Drink. Water Eng. Sci. Discuss.,
https://doi.org/10.5194/dwes-2018-19-AC2, 2018

Open Access
Drinking Water
Engineering and Science
Discussions

Interactive comment on "Hybridisation of brownboost classifier and glowworm swarm based optimal sensor placement for water leakage detection" *by* Rejeesh Rayaroth and Sivaradje Gopalakrishnan

Rejeesh Rayaroth and Sivaradje Gopalakrishnan

rejeeshrayaroth@pec.edu

Received and published: 19 December 2018

General Comments:

1. The results are recommending to put 50 to 500 sensors in a network of 17 km (up to a sensor every 34 meter) are without doubt unfeasible for applying the methodology on real-world systems.

EBBC-GWO method used Barcelona water distribution network. From these source 4645 km of pipeline is considered. It consists of 883 nodes, 927 pipes which distributes

water to 639 consumers. [Addressed in section 3]

Specific Comments:

Introduction And Related Work

2. Page 1 Line 28: The sentence "leaks in WDN are detected through machine learning techniques" is not true, these methods are just one way to detect leaks, in fact, used only in scientific literature and not in practice at all.

In our work, water leakage detection is performed using machine learning techniques. Therefore, 'leaks in WDN are detected through machine learning techniques' is included in introduction. As per your requirement, the introduction is shortened. Therefore, the above said sentence is ignored in introduction.

3. In general, the listed literature is a mix of different methods for leak detection, which are even focusing on different physical effects caused by a leak (e.g. flow, noise, pressure without even mentioning it here) as well as a wild mix between leak localisation, leak detection and optimal sensor placement methods. Therefore, they can't be compared and listed in the way they are presented in introduction and the related works section .

Corrections are addressed in section 1

4. The literature review is insufficient: Only very recent literature is listed . The oldest publication is from 2015. Older key literature as well as novel key literature for this topic is missing. In my opinion, the state of the art is not well described. I recommend the authors to review the literature in the papers that the mention and identify the key papers for this topic.

The state of the art methods are well defined in section 1. older papers are added in section 1.

5. Page 3 Line 8: Throughout the paper it is mentioned that the method allows to detect

leaks with higher accuracy, but it is not clear what is meant. Higher than what?

Higher accuracy means higher classification accuracy. The proposed method classifies the data as normal or abnormal pressure data with classification accuracy (i.e., 20% higher than existing methods).

6. Page 3 Line 23: Related works: Is the work listed in the introduction not also related work? What is the difference of this section to the previous one?

The work listed in the introduction also the related work. The work listed in the introduction is more related to our work. Therefore, they are listed in the introduction and the remaining paper is listed in the related work section.

As per the comment from anonymous referee #1 the introduction and related work sections are combined.

7. -Page 3 Line 13: The paper states the difference between normal data and abnormal data, but it is not clear throughout the paper what this terms actually means and how data is identified as abnormal.

The pressure flow is normal (i.e., maintained by a constant pressure) in the collection pressure data it is classified or identified as normal pressure data. Whereas, the pressure flow is abnormal (not a constant value) in the collection of pressure data is identified as abnormal pressure data. The abnormal pressure occurs when the pipe break, water leakage or fault in the pipeline is identified.

8. Page 3 Line 34: “Optimisation was not carried out in effective manner” is missing an explanation why it is not effective. The same can be found on Page 4 Line 16

But, the optimization was not carried out in effective manner because it failed to consider the real-time operating pressure and flow data in designed approach. [it is addressed in page 3 line 5-6].

The same can be found on Page 4 Line 16

But, the performance of leak identification was not carried out in effective manner because it varied based on the timing and duration of the measurement.[it is addressed in page 3 line 22-23].

9. Page 4 Line 17: What is the difference in this context between optimal sensor placement and sampling design?

Sampling Design

Sampling design (SD) used to perform localization and quantification of pressure sensors in WDS for leak detection. SD was derived based on the criteria of maximization of total leak sensitivity, sensitivity consistence, minimization of information redundancy and sensors number criteria. However, the classification time was not reduced by using SD method.

