J. Korean Soc. Hazard Mitig.
Vol. 20, No. 5 (Oct. 2020), pp.47~56
https://doi.org/10.9798/KOSHAM.2020.20.5.47

ISSN 1738-2424(Print)
ISSN 2287-6723(Online)
www.kosham.or.kr

도시방재

수질-수리 인자를 고려한 상수도 관망 사이버 공격 탐지 모델 개발

Development of a Cyberattack Detection Model for a Water Distribution System using Water Quality and Hydraulic Criteria

민경원* · 최영환 · 김중훈*****

Min, Kyoung Won*, Choi, Young Hwan, and Kim, Joong Hoon*****

Abstract

In recent years, Cyber-Physical Systems (CPSs) have been applied to Water Distribution Systems (WDSs) to facilitate efficient operation and maintenance. Since data are transmitted through the network in such systems, a cyberattack can disrupt the operation of WDSs, for example, by causing water supply reduction, water pollution, and economic losses. In the past few years, cyberattack detection algorithms and various cyberattack scenarios have been proposed. These studies considered either hydraulic factors, such as pipe velocity, nodal pressure, or tank level, or water quality factors. However, an algorithm which considers only one factor cannot prevent the various problems that may arise, such as water quality issues, and the hydraulic and quality factors have a correlation. Therefore, in this study, a framework was developed by considering both hydraulic and water quality factors. The proposed approach was applied to an artificial neural network model. Performance indicators were used to examine the detection performance according to the parameters of the artificial neural network. By comparing the detection performance when only hydraulic factors were considered and the performance when both hydraulic and water quality factors were considered, the effectiveness of the algorithm that consider both hydraulic and water quality factors was demonstrated. A cyberattack detection algorithm that considers both hydraulic and water quality criteria can be applicable in more realistic scenarios and contribute to the establishment of safe infrastructure for the entire process of designing and operating WDSs with CPSs.

Key words : Cyber Physical System, Cyberattack, Water Quality, Detection Algorithm, Artificial Neural Network

요 지

상수도 관망 사이버 물리 시스템은 효율적인 운영관리를 위해 기존의 물 공급 인프라에 적용되고 있다. 상수도 관망 사이버 물리 시스템은 네트워크를 통해 데이터들이 전송되기 때문에 사이버 공격을 받을 시, 물 공급이 원활하게 이루어지지 않거나, 펌프와 밸브의 오작동, 탱크의 월류 등의 문제가 발생한다. 이런 문제와 관련하여 사이버 공격 시나리오, 사이버 공격을 탐지하기 위한 통계적 기법, 머신러닝 알고리즘 등을 적용한 연구가 많이 수행되었으나 이 연구들은 상수도 관망의 수리학적 요인 (절점의 압력, 관의 유량, 탱크 수위 등) 또는 수질학적 요인 중에 하나만을 고려하였다. 하나의 요인만 고려한 알고리즘들은 관련된 다양한 비정상 상황에 대해 탐지할 수 없다. 따라서 본 연구에서는 효과적으로 상수도 관망의 사이버 공격을 탐지하기 위해 수리학적 요인뿐만 아니라 수질학적 요인을 모두 고려한 모델을 제안한다. 제안된 접근법의 탐지 알고리즘으로는 인공신경망을 사용했으며, 인공신경망의 매개변수에 따른 검출 성능을 정량적으로 평가하기 위해 성능지표를 적용하였다. 성능지표를 통해서 수리학적인 요인만을 고려했을 경우와 수질학적 요인도 함께 고려했을 경우의 탐지 성능을 비교하여 수질학적 요인도 함께 고려한 알고리즘의 우수성을 입증했다. 수리 및 수질학적 요인을 고려한 사이버 공격 탐지 알고리즘을 개발함으로써 보다 현실적인 시나리오를 구성할 수 있고 상수도 관망 사이버 물리 시스템 설계 및 운영의 전 과정에 있어서 안전한 인프라 구축에 기여할 수 있다.

핵심용어 : 사이버 물리 시스템, 사이버 공격, 수질, 탐지 알고리즘, 인공신경망

*정회원, 고려대학교 건축사회환경공학과 석박통합과정(E-mail: isbella7977@gmail.com)
Member, Ph.D. Candidate, Department of Civil, Environmental and Architectural Engineering, Korea University
**정회원, 국립경남과학기술대학교 토목공학과 조교수(E-mail: yh.choi@gntech.ac.kr)
Member, Assistant Professor, Department of Civil Engineering, Gyeongnam National University of Science and Technology
***교신저자, 정회원, 고려대학교 건축사회환경공학부 교수(Tel: +82-2-3290-3316, Fax: +82-2-3290-4722, E-mail: jaykim@korea.ac.kr)
Corresponding Author, Member, Professor, Department of Civil, Environmental and Architectural Engineering, Korea University

1. 서 론

4차 산업혁명 시대에 들어서면서 사이버 물리 시스템은 상수도 관망 시스템을 포함한 다양한 사회기반시설(전력시스템, 통신시스템 등)에 적용되고 있다. 이에 따라 상수도 관망의 실시간 감독이 가능하여 효율적인 운영 및 관리를 할 수 있다. 사이버 물리 시스템은 사이버 네트워크와 물리적 장치의 조합으로 정의할 수 있다. 사이버 물리 시스템은 Programmable Logic Controllers (PLCs)와 Supervisory Control And Data Acquisition (SCADA) 시스템의 두 가지의 중요한 구성요소로 이루어져 있다. 상수도 관망에 적용될 경우 PLCs은 각종 센서, 밸브, 액추에이터들을 자동 제어 및 감시에 사용하는 제어 장치이고 SCADA 시스템은 PLCs를 포함하여 상수도 관망 운영을 전체적으로 감독하고 실시간으로 데이터를 저장 및 분석하는 컴퓨터 시스템이다. 사이버 물리 시스템은 PLCs과 SCADA 시스템을 기반으로 실시간으로 감독하고 제어할 수 있으므로 시스템의 운영 및 관리 측면에서는 효율적이지만 정보통신의 사소한 데이터의 입·출력 오류 및 오차에도 심각한 피해를 일으키는 사이버 공격에 취약하다. 사이버 공격이란 컴퓨팅 시스템, 네트워크 자체, 또는 네트워크를 통해 전송되거나 저장되는 정보를 의도적으로 변경, 방해 또는 파괴하는 행동이다. 일반적으로 사이버 공격은 물리적 장치 및 통신 시설(PLCs, SCADA, 센서 및 액추에이터)에 직·간접적인 피해를 통해 운영 및 관리와 관련된 잘못된 정보를 제공한다.

2013년 미국 뉴욕 인근의 댐 관리 시스템의 사이버 공격 사건이 있었으며 2019년 2월에는 콜로라도 북부의 정수시설에 대해 랜섬웨어를 통한 사이버 공격이 발생하는 등 최근 사이버 물리 시스템이 적용된 사회기반 시설들의 사이버 공격 사례가 증가하고 있어서 심각한 사회문제를 초래하고 있다.

