Sensors Placement in Water Distribution Systems Based on Co-evolutionary Optimization Algorithm

Cheng-yu HU, Di-jun Tian, Chao Liu and Xuesong Yan*

School of Computer Science, China University of Geosciences, Wuhan, China, E-mail: huchengyu@cug.edu.cn

Abstract—In recent years, water pollution incidents happen frequently, causing serious disasters and society impact. It is advocated that water quality monitoring sensors shall be deployed in water distribution systems to realize the real-time pollution detection such that we can effectively detect the water pollution event to reduce the risk. However, how to deploy water quality sensors in water distribution systems (WDS) is a non-trivial and challenging task. Sensors placement in WDS is characterized by its extremely high computation complexity, uncertainty because of large-scale water distribution system and dynamic water demand by consumers. Aiming to minimize the average time of detection over all contamination events by placing a limited number of sensors into the water network, we have developed a co-evolution optimization algorithm, which using multiple populations to evolve simultaneously. Results indicate that our proposed algorithm performs better comparing to genetic algorithm and particle swarm algorithm.

I. INTRODUCTION

Recently accidental and intentional contamination of a water distribution system occurs frequently. Sudden water pollution incident, malicious attacks on the water distribution systems causes significant economic losses and bad social influence. In order to prevent major disasters and losses, which is caused by water contaminant incidents, installing water quality sensors in WDS to monitor contaminant event is necessary and valid. However, given a limited number of sensors, how to deploy them effectively is a challenging task because of the large-scale size of urban water network and varied water demand.

Previous studies have generally proposed many optimization models and algorithms for sensor deployment in the drinking WDS. There are three main methods to solve the problem of optimal sensor placement: (1)expert opinion, which use the opinion of engineer who engaged in water supply and drainage engineering to arrange the monitoring sensors [1]. (2) empirically-based method, which refers to the ranking of potential sensor locations based expert information(for example, data from geographical information system) [2]. (3) optimization methods, such as simulation-optimization model, which couples simulation model with optimization algorithm. Among these methods, optimization methods are the most advocated and investigated ones [3].

In the battle of the water sensor networks(BWSN), ostfeld et [4] have collected and compared 15 different approaches for sensor deployment, most researchers focus on optimization methods such as mixed-integer programming [5], heuristic-based algorithms [6] and genetic algorithm [7]. Although these methods can find near optimal sensor layouts, there are still

some problems. One of the most noteworthy problems is the computational efficiency. On the one hand, WDS is a complex system with large scale network nodes, on the other hand, many meta-heuristic algorithm require a high computational resources. Consequently, how to balance the accuracy and efficiency is a key problem.

As one of the most powerful tools for solving optimization problems, co-evolution has attract many researchers' attentions. Co-evolution provides a framework to implement search heuristics that are more efficient than those canonical evolutionary algorithms. By keeping and adjusting the diversity, co-evolutionary can maintain the balance of exploration and exploitation well. To the best of our knowledge, we firstly introduce co-evolution optimization algorithm solving sensors placement in drinking WDS.

In this paper, we aim to minimize the average detecting time over all contamination events, and propose a co-evolution algorithm, which using multiple particle swarms to find optimal sensors layouts scheme simultaneously. Results indicate that the proposed algorithm can quickly find the solution comparing to traditional heuristic approaches and deliver a high performance.

The remainder of this paper is organized as follows: Section II gives the related work on sensor placement in WDS and co-evolution algorithms. Section III address system model and mathematical formulation. Section IV presents our proposed co-evolutionary optimization algorithm. Section V discusses the simulation and the results. Section VI concludes the paper with a summary of this study.

II. RELATED WORK

A. Sensor placement problem

In the past several years, there has been a large number of papers on sensor placement for WDS, Hart and Murray [2] have reviewed more than 90 papers related to this optimization problem. In essence, sensor placement problem is related to the well-known p-median problems or covering problems for facility location. For the p-median problems in WDS, the number of sensors are fixed, the objective of arranging the sensors is to minimize the impacts over all contaminant incidents. For the covering problems, the cost of sensors is mainly concerned.

Several design objectives are used for sensor placement, such as minimizing the detection time, public health impacts, maximizing the likelihood of detecting contamination incidents [4].

Guan et al. [7] first proposed simulation-optimization model and developed an improved genetic algorithm (IGA) to find the optimal design of water sensor placement in WDS, simulation shows the model and algorithm are effective in solving this problem. Li et al [8] applied the non-dominated sorting genetic algorithm to solve the optimization problem.

