedure is performed after the demand estimation, thus if high uncertainty is expected in the demand estimates, the calibration procedure can tune the thresholds of the CDS to account for this high uncertainty. The performance index (η) is 0.98, 0.97, 0.96 for CV of 0.05, 0.15, 0.25, respectively.

4. Conclusions

The present study offered a model-based FD methodology which utilizes a physically-based water hydraulics simulation model (EPANET) for detecting cyberattacks on WDSs. The method utilizes a three-phase approach in which: 1) the demand is estimated based on part of the SCADA readings; 2) a hydraulic model is used for checking whether the SCADA's hydraulic data corresponds to the estimated demand; 3) A multilevel classification approach is then implemented in order to classify the obtained errors into outlier and normal errors. The results show that the proposed methodology was able to raise a timely alarm for all the labeled simulated cyberattack events. However, the results presented herein assumed (unrealistically) perfectly calibrated hydraulic model. Thus, the only source of uncertainty was the network demand data which was not provided in BATADAL datasets. A more generalized case would require us to account for the uncertainty in the calibration of an imperfect model as well as for other sources of uncertainty such as sensor noise. Thus, future research should examine the required calibration's goodness-of-fit and its impact on the detection algorithm's performance.

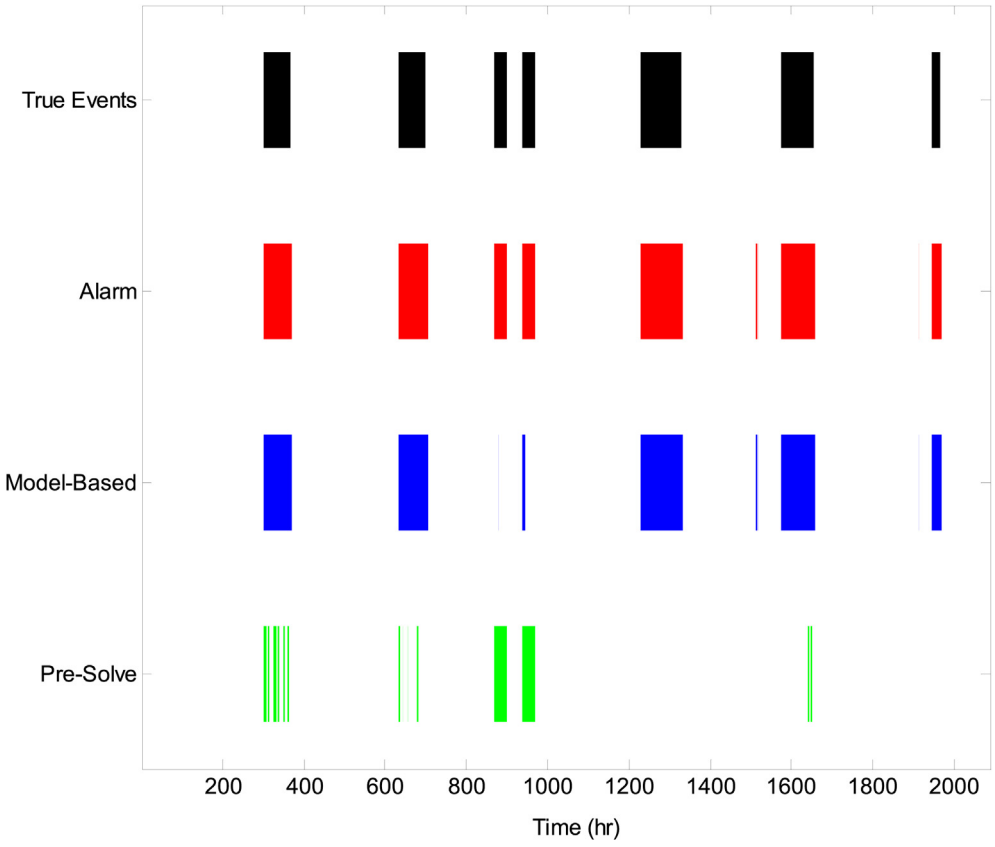

Fig. 9. Detection system performance on the test dataset (i.e. Dataset 3).

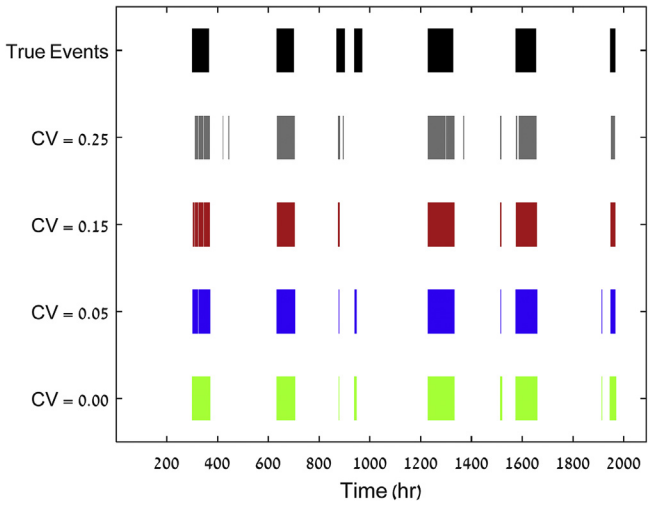

Fig. 10. Model-Based detection performance for different level of noise in demand estimation. CV=Coefficient of Variation.

