Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 48-8 (2015) 1264–1269

Robust Leakage Detection and Interval Estimation of Location in Water Distribution Network

Yeonsoo Kim, Shin Je Lee, Taekyoon Park * **Gibaek Lee** ** **Jung Chul Suh** *** **Jong Min Lee** *,[1]

* *School of Chemical and Biological Engineering, Seoul National University, Seoul, Republic of Korea.*
** *Department of Chemical and Biological Engineering, Korea National University of Transportation, Chungju-si, Republic of Korea.*
*** *Samchully Corporation, Seoul, Republic of Korea.*

Abstract: The water supply network has a complex structure especially in cities with high population density. A damage to the water pipe can occur in the form of a leakage or a burst and the technique for early detection of the occurrence and for the exact determination of the location is required. In this paper, we propose a novel method that can detect the leakage of the water supply network using the pressure data. After the noise is eliminated using the Kalman Filter, the mean of normal state pressure is calculated and deviation with the mean is obtained. By calculating the cumulative integral of the pretreated data and applying a floor function, the leakage can be detected. Once the leakage is detected, the time of occurrence is refined by radius of curvature and the location is estimated by using that time and a statistical method. The verification test is conducted with respect to the two different field data. It is found that the prorposed method is more robust and practical to implement and shows a higher precision compared to the previous methods.

Keywords: Fault detection, Fault location, Statistical analysis, Kalman filters, Cumulative Integrals, Pipe networks,

1. INTRODUCTION

Water supply network facilities are increased and become complex as people gather in a big city. Once pipelines are installed underground, the fault detection becomes very difficult owing to the inaccessibility and complex structure. Therefore if a fault such as the rupture, leakage, and burst occurs, solutions are usually made for the post management. As the time for fault diagnosis and trouble shooting is delayed, direct and indirect losses including the water loss, pipeline network repair cost and damage to the surrounding facilities increase exponentially. Therefore, it is necessary to develop a proactive pipeline network management system to prevent the accidents and minimize the losses.

Methods for monitoring the leakage or the rupture of the water distribution network is classified into two methods; model-based method(Wu et al. (2010); Perez et al. (2014)) and measurement-based method(Covas and Ramos (2010); Mulholland et al. (2014)). The measurement-based method is further classified into two types; the volume balance method and the pressure point analysis method.

Jung and Lansey (2014) propose a method of detecting the burst based on model, using flow data and Kalman Filter. After calculating the flow rate requirements using Kalman filter, the prediction to estimate the flow rate at the point of flow meter is made by hydraulic model. The decision on burst occurrence is made based on the difference between the estimated flow rate and the real measurements.

It is shown that the burst could be detected by fuzzy analysis using the residual value of flow measurement by Ragot and Maquin (2006). They calculate the error between residual value of the predicted flow rate and actual residual value. After that, the highest potential leak points are determined using a fuzzy-based isolation method. Including these two studies, some algorithms use a flow rate.

However, it is easier to install and manage the pressure gauge than the flow meters in the real water distribution network and the number of flow meter is much smaller than that of pressure gauges. The leak detection algorithms based on the pressure measurement have been suggested from these reasons (Ponce et al. (2014)).

Misiunas (2005) proposes the algorithm based on the cumulative summation(CUSUM). Using this, the abrupt changes in pressure can be expressed as peaks. If those peaks exceed a threshold value, it is concluded that the burst event occurs. Srirangarajan et al. (2013) use the pressure measurement and multiscale wavelet analysis(MWA) to detect the burst. The pressure data is decomposed into approximation and detail coefficients up to level 4. First, the time when the coefficients at level 3 and level 4 are large are recorded. If the coefficient magnitudes at higher level are larger than those at lower level around that time,

[1] Associate Professor, Corresponding author(e-mail:jongmin@snu.ac.kr)

10.1016/j.ifacol.2015.09.142

the algorithm indicates that the burst occurs.
However, these two algorithms can only detect large and abrupt pressure change and cannot detect sustained pressure reduction. If the small scale leakage occurs, these two algorithms cannot detect the event. Because there are unknown disturbances such as water demand in real water network, it is not enough to find the abrupt pressure change.
To overcome these problems, this study suggests a new robust algorithm. It uses only the pressure data without flow rate data and detects a small scale leakage. First, the measured pressure is filtered by Kalman Fliter and the cumulative integral is applied to amplify the effect of leak. Second, a floor function and a curvature function are introduced based on the characteristics in the pressure data when the leakage occurs. Finally the statistical method is applied to estimate the leakage location with a specific mathematical confidence. This proposed algorithm was validated by false alarm and leakage detection tests.