Although heuristic algorithms are adapted to solve sensors placement in WDS, their performance deteriorate as the number of dimensions increases, new search schemes or algorithms are urgent in need to accelerate the search speed. In the paper, co-evolution algorithm is proposed and designed to solve the problem because of its natural search ability.

B. Co-evolution Optimization

In biology, co-evolution is "the change of a biological object triggered by the change of a related object". In other words, when changes in at least two species genetic compositions reciprocally affect each others evolution, co-evolution has occurred [9].

In the field of intelligent computation, co-evolutionary algorithm is widely studied and a rich of authors have proposed many heuristic algorithm in the frame of co-evolution.

Potter and De Jong [10] have firstly proposed cooperative co-evolution to tackle complex problems with many dimensions, this method employs the divide-and-conquer technique, and divide an original problem into smaller problems with fewer parameters. Yao et al. [11] proposed a new version of co-evolution algorithm with random grouping and adaptive weighting with which they successfully solved benchmark functions with up to 1000 dimensions. By coupling hydraulic simulation model(EPANET) with co-evolutionary optimization, we first adapt co-evolutionary algorithm to water sensor placement problem in WDS.

III. MATHEMATICAL MODEL

A. System model

In this section, we give a detailed description of the system model. A water distribution network shown as Fig 1 modeled as a graph $G = (V, E)$. E is a set of edges representing pipes, V is a set of vertices, or nodes, where pipes meet. Vertices can represent sources, such as reservoirs or tanks, where water introduced, and sink(demand point) where water is consumed. Each pipe connects two vertices v_i and v_j and is usually denoted(v_i, v_j).

General speaking, the flow patterns varies as the water demand changes, in this paper, we consider contaminant risks under a fixed flow pattern, where we require the direction of the flow on each edge. If there is a positive flow f_{ij} along the edge $e = (v_i, v_j)$, then $f_{ij} = 1$, otherwise $f_{ij} = 0$. Water cannot flow in both directions of a pipe.

In water distribution network, an attack δ_i means conservative contaminants, which will not react with other substances in the network, are injected at node i from the start time t_{ib} to the end time t_{ie} with the mass fraction M_i, M_i is a real vector which represents contaminant mass loading. The contaminant will gradually diluted with the current, its propagation speed
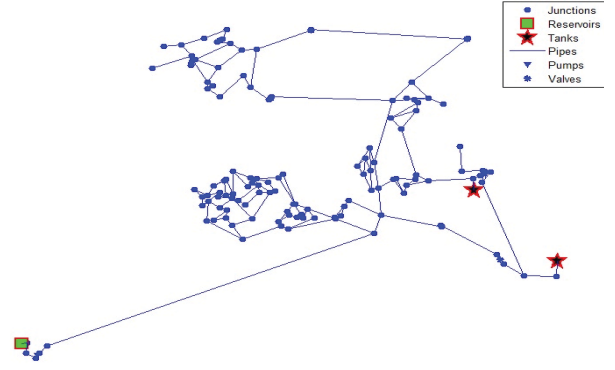

Fig. 1. system model

in accordance with average flow velocity; Each contaminant event means an attack on one node of the water network. A denotes the set of contaminant scenarios against which one sensor layout consisting of p sensors is intended to protect.

We assume that a fixed budget of p sensors are deployed at any junction in a water distribution network, however, the sensors are not allowed to be installed on piles because the water quality simulation software (EPANET) which we used in experiment cannot provide this information. We also assume that each sensor can detect any minimum concentration of contaminant and a general alarm is raised when contaminant is first detected by a sensor.

B. Mathematical model

The purpose of placement of water sensors in WDS is to detect contamination events efficiently such that the risk for the population exposure to contaminants is reduced. In this paper, our formulation models the placement of p sensors on a set $P \subseteq V$ vertices, $X \in P$, $X(j)$ mean the ith sensor in Xth scheme of sensor placement. we use the expected time of detection for all the contaminant event A as optimization objective. For a contamination event $A(i)$, the time of detection by sensor $X(j)$ is the elapsed time $t_{X(j)}$ from the start of the contamination event to the first identified presence of a nonzero contaminant concentration. Given a fixed number of p sensors, the time of first detection is described as follows:

$$t_d = \min_{j=1,...,p} \{t_{X(j)}\}$$

equation (1) represent the objective of minimizing the expected value by computing over the all the contaminant events.

