Leak Detection in Water Distribution Networks via Pressure Analysis Using a Machine Learning Ensemble

Vivencio C. Fuentes Jr.(✉) and Jhoanna Rhodette I. Pedrasa

University of the Philippines, Diliman, 1101 Quezon City, Philippines
vivencio.fuentes@eee.upd.edu.ph, jipedrasa@up.edu.ph

Abstract. Water distribution networks (WDNs) are vital infrastructure which serve as a means for public utilities to deliver potable water to consumers. Naturally, pipelines degrade over time, causing leakages and pipe bursts. Damaged pipelines allow water to leak through, incurring significant economic losses. Mitigating these losses are important, especially in areas with water scarcity, to allow consumers to have adequate water supply. Globally, as the population increases, there is a need to make water distribution efficient, due to the rising demand. Thus, leak detection is vital for reducing the system loss of the network and improving efficiency.

Monitoring WDNs effectively for leakage is often a challenging task due to the size of the area it covers, and due to the need to detect leaks as early as possible. Traditionally, this is done via pipeline inspection or physical modeling. However, such techniques are either time-consuming, resource intensive, or both. An alternative is machine learning (ML), which maps the relationship between pipeline data to detect leakages. This allows for a faster, yet reasonably accurate model for detection and localization. Machine learning techniques could be utilized together as an ensemble, which allows these techniques to work in conjunction with each other. Wavelet decomposition will be performed on the data to allow for a smaller dataset, as well as utilizing possible hidden features for the machine learning model.

Keywords: Water distribution networks · Leak detection · Machine learning

1 Introduction

1.1 Water Distribution Networks

Water distribution networks (WDNs) are systems designed and implemented to deliver potable water from a source to a consumer. However, pipeline

This work is supported by a Commission on Higher Education (CHED) Philippine California Advanced Research Institues (PCARI) grant under Project IIID54 - Resilient Cyber Physical Societal Scale Systems.

P. Pereira et al. (Eds.): SC4Life 2019, LNICST 318, pp. 31–44, 2020.
https://doi.org/10.1007/978-3-030-45293-3_3

infrastructure gradually experience deterioration, which could be due to natural aging, environmental damage, or unauthorized human interference [1]. Damaged pipelines allow water to leak through, which would cause significant economic losses. As the world population grows, there is a need for distributing water efficiently due to scarcity in water supply. Mitigating losses is key, especially in areas with water scarcity, making leak detection an important part for improving the efficiency of a network.

1.2 Leak Detection

Without using automated sensing or analytical methods, leakages are traditionally reported via water utility personnel examining above-ground meters along the WDN [2]. Due to the nature of human involvement, this is a time-consuming method that could possibly incur significant losses before it is addressed.

As leakage events follow hydraulic principles, sensors could be used to monitor characteristic changes within the pipeline. Pressure drops and flow imbalance in certain areas could imply disruption in normal pipeline operations [1].

Data acquired from sensors in pipelines could be studied to develop analytical models. By combining known hydraulic equations and state analysis of the pipeline, physical modeling could be performed for a network [3]. Further analyses of data could be done by applying machine learning techniques to establish correlation between different factors, and distinguish false positives from actual leaks [4]. More often than not, a combination of certain aspects of these techniques are used to improve accuracy of leak detection [5].

1.3 Data Analysis

Datasets containing real-world data could be used to model existing water distribution networks and possible leakage occurrences in conjunction with actual demand patterns [6]. However, in the absence of real-world data, EPANET software would be used to simulate leakage in water distribution networks [7]. Components such as pipes and valves could be implemented in simulations to reflect physical models. Datasets are constructed from such simulations, and could be used for analysis. For this study, the datasets are generated with emphasis on varying sensor density and topology. Datasets from these networks would then be evaluated to test the machine learning model. Using machine learning, leakage detection in these networks could be performed [8].

1.4 Overview

The main objective of this research is to perform simulations on different water distribution networks, and develop a machine learning approach for leak detection. Factors affecting WDN operations such as water demand and physical characteristics of the WDN will be taken into consideration. Performance of the machine learning model will be evaluated on different network topologies and sensor density.

