Reducing Vulnerability to Cyber-physical Attacks in Water Distribution Networks

Nicolas Nicolaou, Demetrios G. Eliades, Christos Panayiotou, Marios M. Polycarpou
KIOS Research and Innovation Center of Excellence
Dept. of Electrical and Computer Engineering
University of Cyprus, Cyprus
Email: {nicolasn, eldemet, christosp, mpolycar}@ucy.ac.cy

***Abstract*—Cyber-Physical Systems (CPS), such as Water Distribution Networks (WDNs), deploy digital devices to monitor and control the behavior of physical processes. These digital devices, however, are susceptible to cyber and physical attacks, that may alter their functionality, and therefore the integrity of their measurements/actions. In practice, industrial control systems utilize simple control laws, which rely on various sensor measurements and algorithms which are expected to operate normally. To reduce the impact of a potential failure, operators may deploy redundant components; this however may not be useful, e.g., when a cyber attack at a PLC component occurs.**

In this work, we address the problem of reducing vulnerability to cyber-physical attacks in water distribution networks. This is achieved by augmenting the graph which describes the information flow from sensors to actuators, by adding new connections and algorithms, to increase the number of redundant cyber components. These, in turn, increase the *cyber-physical security level*, which is defined in the present paper as the number of malicious attacks a CPS may sustain before becoming unable to satisfy the control requirements. A proof-of-concept of the approach is demonstrated over a simple WDN, with intuition on how this can be used to increase the cyber-physical security level of the system.

I. Introduction

Water distribution systems are cyber-physical systems, whose objective is to transfer drinking water to consumers through a complex network comprised of structural elements (such as tanks, pipes and junctions), actuators (such as pumps and valves), as well as sensors (e.g., measuring flows, pressures, quality etc.). For their monitoring and control, Supervisory Control And Data Acquisition (SCADA) systems are used, composed of Programmable Logic Controllers (PLCs) and Communication Modules. Due to their large-scale structure, Water Distribution Systems are exposed to thefts and event attacks. Moreover, as new internet-enabled (IoT) devices are added to the system, new threats can appear, such as cyber-security attacks. From the cybersecurity literature, there has been significant interest in understanding of the challenges of industrial control systems [1], especially after the sophisticated Stuxnet attack to an Iranian power plant in 2010.

In addition to sensitive critical infrastructures, research has demonstrated that the risk of cyber attacks exists also for water distribution networks, e.g., in honeypot studies [2]. The cyber-security issues of water distribution systems has received attention in the previous years [3], and specifically through the "Battle of the Attack Detection Algorithms" (BATADAL), an algorithmic competition which was organized in 2017 in which participants aimed to detect cyber-physical attacks on a realistic water distribution system [4]. Within BATADAL, model-based [5] and data-driven [6], [7] approaches have been proposed for detecting cyber-physical attacks, using optimization, statistical outlier detection and neural networks.

Previous research has examined methods for detecting whether an event has occur. However, detection of an attack does not protect the system from instability or loss of control; this would require hardware and software redundancy which may not be available due to the initial design. The contribution of this work is the introduction of a methodology which aims to reduce the vulnerability to cyber-physical attacks in water distribution networks, by exploiting the analytical redundancies of the physical properties of the system. In specific, the contributions are:

- We introduce a methodology to compute the *cyber-physical security level* of a system, i.e, the number of cyber elements that need to be compromised, in order to affect the control of the physical system.
- We examine how the use of analytical redundancies of the physical properties of the system may enhance the system's cyber-physical security level.

This paper is structured as follows: Section II presents the problem formulation and introduces the cyber-physical security level; Section III, gives a solution methodology which exploits system-specific knowledge to maximize the cyber-physical security level is discussed, through a simple use-case. Section IV concludes the paper and future work is discussed.

II. Problem Formulation

The set of natural numbers is denoted by $\mathbb{N}$; additionally, $\bar{\mathbb{N}} = \mathbb{N} \cup \{0\}$. For any number $n \in \mathbb{N}$ we define $\mathbb{N}_n = \{x \in \mathbb{N} : x \leq n\}$ $\left(\text{resp. } \bar{\mathbb{N}}_n = \{x \in \bar{\mathbb{N}} : x \leq n\}\right)$ as the subset of numbers in $\mathbb{N}$ $\left(\text{resp. } \bar{\mathbb{N}}\right)$ up to and including n. Let k be the discrete time with sampling time Δt. In the notation used, $\wedge$ is the logical 'and', whereas $\vee$ is the logical 'or'.

