Optimal layout of pressure monitoring points in water supply network based on Optics

ChenLei Xie[1,*], QianSheng Fang[1], HongYan Zhang[1], JiXin Zhang[1]

[1]Anhui Province Key Laboratory of Intelligent Building and Building Energy Saving, Anhui JianZhu University, Hefei, China

Abstract. It is necessary to monitor the pressure in the networks in real time, when we face the problem of pipe burst and leakage in urban water supply network. Therefore, it is particularly important to arrange pressure monitoring points in appropriate places in the pipeline networks. The pressure monitoring point layout is often based on similar degree of the node pressure data in the current stage. A method of optimal pressure monitoring point location in the urban water distribution networks was proposed in this paper. Since the above method did not consider spatial properties of network node. The original feature matrix data is constructed by acquiring the spatial attributes of the pressure monitoring nodes and calculating the non-spatial attributes of the nodes. The original feature matrix data will be normalized. Then the Optics clustering algorithm is used to cluster the normalized node feature matrix data to determine the location and number of pressure monitoring points in monitoring area of urban water distribution networks. Experimental results show that the method effectively ensures that the pressure monitoring points can grasp the pressure information of the whole water supply network more comprehensively and rationally, improves the economy of the pressure monitoring points layout, and provides good guidance for the actual layout of pressure monitoring points in municipal water distribution networks.

1 Introduction

The pressure in urban water distribution networks directly reflects the distribution of water pressure in urban water distribution networks, and also reflects the possible problems of pipe burst and leakage, which is of great significance to control the leakage of pipeline network and can reduce unnecessary losses[1]. Therefore, the optimal layout of pressure monitoring points plays a vital role in the stable operation of the pipeline network. However, there are many difficulties in the layout of sensors. It will cause economic waste if the number of pressure monitoring points is too large. It may not be able to fully reflect the real-time change of pressure in the pipeline network while the number is too small. How to arrange pressure monitoring points in order to reflect the change of pressure in pipeline network efficiently and economically is particularly important. For the problem of optimal layout of pressure monitoring points, some existing methods of optimal layout of pressure monitoring points in pipeline network include sensitivity analysis[2], genetic algorithm[3-4], ant colony algorithm[5], particle swarm optimization[6], clustering algorithm[7-8], fault diagnosis[9] and so on. However, those methods usually arrange the pressure monitoring points according to the similarity of the pressure values of each node, without considering the spatial attributes of the network nodes, so they can only simply explain the similarity of the pressure changes between the nodes. After clustering, nodes scattered in different geographic locations of

pipeline network may be classified into the same category, so the clustering results have little practical significance. It does not show the similarity between adjacent geographic nodes in the pipeline network, and it is difficult to effectively reflect the regional distribution of pressure in the pipeline network. It is not very instructive for the optimal layout of pressure monitoring points in the actual pipeline network.

The method proposed in this paper is to construct a node feature matrix containing spatial and non-spatial attributes of nodes, and to find out which nodes have similarities in spatial and non-spatial attributes through clustering algorithm. The Optics (Ordering Points to identify the clustering structure) algorithm used in this paper is a density-based clustering algorithm. Compared with the existing K-MEANS clustering method, the Optics algorithm does not need to set the initial number of clusters. The shape of clusters can be arbitrary, and the parameters of filtering noise can be input when needed. Compared with using DBSCAN algorithm for clustering, DBSCAN algorithm has two initial parameters, E(Neighbourhood Radius) and MinPts(Neighbourhood Minimum Points)[10].Users need to input them manually, and the clustering results are very sensitive to the values of these two parameters. Different parameters will produce different clustering results. Optics algorithm does not show the result-producing cluster, but generates an augmented cluster ranking for clustering analysis, which represents the density-based clustering structure of each sample point.