2 Related Work

2.1 Leak Detection Methods

Without the usage of other leak detection techniques, on-site inspections by water utility personnel are relied upon [2]. Above-ground sensor reading equipment must be deployed along the WDN for a technician or engineer to be able to note sensor data. As the response highly involves human intervention, the system could possibly undergo significant losses before the appropriate action is implemented [1].

Acoustic sensors could be installed along the WDN to detect sound signals, which could be used with signal processing techniques to detect and localize leakages [2]. Signals would be higher in amplitude near the leaks, allowing for localization. This method has the advantage of not needing a complex mathematical model. However, acoustic sensing techniques are easily affected by the physical environment around it; external noise might interfere with sensor readings [9]. The cost of implementing the hardware for the system would be high as well, especially for larger networks [1].

Another technique used for leak detection are balancing methods. Under normal operation, input and output mass flow rate of each pipeline are equal [2], which is supported by the law of conservation of mass. Thus, imbalances of these metrics would imply leakage within the pipeline. The balance between input and output flows could be represented by the equation

$$M_I - M_O = \varDelta M_{pipe} \tag{1}$$

where M_I and M_O represent mass in the inlet and outlet of the pipe, respectively, while $\varDelta M_{pipe}$ represents the imbalance between the two. As mass is proportional to the flow rates, input and output flows could also be represented by the equation

$$Q_I - Q_O = \varDelta Q_{pipe} \tag{2}$$

where Q_I and Q_O similarly represent flow rates in the inlet and outlet of the pipe, respectively, while $\varDelta Q_{pipe}$ represents the imbalance between the two [1]. To be within range of possible errors in sensor readings, an alarm threshold is implemented based on pressure or flow values. While rather simplistic in theory, utilizing pressure and flow data is a fundamental basis for techniques that rely on data processing of hydraulic models.

Transient analysis could also be utilized as a leak detection method [10]. Transient information generated by leakage scenarios could be extracted and then evaluated accordingly. The main idea of transient analysis is to compare signals received by a sensor in a leakage event to signals under normal operation. Focusing on the transient signals can offer more information rather than simply monitoring the WDN at its steady state [10,11].

Model-based methods could be assessed, to determine the state of subsections of a WDN based on water hammer equations [3]. The underlying fundamental principle relating flow and pressure could be summarized in equations

$$\frac{\partial H}{\partial t} = -\frac{c^2}{gA}\frac{\partial Q}{\partial z} \tag{3}$$

$$\frac{\partial Q}{\partial t} = -Ag\frac{\partial H}{\partial z} - \frac{fQ|Q|}{2DA} \tag{4}$$

where t and z represent coordinates in time and space, respectively, H representing pressure, c representing wave speed, g representing Earth's gravity, Q representing flow rate, D representing the diameter of the pipeline, A representing the cross-sectional area of the pipe, and f representing the friction coefficient [3].

While model-based methods based on physical properties of a WDN yield high accuracies in fault detection, it is not feasible to use in large-scale networks due to the amount of parameters and data involved [12,13]. An alternative is to use machine learning techniques, which generate similarly high accuracies, but require less computational load [8,14]. Machine learning maps the dependent and independent parameters of a given system, with little prior process knowledge. Performance of machine learning techniques are based on the design of the technique, which could be improved over time with data [13].

Machine learning techniques and their properties could be incorporated into a single learning algorithm called an ensemble method [15]. Ensemble methods are multiple machine learning techniques that are trained cooperatively to produce better predictive performance. In the field of hydrology, there have been proposals to utilize ensemble methods to improve the performances of WDNs [14,16]. Pairing up single methods with each other could help decrease unwanted levels in bias and variance, which could ultimately increase predictive performance.

2.2 EPANET

EPANET is a simulation software for modeling water distribution networks and examine hydraulic behavior in pipe networks. It is able to simulate different WDN entities, such as pipes, tanks, valves, and reservoirs. Factors such as water pressure, water flow, and chemical concentrations could be monitored within the system. EPANET also provides a visual simulator to help the user build pipe networks and edit properties such as pipe diameter or valve function [7].