Optimal Sensor Placement

Optimal sensor placement is also used to perform localization for leak detection. Here, sensor placement problem was formulated as an integer optimization problem. The optimization criterion was based on minimizing the number of non-isolable leaks according to the isolability criteria. However, sensor placement approach was not effective in detecting the water leak Methodology

10. It is not clear if the paper deals with leak detection or leak localisation, since the literature review deals one time with detection and then switches to localisation and vice versa. For example, sentences in the methodology section like on Page 4 Line 29:“in order to detect the leak location” are confusing

As per your requirement, in order to detect the leak location sentence is changed as In order to perform the leak detection, EBBC-GWO Method is introduced. [it is addressed in page 4 line 17]. The objective of the research work is to perform optimal sensor placement for water leakage detection. Therefore, the literature review deals optimal sensor placement/localization and leak detection concepts.

11. Page 5 Line 2: What is a water distributer system?

Typing error water distribution system [correction addressed in page 4 line 28]

12. Page 5 Line 1-5: Why do the authors introduce a graph description of a water distribution system if it is not used later on? Additionally, a reference is missing to previous literature of how to describe a water distribution system as a mathematical graph. For sure, this is not the idea of the authors.

A water distribution system comprises the set of graph branches (i.e.,G=(V,E)) where E denotes set of edges (i.e., pipes) and V denotes the set of vertices (i.e., nodes) explaining about the pipe connections and endings in (R. Sarrate et al. 2014). A water distribution system is a collection of pipes (i.e., links) connected to the nodes (junctions, tanks and reservoirs). The water flows are computed through hydraulic balancing, and through solving the equations at every node and links. EPANET software works out the network hydraulic equations automatically (Rossman 2000). By using the graph model, EPANET software obtains the pressure and flow rate of every node and links from hydraulic simulation. [Corrections is addressed in page 5]

13. Page 5 Line 5: Reference to EPANET is missing (Rossman 2000)

Reference included in page 4 line 32

14. Page 5 Line 8: How are leaks simulated in this paper with EPANET? Why did the authors use extended-period simulations, it seems there is not need for this.

EPANET software tool is simulated at without leak condition and pressure data readings are collected. During the simulation, it detects the flow of water in pipe, the pressure at each node and height of the water in container

15. An important parameter is the leak's size, because this parameter has an effect on the size of the pressure drop and hence the detectability, but the leak size is not mentioned throughout the paper at all. In fact, while reading the paper, it is not clear if there are any leak simulations performed at all. If that is the case, the whole method

proposed by the authors is very questionable, because the definition of normal and abnormal pressures does not make sense at all. Can the authors please clarify this point, because it is crucial for the whole publication?

Leak's size parameter is included in section 4.5. The size of the leakage was 142meter, 0.99meter, 0.99meter, 12.31meter. [addressed in section 3]

The pressure flow is normal (i.e., maintained by a constant pressure) in the collection pressure data it is classified or identified as normal pressure data. Whereas, the pressure flow is abnormal (not a constant value) in the collection of pressure data is identified as abnormal pressure data. The abnormal pressure occurs when the pipe break, water leakage or fault in the pipeline is identified. [Corrections is addressed in page 6].

16. Figure 1: Besides the bad resolution and that some of the text in the figure is cut away, the figure is not very informative. What does “abnormal pressure data nodes are distributed” even mean?

Figure corrected in page 5.

17. It is not clear why the authors use a brownboost classifier at all, since it is invented for noisy environments. The authors are testing their method on simulations, which are not noisy. What is the reason why this classifier was chosen and no other one?

Brownboost classifier is a robust boosting and higher accurate classifier. It boost all weak (i.e., base) classifiers and combine to make strong one. Besides, the time parameter and error rate is highly concentrated in brownboost classifier. This in turns, the performance of classification is improved with minimum. Therefore, Enhanced Brownboost classifier is used in EBBC-GWO method for classifying the collection of pressure data

18. It is not clear why a k-NN classifier is used before the brownboost classifier. Is it a k-NN classifier or is it just the application of equation 1 on the pressure data?

The brownboost classifier performed to boost all the base classifiers., In EBBC-GWO, k-NN classifier is considered as the base classifier since the data in the network are not linearly separable. Therefore, k-NN classifier is used before the brownboost classifier

19. It is not clear why the outcome of Equation 1 on Page 7 is binary (0 or 1 as stated on Page 7 Line 19). Looking at the equation, the outcome is supposed to be a floating point number between 0 and 1.

In proposed work, the collection of data is classified as either normal or abnormal. The floating point value is rounded off in proposed work (i.e., 0.4 is 0 and 0.6 is 1).