2. PROPOSED ALGORITHM

2.1 Basic Assumption

In this paper, only the pressure data is used. Although there are flow meters in water distribution network, the number of the flow meters is very small compared to that of the pressure gauge. Because the flow meter is relatively difficult to install, maintain, and repair(Ponce et al. (2014)).

2.2 Data Generation

Because only the burst case data of real water distribution network can be obtained from the field test we performed, it is necessary to generate the leakage case data. Based on the real burst data, the leakage data is generated.

$$P_{leak} = K_{factor} \times P_{burst,e} + (1 - K_{factor}) \times P_{burst,e}[k_0] + noise \tag{1}$$

The specific process is as follows:

(1) Using the Kalman Filter (KF), separate the real burst pressure data into the pressure estimation ($P_{burst,e}$) and noise.

(2) Multiply a scaling factor K_{factor} by $P_{burst,e}$.

(3) Correct the initial point by adding $((1 - K_{factor}) \times P_{burst,e}[k_0])$.

(4) Add the noise separated in step (1).

2.3 Data Filtering

To reduce the computational load, only the one-tenth of generated data was used resulting in the change of the sampling frequency from 250 Hz to 25 Hz. These data is filtered using KF and we obtain the estimated pressure value($P_{leak,e}$). The period of sampling(T) is 0.04 s. Let $P[k]$ be the pressure of leakage case, $w[k]$ be the model error, $z[k]$ be the measurement, and $v[k]$ be the sensor noise at kT s. In model equation, the current pressure is assumed equal to the pressure of previous time and Q which is the covariance of $w[k]$ is set small because the pressure is sampled in a very short interval. When observing the pressure data in the normal state, noise due to the pressure sensor is shown about ± 5 kPa. Based on this, R which is the covariance of $v[k]$ is set 3. C_0 is the error covariance of the predicted initial value and is set to be same as that of R, because initial pressure measurement is assigned as the initial value.

$$\begin{aligned} P[k+1] &= P[k] + w[k] \\ z[k] &= P[k] + v[k] \\ Q &= 0.001 \\ R &= 3 \\ C_0 &= 3 \end{aligned} \tag{2}$$

2.4 Data Shifting

Because the range of the measured pressure is different for each sensor, it is required to correct the average pressure value to near zero in normal state before the leakage occurs. For this purpose, the average value is subtracted from the measured pressure.

$$\begin{aligned} P_m &= \sum_{i=1}^{j} P_{leak,e}[i] \\ P_{shift}[k] &= P_{leak,e}[k] - P_m \end{aligned} \tag{3}$$

Here j is obtained by calculating $\frac{50s}{0.04s}$, which means the sampling time is 0.04s and the interval for obtaining the average value is 50s. Note that the value of 50 s can vary but it does not cause significant change in the result because it means a simple shifting. The data shifting and the subsequent process are applied in a receding horizon fashion with the time window size of W and update interval T_u. For example, if the first window for observation is between t_0 and $t_0{+}W$, then the second window will be between $t_0{+}T_u$ and $t_0{+}T_u{+}W$. It is necessary because an aperiodic fluctuation exists in data of the day caused by water usage. The size of W and T_u can be chosen randomly by users because they don't have effect on the performance of the proposed algorithm unless the size of W is too large.