상수도 관망 역시 대표적인 사회 기반 시설로서 사이버 공격으로 인한 PLCs과 SCADA의 오작동으로 인해 비정상적인 물 공급, 수질 악화 등의 결과를 통해 인적·사회적·경제적 문제를 일으킨다. 따라서, 상수도 관망에 발생할 수 있는 사이버 공격을 신속하게 탐지하고 대응할 수 있는 기법들에 관한 연구는 필수적이다. 최근 상수도 관망 분야에서 사이버 공격의 탐지 및 대응을 위한 다양한 연구들이 수행되고 있다.

Witkowski (2017)는 사물인터넷, 빅데이터 등이 산업에 적용되는 방법과 4차 산업혁명 시대의 성공을 위한 핵심 조건이 사이버 물리 시스템이라고 분석했다. Cho et al. (2012)은 미래 상하수도 시설은 IT 등의 첨단기술을 접목한 융합형 실시간 모니터링, 관망 최적화 시스템 구축 등 네트워크 기반의 지능형 시스템으로써 기존의 운영관리 기술보다 더 높은 전문적인 운영기술 및 스마트 시스템을 능숙하게 운전할 수 있는 인력 양성의 중요성을 강조한다. Kim et al. (2014)은 지능형 상수도 관망의 스마트 미터 및 다항목 수질측정 장비를 이용하여 효율적으로 상수도 관망을 운영 관리 할 수 있는 기법을 제안했으며 Baik et al. (2005)은 테러, 재난 등과 같은 상황에서 용수공급시설의 안정성을 확보를 위해 국내 실정에 적합한 통합시스템의 구조를 개발했다.

Rasekh et al. (2016)은 스마트 워터 네트워크 보안의 중요성과 함께 보안 메커니즘 구축, 보안 분석 등의 중요성에 관해서 설명한다. Hassanzadeh et al. (2020)은 수자원 인프라에서 사이버 공격을 받은 여러 사례에 관해서 설명한다. Ntuli and Abu-Mahfouz (2016)는 기존 보안 솔루션과 설계 패턴을 활용하여 스마트 물 관리 시스템을 위한 보안 아키텍처를 제안했다.

Lin et al. (2009)은 사이버 물리 시스템이 적용된 상수도 관망을 EPANET과 MATLAB을 연동하여 모의하는 시스템을 개발했다. Taormina et al. (2017)은 사이버 물리 시스템이 적용된 상수도 관망이 사이버 공격을 받았을 때 탱크의 수위 변화나 밸브의 오작동을 중심으로 시나리오를 개발하고 해당 시나리오를 모의할 수 있는 epanetCPA를 개발하였다. 또한, Mishra et al. (2019)도 사회 인프라에 대한 사이버 공격의 프레임워크를 제안하여 현재 운영 중인 상수도 관망에 적용하여 검증했다. Taormina and Galelli (2018)는 딥러닝 기법을 사용하여 사이버 물리적 공격을 탐지하고 공격받은 요소를 식별하는 기법을 개발했으며 Abokifa et al. (2017)은 통계기법, 인공신경망과 주성분 분석을 통해 사이버 공격을 탐지하는 기법을 개발했다. Housh and Ohar (2018)는 Model-based 접근을 통해 사이버 공격을 탐지하는 기법을 제안했으며 Perelman et al. (2012)은 상수도 관망의 수리학적인 인자들은 고려하지 않고 수질학적인 인자들만을 고려하여 인공신경망을 활용한 이상 탐지기법을 개발했다. Panchal et al. (2011)은 인공신경망의 성능은 은닉층의 개수, 한 은닉층에서의 뉴런의 개수 등에 의해서 달라지며 Input과 Output의 데이터 수에 따라 은닉층의 개수, 한 은닉층에서의 뉴런의 개수를 결정하는 방법을 소개했다. 그러나 이러한 연구들은 수리 및 수질학적인 요소 중 하나만 고려하여 사이버 공격 탐지기법을 개발하였다. 상수도 관망의 역할은 수요자가 필요로 하는 양의 물을 수압 및 수질 기준을 만족시키는 동시에 안정적으로 공급하는 것이기 때문에 수리학적인 요소뿐만 아니라 수질학적 요소 또한 동등한 중요도로 고려되어야 한다.

따라서, 본 연구에서는 상수도 관망의 수리·수질 요소를 동시에 고려하여 상수도 관망 사이버 물리 시스템의 사이버 공격 시나리오를 구성하고, 각 시나리오는 수리·수질해석을 통해 얻은 절점의 압력, 탱크의 수위 등의 수리학적 요소와 수질 요소인 Water Age를 고려하였다. 인공신경망 기법을 적용하여 사이버 공격 탐지 알고리즘을 개발하였으며, 최적의 탐지 성능을 위해 인공신경망의 매개변수인 은닉층 수와

뉴런 수에 대한 민감도 분석을 수행하였다. 개발된 상수도 관망 사이버 공격 탐지 알고리즘의 정량적인 성능 비교를 위해 7가지 성능지표를 적용하였다.

2. 사이버 공격 시나리오 개발

본 연구에서는 상수도 관망 사이버 물리 시스템을 구축하고 사이버 공격이 가능한 네트워크 계층 및 프레임워크를 구성하여 수리학적 요인과 수질 요인을 동시에 고려한 사이버 공격 시나리오를 개발했다.

Fig. 1과 같이 사이버 물리 시스템이 상수도 관망에 적용되었을 때, 상수도 관망의 시스템 요소(펌프, 밸브, 탱크)의 감독 및 제어를 각 PLC에서 수행하게 된다. 이후 SCADA는 PLCs의 데이터들을 전송받아 데이터들의 저장 및 분석을 통해 PLCs을 제어하여 상수도 관망의 전체적인 운영을 감독한다.

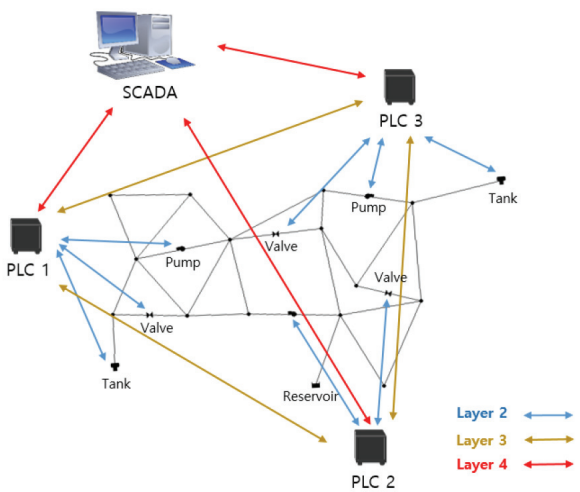

Fig. 1. Schematic Diagram of Water Distribution System with Cyber Physical System Applied

정보 보안 관점에서 상수도 관망에서의 사이버 공격은 네트워크 계층에 따라 시나리오를 개발할 수 있다. Zimmermann (1980)이 개발한 컴퓨터 네트워크의 계층 OSI 모형을 기반으로 상수도 관망 SCADA 시스템에서의 사이버 공격이 가능한 네트워크 계층은 Fig. 2와 같이 총 5개로 구성된다.