By simulating emissions within a water distribution network, EPANET is able to model leakage events, and the properties of the network during such an event could be studied. EPANET has been used in modeling and detecting leakages in WDNs of real-life communities and small-scale testbeds [17,18] (Fig. 1).

EPANET takes advantage of combining demand patterns and leakage properties to model leaks within a specified time frame. This would help in detecting the possibility of leakage within the WDN [18]. A relation used in EPANET to model leakage is as follows [17]:

$$Q = E_c * P^{P_{exp}}, \tag{5}$$

where Q represents flow, E_c represents the emitter coefficient, P represents pressure, and P_{exp} represents the pressure exponent, which is usually set to 0.5 for water networks.

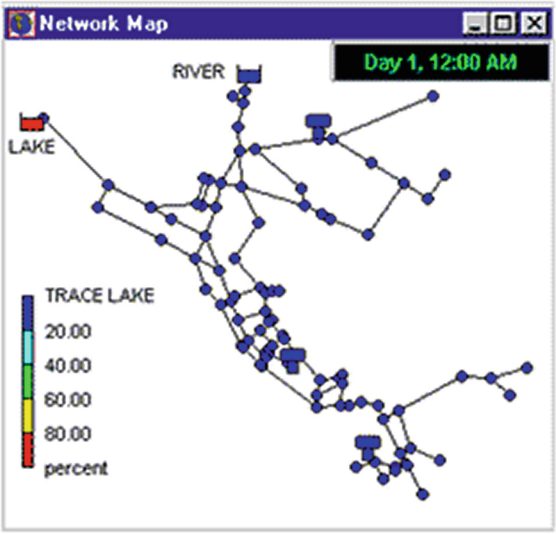

Fig. 1. EPANET window simulation

3 Methodology

The methodology is composed of three major tasks: simulations, feature extraction, and classification. Simulation modeling includes the construction of a functional and realistic water distribution network in simulation software. Simulations involving leakage events are carried out afterwards, done on WDNs with varying topology and sensor density.

The proposed solution is to apply wavelet series decomposition on pressure data, and to use the energy of each wavelet as a feature for the machine learning model. Specifically, a CNN ensemble is implemented as part of the machine learning model to analyze the data. The CNN ensemble would effectively classify the pressure profiles from the sensors as to whether or not they have a leak. Instead of using each recording from each sensor node as a feature, the energy of each wavelet is used, effectively reducing the size of the dataset. This allows the machine learning model to learn from another feature, which could result in better detection.

3.1 Hydraulic Simulations

In the absence of datasets for real-life water distribution networks, simulations are performed in EPANET to produce data from networks with similar characteristics. Water distribution networks are constructed in EPANET by implementing elements such as pipes, junctions, and reservoirs. These networks in the EPANET program reflect certain parts, if not the whole, of actual WDNs.

For the analysis, three distinct EPANET networks are used: the Cherry Hills/Brushy Plains network, the New York City Tunnel network, and the

Fossolo network. These networks are derived from actual implementations of existing water distribution networks [19,20], and are distributed with EPANET releases [7].

The simulations would incorporate hourly water demand profiles used by households [18]. A profile could be constructed via analysis of previous consumption data [21], or sociological analysis [22]. For example, usage in residential areas is typically at lowest after midnight, and would peak during early mornings or early evenings [23]. Accuracy of leak detection would depend on how much volume of water is taken into consideration within pipelines.

Pipes, junctions, and other elements close to the emitter nodes are monitored for changes. The model for the water distribution networks is able to capture key characteristics, and these will be the basis for leak detection. Such characteristics include pressure, water flow, water head, and water quality. Out of these, pressure changes are most reflective of leakage scenarios [4], making it the ideal data input for the proposed solution. Pressure data are recorded accordingly based on EPANET's time reporting step, which was set to every 30 s. Changes in the pressure profile is then analyzed to determine and localize leaks.

Changes on leak sizes and leak locations are then performed. These combinations of changes are used to create datasets for machine learning. To introduce leakage events, the emitter coefficients of the nodes within the network are be increased. Shown in Fig. 2 are the differences between the pressure profile of a node in the Cherry Hills/Brushy Plains network within a 24-h period upon increasing the emitter coefficient. The drop in pressure is found to be consistent with hydraulic principles.