Acknowledgments

This research was made possible by the financial support of the Israeli Water Authority (grant #4501284516), The Center for Cyber Law & Policy and The Minerva Center for the Rule of Law. The authors would also like to thank the BATtle of the Attack Detection ALgorithms (BATADAL) organizers: Dr. RICCARDO TAORMINA, Prof. STEFANO GALELLI, Dr. NILS OLE TIPPENHAUER, Prof. AVI OSTFELD, Mr. ELAD SALOMONS and Dr. DEMETRIOS ELIADES for organizing the battle and providing the data.

References

Amin, S., Litrico, X., Sastry, S.S., Bayen, A.M., 2013b. Cyber security of water SCADA systems—Part II: attack detection using enhanced hydrodynamic models. IEEE Trans. Contr. Syst. Technol. 21 (5), 1679–1693.

Amin, S., Litrico, X., Sastry, S., Bayen, A.M., 2013a. Cyber security of water SCADA systems—Part I: analysis and experimentation of stealthy deception attacks. IEEE Trans. Contr. Syst. Technol. 21 (5), 1963–1970.

Arad, J., Housh, M., Perelman, L., Ostfeld, A., 2013. A dynamic thresholds scheme for contaminant event detection in water distribution systems. Water Res. 47 (5), 1899–1908.

BATADAL, 2016. BATtle of the Attack Detection ALgorithms (BATADAL). http://www.batadal.net (accessed 26 September 2017).

Bobat, A., Gezgin, T., Aslan, H., 2015. The SCADA system applications in management of Yuvacik Dam and Reservoir. Desalination and Water Treatment 54 (8), 2108–2119.

Bobba, R.B., Rogers, K.M., Wang, Q., Khurana, H., Nahrstedt, K., Overbye, T.J., 2010. Detecting false data injection attacks on dc state estimation. In: Proceedings of the First Workshop on Secure Control Systems, CPSWEEK, vol. 2010.

Cheung, S., Dutertre, B., Fong, M., Lindqvist, U., Skinner, K., Valdes, A., 2007. Using model-based intrusion detection for SCADA networks. In: Proceedings of the SCADA Security Scientific Symposium, vol. 46, pp. 1–12.

Cleveland, F.M., 2008. Cyber security issues for advanced metering infrastructure (AMI). In: Power and Energy Society General Meeting-conversion and Delivery of Electrical Energy in the 21st Century. IEEE, pp. 1–5, 2008 IEEE.

Dakin, R., Newman, R., Groves, D., 2009. The case for cyber security in the water sector. American Water Works Association. J. 101 (12), 30–32.

Ericsson, G.N., 2010. Cyber security and power system communication—essential parts of a smart grid infrastructure. IEEE Trans. Power Deliv. 25 (3), 1501–1507.

Esmalifalak, M., Shi, G., Han, Z., Song, L., 2013. Bad data injection attack and defense in electricity market using game theory study. IEEE Transactions on Smart Grid 4 (1), 160–169.

Freeman, P., Pandita, R., Srivastava, N., Balas, G.J., 2013. Model-based and data-driven fault detection performance for a small UAV. IEEE/ASME Transactions on Mechatronics 18 (4), 1300–1309.

Gertler, J., 1998. Fault Detection and Diagnosis in Engineering Systems. CRC press.

Giani, A., Bitar, E., Garcia, M., McQueen, M., Khargonekar, P., Poolla, K., 2013. Smart grid data integrity attacks. IEEE Transactions on Smart Grid 4 (3), 1244–1253.

Horta, R., 2007. The city of Boca Raton: a case study in water utility cybersecurity. Am. Water Works Assoc. J. 99 (3), 48–50.