$$f = \min \frac{1}{V} \sum_{i=1}^{V} t_d \qquad (1)$$

where f denotes the mathematical expectation of the minimum detection time t_d. Because each contamination event involves a single injection, which may occur at any network node, so the number of all contamination events is V, if one
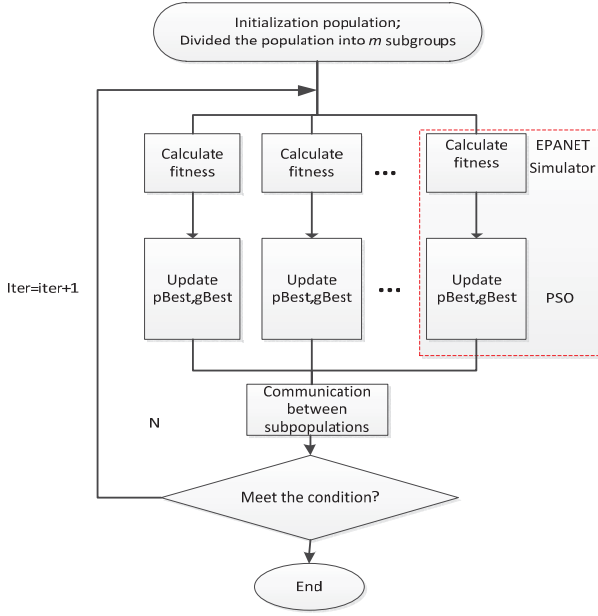
Fig. 2. The frame of co-evolutionary for sensor placement

contaminant event is not detected by any sensor, then t_d is equal to the simulation time.

$t_{X(j)}$ is calculated by EPANET simulator. For purposes of evaluation, contaminant concentrations were evaluated using a 5-min time step.

IV. CO-EVOLUTIONARY OPTIMIZATION FOR SENSOR PLACEMENT(COSP)

Co-evolutionary optimization is a framework that multiple populations cooperate to find the optimal solutions, In the paper, we use multiple particle swarms to search different areas in the solution space. The framework is shown as figure 2 :

The work flow of the COSP consists of the following stages:

1) Set up the maximum iteration and the first iteration $iter \leftarrow 0$;
2) Generate original population $\boldsymbol{P}(iter)$ made up of m initial individuals, then divide the population into several subgroups in the form of $\boldsymbol{P}(iter)=\{P_1(iter), P_2(iter),..., P_n(iter)\}$, n is an integral number of subgroup, initialize the population speed $V(t)$.
3) Calculate each individual fitness function $F_j(j = 1, 2, ..., m/n)$ by EPANET in each subgroup $P_i(t)(i = 1, 2, ..., n)$.
4) Execute independent evolution among each group $P(iter)=\{P_1(iter), P_2(iter), ..., P_n(iter)\}$;
   a) Independently initialize evolutionary counter $t \leftarrow 1$ and make an evaluation of each individual in each subgroup using EPANET;
   b) sort the individual based the fitness value, record the global best individual as $gBest$ and local best individual $pBest$;

   c) Update each particle's speed using the eqation $V_i(t+1) = V_i(t) + C_1 * rand * (pBest - P_i(t)) + C_2 * rand * (gBest - P_i(t));$
   d) Update each particle's velocity according to the equation $P_i(t+1) = P_i(t) + V_i(t+1)$;
   e) $t = t + 1$, t is equal to a pre-defined threshold ρ, stop evolving.
5) Information exchange among the sub-populations. By replacing the worst individual in the subgroups with the global best individual $gbest$, every subgroup can share the best information of the whole population, thus search speed is accelerated.
6) Stop evolving the whole population when results fit with ending conditions, otherwise, renew independently iteration counter $iter \leftarrow iter + 1$ and turn to step (4).

V. EXPERIMENT SIMULATION AND ANALYSIS

A. Parameters settings of WDS and algorithm

This paper uses a standard test network [4]. This network consists of 127 nodes, including two tanks, two sources and 171 pipes. The hydraulic time step is 10 minutes, and the water quality time step is 5 minutes. The total simulation time is 34 hours. A nonreactive contaminant is introduced into the network at any node at the beginning of the simulation and the mass loading of 300mg/L.