2.5 Cumulative Integral

The major difference between the burst and leakage is the degree of pressure change. The change in leakage is much smaller than that in burst(Rashid et al. (2014)). In other words, it is not easy to separate the pressure changes caused by leakage from by others in the presence of disturbances such as water consumption. To overcome this obstacle, we focus on the tendency to maintain the reduced pressure value after the leakage occurrence. By using cumulative integral, it is possible to make that tendency visible.

$$P_{CI}[k] = \sum_{i=1}^{k} P_{shift}[i]\Delta(iT)$$

2.6 Floor Function

Although we modify the measured pressure data with various methods, the method which can identify the reduced pressure value which is maintained is needed. Most researches apply the methods to determine the time of the instantaneous pressure drop, which can cause false alarms if the amount of the leakage is small in the noisy situation. The floor function with following method can filter out the small pressure drop by the disturbances and can identify the decreasing tendency of P_{CI} caused by leakage.

$$P_{floor}(k) = \left[\frac{P_{CI}(k)}{\frac{P_m}{10}} \right] \tag{4}$$

ϕ $\equiv$ the threshold value of the time where the floor function is constant.

$N \equiv$ the threshold number of the consecutive decrease in the floor function.

By observing P_{floor} with ϕ and N which is selected by normal pressure data, leakage is detected and an alarm is generated if the continuous decreasing tendency of P_{floor} occurs C times with the time interval ψ_i $(i = 1, 2, \cdots, C)$ such that $C \geq N$ and $\psi \leq \phi$ for all i. And the first time of P_{floor} decrease(t_{drop}) is recorded. Three parameters are closely related: the denominator $(\frac{P_m}{10})$, ϕ, and N. The denominator affects the frequency of the segment and it is determined by the size of leakage or burst that user want to detect. In this study, we set the denominator as a form of $\frac{P_m}{10}$ to make the P_{CI} over the 10% of P_m to be recorded. After setting the denominator, ϕ and N is determined by using the normal pressure data. The maximum time duration and the maximum number of continuous decrease in P_{floor} are respectively M' and N' in normal pressure data. Then user can set M which is lager than M' by considering margin and M means $\phi \times N$. User also set N which is larger than N'. By M and N, ϕ is determined automatically.

2.7 Curvature Function

Although t_{drop} is recorded, the arrival time of the leakage effect to sensor can be refined by finding the cusp in P_{CI} near the t_{drop}. For this, we introduce a curvature function, $\kappa(k)$, which can present the smoothness of the graph. It is possible to find sudden cusp points based on the curvature. Unlike MWA method, the curvature function has no loss of time information resulting in the improvement in the accuracy of the estimation results.

$$\kappa(k) = \frac{1}{R(k)} \tag{5}$$

Here, $R(k)$ is the radius of the circle which is constructed by adjacent three points of P_{CI}. Applying the curvature function to P_{CI}, the cusp can be detected by finding the maximum of the curvature near the t_{drop}, which is used as wave arrival time to sensor.

2.8 Location Estimation

Previous studies estimate the burst location using the estimated wave speed and the result is presented as a point with distance error. However, it is unrealistic to investigate the specific point obtained from the leak detection algorithm even if the error is very small. When the leak is observed, the section of the water pipeline rather than the point is investigated. Considering these, a new method for the location estimation is proposed.

Time of Occurrence The leakage in the water distribution system can be expressed by two variables : occurrence time(t_{occ}) and the location of leakage(L_{leak}). Instead of calculating the wave speed, if we assume the wave speed due to the leakage has a constant value, the degree of freedom can be reduced to one and we only need to find one of t_{occ} and L_{leak}.

Confidence Bound Estimation Based on the detection time in each sensor, confidence bound is estimated.

a. Node Generation

The GIS data for the water distribution network is used. The distances between the adjacent nodes, D_i, are divided into $[\frac{D_i}{10}] + 1$ segments with an equal length if D_i is larger than 10 m for precision of location estimation.

b. Selection of Basis Sensors

Let n be the number of the sensors ($n \geq 4$). Two sensors are needed to calculate the time of occurrence($t_{occ,s}$) at node s where it is assumed that the leakage occurs. Thus, ${}_nC_2$ is the possible number of bases. The time of occurrence at the selected node is calculated as follows.

$$d_{is} : d_{js} = (t_{i,basis} - t_{occ,s}) : (t_{j,basis} - t_{occ,s}),\ (i < j) \tag{6}$$