---

*Corresponding author: xcl@ahjzu.edu.cn

This augmented cluster ranking is about reachable distance and ranking of nodes. Based on the reachable distance and the ranking of nodes, the appropriate clustering results are selected. The procedure of the proposed method can be described as follows: The spatial and non-spatial attributes of each node are collected to construct the original node feature matrix. After normalization, the original node feature matrix is input into Optics clustering algorithm to get the clustering results. According to the clustering results and the influence degree of nodes, the location and number of pressure monitoring points in the monitoring area of water distribution networks are finally determined. The experimental results show that compared with clustering analysis based on the similarity of node pressure changes, the method solves the problem that other methods do not consider the spatial attributes of node in the layout of pressure monitoring points in water distribution networks.

2 The Introduction of Optics algorithm

Optics algorithm is a density-based clustering algorithm. It is an improved DBSCAN algorithm. It improves the problem that the clustering result of DBSCAN algorithm is too sensitive to initialization parameters (E-neighbourhood radius and MinPts (E-neighbourhood minimum point), which leads to different clustering results due to the disadvantage that different parameter values. The procedure of Optics algorithm can be described as follows:

Input： Sample Set U, Minimum domain points of core objects for each node, MinPts.

Output： Sorting of nodes and reachable distance of each nodes.

Step1 Two queues are created the *ordered* queue and *result* queue. The *ordered* queue are used to store the core node object and its direct reachable node object, and are arranged in ascending order of reachable distance. The *result* queue is used to store the output order of the node;

Step 2 Judging whether all the nodes in the node feature matrix data set U have been processed, and if so, jump to *Step 8*. Otherwise, select a node that is not processed, that is, not in the *result* queue and is the object of the core node, and find all the direct density reachable nodes. If the core node does not exist in the *result* queue, the core node will be selected. And all its direct density reachable nodes are placed in the *ordered* queue and sorted by reachable distance;

Step 3 Judging whether the *ordered* queue is empty, if so, jump to *Step 2*. Otherwise, take the first node object with the minimum distance out from the *ordered* queue to expand. If the extracted node does not exist in the *result* queue, it will be saved in the *result* queue;

Step4 Judging whether the first node is a core node object, if not, jump to *Step 3*, or find all the direct density reachable nodes of the first node;

Step 5 Judging whether the direct density reachable node already exists in the *result* queue, if so, do not process and jump to *Step 3*, otherwise proceed to the next step;

Step 6 If the direct density reachable node already exists in the *ordered* queue, if the new reachable distance is less than the old reachable distance, the new reachable distance is used to replace the old reachable distance, and the *ordered* queue is reordered and jumped to *Step 3*;

Step 7 If there is no direct density reachable node in the *ordered* queue, the node is inserted and the *ordered* queue is reordered and jumped to *Step 3*;

Step 8 The algorithm ends and outputs are the *ordered* queue and the *result* queue.

3 The research on the layout of the pressure monitoring points in water distribution networks

3.1 Process of Pressure Monitoring Point Layout

If only the pressure data (non-spatial attributes) are used to analyze the operation state of pipe network, it often can't reflect the distribution of the pipeline network pressure in the space.In order to fully reflect the correlation between the nodes in the pipeline network, this paper collects the spatial and non-spatial attributes information of each node as the parameters to describe the features of the nodes, which can not only completely describe the similarities and differences of nodes in non-spatial attributes, but also completely describe the similarities and differences of nodes in spatial attributes.

The paper determines the position and quantity of pressure monitoring points in the monitoring area of urban water distribution networks as shown in Figure 1, which is divided into the following five steps:

Step 1: The pressure data of all nodes (N nodes) in the water distribution networks under normal conditions are collected, as well as the pressure data of all nodes in the network when the water consumption of each node changes.

Step 2: The spatial attributes of each pressure monitoring node in the pipeline network (namely geographic coordinate data) are acquired, and then the non-spatial attributes of each node are calculated in turn, including the pressure influence degree of each node and the variance of the pressure value of each node. The above data constitute the original feature matrix data of each node in the pipeline network.The original feature matrix data format of each node is $U=\{(X_i, Y_i, EF_i, D_i)$, i=1, 2, 3... N$\}$, X_i and Y_i represent the coordinates of each node, EF_i is the influence degree, D_i represents the variance of each node's pressure value respectively.