20. Page 9 Line 10: How and to what extend does the brownboost classifier improve the classification accuracy and compared to what?

Brownboost classifier includes the ability to learn a collection of pressure data into a fixed level of accuracy thus classifies the data with classification accuracy than the state-of-the-art methods. [Correction addressed in page 9 line 12] Brownboost classifier is able to convert weak classifier into strong classifier. In addition voting process is carried out to classify the pressure data of node thus increases the classification accuracy in brownboost classifier

21. The enhanced brownboost classifier method is missing a crucial citation to the original paper by (Freund 2000), who invented this classification method. This is a clear case of plagiarism. This situation is further aggravated by the fact that this is one of the two key methods in this paper

Reference Included in page 6 line 20

22. Similarly to the brownboost classifier method, once again, the glowworm swarm optimization model is missing a crucial citation to the original work by (Krishnanand andGhose 2006), the inventors of this algorithm.

Reference Included in page 10 line 2

23. Page 11 Equation 11: Parameter gamma is not defined or mentioned in the text. Furthermore, maybe the most important part of a sensor placement algorithm is how to compute the objective function. It is not clear through the whole paper how the authors actually compute this function nor what the objective function means in the context of this paper at all.

Parameter gamma is not defined or mentioned in the text. [Correction addressed in page 12]

The objective function calculation is addressed in equation (12)

24. Page 12 Equation 13: There is an error in the equation. L_b(t) is in the subscript of the sum.

Correction addressed in equation 14.

25. Algorithm 2: Since glowworm swarm optimisation is a heuristic method, it cannot be guaranteed that it leads to the optimal solution / optimal node for sensor placement

Glowworm swarm optimization performed to provide the best corrective measures to designers when assigning correct heuristic. In proposed Glowworm swarm optimization, the objective function calculation is used to provide value of identifying optimal sensor placement. Moreover, it is cheap, simple and fast method to provide better solution for optimal sensor placement.

Simulation settings:

26. The settings of the constants in the optimisation algorithm (beta, gamma, rho) is not mentioned here, but for optimisation this is of high interest.

Correction addressed in table 1

27. The paper is missing a figure showing the DMA in Barcelona crucial for a further understanding of the results of this paper. Furthermore, it is not clear how the authors get the hydraulic network. Did they get it from researchers in Barcelona? Then it might

be also necessary to cite the publication where this network has been introduced for the first time

Fig. 1 : Map of the portion of the Barcelona WDN Fig. 2: Simulation model of a portion of Barcelona WDN (arino et al.2017) Ramon Ariño, Jordi Meseguer, Ramon Pérez and Joseba Quevedo, “Case Studies”, Springer, Real-time monitoring and operational control of drinking water systems,

28. Page 14 Line 5-11: It is not clear why the authors have chosen the abbreviations, for example, RM for “node with demand”? Flow and pressure are identified at inflow and outflow point. DMA includes 311 nodes with demand (RM type), 60l nodes without demand (EC type), 48 hydrant nodes without demand (HI type), 14 dummy valve nodes without demand (VT type) and 448 dummy nodes without demand (XX type) (R. Sarrate et al. 2014). [Correction addressed in section 3] RM for “node with demand is referred from (R. Sarrate et al. 2014).

29. Figure 4: The resolution of the figure is very bad. This has to be improved. Additionally,the figure shows a standard EPANET network (Net 3). Looking at this figure and the fact that the Barcelona network is missing, it is not clear to the reviewer if the authors actually used the Barcelona network for the simulations in this paper, since important materials (Barcelona network model) is not shown

quality of improved Figure 4 in section 3. In our work we consider District Metered Areas (DMA) in Barcelona network.

From the above Fig. 3: The small area that contains 100 sensor nodes are considered to perform simulation in figure 4 a and b Simulation results:

30. -In my opinion the convergence and convergence speed of the optimization method is of interest, but not mentioned here.Convergence and convergence speed of EBBC-GWO

[Addressed in section 4.3] Fig. 4:

31. It is not clear how the authors decide between normal and abnormal pressure data? What does it even mean in this context? Pressure in WDS is also dependent, where in the system it is measured (elevation, roughness values of pieps, valve settings,: : :) so just classifying points according to their pressure won't result in finding leaks automatically. Did the authors generate data by simulating leaks? How many leaks where simulated? What was the leak size?