Here, d_{is} and d_{js} are the distances from the basis sensors (the i^{th} and j^{th} sensor) to the selected node, respectively. $t_{i,basis}$ and $t_{j,basis}$ are the recorded time at each basis sensors. There is no verification points with two basis sensors for $n = 2$ and the confidence of interval is low for $n = 3$. Therefore, for $2 \leq n \leq 3$ we used the method proposed by Srirangarajan et al. (2013) which estimated the burst location by searching for the node s that minimize (7). There is no solution for $n = 1$.

$$J(s) = \sum_{i,j=1}^{n} |(t_{i,obs} - t_{j,obs}) - (\tau_i - \tau_j)|,\ (i > j) \tag{7}$$

Here, $t_{i,obs}$ means the recorded arrival time of each sensor and τ_i is the calculated arrival time of wave to each sensor by using estimated wave speed.

c. Verification of the Measured Time

From the calculated $t_{occ,s}$ based on two basis sensors, it is possible to obtain the estimated times which the shock wave arrives at each other sensors.

$$d_{is} : d_{ls} = (t_{i,basis} - t_{occ,s}) : (t'_l - t_{occ,s}),\ (l \neq i, j) \tag{8}$$

Since $t_{i,basis}$ and $t_{j,basis}$ are related via (6), d_{is} and $t_{i,basis}$ can be replaced by d_{js} and $t_{j,basis}$. After calculating the t'_l for $(n-2)$ sensors, we can set the objective function, J, as the sum of the squared errors of estimated arrival time.

$$J(s) = \sum_{l=1}^{n} (t'_l - t_{l,obs})^2, \ (l \neq i, j) \quad (9)$$

Here, $t_{l,obs}$ means the recorded arrival time of each sensor. By finding the node s which has minimum of $J(s)$, we can obtain at least one point for each sensor combinations. This means that the result is shown as more than ${}_nC_2$ points in the water distribution system.

d. Statistical Estimation

If the leakage location is estimated by a point, the pipeline in all direction should be investigated. However, if the location is estimated by the interval of pipeline, then only that interval is needed to be observed. Hence, we need to estimate the specific interval where the leakage occurs. Considering the condition of the equipment, we can set the length of the excavation, L_{exc}.

$$L_{exc} = k_{exc} \times L_{element}, \ (k_{exc} \geq 1) \quad (10)$$

Because there is a possibility to replace the buried pipeline element, k should be equal to or greater than 1. After setting L_{exc}, we can find an interval, U, which contains the most estimated points we obtained before and the length of U is L_{exc}.
For $s_i (i = 1, 2, \cdots, I)$ which are included in U, it is possible to define the mean (m_U) and the variance (σ_U^2) by introducing the origin point of the interval U. The origin point is selected as the end of the selected interval, U.

$$\begin{aligned} d'_i \equiv \text{the distance between } S_i \text{ and the origin point} \\ m_U = E(d'_i) \\ \sigma_U^2 = V(d'_i) \end{aligned} \quad (11)$$

Now, we can finally have the estimated interval where the leakage occurs with the following confidence.

$$U_{final} = [m_U - Z \times \frac{\sigma'_U}{\sqrt{I-1}}, m_U + Z \times \frac{\sigma'_U}{\sqrt{I-1}}] \quad (12)$$

Here, Z is the confidence coefficient from the normal distribution. For example, if we estimate the leakage point with a 95% confidence, Z=1.96.

3. RESULTS AND DISCUSSION

To verify efficacy of the prorposed algorithm, two field tests were conducted in different water distribution networks of two cities in South Korea, Yeongwol and Yangsan. For the comparison, CUSUM test(Misiunas (2005)) and MWT(Srirangarajan et al. (2013)), were applied to the same data.

Case I. Yeongwol The distribution network in Yeongwol has six sensors and the total length of pipelines is 7266.4356 m with 861 number of nodes.
For the case I, the parameters for false alarm test and leakage detection test are ϕ=20 s and N=10 which is selected based on 1 h normal state data. The CUSUM test was used with parameter values, $\lambda = 0.7, \nu = 0.01kPa$, and $h = 8kPa$. The wave speed for the location algorithm of MWA is 1200 m/s.

[False Alarm Test]

This test uses another 1 h normal state pressure data with the moving window size of 400 s and the update period of 200 s.