Step 3: Normalization of the original feature matrix.Then the original feature matrix data of each node are normalized.

Step 4: Optics clustering algorithm is used to cluster and analyze the feature matrix data of nodes, and the ordered set of pressure monitoring nodes and the reachable distance of each node are obtained.

Step 5:Based on the orderly set of pressure monitoring nodes and the reachable distance of each node, combined with the influence degree of each node, the location and number of pressure monitoring points in the monitoring area of urban water distribution networks are finally determined.
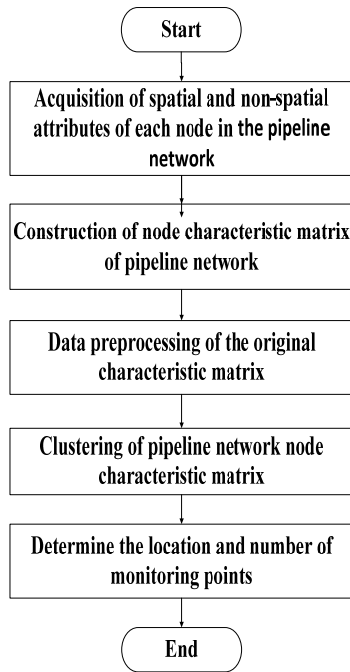

Fig. 1. Implementation procedure of monitoring point layout

3.2 Construction of Node Feature Matrix

In order to fully reflect the correlation between nodes, this paper adopts the spatial clustering method which integrates spatial location and attribute features. The spatial and non-spatial attributes of nodes are combined to describe the differences between the spatial and non-spatial attributes of nodes. In this paper, the X_i and Y_i coordinates of each node are selected as the spatial attributes of the node, *Position=$\{(X_i, Y_i)$, i=1, 2..., N$\}$* and the *EF_i* of hydraulics and the variance *D_i* of the pressure values of each node in statistics are used as the non-spatial attributes of the node. The above data constitute the original feature matrix data *U= (X_i, Y_i, EF_i, D_i)* of each node in the pipeline network.

Pressure data of each node in the pipeline network under normal conditions collected by sensors (P):

P=[P_{11} P_{12}...P_{1N}
  P_{21} P_{22}...P_{2N}
 .....
  P_{M1} P_{M2}...P_{MN}]

(N: Number of nodes in pipeline network, M: Number of data set, Time series collected)

And the pressure data of all nodes in the pipe network when the water consumption of each node changes(P_{b1}, P_{b2},..., P_{bN}):

P_{bt}=[P_{bt11} P_{bt12}...P_{bt1N}
   P_{bt21} P_{bt22}...P_{bt2N}
   ......

P_{btM1} P_{btM2}...P_{btMN}] (t=1,2...,n)

The influence degree of nodes is the sum of the effects of the changes of water consumption on the waterhead of each node in the whole pipeline network. The calculation formula is as follows:

$$EF_i = \frac{\sum_{q=1}^{M} \sqrt{\frac{\sum_{t=1}^{N}(P_{qi} - P_{btqi})^2}{N}}}{M} \quad (i=1,2...,N)$$

Among them, M is the number of time series, namely the total number of samples, N is the number of pressure sensors, P_{qi} is the pressure data collected by the sensors under normal conditions, and P_{btqi} is the pressure data collected by the sensors when the water consumption of each node changes.

The pressure data of each node of the monitoring area in pipeline network under normal conditions collected by the water supply experimental platform are used to calculate the variance data set of each node pressure value one by one($D= \{(D_i)$,i=1,2...,N$\}$), The specific formula is as follows:

$$D_i = \frac{\sum_{q=1}^{M}(P_{qi} - \bar{P}_i)^2}{M} \quad (i=1,2...,n)$$

Among them, M is the number of time series, that is, the total number of data sets collected, and P_{qi} is the pressure data collected by the sensors in the normal operation of the pipeline network.