Layer 1은 펌프, 밸브, 액추에이터 등을 직접 공격하는 시나리오, Layer 2는 펌프, 밸브, 액추에이터들과 PLC 간의 통신 네트워크를 공격하는 시나리오, Layer 3은 PLC 장치를 직접 공격하거나 PLC 간의 통신 네트워크를 공격하는 시나리오, Layer 4는 PLC와 SCADA 시스템의 통신하는 네트워크를 공격하는 시나리오, Layer 5는 SCADA 시스템을 직접 공격하는 시나리오이다. 사이버 공격의 시나리오는 네트워크 계층으로 구분한 후 수리학적인 요소와 수질 요소 두 가지로 나누어 개발된다.

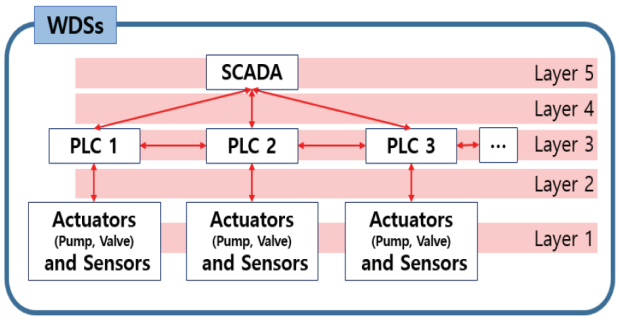

Fig. 2. Classification of Cyberattack Layer in Water Distribution System with Cyber Physical System Applied

2.1 상수도 관망에서의 사이버 공격 프레임워크

사이버 공격의 형태는 Fig. 2의 네트워크 계층에 대한 공격에 따라 직접적인 공격과 간접적인 공격으로 구분된다. 직접적인 공격은 사이버 공격이 Layer 1에 해당하는 펌프, 밸브 등 액추에이터나 각종 센서에 물리적인 공격을 수행하는 경우, Layer 3에 해당하는 PLCs을 직접 물리적으로 공격하는 경우와 Layer 5의 SCADA 시스템을 직접 공격함으로써 전체 시스템의 제어 규칙을 변경할 수 있다. 간접적인 공격은 Layer 2의 PLC와 액추에이터나 센서와의 통신을 공격함으로써 잘못된 정보를 전달해 오작동의 발생을 일으키거나, Layer 3의 PLC 자체를 공격함으로써 잘못된 제어 명령을 통해 시스템 과부하 등을 발생시키는 것이다. 또한, Layer 4의 PLC와 SCADA 간의 네트워크를 공격함으로써 전체적인 시스템에 잘못된 명령이 입력될 수 있으며 잘못된 명령이 PLC에게 전달되어 액추에이터나 센서들의 오작동을 일으킬 수 있다.

상수도 관망이 사이버 공격을 받을 경우, 센서 및 액추에이터들의 오작동, PLCs과 SCADA 시스템의 오작동, 상수관도 관망 구성요소들의 오작동 및 파괴 등으로 이어진다. 이러한 결과를 통해 탱크의 월류, 관로의 누수 및 파괴, 염소 투입 및 농도 통제 실패 등이 발생하여 수리학 및 수질학적인 요소들의 비정상적인 상황이 발생한다(Fig. 3).

2.2 수리/수질 인자를 고려한 사이버 공격 시나리오

본 연구에서는 수리학적 요인과 수질학적인 요인을 동시에 고려한 사이버 공격 시나리오를 개발하였다. 개발된 시나리오는 Fig. 2의 네트워크 계층에 대한 공격 형태에 따라 6가지 형태로 구성된다.

- 시나리오 1: 펌프, 밸브, 센서 등의 물리 시스템을 직접 공격한다. 공격자는 직접 물리 시스템을 손상하거나 다른 물리 시스템으로 교체할 수 있다. 공격으로 인해 손상당한 물리 시스템을 통제하는 PLCs은 NULL 값 또는 변경된 값으로 전송받는다.

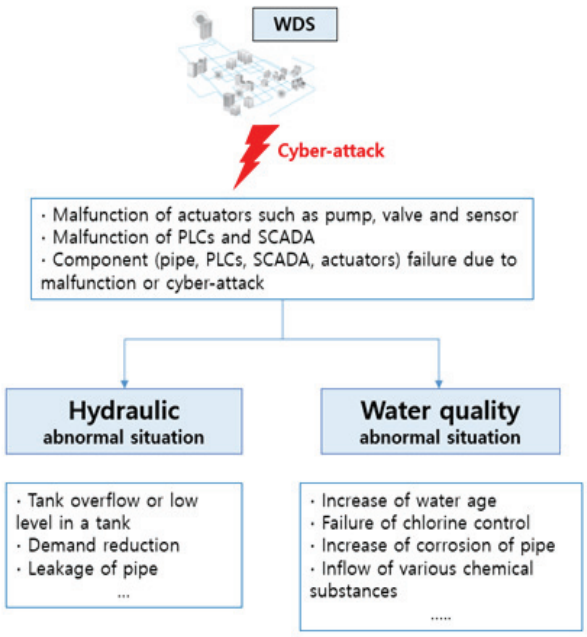

Fig. 3. Classification of Cyberattack in Water Distribution System with Cyber Physical System Applied

- 시나리오 2: PLC와 액추에이터 및 센서가 통신하는 네트워크를 공격한다. 전송되는 데이터를 조작 및 손상, 도청, 서비스 거부(Denial of service, Dos) 등의 공격을 할 수 있다.
- 시나리오 3: PLC 간의 통신 네트워크를 공격한다. 예를 들어, PLC 1은 탱크 1의 수위의 정보를 받아 통제하고 있으며, PLC 2는 PLC 1에게 탱크 1의 수의 정보를 받아 펌프 1의 운영을 통제하고 있다고 가정할 때, 이 2개의 PLC 간의 통신 네트워크 공격으로 펌프의 잘못된 운영에 따라 탱크가 넘치는 등의 상수도 관망의 운영에 차질이 생긴다.
- 시나리오 4: PLC 자체를 공격한다. PLC의 정상적인 작동을 완전히 중지시킬 수 있고, PLC 안의 제어 로직들을 변경할 수 있으며, SCADA 시스템에 보내는 데이터를 변경 및 손상 등의 공격을 할 수 있다.
- 시나리오 5: PLC와 SCADA 시스템의 통신 네트워크를 공격한다. 전송되는 데이터를 조작, 도청 또는 서비스 거부(Denial of service, Dos) 등의 공격을 할 수 있다.
- 시나리오 6: SCADA 시스템 자체를 공격한다. 상수도 관망 운영 로직 변경, 상수도 관망 전체 운영 중단, 상수도 관망 운영 도청 등의 공격을 할 수 있다.