To facilitate leak detection, sensor nodes are implemented into every EPANET simulation of a water distribution network. These are assigned out of existing junction nodes, and the placement of these would have an impact in detecting leaks. This allows for a mechanism to monitor and record pressure profiles within the network, which would then be analyzed accordingly [18].

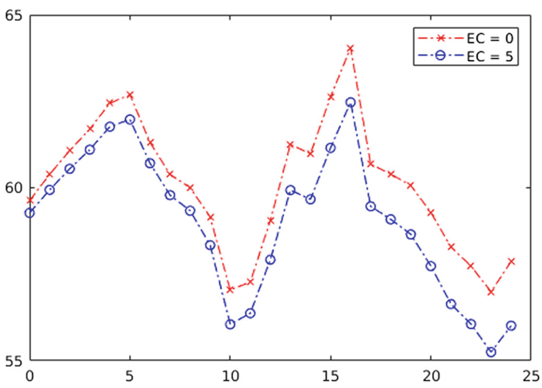

Fig. 2. Pressure profile of the same node at $E_c = 0$ and at $E_c = 5$

To test the effect of sensor density, sensor placement is done differently on each simulation. Specifically, different simulations have sensors that are placed on locations that are 1-, 2-, and 3-hop distances from non-sensor junction nodes, without overlap.

3.2 Feature Extraction and Wavelet Decomposition

The data streams that will be observed are pressure streams coming from sensor nodes. Intuitively, pressure data observed from components away from the leakage event would be inversely proportional to the distance from the leak. Generally, leakage within the vicinity would result in more drastic pressure drops.

Furthermore, wavelet transforms could be applied to these signals, where they would be decomposed into their respective approximation coefficients and detail coefficients [4, 24]. The approximation coefficient corresponds to the lower frequency of the original signal, while the detail coefficients correspond to the higher frequency of the original signal. Different levels of the detail coefficient could possibly detect transients in the signal, which could be correlated with detecting leakages in the pipeline, as leakage events come with changes of pressure. The detail coefficients also capture the noise element of the signal, while the approximation coefficient resemble the original signal (Fig. 3).

The energy of the coefficients would be computed, and used as the feature vectors for each sensor, effectively reduceing the input size of the dataset. Utilizing the energy of the pressure signals instead of raw pressure data allow for minimizing energy consumption of the sensors in sending data over a network [18], as well as allowing the CNN ensemble to possibly find more hidden features [4].

These data would then be analyzed via machine learning by a CNN ensemble. It is expected that the different networks would provide different pressure profiles for the CNN ensemble to learn.

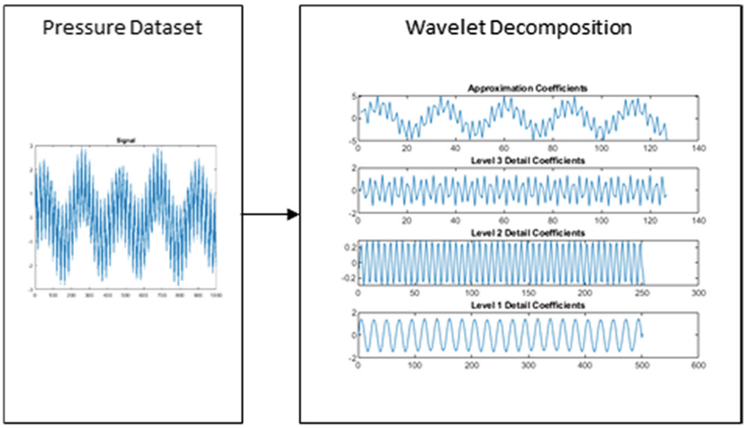

Fig. 3. Wavelet Decomposition

3.3 Classification via Machine Learning Ensemble

The main concern of the study is to identify the occurrence of leakage scenarios given a profile of a water distribution network. The study also aims to compare the effectiveness of the proposed solution on different network topologies and different sensor densities.