In Table 1, CUSUM test shows the number of false alarms. Because CUSUM test needs a threshold value, h, it is not suitable to the networks with unknown disturbances such as water usage. Figure 1 shows the first 3 window CUSUM results with T_u=200 s and W = 400 s for the sensor 1.

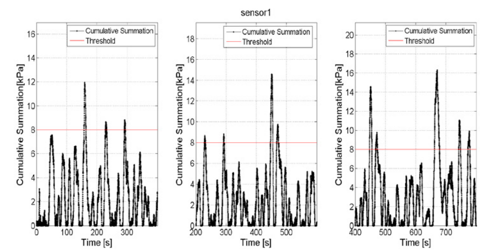

Fig. 1. CUSUM test for the normal state data in Yeongwol ($\lambda = 0.7, \nu = 0.01kPa, h = 8kPa$)

The MWA was applied and the number of the false alarms is 53 in six sensors during 1 h, which is written in Table 1. Because the normal state data oscillates with the disturbances, the false alarms are detected in normal situations. Figure 2 shows the first 3 window MWA results with T_u=200 s and W = 400 s for the sensor 1.

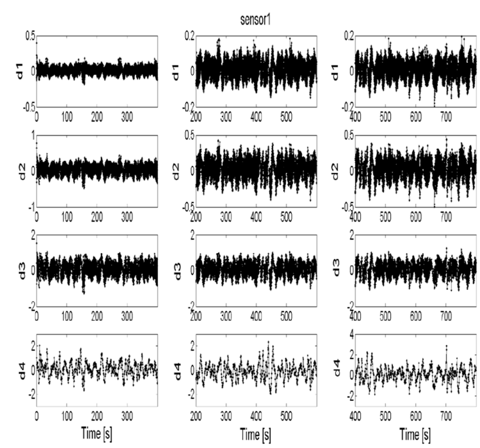

Fig. 2. MWA for the normal state in Yeongwol

There is a small number of false alarms in normal state situation when the proposed algorithm is applied because the sudden and continuous decreasing tendency is not frequently found in the normal state data. And the first 3 window results with T_u=200 s and W = 400 s for the sensor 1 are shown in Figure 3

Table 1. False alarm test of 6 sensors for 1 h in Yeongwol

	CUSUM	MWA	Proposed Algorithm
False Alarm	47	53	9

[Leakage Detection Test]

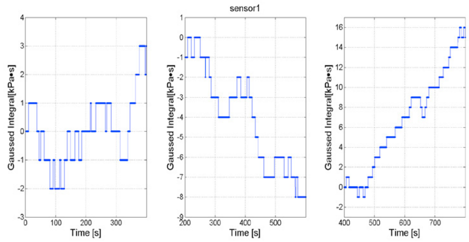

Fig. 3. Proposed method for the normal state in Yeongwol

This test uses the 400 s leakage data which was constructed from the real burst data in Yeongwol by (1) with the scaling factor of 0.1. The leakage occured in 190 s - 210 s.

CUSUM test cannot estimate the location of the leakage because two sensors have missed alarms and other two sensors have false alarms. The missed alarm means that there is no alarm when leakage occurs and the false alarm means that the alarm occurs long before the leakage occurs.

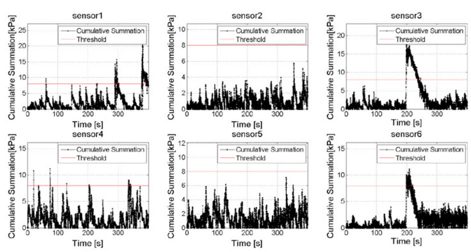

Fig. 4. CUSUM test for the leakage in Yeongwol ($\lambda = 0.7$, drift=0.01, h=8)

When the MWA is applied, the location of the leakage point was obtained with the error of 110.1695 m. Although there are many false alarms in the normal state data, the MWA method can detect the sudden cusp point resulting in the relatively small error in the location estimation while the CUSUM method cannot find the leakage location. Figure 5 shows the first window with W = 400 s for the 6 different sensors.