The pressure flow is normal (i.e., maintained by a constant pressure) in the collection pressure data it is classified or identified as normal pressure data. Whereas, the pressure flow is abnormal (not a constant value) in the collection of pressure data is identified as abnormal pressure data. The abnormal pressure occurs when the pipe break, water leakage or fault in the pipeline is identified. [Correction is addressed in page 6].

Pressure in WDS is also dependent where in the system it is measured, so just classifying points according to their pressure. By considering the classification result, the pressure data is identified as normal or abnormal. From that the leak is detected. The authors generate data by simulating leaks. Five leaks were simulated. The size of the leakage was 142meter, 0.99meter, 0.99meter, 12.31meter. [addressed in section 3]

32. It is not clear why the authors have chosen the two methods (SVM and Graph-partitioning) for comparison of their method? It seems that these methods are chosen at random from literature? Why haven't the authors chosen other methods that might perform better?

Based on the objective of the proposed EBBC-GWO framework (i.e., increase the classification accuracy of water leakage detection with minimum time), the existing methods such as 1D-CNN-SVM model and Multi-Stage Graph Partitioning Approach are taken as base paper. The proposed work concept is derived by considering problems of these base papers. The drawbacks of these methods are effectively convinced by implementing proposed work. These two base papers are explained to understand the proposed work and these are more related to the proposed work. Therefore, 1D-CNN-SVM model and Multi-Stage Graph Partitioning Approach are chosen for comparison purpose.

33. -In general, it is not clear throughout the paper how the results are generated. The paper shows only sample calculations without detailed explanation. After the sample calculations, tables are listed with numbers and it is not clear, where this numbers come from.

The results are generated form MATLAB simulations and these results (i.e., numbers) are tabled in the paper. Sample calculation with detailed explanations is given in section 4.1.

34. -Using classification time as a measure for the performance of the algorithms is in context of sensor placement very questionable. Furthermore, the reviewer does not see the benefit of a classification time being 36 ms in contrast to 72 ms, since both are very fast. The interesting question would be the convergence time of the optimal sensor placement method, which is not listed in this paper.

The convergence time of optimal sensor placement method using EBBC-GWO Method is 1.3ms [addressed in section 4.2]

35. Does the number of pressure data in Figures 5 to 8 correspond to the number of sensors? If that is the case the method would be useless, because deploying 50-500pressure sensors in a water distribution system of total pipe length of 17 km results in a pressure sensor every 340 to 34 meter. This is a highly unrealistic number of sensors for such a small distribution system. Certainly, no water utility would be able to afford that number of sensors.

The number of pressure data in Figures 5 to 8 is not correspond to the number of sensors. The one sensor node sense the number of pressure data. EBBC-GWO method used Barcelona water distribution network. From these source 4645 km of

pipeline is considered. It consists of 883 nodes, 927 pipes which distributes water to 639 consumers. [Addressed in section 3]

36. For optimal sensor placement algorithms the most interesting outcome is the location where sensors should be placed. The optimal sensor positions are not shown in this paper.?

Glowworm swarm optimization calculates the objective function for neighboring node and current node. If the value of objective function is higher than the current node, the neighboring node is considered as optimal node to place the sensor for leakage detection. Fig. 5:

As shown Fig. 5: red color denotes the optimal sensor placement. [addressed in section 3]

37. A final comment about the use of abbreviations: The authors define abbreviations like EBBC-GWO multiple times in the paper without using it. Basically, in each section the abbreviations are defined again, which is certainly not the purpose of abbreviations at all.

Corrections are addressed throughput the paper

38. Finally, there are a lot of repetitions of paragraphs, hence, the paper is not concise

Corrections are addressed throughout the paper.

Please also note the supplement to this comment:
https://www.drink-water-eng-sci-discuss.net/dwes-2018-19/dwes-2018-19-AC2-supplement.pdf

Interactive comment on Drink. Water Eng. Sci. Discuss., https://doi.org/10.5194/dwes-2018-19, 2018.