We focus in Cyber-Physical systems (such as water distribution systems) that are composed of four main sets of elements:

1) a set $\mathcal{P} = \{p_1, p_2, \ldots, p_{N_p}\}$ of N_p plant components (such as water source, pipes, water tanks and pumps),
2) a set $\mathcal{S} = \{s_1, s_2, \ldots, s_{N_s}\}$ of N_s sensing components (such as pressure or flow sensors),

3) a set $\mathcal{C} = \{c_1, c_2, \ldots, c_{N_c}\}$ of N_c actuating components (such as contactors/relays), and
4) a set $\mathcal{L} = \{a_1, a_2, \ldots, a_{N_a}\}$ of N_a software agents (such as control algorithms and estimators), which are implemented by the ICT components of the system (such as PLCs, micro-controllers and communication modules).

We assume that, at discrete time k, $s_i(k) \in \mathbb{R}$ is the measurement of the i-th sensor, and $c_i(k)$ the i-th control signal. Furthermore, agent a_i is defined as a software algorithm which takes as arguments a subset of the sensing and outputs of other agents, as well as the discrete time, which returns a real value, such that $a_i : 2^{\mathcal{S}} \times 2^{\mathcal{L}} \times \mathbb{N} \mapsto \mathbb{R}$. Note that in this notation, $2^{\mathcal{S}}$ and $2^{\mathcal{L}}$ is the power set which corresponds to all the subsets of $\mathcal{S}$ and $\mathcal{L}$ respectively. Furthermore, we define $C_i \subseteq \mathcal{C}$ as the subset of actuating components that are controlled by a_i, $M_i \subseteq \mathcal{S}$ as the subset of sensing components that provide input to a_i, and $L_i \subseteq \mathcal{L}$, as the subset of agents that provide input to a_i.

Given the system with the physical and cyber components which were introduced above, we can model the Cyber-Physical system as a directed graph, defined as *logic graph* $G = (V, E)$, where:

$$
\begin{aligned}
V &= \{v : v \in \mathcal{C} \cup \mathcal{S} \cup \mathcal{L}\},\\
E &= E_c \cup E_s \cup E_l,\\
E_c &= \{(a_i, c_j) : a_i \in \mathcal{L} \ \wedge \ c_j \in C_i\},\\
E_s &= \{(s_j, a_i) : a_i \in \mathcal{L} \ \wedge \ s_j \in M_i\},\\
E_c &= \{(a_j, a_i) : a_i \in \mathcal{L} \ \wedge \ a_j \in L_i.\}
\end{aligned}
$$

From the definition, the logic graph G contains the interconnections between sensor measurements with agents, agents with other agents, and agents with control actuators. We say that a sensor s_i *controls* an actuator c_j in the logic graph G, if G contains a path that starts from sensing node s_i and ends at actuating node c_j. In practice, this graph may be constructed by processing the model of a WDN and its ICT infrastructure.

A. Control Graphs

In this section we define the concept of *control graphs*, which correspond to the set $\mathcal{G}_{\Sigma_j,c_j}$ of all the connected subgraphs in G, which contain sensing nodes from a set $\Sigma_j \subseteq \mathcal{S}$ that control the actuating node c_j, and computed though a function $g(\cdot)$. In specific, we consider that $\mathcal{G}_{\Sigma_j,c_j} = g(\langle \Sigma_j, c_j \rangle, G)$ computes the *control graphs* of c_j.