A machine learning ensemble is implemented by training multiple machine learning sub-models then stacking them together to form an ensemble. Specifically, an ensemble consisting of a stacked convolutional neural network (CNN) sub-models is used to classify leakage events by learning patterns from the generated wavelets. A CNN is a machine learning implementation used for feature extraction [25] and has been used extensively in the field of hydrology [8,26]. This could be utilized to reveal hidden features relevant for leak detection. The CNN ensemble would act as a binary classifier for leak detection, where its output would be to determine whether or not a leak exists within the network. Adjustments in configurations are made to maximize performance for a specific network model.

Data generated from EPANET simulations are used as input for the machine learning model. The models are implemented in Python using Keras and Tensorflow as framework, with a 70/30 split of the dataset for the training and test set.

3.4 Evaluation Metrics

The techniques above will be evaluated with respect to certain metrics: classification accuracy, true positive rate (TPR), false positive rate (FPR), and the area under the curve (AUC) of the receiver operating characteristic (ROC) [17,27]. A summary of these metrics is as follows:

$$Accuracy = \frac{correct}{total} * 100 \tag{6}$$

$$TPR = \frac{TP}{TP + FN} \tag{7}$$

$$FPR = \frac{FP}{FP + TN} \tag{8}$$

where TP, TN, FP, and FN represent the total amount of true positives, true negatives, false positives, and false negatives respectively. A true positive (TP) is defined as correctly identifying a leakage scenario occurring in the WDN, while a true negative (TN) is defined as correctly identifying that the WDN is operational without leaks. A false positive (FP) is defined as incorrectly noting a leak in the WDN.

The AUC-ROC is an often used evaluation metric for machine learning [27], which is generally characterized by plotting the true positive rate against the false positive rate. The area is a measure of how much the model is able to distinguish between classes. A higher AUC corresponds to a better classifier.

4 Results and Discussion

Shown in Table 1 are the physical attributes of the tested networks. These serve as a baseline on how network topologies physically differ from one another.

Table 1. Physical characteristics of different network topologies

Metric	New York tunnel	Cherry plains	Fossolo
Nodes	19	34	36
Avg. Node Connections	2.105263	2.294118	3.194444

Tables 2, 3, 4 present the performance of the proposed leak detection method. The ensemble CNN method was able to reach accuracies of more than 90% in each topology, even when considering the differences in sensor density. While this is the case, the accuracy of the model is dependent on how complex or large the network; leak detection on networks with more connections between nodes tend to be slightly less accurate than networks with fewer connections. Similarly, the true positive rate for all topologies and sensor densities reach an adequate percentage for leak detection. This metric is relevant as there is a need to correctly determine the presence of leaks in the event of one occurring. In all region sizes, it is seen that the Fossolo network has the lowest true negative rate, while the Cherry Plains network has the highest. The true negative rate of the Fossolo network is at its lowest when using data from the 3-hop node distance region. This may be due to the fact that the dataset is mostly comprised of leakage scenarios, with non-leakage scenarios only taking up roughly 14% of the datasets, and compounded further by having generated less data from a network configuration with less sensors. It is likely that the CNN ensemble fails to find more hidden features in these networks, where the pressure profiles are comprised of more nodal connections; hence the low true negative rate. Generally, this is more acceptable than the opposite, where there is a low true positive rate, and a high negative rate, as inaction for an existing leak would entail more economic losses than routinely checking on pipe operations.

Table 2. Performance of the model for leak detection of different network topologies, with the sensor density based on 1-Hop node distances

Sensor density based on 1-Hop node distances			
Metric	New York tunnel	Cherry plains	Fossolo
Regions	8	14	10
Accuracy	0.987737584	0.98560658	0.96956944
TPR	0.994318182	0.98458498	0.990661188
TNR	0.946188340	0.992268041	0.832524272

Table 3. Performance of the model for leak detection of different network topologies, with the sensor density based on 2-Hop node distances

Sensor density based on 2-Hop node distances			
Metric	New York tunnel	Cherry plains	Fossolo
Regions	4	6	4
Accuracy	0.984671980	0.977039068	0.966655876
TPR	0.994318182	0.973913043	0.988793425
TNR	0.923766816	0.99742268	0.822815534