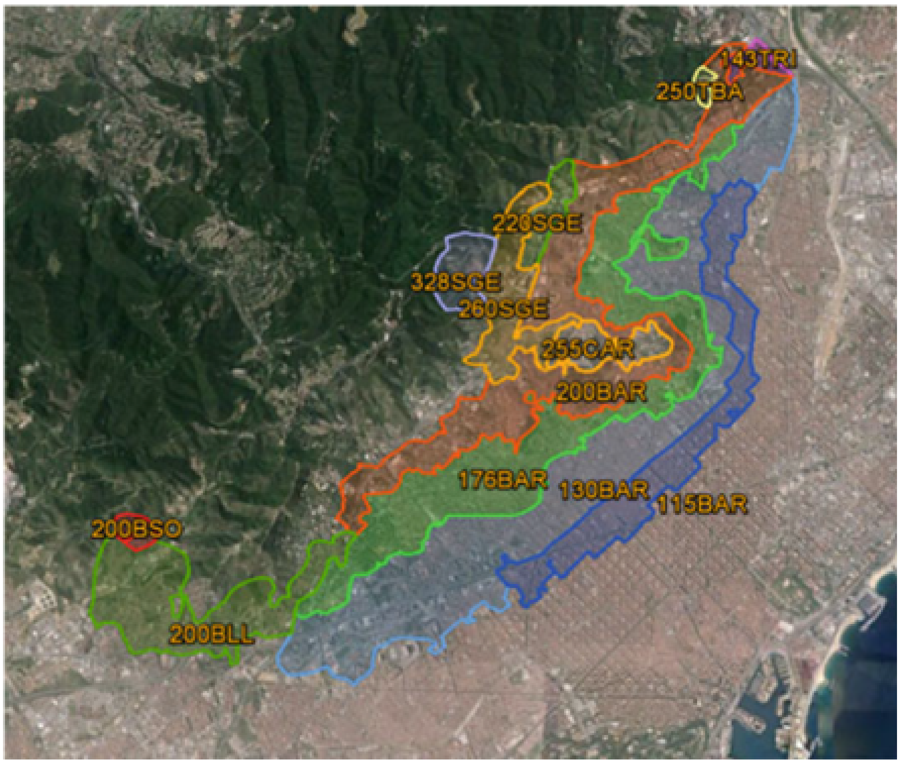

Fig. 1. Map of the portion of the Barcelona WDN

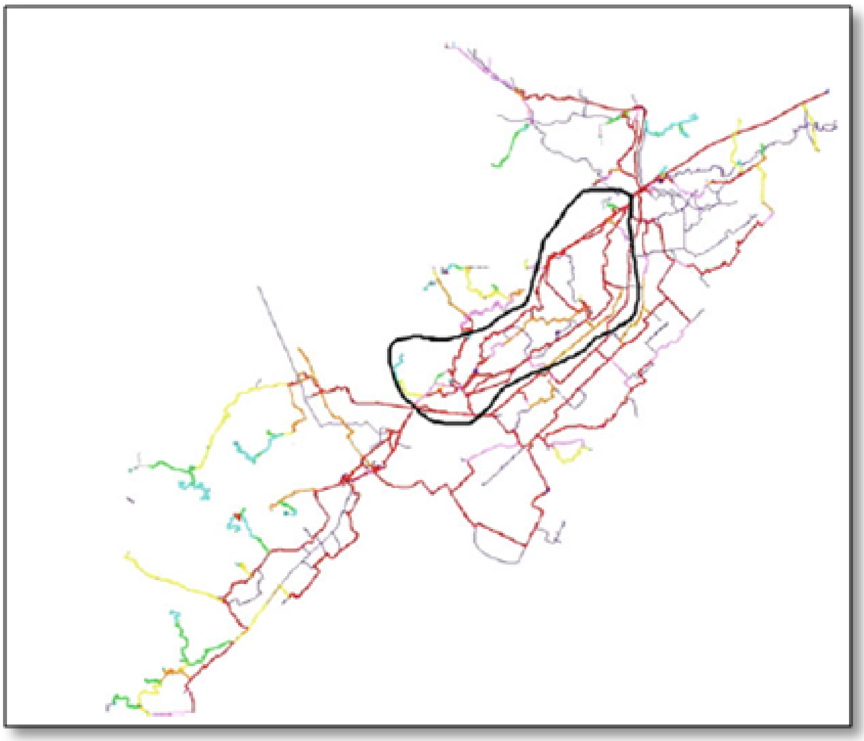

Fig. 2. Simulation model of a portion of Barcelona WDN (arino et al.2017)