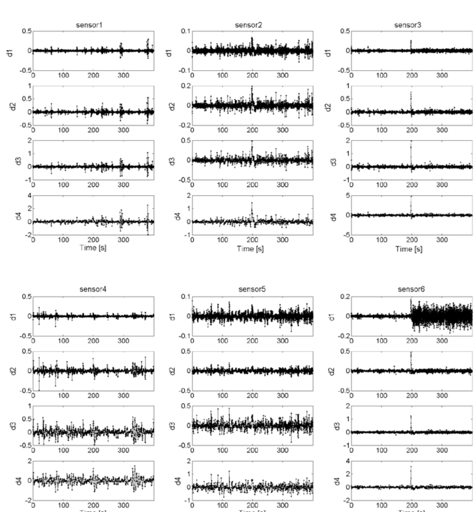

Fig. 5. MWA for the leakage in Yeongwol

Using the proposed algorithm, the average point of the leakage points and the interval around that point was estimated. The result is shown in Figure 6. The length between the average point and real leakage point is 54.7365 m and the leakage point is included in the estimated interval. The summarized result for the case I leakage

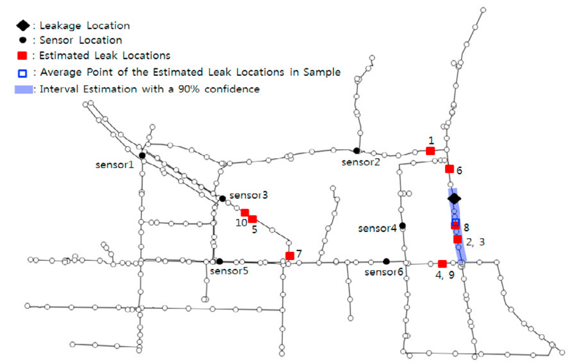

Fig. 6. The estimated interval and average point for the leakage in Yeongwol

detection test is shown in Table 2.

Table 2. Leakage detection test in Yeongwol

	CUSUM	MWA	Proposed Algorithm
False Alarm	2	1	0
Missed Alarm	2	1	1
Error[m]	N/A	110.1695	54.7365

Case II. Yangsan The pressure data from Yangsan was used to compare the three different algorithms. Three sensors are located in the network and the total length of pipelines is 10167.9709 m with 1154 number of nodes. For the case II, the parameters for false alarm test and leakage detection test are ϕ=24 s and N=10. CUSUM test was used with parameter values, $\lambda = 0.7, \nu = 0.01kPa$, and $h = 10kPa$. The wave speed for location algorithm in MWA is 1200m/s.

[False Alarm Test]

Using CUSUM Test, many false alarms were detected in the normal state.
The MWA method shows 16 false alarms in 3 sensors during 0.5 h. It shows the largest number of the false alarms because there exist disturbances.
The proposed algorithm has a small number of false alarms in normal state situation as with the case I. The results are summarized in Table 3 and the first 3 window results with T_u=200 s and W = 400 s for the sensor 1 are shown in Figure 7.

Table 3. False alarm test of 3 sensors for 0.5 h in Yangsan

	CUSUM	MWA	Proposed Algorithm
False Alarm	11	16	1

[Leakage Detection Test]

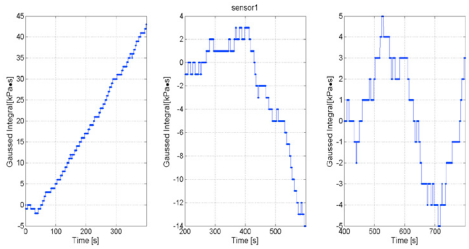

Fig. 7. Proposed method for the normal state in Yangsan

This test uses the 400 s leakage data which is made from real burst data in Yangsan by (1) with scaling factor 0.5. The leakage occurs about 190 s - 210 s.

CUSUM test cannot estimate the location of the leakage because two sensors alarm long before the leakage occurs The location of the leakage point is obtained with the error of 870.2474 m when the MWA is applied.

Using the proposed algorithm, the leakage point is estimated. The result is shown in Figure 8. The length between the estimated point and real leak point is 161.2870 m. Although the interval estimation is not possible because of the number of the sensors which is less than 4, the error is relatively small compared to that of MWA and CUSUM cannot find the leak point. The summarized result for the case II leakage detection test is shown in Table 4.