The z-th subgraph given a set of sensors Σ_j and an actuating node c_j is therefore defined as $G^z_{\Sigma_j,c_j} = (V^z_{\Sigma_j,c_j}, E^z_{\Sigma_j,c_j}) \in \mathcal{G}_{\Sigma_j,c_j}$, where $V^z_{\Sigma_j,c_j} = \Sigma_j \cup \{c_j\} \cup \Lambda$ such that Λ is a set of agents, and $\forall a_i \in \Lambda$, then $C_i \cup M_i \cup L_i \subseteq V_{\Sigma_j,c_j}$ is the set of nodes in the subgraph, and $E^z_{\Sigma_j,c_j}$ the set of 2-tuples which correspond to the edges between the nodes. Let, $G^0_{\Sigma_j,c_j}$ be the smallest, or so called *core*, control graph of c_j, i.e. $|V^0_{\Sigma_j,c_j}| \leq |V^z_{\Sigma_j,c_j}|$, for $z \in \mathbb{N}_{|\mathcal{G}_{\Sigma_j,c_j}|}$. Also, let $G_{\Sigma_j,c_j} = (V_{\Sigma_j,c_j}, E_{\Sigma_j,c_j})$ be the *complete* control graph of c_j, such that, $V_{\Sigma_j,c_j} = \bigcup_{z \in \mathbb{N}_{|\mathcal{G}_{\Sigma_j,c_j}|}} V^z_{\Sigma_j,c_j}$, and $E_{\Sigma_j,c_j} = \bigcup_{z \in \mathbb{N}_{|\mathcal{G}_{\Sigma_j,c_j}|}} E^z_{\Sigma_j,c_j}$. For simplicity, we consider networks that for each $c_j \in \mathcal{C}$ there is exactly a single control graph $G^0_{\Sigma_j,c_j}$ from some sensor $\Sigma_j = \{s_i\}$. In such a case we denote such graph $G^0_{s_i,c_j}$. It is easy to see that the above function may be used to return all the connected subgraphs between any set of sensors S and an element $v_j \in V$.

B. Adversarial Model

We assume an adversary $\mathcal{A}$ that is able to corrupt any node (physical, sensing, actuating or local) in graph G. We classify the adversaries in terms of the number of nodes they can infiltrate. For a set of subgraphs $\mathcal{G}_{\Sigma_j,v_j}$, let $\mathcal{A}^z_{\Sigma_j,v_j} = \{A : A \subseteq V_{\Sigma_j,v_j} - \{v_j\} \ \wedge \ |A| = z\}$, for $z \in \mathbb{N}_{|V_{S,v_j}|}$. That is, we assume that an adversary $\mathcal{A}^z_{\Sigma_j,v_j}$ can infiltrate any subset of z nodes in V_{Σ_j,v_j}.

C. Cyber-physical Security Level

Given a core control graph $G^0_{\Sigma_j,c_j}$ for a sensor-actuator pair $\langle S, c_j \rangle$ of a system graph G, we can define the *cyber-physical security level* of each element $v \in V^0_{S,c_j}$, that specifies the number of elements an adversary should infiltrate in order to disconnect the control graph $G_{\Sigma_j,v}$. Before, proceeding to the definition, let us define, for any $V' \subseteq V$, the operation

$$G \ominus V' = (V - V', E - \{(v_i, v_j) : v_i \in V' \ \vee \ v_j \in V'\}),$$

that returns a graph by removing the vertices in V' from G.

Definition 2.1: The cyber-physical security level of an element $v \in V^0_{\Sigma_j,c_j}$ is given by:

$$
\begin{aligned}
d(v, G_{\Sigma_j,c_j}) = \min\{k : \exists A \in \mathcal{A}^k_{\Sigma_j,c_j}\}\\
\text{s.t. } g(\langle \Sigma_j, v \rangle, G_{\Sigma_j,c_j} \ominus A) = \emptyset.
\end{aligned}
$$

Definition 2.2: The cyber-physical security level of a control graph G_{Σ_j,c_j} is:

$$d(G_{\Sigma_j,c_j}) = \min\{d(v, \langle \Sigma_j, c_j \rangle) : \forall v \in V^0_{\Sigma_j,c_j}\}.$$

Definition 2.3: The cyber-physical security level of a system graph G is:

$$d(G) = \min(\{d(G_{\Sigma_j,c_j}) : \forall c_j \in \mathcal{C} \ \wedge \ \forall \Sigma_j \in 2^{\mathcal{S}}\}).$$

Given Definition 2.3, the problem is formulated as follows: design an algorithm which returns a graph G' such that $d(G') > d(G)$.