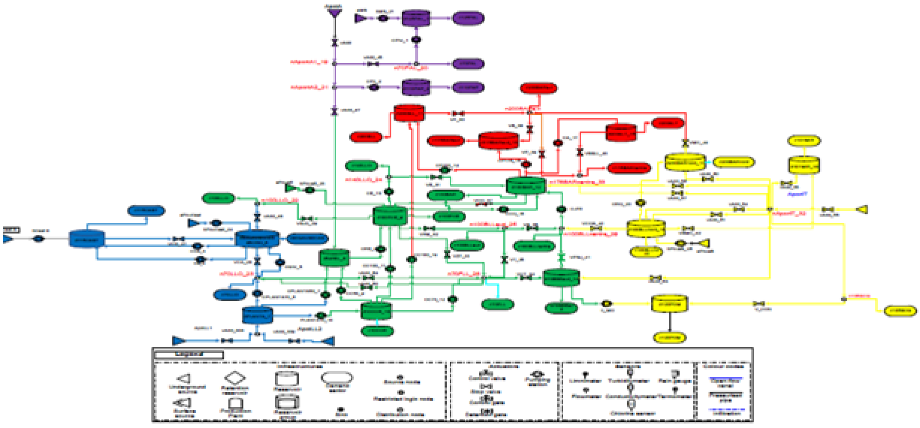

Fig. 3. Barcelona water network

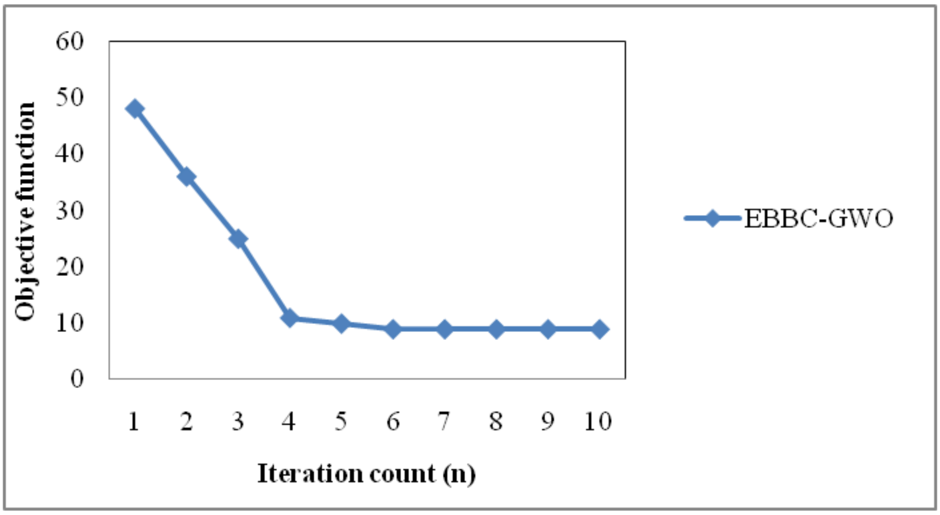

Fig. 4. Objective function vs Iteration Count

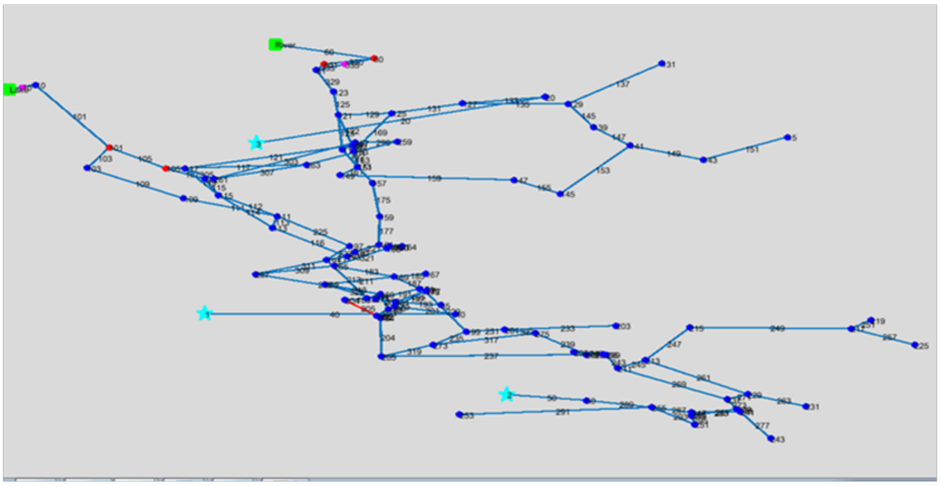

Fig. 5. optimal sensor placement in Water Distribution Network