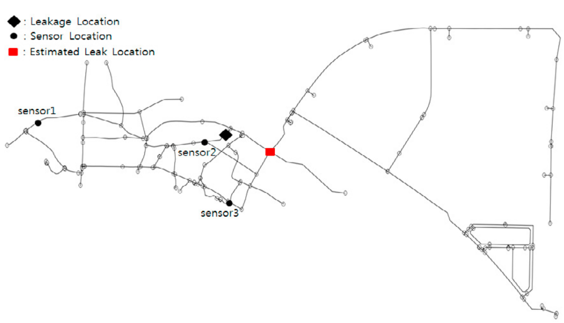

Fig. 8. The estimated location for the leakage in Yangsan

Table 4. Leakage detection test in Yangsan

	CUSUM	MWA	Proposed Algorithm
False Alarm	2	0	0
Missed Alarm	0	1	0
Error[m]	N/A	870.2474	161.2870

4. CONCLUSION

This paper proposes a robust method using cumulative integral, floor function, radius of curvature, and statistical estimation for detecting the leakage and estimating the location of the leakage in water distribution network. The characteristic of the leakage is amplified by using cumulative integral while the small disturbance is eliminated. The floor function make the detection of pressure decrease maintenance possible and the exact detection of the leakage occurrence time can be obtained by observing the curvature value. Finally, the statistical estimation based on the confidence interval makes it possible to estimate the location of the leakage as a interval with a specific confidence. The proposed method shows a better performance compared with two previous methods: CUSUM test and MWA. It makes less false alarms in normal state and detects the leakage with high precision. It is also applicable even if a small amount of the leakage occurs in noisy circumstances.

ACKNOWLEDGEMENTS

This work is supported by the Korea Ministry of Environment as a Project for Developing Eco-Innovation Technologies (GT-11-G-02-001-3). This work was supported by Ministry of Trade, Industry and Energy (MOTIE, Korea) as "Industrial Core Technology Development Program (10038671)".

REFERENCES

Covas, D. and Ramos, H. (2010). Case Studies of Leak Detection and Location in Water Pipe Systems by Inverse Transient Analysis. *Journal of Water Resources Planning and Management*, 136(2), 248-257.

Jung, D. and Lansey, K. (2014). Burst Detection in Water Distribution System Using the Extended Kalman Filter. *Procedia Engineering*, 70, 902 - 906. 12th International Conference on Computing and Control for the Water Industry.

Misiunas, D. (2005). *Failure Monitoring and Asset Condition Assessment in Water Supply Systems.* Lund University, Department of Industrial Electrical Engineering and Automation.

Mulholland, M., Purdon, A., Latifi, M., Brouckaert, C., and Buckley, C. (2014). Leak Identification in a Water Distribution Network using Sparse Flow Measurements. *Computers and Chemical Engineering*, 66, 252-258.

Perez, R., Sanz, G., Puig, V., Quevedo, J., Escofet, M.C., Nejjari, F., Meseguer, J., Cembrano, G., Tur, J.M., and 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]. *Control Systems, IEEE*, 34(4), 24-36.

Ponce, M., Castanon, L., and Cayuela, V. (2014). Model-based Leak Detection and Location in Water Distribution Networks Considering an Extended-horizon Analysis of Pressure Sensitivities. *Journal of Hydroinformatics*, 16(3), 649-670.

Ragot, J. and Maquin, D. (2006). Fault Measurement Detection in an Urban Water Supply Network. *Journal of Process Control*, 16(9), 887 - 902.

Srirangarajan, S., Allen, M., Preis, A., Iqbal, M., Lim, H., and Whittle, A. (2013). Wavelet-based Burst Event Detection and Localization in Water Distribution Systems. *Journal of Signal Processing Systems*, 72(1), 1-16.

Wu, Z., Sage, P., and Turtle, D. (2010). Pressure-Dependent Leak Detection Model and Its Application to a District Water System. *Journal of Water Resources Planning and Management*, 136(1), 116-128.

Rashid, S., Qaisar, S., Saeed, H., and Felemban, E. (2014). A Method for Distributed Pipeline Burst and Leakage Detection in Wireless Sensor Networks Using Transform Analysis *International Journal of Distributed Sensor Networks*, 2014, 1-14.