III. Solution Approach and Example

In this section, we demonstrate through an illustrative example how we can increase the cyber-physical security level of a single control graph. Consider a typical water network depicted in Fig. 1. The network is comprised of a single water source (a borehole) and a controllable pump which at some point adds pressure so that the water is transported to a tank over some long distance. The water stored in the tank is then distributed to consumers though the distribution network, depending on periodic demand requests. This typical setup can be found in most water systems of any size, throughout the world. In this case the finite-state machine control logic can

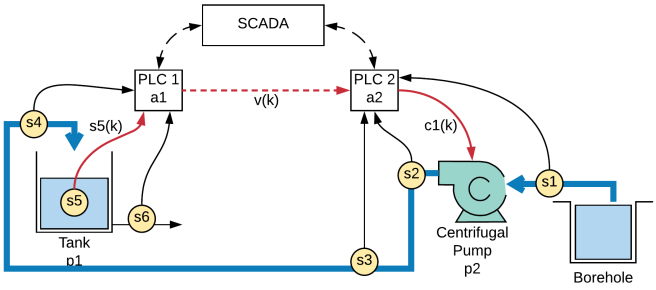

Fig. 1. A typical water transport network composed of a single source connected to a controllable pump, feeding a tank which in turn provides water to consumers. The yellow nodes are sensors of various types. The thick blue line corresponds to the pipe. The dashed line corresponds to remote communication between the components, and the red line shows the flow on information from s_5 to PLC 2 for controlling the pump.

be as follows: (1) When the level of water in the tank is *below* some threshold, *turn on* the pump, (2) When the level of water in the tank is *above* some threshold, *turn off* the pump, (3) Otherwise, do nothing.

This is achieved with 2 PLCs, one connected to the pump, and one to the tank, as well as two sensors, one measuring the water level in the tank and another measuring the pump outflow. In specific, the sensing node s_5 provides the water-level state measurement $s_5(k)$ to the agent in 'PLC 1', a_1. There, the control logic is executed, and the result $v(k)$ is transmitted to 'PLC 2', where another control logic is executed, a_2. This control logic instructs the contactor (i.e., an electrically operated relay) through a signal $c_1(k)$ to turn on (or off) the pump, if the pump flow s_3 is below a threshold (i.e., it is not working). The two PLCs also communicate with the SCADA, using the agents a_4 and a_5.

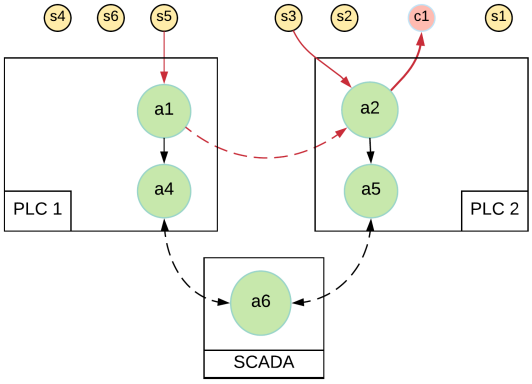

Fig. 2. Logic graph G of the CPS that appears in Fig. 1. The red lines indicate the *core control graph* of the system. The green elements represent existing agents, while yellow nodes are the sensors and red nodes the actuators.

The logic graph of such a system is depicted in Fig. 2. Notice that although the sensors may be physically connected to a PLC, if there is no component that utilizes their measurements, then they are not connected in the logic graph. From the logic graph we can extract the control graphs for c_1. Sensors s_5 and s_3 are used to compute the actuating signal c_1. The control set $\Sigma_1 = \{s_5, s_3\}$ and the graph set $\mathcal{G}_{\Sigma_1,c_1}$ contains one control graph, the core control graph. The core control graph can be expressed as $G^0_{\Sigma_1,c_1} = (\{s_5, a_1, a_2, s_3, c_1\}, \{(s_5, a_1), (a_1, a_2), (s_3, a_2), (a_2, c_1)\})$ (i.e., the red-colored path in Fig. 2).

Let's assume that a malicious party would like to affect the operation of this system. This can be achieved using various attack vectors:

1) Unplug the water level sensor or replace with fake (affects s_5).
2) Replace the control logic in the tank's PLC with a new control logic (affects a_1).
3) Replace the control logic in the pumps's PLC with a new actuation logic (affects a_2).
4) Unplug/destroy the pump actuator (affects c_1).

Each of these single attacks can affect the operation of the system, which is confirmed by the fact that the cyber-physical security level of the graph is $d(G^0_{\Sigma_1,c_1}) = 1$.

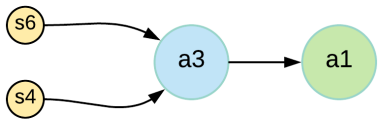

Fig. 3. Creation of agent a_3 and connection to a_1. New agents are represented in blue.