To verify the effectiveness of our proposed algorithm, we compared COSP with genetic algorithm and particle swarm algorithm under the same conditions. Table I lists the parameter settings of the three algorithms

TABLE I
PARAMETER SETTINGS

Algorithm	parameters	individuals	iteration
GA	Pc=0.95, Pm=0.1, Roulette Selection	200	100
PSO	W=1, C1=C2=1	200	100
$COSP$	W=1, C1=C2=1	20×10	100

For the purpose of a fair evaluation of the three algorithms, we perform the same number of the evaluation on the fitness function. For the GA and PSO, the population consist of 200 individuals and evolve 100 generations, the evaluation times equal to 200×100, for COSP algorithm, the population is divided into 20 subgroups and each subgroup consists of 10 individuals, the subgroups evolve concurrently, after 10 generation of evolutions, all subgroups communicate with each other by coping the global best individual $gbest$, every subgroup independently evolves 100 times, so the sum of the evaluation times is $20 \times 10 \times 10 \times 10$ times.

B. Results

The curve of fitness value of the three algorithm is shown as figure 3, we can see that the curve of COSP decrease quickly, which mean COSP has a quick exploration in the search space. The GA behaves worst among the three algorithms.

Figure 4 shows the optimal sensor layouts found by the three algorithms. We find that the result vary considerably ,
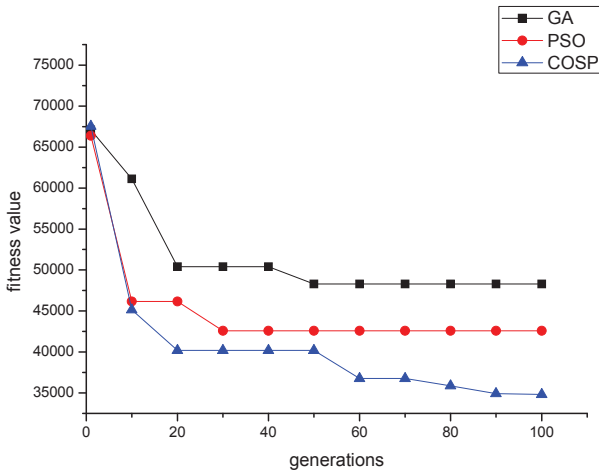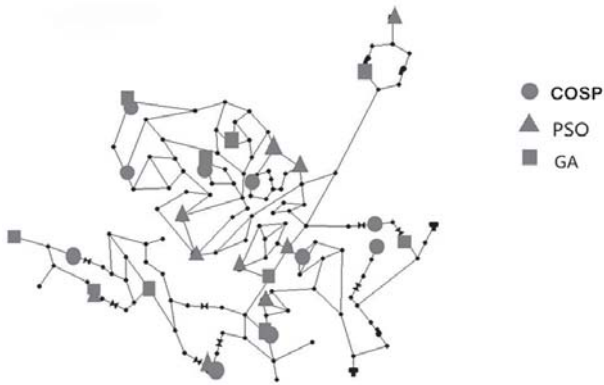
Fig. 3. The fitness curve of three algorithms

Fig. 5. expected time of detection VS the number of sensors

Fig. 4. optimal sensors layouts found by three algorithms
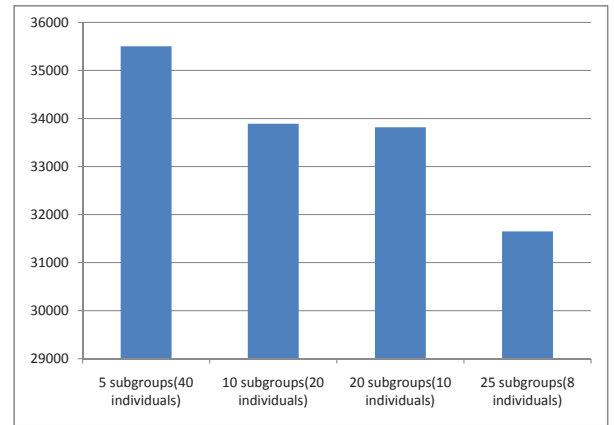
Fig. 6. the expected time of detection VS the number of subgroups

however, the layouts by COSP give a minimum expectation of detect time comparing to the other algorithm, we believe that the contaminant risk is much less under the protection by sensor layouts of COSP.

C. Discussion

1) Discussion for different numbers of sensors: To explain the scalability of COSP and study the effect of different number of sensors, we assume that the number of sensors range from 10 to 50.