While there is not much difference in the accuracy and true positive rate metrics between sensor densities, a noticeable trend is seen in the true negative rate. As the sub-regions of the network become larger, a decline in the true negative rate can be noticed. This is corroborated by the decrease in the AUC-ROC for the Fossolo network, as shown from Figs. 4, 5, 6, as it is affected by the increase in false negatives. It is possible that covering a large area within the

Table 4. Performance of the model for leak detection of different network topologies, with the sensor density based on 3-Hop node distances

Sensor density based on 3-Hop node distances			
Metric	New York tunnel	Cherry plains	Fossolo
Regions	3	4	3
Accuracy	0.984671980	0.979095271	0.946260926
TPR	0.995028409	0.980237154	0.987672768
TNR	0.919282511	0.971649485	0.677184466

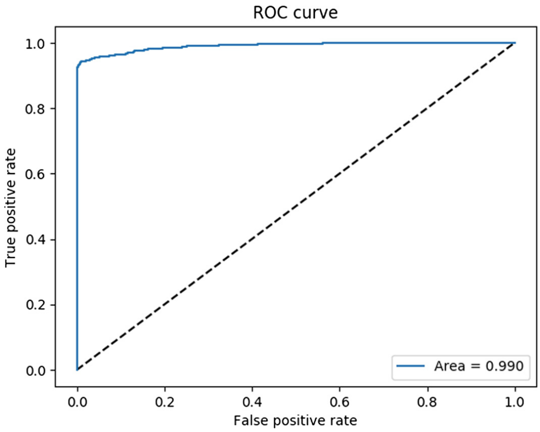

Fig. 4. ROC curve of the Fossolo network, 1-hop region size

WDN makes it more difficult for the machine learning model to learn connections in the pressure profile and correctly classify non-leakage scenarios. Having more connections on average for each junction further exacerbates this problem, as nodal pressures are more affected by neighboring nodes.

In general, the model was able to correctly classify leakage and non-leakage scenarios correctly. From the results gathered, significant increases in accuracy, TPR, and TNR for leak detection can be achieved by allowing the model to be trained further, as each network configuration was trained using the same machine learning model and the same amount of training periods. Networks with higher connections between nodes require more training as well, to find more connections between the different pressure profiles generated by each node.

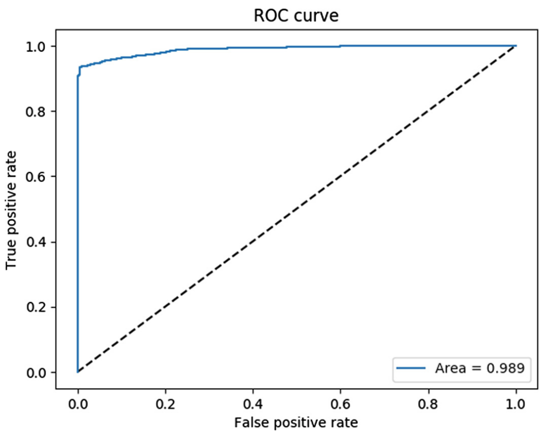

Fig. 5. ROC curve of the Fossolo network, 2-hop region size

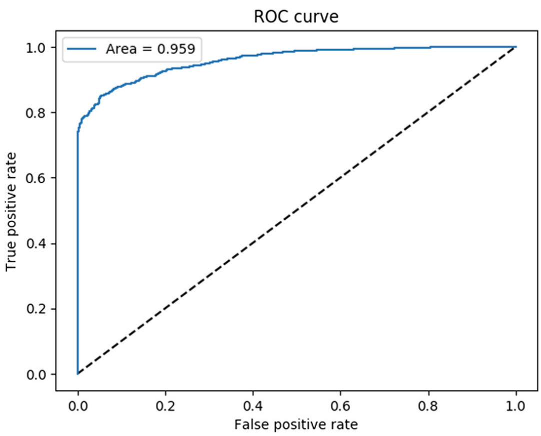

Fig. 6. ROC curve of the Fossolo network, 3-hop region size

5 Conclusions

This study explored the problem of detecting leakages in water distribution systems. The proposed solution of using a CNN ensemble with wavelet decomposition was able to determine occurrences of leaks within the pipeline. Wavelet decomposition allowed the data used for machine learning features to be relatively small, while retaining information regarding changes in the pressure profile. By using a CNN as a classifier, connections in the pressure profiles were analyzed and yielded sufficient performance by the model. Generally, the model is able to perform well on different network topologies, different sensor densities, and certain combinations thereof.