Consider the case of an attack on s_5, which interrupts the transmission of the water level to the 'PLC 1'. Since there are available sensors measuring the inflow/outflow of the tanks (s_4 and s_6 respectively), and since we know from the application domain that there exists a function which can estimate water level based on inflow/outflow sensor measurements, a new agent a_3 can be added to 'PLC 1' linked to the sensors and the a_1, as in Fig. 3. These new relations essentially correspond to a new control graph, $G_{\Sigma'_1,a_1} = (\{s_4, s_6, a_3, a_1\}, \{(s_4, a_3), (s_6, a_3), (a_3, a_1)\})$ (see Fig. 3), for $\Sigma'_1 = \{s_4, s_6\}$. The core control graph $G^0_{s_5,c_1}$ and $G_{\Sigma'_1,a_1}$ can be combined, yielding the control graph $G_{\Sigma_1,c_1} = G^0_{s_5,c_1} \oplus G_{\Sigma'_1,a_1}$, for $\Sigma_1 = \{s_4, s_5, s_6\}$ if they share a common vertex. Specifically, for G_i and G_j, if $\exists v$ s.t. $v \in V_i$ and $v \in V_j$, then the two graphs are composed as follows:

$$G_{k_3} = G_{k_1} \oplus G_{k_2} = (V_{k_1} \cup V_{k_2}, E_{k_1} \cup E_{k_2})$$

Therefore, the new graph is $G_{\Sigma_1,c_1} = (V_{\Sigma_1,c_1}, E_{\Sigma_1,c_1})$, with

$$\begin{aligned} V_{\Sigma_1,c_1} &= \{s_5, s_4, s_6, a_3, a_1, a_2, c_1\} \\ E_{\Sigma_1,c_1} &= \{(s_5, a_1), (s_4, a_3), (s_6, a_3), (a_3, a_1)(a_1, a_2), (a_2, c_1)\} \end{aligned}$$

The new security graph can be seen in Fig. 4. As a result, the cyber-physical security level of a_1 is now $d(a_1, \langle \Sigma_1, c_1 \rangle, G_{\Sigma_1,c_1}) = 2$. Notice that the degree for the

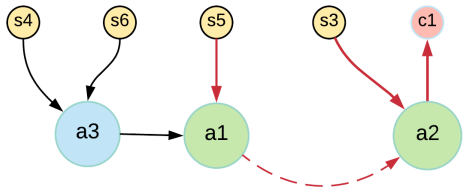

Fig. 4. Graph yielded from the composition of $G^0_{s_5,c_1}$ with $G_{\Sigma'_1,a_1}$

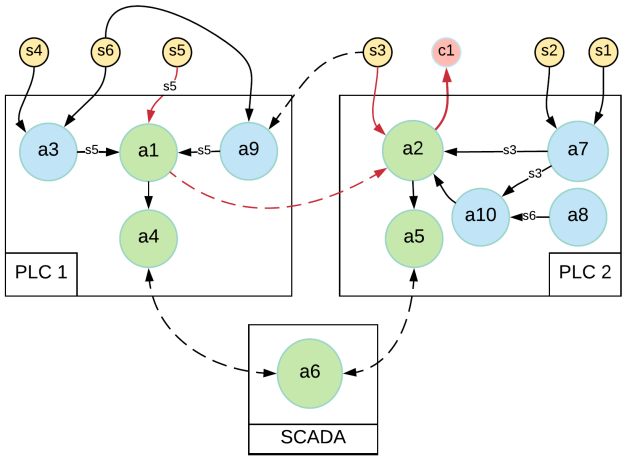

Fig. 5. The control graph of G_{S,c_1}, generated using redundant sensor and software agents. The cyber-physical security level of the graph is $d(G_{S,c_1}) = 2$. As G_{S,c_1} is the complete system we consider then $d(G) = d(G_{S,c_1})$.

element a_2 remained 1 as if either a_1 or s_3 is compromised then a_2 cannot be computed. Hence the degree of the graph remains $d(G_{\Sigma_1,c_1}) = 1$. Figure 5 depicts the final graph with five additional software components and new communication channels. In specific, in 'PLC 1', a_1 (the water tank control law) has cyber-physical security level 2 since it requires two attacked components in order to fail. Overall the new software modules, a_3 and a_9, communicate the estimation of the water tank level. Algorithm a_3 corresponds to the level estimation considering the sensor s_4 and s_6, whereas algorithm a_9 to the sensor s_6 and s_3, which measures the outflow of the pipe into the tank; this connection can be through the SCADA. Moreover in 'PLC 2', the new algorithm a_7 corresponds to an estimator for computing the pump flow using the pressure sensors s_2 and s_1. Tank outflow sensor s_6 can be estimated offline using statistical analysis; therefore, a_8 describes the periodic function which is solved for estimating tank outflow. Finally, a_{10} is essentially a virtualization of a_1, for computing the control signal in case a_1 has been compromized. Overall, the new control graph has a cyber-physical security level 2.