4 Experiment and Analysis

4.1 Introduction of experimental platform

The paper chooses the looped network of water supply experimental platform (Figure 2) of Anhui Province Key Laboratory of Intelligent Building and Building Energy Saving as the research object.The system covers an area of 70 m^2, the length of the pipe section is 120 m, and the diameter of the pipe varies from DN30 to DN50. There are 19 water pressure sensors. The topology of the experimental platform is shown in Figure 3. The green triangle No.1-No.19 represents the location of the sensors, and the blue circle No.1, No.2 represents the water access port for residents.

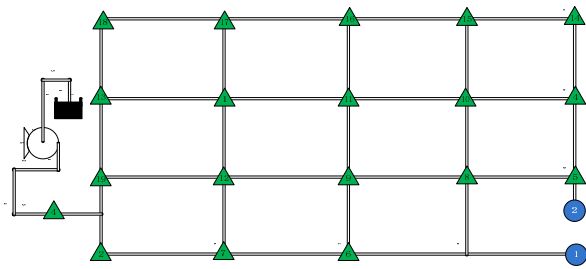**Fig. 2.**Water Supply Experimental Platform
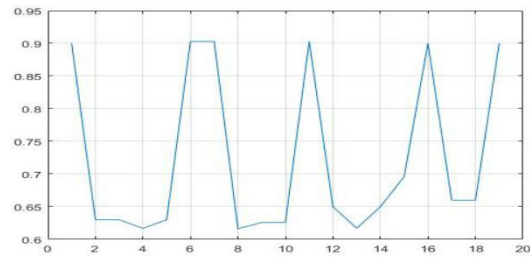
Fig. 3. Topology of Water Distribution Networks Model

4.2 Creating the Node feature Matrix

The pressure sensor installed on the water supply experimental platform is used to acquire the normal pressure value data and the change pressure data of all nodes. The pressure data were collected continuously for 168 hours, and the time interval of data acquisition was 15 minutes. According to thenode feature matrix formula of 3.2, the node feature matrix is solved. Firstly, the position coordinates ($Position=\{(X_i,Y_i),i=1,2...,N\}$) of nodes were obtained from the water supply experimental platform; and then the influence degree($EF=\{(EF_i),i=1,2...,N\}$)of each node and the variance data set ($D=\{(D_i),i=1,2...,N\}$) of individual node pressure values are calculated one by one according to the formula based on the normal node pressure.

Through the calculation of the above steps, the influence degree and variance value of 19 monitoring points are obtained, and the feature attributes of the nodes are constructed, as shown in Table 1.

Table 1. The original feature attributes of each node.

Node	X Coordinates	Y Coordinates	Influence Degree	Standard Deviation
1	44.65	44.48	1.759949	0.447189
2	0	0r	2.313524.	0.443937
3	-44.65	11.12	0.17639627	0.42923
4	-44.65	11.12	0.176396	0.42923
...	...	...	...	...
19	0	22.24	1.603233	0.404891

4.3 Experimental procedure

After the original feature matrix data of each node was created by the method proposed in 4.2, the data was normalized by Z-score method, and clustered by Optics algorithm, the obtained result graph is shown in Figure4.

Fig. 4. The reachable distance graph of each node obtained by Optics algorithm

Table 2. Clustering results of pressure monitoring points

Categories	Node numbers
1	1 4 10 11 13
2	18 17 16
3	5 8 9
4	2 6

The results of Table 2 are reproduced on the topological map. As shown in Figure 5, red, yellow, violet and blue represent the same type of monitoring points respectively.
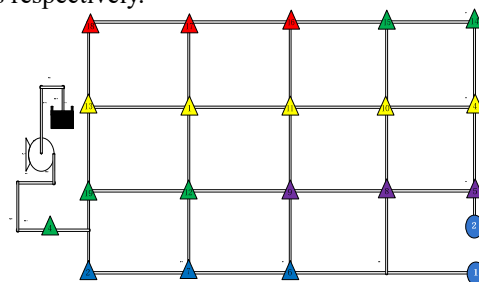