5.1 Limitations of the Study

The research dealt exclusively with simulated data, as real-life datasets were found to be incomplete or hard to come by. Sensor placement within the simulated networks were also not optimized. While the sensors were physically distributed equitably over the network, a study was not done on how specific sensor placements affect the overall performance of the model. Determining where the leak is located is likewise not incorporated into the study. As the study simply aims to detect whether or not a leak exists in the system, the utility of this knowledge depends on the size of the network and how easy it is for the leak to be alleviated. The trained model was not used in actual implementations of water distribution networks. This is primary due to the scale of the simulated networks, and the difficulty in implementing small-scale ones.

5.2 Future Work

Improvements in the implementation of the CNN ensemble could still lead to better performance. From the results, it could be seen that the model performs worse on networks that are relatively large and have a higher number of connections per node; better results could be achieved by optimizing the layers and parameters of the model. A better model would also allow for a better classifier, allowing for the classification of more regions within the network or more dense sensor distributions.

Similar to leak detection, machine learning techniques could also possibly be used for determining the location of the leak within the network. The CNN could be used as a multiple classifier, where outputs are different regions within the network, narrowed down to a smaller area.

References

1. Adedeji, K.B., Hamam, Y., Abe, B.T., Abu-Mahfouz, A.M.: Towards achieving a reliable leakage detection and localization algorithm for application in water piping networks: an overview. IEEE Access **5**, 20272–20285 (2017). https://doi.org/10.1109/access.2017.2752802

2. Wan, J., Yu, Y., Wu, Y., Feng, R., Yu, N.: Hierarchical leak detection and localization method in natural gas pipeline monitoring sensor networks. Sensors **12**(12), 189–214 (2011). https://doi.org/10.3390/s120100189
3. He, Y., Li, S., Zheng, Y.: Distributed state estimation for leak detection in water supply networks. IEEE/CAA J. Autom. Sin. **PP**(99), 1–9 (2017). https://doi.org/10.1109/JAS.2017.7510367
4. Gupta, G.: Monitoring Water Distribution Network using Machine Learning. KTH Royal Institute of Sweden, Stockholm (2017)
5. Xu, Q., Liu, R., Chen, Q., Li, R.: Review on water leakage control in distribution networks and the associated environmental benefits. J. Environ. Sci. (China) **26**(5), 955–961 (2014). https://doi.org/10.1016/S1001-0742(13)60569-0
6. Conejos, M.P., Alzamora, F.M., Alonso, J.C.: A water distribution system model to simulate critical scenarios by considering both leakage and pressure dependent demands. Procedia Eng. **186**, 380–387 (2017). https://doi.org/10.1016/j.proeng.2017.03.234
7. EPANET. https://www.epa.gov/water-research/epanet. Accessed 23 May 2018
8. Kang, D., Lansey, K.: Novel approach to detecting pipe bursts in water distribution networks. J. Water Resour. Plan. Manage. **140**(1), 121–127 (2014). https://doi.org/10.1061/(ASCE)WR.1943-5452.0000264
9. Fang, C.M., Lin, S.C.: 5.4 GHz high-Q bandpass filter for wireless sensor network system. In: IEEE SENSORS 2009 Conference (2009). https://doi.org/10.1109/ICSENS.2009.5398458
10. Colombo, A.F., Lee, P., Karney, B.W.: A selective literature review of transient-based leak detection methods. J. Hydro-Environ. Res. **2**(4), 212–227 (2009). https://doi.org/10.1016/j.jher.2009.02.003
11. Wang, X., Ghidaoui, M.S.: Identification of multiple leaks in pipeline: linearized model, maximum likelihood, and super-resolution localization. Mech. Syst. Signal Process. **107**, 529–548 (2018). https://doi.org/10.1016/j.ymssp.2018.01.042
12. Quiñones-Grueiro, M., Verde, C., Llanes-Santiago, O.: Demand model in water distribution networks for fault detection. IFAC-PapersOnLine **50**(1), 3263–3268 (2017). https://doi.org/10.1016/j.ifacol.2017.08.460
13. Goyal, M.K., Ojha, C.S.P., Burn, D.H.: Machine learning algorithms and their application in water resources management. In: Sustainable Water Resources Management, pp. 165–177. American Society of Civil Engineers (2017). https://doi.org/10.1061/9780784414767.ch06
14. Kang, J., Park, Y., Lee, J., Wang, S., Eom, D.: Novel leakage detection by ensemble CNN-SVM and graph-based localization in water distribution systems. IEEE Trans. Ind. Electron. **65**(5), 4279–4289 (2018). https://doi.org/10.1109/TIE.2017.2764861
15. Dietterich, T.G.: Ensemble methods in machine learning. In: Kittler, J., Roli, F. (eds.) MCS 2000. LNCS, vol. 1857, pp. 1–15. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-45014-9_1
16. Shi, F., Liu, Z., Li, E.: Prediction of pipe performance with ensemble machine learning based approaches. In: 2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC), Shanghai, pp. 408–414 (2017). https://doi.org/10.1109/SDPC.2017.84
17. Nasir, M.T., Mysorewala, M., Cheded, L., Siddiqui, B., Sabih, M.: Measurement error sensitivity analysis for detecting and locating leak in pipeline using ANN and SVM. In: 2014 IEEE 11th International Multi-Conference System Signals Devices, SSD 2014, pp. 7–10 (2014). https://doi.org/10.1109/SSD.2014.6808847