In practice, for computing the graphs, we propose the use of algorithms which translate water distribution models (e.g., EPANET models) into logic graphs. Moreover, domain knowledge and automated reasoning, based on the SEMIoTICS architecture [8], can be used to determine what new sensors and software agents are needed. For this, all the network components (sensors, actuators, agents) must be semantically annotated, and it assumes that a domain knowledge model exists which semantically describes the functions.

IV. Conclusions & Future Work

In this work we introduced a new methodology for reducing the vulnerability to cyber-physical attacks of water distribution, through the addition of new software components, which take advantage of analytical redundancies. In practice, given a graph which describes the cyber and physical components of the system (e.g., modelled in EPANET), a knowledge model and semantic reasoning can be used to infer new connects which will increase the number of computed states in the system. This will assist in reducing the potential impact of an attack at a specific component of the system, as this may be replaced by an estimated value computed on the same or or a different algorithm. The principle has been demonstrated in a proof-of-concept use-case of a simple network. Future work will demonstrate a complete algorithm for computing the augmented graph using semantically-enhanced reasoning.

Acknowledgment

This research is partially funded by the European Union Horizon 2020 programme under grant agreement no. 739551 (KIOS CoE), and by the Interreg V-A Greece-Cyprus 2014-2020 programme, co-financed by the European Union (ERDF) and National Funds of Greece and Cyprus, under the project SmartWater2020.

References

[1] C. Hankin, *Game Theory and Industrial Control Systems*. Springer International Publishing, 2016, pp. 178–190.

[2] K. Wilhoit, "The SCADA That Didn't Cry Wolf," Trend Micro Incorporated, Tech. Rep., 2013.

[3] R. Taormina, S. Galelli, N. O. Tippenhauer, E. Salomons, and A. Ostfeld, "Characterizing Cyber-Physical Attacks on Water Distribution Systems," *ASCE Journal of Water Resources Planning and Management*, vol. 143, no. 5, p. 12, May 2017.

[4] R. Taormina, S. Galelli, N. O. Tippenhauer, E. Salomons, A. Ostfeld, D. G. Eliades, M. Aghashahi, R. Sundararajan, M. Pourahmadi, M. K. Banks, B. M. Brentan, E. Campbell, G. Lima, D. Manzi, D. Ayala-Cabrera, M. Herrera, I. Montalvo, J. Izquierdo, E. Luvizotto, S. E. Chandy, A. Rasekh, Z. A. Barker, B. Campbell, M. E. Shafiee, M. Giacomoni, N. Gatsis, A. Taha, A. A. Abokifa, K. Haddad, C. S. Lo, P. Biswas, M. F. K. Pasha, B. Kc, S. L. Somasundaram, M. Housh, and Z. Ohar, "The battle of the attack detection algorithms: Disclosing cyber attacks on water distribution networks," *ASCE Journal of Water Resources Planning and Management*, Mar. 2018, (in print).

[5] M. Housh and Z. Ohar, "Model based approach for cyber-physical attacks detection in water distribution systems," in *Proc. of World Environmental and Water Resources Congress*, 2017, p. 10.

[6] A. A. Abokifa, K. Haddad, C. S. Lo, and P. Biswas, "Detection of cyber physical attacks on water distribution systems via principal component analysis and artificial neural networks," in *Proc. of World Environmental and Water Resources Congress*, 2017, p. 15.

[7] M. Giacomoni, N. Gatsis, and A. Taha, "Identification of cyber attacks on water distribution systems by unveiling low-dimensionality in the sensory data," in *Proc. of World Environmental and Water Resources Congress*, 2017, p. 15.

[8] G. M. Milis, C. G. Panayiotou, and M. M. Polycarpou, "SEMIoTICS: Semantically-enhanced iot-enabled intelligent control systems," *IEEE Internet of Things Journal*, p. 10, 2017, (in print).