3. 모형 개발

상수도 관망 사이버 물리 시스템의 사이버 공격 탐지 알고리즘으로 인공신경망을 사용하였으며, 2장에서 개발한 수리·수질 요인을 고려한 사이버 공격 시나리오를 대상 관망에 적용하였다. 대상 관망을 통해서 인공신경망의 훈련 및 Test 데이터를 생성했다. 훈련 데이터를 통해 인공신경망을 훈련 시킨 후 Test 데이터를 통한 인공신경망의 결과를 성능지표를 통해 성능을 평가했다. 구체적인 방법은 다음과 같다.

3.1 인공신경망

인공신경망(McCulloch and Pitts, 1943)은 인간 두뇌에 관한 최초의 논리적 모델링 기법으로 신경망에서 영감을 받아 개발된 알고리즘이다. 뇌 신경의 시냅스 결합을 통해 네트워크를 형성한 인공 뉴런이 학습을 통해 시냅스의 결합 세기를 변화시켜 문제 해결 능력을 갖추는 모델을 말한다. 이러한 인공신경망의 성능은 은닉층의 수(Number of Hidden Layer, N_{HL}), 뉴런의 수(Number of Neuron, N_N), 뉴런의 가중치에 따라 성능이 달라진다.

본 연구에서 정상 상태와 사이버 공격을 받은 비정상 상태의 데이터를 생성하여 사이버 공격 탐지 알고리즘에 적용하고 검증하였다. 사이버 공격 탐지 알고리즘은 은닉층의 수와 뉴런의 수의 민감도 분석을 통해 최적의 성능을 갖는 인공신경망 매개변수를 결정한다.

3.2 사이버 공격 탐지 성능지표

본 연구에서의 인공신경망을 통한 탐지결과는 사이버 공격 여부에 따라 2가지의 형태로 도출된다. 도출된 탐지결과는 실제 사이버 공격(Observed Attack)과 비교하여 성능을 평가한다. 본 연구에서 사용한 성능지표(Powers, 2011)는 Eqs. (1)~(7)과 같다.

Table 1에서 True Positive (TP)는 실제 사이버 공격을 받았을 때 인공신경망이 공격 탐지를 동일하게 했을 경우의 데이터의 수를 의미하며, False Positive (FP)는 실제로 사이버 공격을 받지 않았을 때 인공신경망이 공격 탐지를 다르게 했을 경우의 데이터의 수를 의미한다. True Negative (TN)는 실제로 사이버 공격을 받지 않았을 때 인공신경망이 공격 탐지를 동일하게 했을 경우의 데이터의 수를 의미하며, False Negative (FN)는 실제로 사이버 공격을 받았을 때 인공신경망이 공격 탐지를 다르게 했을 경우의 데이터의 수를 의미한다. Eq. (1)은 실제로 사이버 공격을 받은 데이터에 대해서 인공신경망이 공격 탐지를 동일하게 했을 경우의 데이터의 비율을 의미하며 1에 가까울수록 모델 성능이 좋다. Eq. (2)는 실제로 사이버 공격을 받은 데이터에 대해서 인공신경망이 공격 탐지를 다르게 했을 경우의 데이터의 비율을 의미하며 0에 가까울수록 모델 성능이 좋다. Eq. (3)은 실제로 사이버 공격을 받지 않은 데이터에 대해서 인공신경망이 공격 탐지를 다르게 했을 경우의 데이터의 비율을 의미하며 0에 가까울수록 모델 성능이 좋다. Eq. (4)는 실제로 사이버 공격을 받지 않은 데이터에 대해서 인공신경망이 공격 탐지를 동일하게 했을 경우의 데이터의 비율을 의미하며 1에

Table 1. Performance Indicators for Classification Problems

Total Data		True Condition		Performance Indicator
		In Cyberattack	Not in Cyberattack	
Predicted Condition (Proposed in this study)	In Cyberattack	True Positive	False Positive	$\text{Precision}(P_p) = \frac{\sum True\ positive}{\sum \text{Predicted condition positive}}$ (5)
	Not in Cyberattack	False Negative	True Negative	
Performance Indicator		$Sensitivity(Recall, P_s) = \frac{\sum True\ positive}{\sum Condition\ positive}$ (1)	$Fall-out\ (P_f) = \frac{\sum False\ positive}{\sum Condition\ negative}$ (3)	$Accuracy\ (P_a) = \frac{\sum True\ positive + \sum True\ negative}{\sum Total}$ (6)
		$Miss\ rate\ (P_m) = \frac{\sum False\ negative}{\sum Condition\ positive}$ (2)	$Specifivity(Selectivity, P_{ss}) = \frac{\sum True\ negative}{\sum Condition\ negative}$ (4)	$F_1\ score\ (P_{f1}) = 2 \cdot \frac{\text{Precision} \cdot Recall}{\text{Precison} + Recall}$ (7)

가까울수록 모델 성능이 좋다. Eq. (5)는 모델의 정확도를 의미하며 직관적으로 성능을 평가하는 지표이다. 전체 데이터에 대해서 TP와 TN의 비율을 의미한다. Eq. (6)은 인공신경망이 사이버 공격을 탐지한 데이터에 대해서 실제로 사이버 공격을 받은 데이터의 비율을 의미한다. Eq. (7)은 Precision과 Sensitivity (Recall)의 조화평균을 의미하며 모델의 전체적인 성능을 평가할 수 있는 지표이다. Eqs. (1)~(7)은 모두 0에서 1 사이의 값을 가지며 Eqs. (1), (4)~(7)은 1에 가까울수록 Eqs. (2), (3)은 오경보율로서 0에 가까울수록 정확한 탐지 확률을 의미한다. Eq. (1)과 Eq. (2), Eq. (3)과 Eq. (4)의 합은 모두 1이다. Eq. (1)의 값이 1에 가까울수록 Eq. (2)의 값은 0에 가까워진다. 마찬가지로 Eq. (4)의 값이 1에 가까울수록 Eq. (3)의 값은 0에 가까워진다. 즉, Eq. (1)과 Eq. (2), Eq. (3)과 Eq. (4)는 서로 높은 상관관계를 가지며 Eq. (5)는 Eq. (7)에서 고려되기 때문에 본 연구에서는 Eq. (1), Eq. (4), Eq. (6), Eq. (7)의 평균 성능(P_{av})을 통해 모델의 성능을 비교 및 평가했다.

3.3 C-town 관망

Fig. 4의 C-town 네트워크(Ostfeld et al., 2012)는 실제 네트워크를 기반으로 만들어진 상수도 관망이다. 388개의 수요 절점, 429개의 관로, 7개의 탱크, 11개의 펌프, 4개의 밸브, 9개의 PLC와 1개의 SCADA로 구성된 C-town 관망은 The BATtle of the Attack Detection Algorithms (BATADAL)에서 소개되었다(Taormina et al., 2018). PLC가 통제하는 센서와 액추에이터들은 Table 2와 같다.