Housh, M., Ostfeld, A., 2015. An integrated logit model for contamination event detection in water distribution systems. Water Res. 75, 210–223. https://doi.org/10.1016/j.watres.2015.02.016.
Housh, M., Ohar, Z., 2017a. Integrating physically based simulators with Event Detection Systems: multi-site detection approach. Water Res. 110, 180–191.
Housh, M., Ohar, Z., 2017b. Model based approach for cyber-physical attacks detection in water distribution systems. In: World Environmental and Water Resources Congress, pp. 727–736, 2017.
Hug, G., Giampapa, J.A., 2012. Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks. IEEE Transactions on Smart Grid 3 (3), 1362–1370.
International Business Machines Corporation, 2014. IBM ILOG CPLEX V12.6.1: Users Manual for CPLEX. IBM, New York.
ICS-CERT, 2016. NCCIC/ICS-CERT Year in Review: FY 2015. Report No. 15–50569. U.S. Department of Homeland Security – Industrial Control Systems-Cyber Emergency Response Team, Washington, D.C.
Kim, T.T., Poor, H.V., 2011. Strategic protection against data injection attacks on power grids. IEEE Transactions on Smart Grid 2 (2), 326–333.
Kosek, A.M., Gehrke, O., 2016. Ensemble regression model-based anomaly detection for cyber-physical intrusion detection in smart grids (2016, October). In: Electrical Power and Energy Conference (EPEC). IEEE, pp. 1–7. IEEE.
Kosut, O., Jia, L., Thomas, R.J., Tong, L., 2010a. Limiting false data attacks on power system state estimation. In: Proceedings of the 44th Conference on Information Sciences and Systems.
Kosut, O., Jia, L., Thomas, R.J., Tong, L., 2010b. On malicious data attacks on power system state estimation. In: Proceedings of the 45th International Universities' Power Engineering Conference (UPEC'10).
Kosut, O., Jia, L., Thomas, R.J., Tong, L., 2010c. Malicious data attacks on smart grid state estimation: attack strategies and countermeasures. In: Proceedings of the IEEE Conference on Smart Grid Communications.
Lee, E.A., 2008. Cyber physical systems: design challenges. In: 2008 11th IEEE International Symposium on Object and Component-oriented Real-time Distributed Computing (ISORC), pp. 363–369.
Lewis, J.A., 2002. Assessing the Risks of Cyber Terrorism, Cyber War and Other Cyber Threats. Center for Strategic & International Studies, Washington, DC.
Lofberg, J., 2004. YALMIP: a toolbox for modeling and optimization in MATLAB. In: Computer Aided Control Systems Design, 2004 IEEE International Symposium on. IEEE, pp. 284–289.
Liu, Y., Ning, P., Reiter, M.K., 2011. False data injection attacks against state estimation in electric power grids. ACM Trans. Inf. Syst. Secur. 14 (1), 13.
Liu, X., Liu, X., Li, Z., 2015. Cyber risk assessment of transmission lines in smart grids. Energies 8 (12), 13796–13810.
Meseguer, J., Mirats-Tur, J.M., Cembrano, G., Puig, V., Quevedo, J., Pérez, R., Ibarra, D., 2014. A decision support system for on-line leakage localization. Environ. Model. Software 60, 331–345.
Negrete-Pincetic, M., Yoshida, F., Gross, G., 2009. Towards quantifying the impacts of cyber attacks in the competitive electricity market environment. In: PowerTech, 2009 IEEE Bucharest. IEEE, pp. 1–8.
Perelman, L., Amin, S., 2014. A network interdiction model for analyzing the vulnerability of water distribution systems. In: Proceedings of the 3rd International Conference on High Confidence Networked Systems, HiCoNS '14. ACM, New York, NY, USA, pp. 135–144.
Perez, R., Sanz, G., Puig, V., Quevedo, J., Escofet, M.A.C., Nejjari, F., Sarrate, R., 2014. Leak localization in water networks: a model-based methodology using pressure sensors applied to a real network in Barcelona [applications of control]. IEEE Contr. Syst. Mag. 34 (4), 24–36.
Rasekh, A., Hassanzadeh, A., Mulchandani, S., Modi, S., Banks, M.K., 2016. Smart water networks and cyber security. J. Water Resour. Plann. Manag. 142 (7).
Slay, J., Miller, M., 2008. Lessons learned from the maroochy water breach. In: Goetz, E., Shenoi, S. (Eds.), Critical Infrastructure Protection. Springer US, Boston, MA, pp. 73–82.
Spellman, F.R., 2013. Handbook of Water and Wastewater Treatment Plant Operations. CRC Press.
Taormina, R., Galelli, S., Tippenhauer, N.O., Salomons, E., Ostfeld, A., 2017. Characterizing cyber-physical attacks on water distribution systems. J. Water Resour. Plann. Manag. 143 (5).
Taormina, R., Galelli, S., Tippenhauer, N.O., Salomons, E., Ostfeld, A., 2016. BATtle of the Attack Detection ALgorithms (BATADAL) - Detailed Problem Description and Rules. Retrieved from: https://www.batadal.net/images/rules.pdf.
USEPA, 2013. EPANET 2.00.12. U.S. Environmental Protection Agency, Cincinnati, Ohio. http://www.epa.gov/nrmrl/wswrd/dw/epanet.html (accessed 10 August 2014).
Weimer, J., Araujo, J., Amoozadeh, M., Ahmadi, S.A., Sandberg, H., Johansson, K.H., 2013. Parameter-invariant actuator fault diagnostics in cyber-physical systems with application to building automation. In: Control of Cyber-physical Systems. Springer International Publishing, pp. 179–196.
Xie, L., Mo, Y., Sinopoli, B., 2010. False data injection attacks in electricity markets. In: Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on. IEEE, pp. 226–231.
Yang, D., Usynin, A., Hines, J.W., 2006. Anomaly-based intrusion detection for SCADA systems. In: 5th Intl. Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technologies, pp. 12–16.
Yuan, Y., Li, Z., Ren, K., 2011. Modeling load redistribution attacks in power systems. IEEE Transactions on Smart Grid 2 (2), 382–390.