Fig. 5. Clustering result by adding spatial attributes

In order to compare the superiority and rationality of the proposed method with respect to the layout of pressure monitoring points in the water distribution networks with respect to the non-spatial attributes algorithm. In this paper, the original feature matrix which does not contain the spatial attributes of each node is constructed. The format of original feature matrix which does not contain the spatial attributes of each node is like $U=(EF_i, D_i)$. The original feature matrix contains only two attributes of influence degree and standard deviation. The original feature matrix $U=(EF_i, D_i)$ is normalized and clustered by Optics. The results are shown in Table 3 and Figure 6. The results obtained by using the same clustering method are compared with the results obtained by the method proposed in this paper.

Table 3. Clustering results without adding spatial attributes.

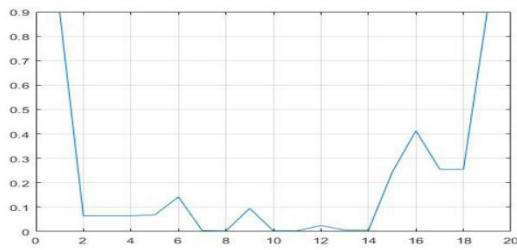| Categories	Node numbers
1	4, 5, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18
2	2, 6

Fig. 6. Reachable distance of nodes without adding spatial attributes.

The clustering result which original feature matrix is $U= (EF_i, D_i)$, is reproduced on the topology map, as shown in Figure 7.
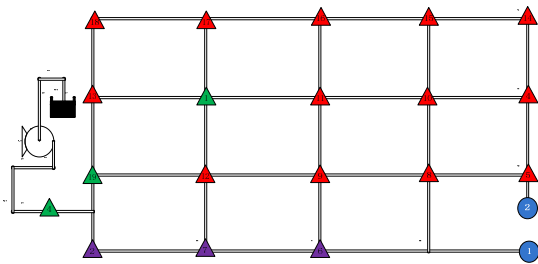

Fig. 7. Clustering result graph of without adding spatial attributes.

According to the principle of selecting the node with the most influence as pressure monitoring point in the same category as pressure monitoring point. It can be seen from Figure 7 that the nodes belonging to the same class after clustering are too congregate to reflect pressure status of the entire pipe network, and the clustering results only reflect the approximation degree of pressure changes of each node.

4.4 Creating the Node Feature Matrix

By comparing the clustering results of adding node spatial attributes to the node feature matrix and without adding node spatial attributes to the node feature matrix, we can find that the clustering result of adding a node spatial attributes is more accurate and reasonable. In this paper, the following experimental results are obtained: the results obtained by the pressure monitoring point layout method of water distribution networks proposed in the paper are better. Finally, the 19 nodes in the water distribution networks will be clustered into 4 categories. The experimental results are shown in Table 4.

Table 4. Clustering results by adding spatial attributes

Categories	Node position	Node position after optimization
1	14 10 11 13	11
2	18 17 16	16
3	5 8 9	8
4	2 6	2

How to select a representative pressure monitoring point based on the clustering result, we must consider whether the arranged pressure monitoring point can truly reflect the running state of the entire water supply network. Because if the degree of influence of the node is higher, it indicates that the pressure change of the node can more accurately represent the pressure change of the area. As shown in Table 3, all the nodes were copolymerized into 4 categoris. According to the principle of selecting the node with the most influence as the pressure monitoring point, the final pressure monitoring point is 2, 8, 11, and 16. As shown in Figure 8, the red, yellow, purple, and blue triangles represent the optimized pressure layout points.
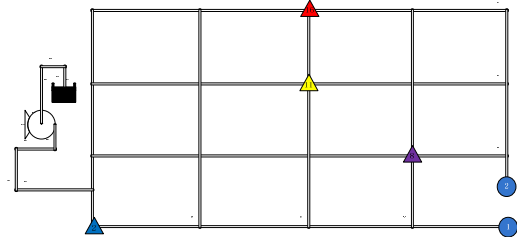

Fig. 8. Monitoring point layout map by adding space attributes

The experimental results obtained by the method proposed in this paper are shown in Figure 7. The nodes belonging to the same class after clustering constitute the node set of the region. The selected pressure monitoring points are evenly distributed in different areas of the pipe network, reflecting the pressure distribution of the entire water supply network. As shown in Figure 5, the nodes belonging to the same class after clustering are so congregate that can not be selected as the pressure monitoring pointwhen the spatial attributes is not added.