3.4 데이터 생성

본 연구에서 수리·수질 인자를 고려한 사이버 공격을 탐지하기 위해 EPANET (Rossman, 2000)을 이용하여 정상 및 비정상 데이터를 생성하였다. 상수도 관망 설계 및 운영 관련 선행 연구(Tamminen et al., 2008; Shamsaei et al., 2013)에서 물의 정체(Stagnation)를 통해 상수도 수질을

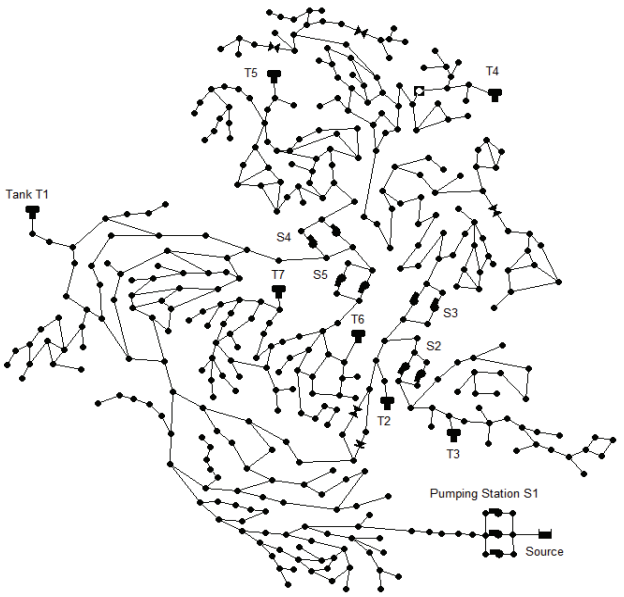

Fig. 4. C-town Water Distribution System

Table 2. Sensors and Actuators Controlled by PLC

PLC	Sensor	Actuator
PLC1	-	PU1 (T1), PU2 (T1)
PLC2	T1	-
PLC3	T2	V2 (T2), PU4 (T3), PU5 (T3), PU6 (T4), PU7 (T4)
PLC4	T3	-
PLC5	-	PU8 (T5), PU9 (-), PU10 (T7), PU11 (T7)
PLC6	T4	-
PLC7	T5	-
PLC8	T6	-
PLC9	T7	-

Table 3. Abnormal State Scenario in Cyberattack

Scenario	Normal/ Abnormal state	Control rule
Scenario 1	Normal state	LINK PU1 OPEN IF NODE T1 BELOW 4 LINK PU1 CLOSED IF NODE T1 BELOW 6.3 LINK PU2 OPEN IF NODE T1 BELOW 1 LINK PU2 CLOSED IF NODE T1 BELOW 4.5
	Abnormal state	The control rule of PU1 and PU2 is changed by attacking C-town's PLC 1
		LINK PU1 OPEN IF NODE T1 BELOW 4 LINK PU1 CLOSED IF NODE T1 BELOW 6.5 LINK PU2 OPEN IF NODE T1 BELOW 1 LINK PU2 CLOSED IF NODE T1 BELOW 6.5
Scenario 2	Normal state	LINK V1 OPEN IF NODE T2 BELOW 0.5 LINK V1 CLOSED IF NODE T2 BELOW 5.5
	Abnormal state	The control rule of V2 is changed by attacking C-town's PLC 3
		LINK V1 OPEN IF NODE T2 BELOW 6.4999 LINK V1 CLOSED IF NODE T2 BELOW 6.4999

평가하는 인자로 Water age를 고려하여 본 연구에서도 같은 수질 인자를 고려했다. EPANET 프로그램을 통해 C-town 네트워크의 탱크의 수위 및 Water Age 데이터(14개), 펌프와 밸브 2의 유량, 상태(on/off) 및 Water Age 데이터(36개), 펌프와 밸브 2에서 유량이 유입 및 유출되는 절점의 압력과 Water Age 데이터(24개)를 생성했다. 즉, 총 74개의 변수(수리학적 요인(43개), 수질학적 요인(31개))에 대해서 15분 단위로 720시간에 대해서 데이터를 생성했다. 비정상 데이터는 사이버 공격 시나리오(Table 3)를 적용하여 생성하였다. 훈련 데이터와 Test 데이터의 비율은 각각 75%, 25%의 비율로 무작위로 생성했다. Test 데이터에서 Time Step 1000 기점으로 전은 시나리오 1에 대한 비정상 상황이며 후는 시나리오 2에 대한 비정상 상황이다.

4. 적용 및 결과

본 연구에서 개발한 사이버 공격 탐지 알고리즘의 탐지결과는 Table 4, Table 5, Fig. 5와 같다. Table 4는 수리학적인 요소를 고려한 경우이며 Table 5는 수질학적인 요소도 함께 고려한 사이버 공격 탐지 알고리즘의 탐지결과 성능이다.

Table 4. Comparison Cyberattack Detected by Artificial Neural Network with Actual Cyberattack (Hydraulic Criteria)

Case	Number of Hidden Layers (N_{HL})	Number of Neurons (N_N)	TP	FP	TN	FN	P_s	P_m	P_f	P_{ss}	P_a	P_p	P_{fl}	P_{av}
1	1	4	1,147	253	409	291	0.82	0.18	0.42	0.58	0.74	0.80	0.81	0.73
2		5	1,330	70	23	677	0.95	0.05	0.97	0.03	0.64	0.66	0.78	0.60
3		7	1,226	174	466	231	0.88	0.12	0.33	0.67	0.81	0.84	0.86	0.80
4		**10**	**1,262**	**138**	**486**	**214**	**0.90**	**0.10**	**0.31**	**0.69**	**0.83**	**0.86**	**0.88**	**0.82**
5		**15**	**1,221**	**179**	**557**	**143**	**0.87**	**0.13**	**0.20**	**0.80**	**0.85**	**0.90**	**0.88**	**0.84**
6	2	[4,4]	1,251	149	397	303	0.89	0.11	0.43	0.57	0.78	0.81	0.85	0.77
7		[5,5]	1,273	127	149	551	0.91	0.09	0.79	0.21	0.68	0.70	0.79	0.64
8		[7,7]	1,251	149	486	214	0.89	0.11	0.31	0.69	0.83	0.85	0.87	0.82
9		[10,10]	1,180	220	221	479	0.84	0.16	0.68	0.32	0.67	0.71	0.77	0.64
10		[15,15]	1,172	228	524	176	0.84	0.16	0.25	0.75	0.81	0.87	0.85	0.81
11	3	[4,4,4]	1,171	229	362	338	0.84	0.16	0.48	0.52	0.73	0.78	0.81	0.72
12		[5,5,5]	1,059	341	538	162	0.76	0.24	0.23	0.77	0.76	0.87	0.81	0.77
13		[7,7,7]	1,244	156	403	297	0.89	0.11	0.42	0.58	0.78	0.81	0.85	0.77
14		**[10,10,10]**	**1,246**	**154**	**544**	**156**	**0.89**	**0.11**	**0.22**	**0.78**	**0.85**	**0.89**	**0.89**	**0.85**
15		[15,15,15]	1,190	210	530	170	0.85	0.15	0.24	0.76	0.82	0.88	0.86	0.82