18. Gupta, K., Kishore, K., Jain, S.C.: Modeling and simulation of CEERI's water distribution network to detect leakage using HLR approach. In: 2017 6th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions) (2017). https://doi.org/10.1109/ICRITO.2017.8342440
19. Rossman, L.A., Clark, R.M., Grayman, W.M.: Modeling chlorine residuals in drinking-water distribution systems. J. Environ. Eng. **120**(4), 803–820 (1994). https://doi.org/10.1061/(ASCE)0733-9372(1994)120:4(803)
20. University of Exeter Centre for Water Systems. http://emps.exeter.ac.uk/engineering/research/cws/resources/benchmarks/. Accessed 23 May 2018
21. Delorme-Costil, A., Bezian, J.J.: Forecasting domestic hot water demand in residential house using artificial neural networks. In: 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), Cancun, pp. 467–472 (2017). https://doi.org/10.1109/ICMLA.2017.0-117
22. Domene, E., Sauri, D.: Urbanisation and water consumption: influencing factors in the metropolitan region of Barcelona. Urban Stud. **43**(9), 1605–1623 (2006). https://doi.org/10.1080/00420980600749969
23. Letting, L.K., Hamam, Y., Abu-Mahfouz, A.M.: Estimation of water demand in water distribution systems using particle swarm optimization. Water, 9(8) (2017). https://doi.org/10.3390/w9080593
24. Fagiani, M., Squartini, S., Severini, M., Piazza, F.: A novelty detection approach to identify the occurrence of leakage in smart gas and water grids. In: Proceedings of the International Joint Conference Neural Networks, vol. 2015, September 2015. https://doi.org/10.1109/IJCNN.2015.7280473
25. Wu, Z.Y., El-Maghraby, M., Pathak, S.: Applications of deep learning for smart water networks. Procedia Eng. **119**, 479–485 (2015). https://doi.org/10.1016/j.proeng.2015.08.870
26. Xu, Y., Zhang, J., Long, Z., Chen, Y.: A novel dual-scale deep belief network method for daily urban water demand forecasting. Energies **11**(5), 1068 (2018). https://doi.org/10.3390/en11051068
27. Sokolova, M., Lapalme, G.: A systematic analysis of performance measures for classification tasks. Inf. Process. Manage. **45**(4), 427–437 (2009). https://doi.org/10.1016/j.ipm.2009.03.002