In summary, the following conclusions are drawn: The water distribution networks pressure monitoring point layout method proposed in this paper adds the node spatial attributes,the selected pressure monitoring point has accurately representative, which can monitor the pressure status of the entire pipe network in real time. This method can be applied to actual engineering.

5 Conclusions and Prospects

In order to solve the problem of optimal layout of pressure monitoring points in water supply networks, this paper proposes an optimal layout method for pressure monitoring points of water supply networks.
Firstly, the spatial and non-spatial attributes of each pressure monitoring point are obtained through measurement and acquisition, and then the original feature matrix data of each node in the pipeline network are constructed by using the above attributes. Secondly, the feature matrix data of nodes are obtained by normalization processing. Thirdly, the Optics algorithm is used to realize the cluster analysis of the node feature matrix data, then the reachable distance and node sequence of the pressure monitoring node are obtained. Based on the clustering results and influence degree of each node, the location and number of pressure monitoring points in the monitoring area of municipal water supply network were finally determined. The

optimized sensor layout by this method can effectively reflect the regional distribution of pipe network pressure, and effectively solve the problem of optimal layout of pressure monitoring points of municipal water supply network.

Acknowledgments

References

1. ZY Wu, P Sage, D Turtle. Pressure-dependent leak detection model and its application to a district water system.Journal of Water Resources Planning & Management , 2010 , 136 (1) :116-128

2. Doosun Kang, Kevin Lansey. Novel Approach to Detecting Pipe Bursts in Water Distribution Networks. Journal of Water Resources Planning and Management, 2014, p(33):121-127

3. Liu ShuMing, Wang HuanHuan, Xu Peng et al. Optimal location of monitoring points for large-scale multi-purpose water supply network. Journal of Tsinghua University (NATURAL SCIENCE EDITION),2013,01:78-83.(in Chinese)

4. Casillas MV, Puig V et al. Optimal sensor placement for leak location in water distribution networks using genetic algorithms. Sensors (Basel), 2015, 13 (11) :14984-15005

5. XU Gang，ZHANG Tu-qiao，LV Mou. Optimized Location of Monitoring Points for Water Distribution System Based on Sensitivity Analysis and Ant Colony Algorithm[J]. China Water & Wastewater,2007,11:94-96.(in Chinese)

6. CN Nyirenda, P Makwara, L Shitumbapo.Particle Swarm Optimization Based Placement of Data Acquisition Points in a Smart Water Metering Network. Sai Intelligent Systems Conference, 2016 :905-916

7. Lu H , Shuping L , Weiwei Z , et al. Study on optimized layout of pressure monitoring points in water supply network based on existing pressure monitoring points[J]. Water & Wastewater Engineering, 2013, 39(2):119-122.(in Chinese)

8. Jing-Qing L , Dong-Jin G , Ping Y E . Improved K average spatial clustering method for nodes of water distribution system[J]. Journal of Zhejiang University, 2015.(in Chinese)

9. Dong L , Huifeng X , Wenyu Z , et al. Optimal layout method of water pressure monitoring points for water supply network based on fault diagnosis[J]. Journal of Civil Architectural & Environmental Engineering, 2018.(in Chinese)

10. Zhiping Z , Jiefeng W , Shuwei Z , et al. An improved adaptive and fast AF-DBSCAN clustering algorithm[J]. Caai Transactions on Intelligent Systems, 2016.(in Chinese)