Table 5. Comparison Cyberattack Detected by Artificial Neural Network with Actual Cyberattack (Hydraulic Criteria + Water Quality Criteria)

Case	Number of Hidden Layers (N_{HL})	Number of Neurons (N_N)	TP	FP	TN	FN	P_s	P_m	P_f	P_{ss}	P_a	P_p	P_{fl}	P_{av}
1	1	4	1,278	122	436	264	0.91	0.09	0.38	0.62	0.82	0.83	0.87	0.80
2		5	1,083	317	155	545	0.77	0.23	0.78	0.22	0.59	0.67	0.72	0.57
3		**7**	**1,160**	**240**	**637**	**63**	**0.83**	**0.17**	**0.09**	**0.91**	**0.86**	**0.95**	**0.88**	**0.86**
4		10	1,130	270	597	103	0.81	0.19	0.15	0.85	0.82	0.92	0.86	0.83
5		15	1,199	201	515	185	0.86	0.14	0.26	0.74	0.82	0.87	0.86	0.81
6	2	[4,4]	1,189	211	578	122	0.85	0.15	0.17	0.83	0.84	0.91	0.88	0.84
7		[5,5]	1,201	199	527	173	0.86	0.14	0.25	0.75	0.82	0.87	0.87	0.82
8		[7,7]	1,229	171	526	174	0.88	0.12	0.25	0.75	0.84	0.88	0.88	0.83
9		**[10,10]**	**1,167**	**233**	**610**	**90**	**0.83**	**0.17**	**0.13**	**0.87**	**0.85**	**0.93**	**0.88**	**0.85**
10		[15,15]	1,076	324	530	170	0.77	0.23	0.24	0.76	0.76	0.86	0.81	0.77
11	3	[4,4,4]	1,110	290	566	134	0.79	0.21	0.19	0.81	0.80	0.89	0.84	0.80
12		[5,5,5]	1,125	275	532	168	0.80	0.20	0.24	0.76	0.79	0.87	0.84	0.79
13		[7,7,7]	1,266	134	446	254	0.90	0.10	0.36	0.64	0.82	0.83	0.87	0.80
14		**[10,10,10]**	**1,190**	**210**	**604**	**96**	**0.85**	**0.15**	**0.14**	**0.86**	**0.85**	**0.93**	**0.89**	**0.86**
15		[15,15,15]	1,114	286	481	219	0.80	0.20	0.31	0.69	0.76	0.84	0.82	0.76

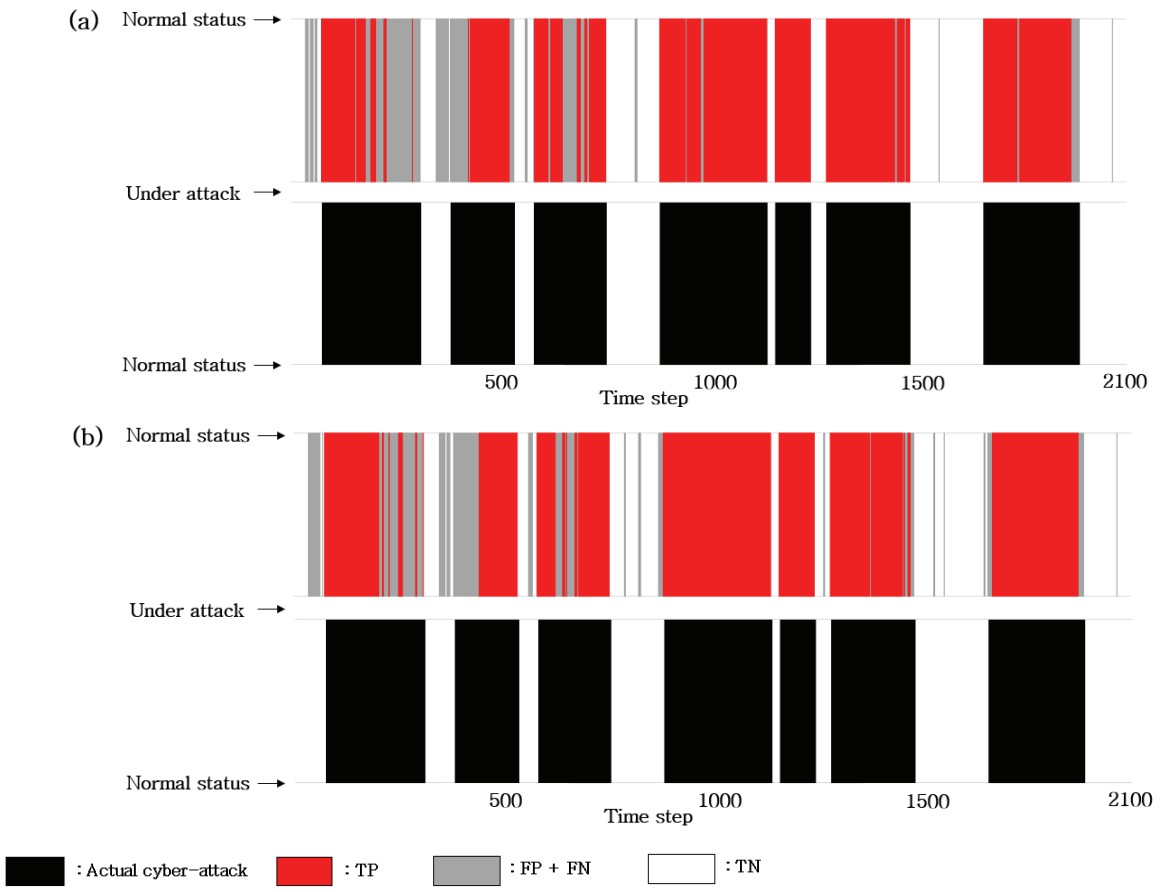

Fig. 5. Comparison Cyberattack Detected by Artificial Neural Network with Actual Cyberattack (Scenario 1 : Cyberattack before time step 1000, Scenario 2 : Cyberattack after time step 1000) (a) Case 3 (N_{HL} = 1, N_N = 7), (b) Case 14 (N_{HL} = 3, N_N = [10,10,10], (c) Case 9 (N_{HL} = 2, N_N = [10,10])

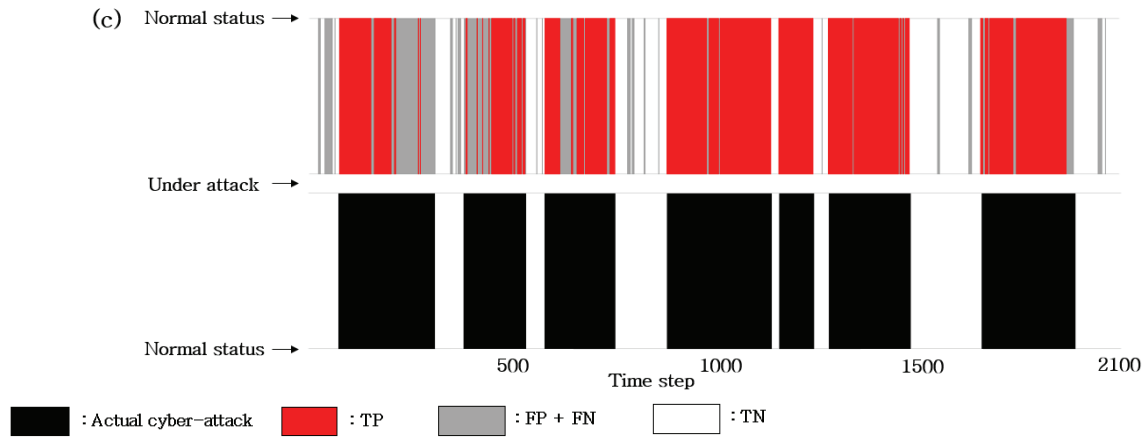

Fig. 5. (Continued)

본 연구에서 제안한 사이버 공격 탐지 알고리즘은 인공신경망을 기반으로 개발되었기 때문에 N_{HL}와 N_N에 따라 성능의 차이를 보인다. 따라서, 본 연구에서는 N_{HL}의 수와 N_N에 대해 민감도 분석을 수행하여 최적의 탐지 성능을 도출할 수 있는 매개변수를 결정한다. 또한, 성능의 정량적인 비교를 위해 3.2장의 성능지표를 사용하였으며, 그 결과는 Table 4, Table 5와 같다.

수리학적인 요소만 고려한 사이버 공격 탐지 알고리즘의 결과인 Table 4에서 P_{av}를 분석해 본 결과 N_{HL}가 3일 때 N_N가 [10,10,10]일 때가 0.85로 가장 우수했다. N_{HL}에 따라 성능의 f1_score 평균(e.g., N_{HL} = 1 일때, N_N = 4, 5, 7, 10, 15의 f1_score 평균)을 비교하였을 경우 N_{HL}가 1, 2, 3개일 때 각각 0.763, 0.740, 0.788로 N_{HL}가 3개일 때가 가장 좋은 탐지 성능을 보였다. 또한, N_{HL}에 상관없이 N_N가 4, 5, 7, 10, 15로 구성되어 있을 때 평균 성능은 각각 0.74, 0.67, 0.79, 0.77, 0.82로 N_N가 15개로 구성되어 있을 때가 가장 좋은 탐지 성능을 보였다.

수질 및 수리학적인 요소를 모두 고려한 사이버 공격 탐지 알고리즘의 결과인 Table 5에서 P_{av}를 분석해 본 결과 N_{HL}가 1일 때 N_N가 7일 때가 0.869로 가장 우수했다. N_{HL}에 따라 성능의 f1_score 평균을 비교하였을 경우 N_{HL}가 1, 2, 3개일 때 각각 0.78, 0.82, 0.80으로 N_{HL}가 2개일 때가 가장 좋은 탐지 성능을 보였다. 또한, N_{HL}에 상관없이 N_N가 4, 5, 7, 10, 15로 구성되어 있을 때 평균 성능은 각각 0.82, 0.73, 0.83, 0.85, 0.78로 N_N가 10개로 구성되어 있을 때가 가장 좋은 탐지 성능을 보였다.

N_{HL}와 N_N의 구분에 따라 평균 성능을 분석해 본 결과 수질 및 수리학적인 요소를 고려한 사이버 공격 탐지 알고리즘의 15개의 결과 중에서 11개가 수리학적인 요소만 고려했을 경우보다 탐지 결과가 우수하게 나왔다. 동일한 조건에서 수질학적인 요인도 같이 고려해야 이상 탐지하는 데 있어서 탐지성능이 더 높다는 의미이다. 수리학적인 요소만 고려했을 때 탐지 성능이 더 우수한 결과를 나타난 모형은(N_{HL} = 1, N_N = 5), (N_{HL} = 1, N_N = 15), (N_{HL} = 2, N_N = [15,15]), (N_{HL} = 3, N_N = [15,15,15]) 일 때이다.

Fig. 5에서는 Table 5에서 우수한 성능을 보인 결과 3개를 그래프로 표현하였다. Fig. 5(a)는 Case 3 (N_{HL} = 1, N_N = 7)일 때, Fig. 5(b)는 Case 14 (N_{HL} = 3, N_N = [10,10,10])일 때, Fig. 5(c)는 Case 9 (N_{HL} = 2, N_N = [10,10])일 때의 결과 그래프이다. (a), (b), (c)의 P_{av}의 값은 각각 0.869, 0.863, 0.857이다. 그래프의 X축은 Time step을 의미하며 Y축은 사이버 공격의 여부를 의미한다. 검은색 부분이 실제로 사이버 공격을 의미하고, 빨간색 부분은 실제로 사이버 공격을 받은 기간에 대해서 사이버 공격 탐지 알고리즘도 동일하게 탐지한 결과로써 TP를 의미한다. 회색 부분은 오탐지 결과로써 FP와 FN을 더한 것을 의미하며 하얀색 부분은 TN을 의미한다. Fig. 5를 통해서 Time Step에 따른 성능 확인이 가능하며 Time Step 1000 기점으로 전에는 오탐지가 많고 후에는 적은 것을 확인할 수 있다. Time Step 1000 기점으로 전의 비정상 상황은 시나리오 1에 대한 사이버 공격이며 후의 비정상 상황은 시나리오 2에 대한 사이버 공격이다. 시나리오 1일 경우 PU1과 PU2의 가동을 중지하는 기준은 T1의 수위가 6.3, 4.5에서 모두 6.5 미만으로 수위 기준에 미세한 변경이 있었다. 결국 수리 및 수질학적인 데이터들에서도 미세한 차이가 나기 때문에 정상 상태와 비정상 상태의 구분이 불분명하여 오탐지가 많이 나온 것으로 보인다. 시나리오 2일 경우 V1의 열고 닫는 T2의 수위 기준이 0.5, 5.5에서 각각 6.4999, 6.4999로 큰 변동이 있었기 때문에 수리 및 수질학적인 데이터들에서도 큰 차이가 있어 정상 상태, 비정상 상태의 구분이 분명하여 오탐지가 적은 것으로 보인다.

5. 결 론

제4차 산업 혁명 시대에 들어서면서 사이버 물리 시스템이 중요한 국가 기반 시설인 상수도 관망에 적용되고 있다. 사이버 물리 시스템은 정보통신기술 기반으로 통신이 되기

때문에 사이버 공격의 우려가 된다. 사이버 물리 시스템이 적용된 상수도 관망이 사이버 공격을 받으면 공급량 부족, 관 파단, 수질 악화 등 비정상적인 상황을 초래한다. 따라서 상수도 관망의 보안 문제에 대해서도 대비할 필요가 있다. 사이버 보안과 관련된 이전의 연구들은 상수도 관망의 사이버 공격 시나리오를 구성하고 수리학적 기준만을 고려하여 사이버 공격 탐지 알고리즘을 개발했다. 그러나 상수도 관망의 목적은 언제 어디서나 양질의 물을 충분히 공급하는 것이다. 따라서 본 논문에서는 수질 기준을 고려하여 시나리오를 확장함으로써 사이버 공격 탐지 알고리즘의 효율성과 정확도를 향상했다. 그러나 본 연구에서는 수질 요인으로 Water Age만을 고려하였기 때문에 향후 연구로는 탁도, 잔류염소, pH 등 시나리오를 확장하여 데이터를 구성하고, 인공신경망뿐만 아니라 다양한 딥러닝 기법을 사용하여 비교 분석할 계획이다. 딥러닝 기법에 사용되는 Input Data로써 실제와 유사한 가상의 데이터를 구축할 때 EPANET 2.0에서 다양한 Demand Pattern을 고려하여 데이터를 얻어 수리 및 수질 데이터들의 상관성 분석을 통해 중요한 인자만을 도출하여 사용할 예정이며, 탐지뿐만 아니라 상수도 관망 어떤 파이프 부분에서 이상이 발생했는지 추적하는 기법도 연구할 계획이다. 이러한 연구를 통하여 사이버 물리 시스템이 적용된 상수도 관망을 설계하고 운영하는 전 과정에서 안전한 물 공급 인프라를 구축에 기여할 것으로 기대된다.

감사의 글

본 연구는 환경부의 “글로벌탑 환경기술개발사업(2016002120004)”으로 지원받은 과제입니다.

References

Abokifa, A.A., Haddad, K., Lo, C.S., and Biswas, P. (2017). Detection of cyber physical attacks on water distribution systems via principal component analysis and artificial neural networks. *Proceedings of World Environmental and Water Resources Congress 2017,* pp. 676-691.

Baik, C.W., Kim, E.S., Jun, H.D., and Kim, J.H. (2005). Development of SCADA system for security of water network. *Proceedings of 2005 Conference*, Korean Society of Civil Engineers, pp. 392-395.

Cho, E.S. et al. (2012). *Waterworks and sanitation strategy to promote future growth and development.* Green Growth Research Report No. 2012-09, Korea Environment Institute.

Hassanzadeh, A., Rasekh, A., Galelli, S., Aghashahi, M., Taormina, R., Ostfeld, A., et al. (2020). A review of cybersecurity incidents in the water sector. *Journal of Environmental Engineering*, Vol. 146, No. 5, 03120003. doi:10.1061/(ASCE)EE.1943-7870.0001686

Housh, M., and Ohar, Z. (2018). Model-based approach for cyber-physical attack detection in water distribution systems. *Water Research*, Vol. 139, pp. 132-143.

Kim, D.H., Choe, D.Y., and Kim, J.H. (2014). Optimization of water distribution system operation management based on smart meter and sensor network. *Water for Future*, Vol. 47, No. 5, pp. 22-27.

Lin, J., Sedigh, S., and Miller, A. (2009). Towards integrated simulation of cyber-physical systems: A case study on intelligent water distribution. *Proceedings of 2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing,* IEEE, pp. 690-695.

McCulloch, W.S., and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, Vol. 5, No. 4, pp. 115-133.

Mishra, V.K., Palleti, V.R., and Mathur, A. (2019). A modeling framework for critical infrastructure and its application in detecting cyber-attacks on a water distribution system. *International Journal of Critical Infrastructure Protection*, Vol. 26, 100298. doi:10.1016/j.ijcip.2019.05.001

Ntuli, N., and Abu-Mahfouz, A. (2016). A simple security architecture for smart water management system. *Procedia Computer Science*, Vol. 83, pp. 1164-1169.

Ostfeld, A., Salomons, E., Ormsbee, L., Uber, J.G., Bros, C.M., Kalungi, P., et al. (2012). Battle of the water calibration networks. *Journal of Water Resources Planning and Management*, Vol. 138, No. 5, pp. 523-532.

Panchal, G., Ganatra, A., Kosta, Y.P., and Panchal, D. (2011). Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. *International Journal of Computer Theory and Engineering*, Vol. 3, No. 2, pp. 332-337.

Perelman, L., Arad, J., Housh, M., and Ostfeld, A. (2012). Event detection in water distribution systems from multivariate water quality time series. *Environmental Science & Technology*, Vol. 46, No. 15, pp. 8212-8219.

Powers, D.M. (2011). Evaluation: From precision, recall

and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, Vol. 2, No. 1, pp. 37-63.

Rasekh, A., Hassanzadeh, A., Mulchandani, S., Modi, S., and Banks, M. K. (2016). Smart water networks and cyber security. *Journal of Water Resources Planning and Management*, Vol. 142, No. 7, 01816004. doi:10.1061/(ASCE)WR.1943-5452.0000646

Rossman, L.A. (2000). *EPANET 2. Users manual*. EPA/600/R-00/057, U.S. Environmental Protection Agency (EPA), Washington, D.C., USA.

Shamsaei, H., Jaafar, O., and Basri, N.E.A. (2013). Effects residence time to water quality in large water distribution systems. *Engineering*, Vol. 5 No. 4, 2013, pp. 449-457. doi:10.4236/eng.2013.54054

Tamminen, S., Ramos, H., and Covas, D. (2008). Water supply system performance for different pipe materials Part I: Water quality analysis. *Water Resources Management*, Vol. 22, No. 11, pp. 1579-1607.

Taormina, R., Galelli, S., Tippenhauer, N.O., Salomons, E., and Ostfeld, A. (2017). Characterizing cyber-physical attacks on water distribution systems. *Journal of Water Resources Planning and Management*, Vol. 143, No. 5, 04017009. doi:10.1061/(ASCE)WR.1943-5452.0000749

Taormina, R., Galelli, S., Tippenhauer, N.O., Salomons, E., Ostfeld, A., Eliades, D.G., et al. (2018). Battle of the attack detection algorithms: Disclosing cyber attacks on water distribution networks. *Journal of Water Resources Planning and Management*, Vol. 144, No. 8, 04018048. doi:10.1061/(ASCE)WR.1943-5452.0000969

Taormina, R., and Galelli, S. (2018). Deep-learning approach to the detection and localization of cyber-physical attacks on water distribution systems. *Journal of Water Resources Planning and Management*, Vol. 144, No. 10, 04018065. doi:10.1061/(ASCE)WR.1943-5452.0000983

Witkowski, K. (2017). Internet of things, big data, industry 4.0 – Innovative solutions in logistics and supply chains management. *Procedia Engineering*, Vol. 182, pp. 763-769.

Zimmermann, H. (1980). OSI reference model - The ISO model of architecture for open systems interconnection. *IEEE Transactions on Communications*, Vol. 28, No. 4, pp. 425-432.

Received	July 16, 2020
Revised	July 20, 2020
Accepted	September 3, 2020