From the figure 5, We can find that the expected time of detection for all the contaminant events decreases as the number of the sensors increase, which is consistent with our commonsense.

2) Discussion for co-evolutionary algorithms:

For co-evolutionary algorithms, there are two critical issues (1) how many subgroups are suitable for the optimization problem , (2) how to set the frequency of communication among subgroups. To find the fittest size of subgroup, We divide the whole population into 5, 10, 20, 25 subgroups which consist of 40, 20, 10, 8 individuals, the total number of individuals is 200.
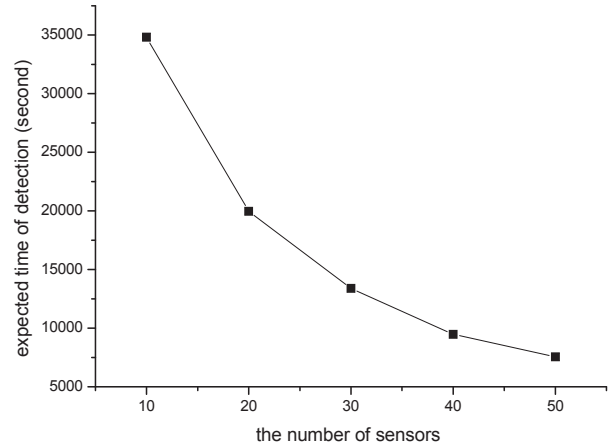
From the figure 6, we can see that the expected time of detection decrease as the number of subgroup increase, it means that more subgroups and less individuals will enhance the performance of the Co-evolutionary algorithm, however, for the demand of fully exchange information among the individuals, the size of subgroup cannot be too small.

The frequency of communication is much critical for co-evolutionary algorithm, if the interval time between two suc-cessional communication is too short, extra computation over-head will increase. In this paper, aiming to get the appropriate frequency among subgroups, we set 5, 10, 20, 50 generation as the frequency, which means the subgroups evolve 5, 10, 20, 50 generations respectively, then all the subgroups communicates by importing the global individual $gbest$.

From figure 7 , we can see that the subgroups exchange the information with 10 generations independently evolution can achieve the best result .

VI. CONCLUSION

In this paper, we present a co-evolutionary optimization algorithm to solve the sensor placement problem. First, On the
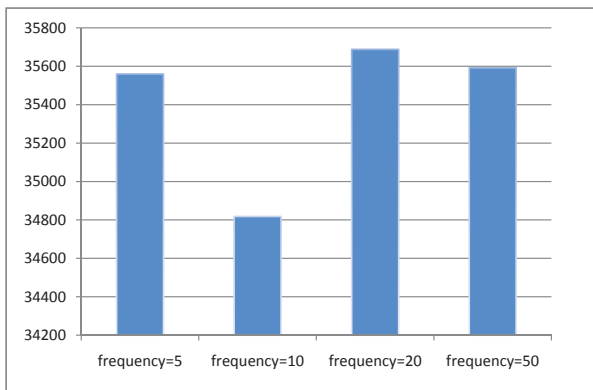
Fig. 7.   the expected time of detection VS communication frequency

basis of the works of pioneer contributors, we give a detailed description of optimization model. Second, to verify the effectiveness of the algorithm, We compare the co-evolutionary algorithm to genetic algorithm and particle swarm algorithm. Results shows that our approach behaves better than the other algorithms.

we also made a discussion on the stability and the parameters of our approach. However, a lot of work still remains to be done for sensor placement in water distribution systems. For example, multi-objective optimization , and multiple flow patterns for the problem should be considered in the future study.

VII. Acknowledgement

References

[1] G. Trachtman, *A Strawman, Common Sense Approach for Water Quality Sensor Site Selection.*, 2006, ch. 105, pp. 1–13. [Online]. Available: http://ascelibrary.org/doi/abs/10.1061/40941

[2] W. Hart and R. Murray, "Review of sensor placement strategies for contamination warning systems in drinking water distribution systems," *Journal of Water Resources Planning and Management*, vol. 136, no. 6, pp. 611–619, 2010.

[3] K. Diao and W. Rauch, "Controllability analysis as a pre-selection method for sensor placement in water distribution systems," *Water Research*, vol. 47, no. 16, pp. 6097 – 6108, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0043135413006003

[4] A. Ostfeld and l. e. a. Uber, "The battle of the water sensor networks (bwsn): A design challenge for engineers and algorithms," *Journal of Water Resources Planning and Management*, vol. 134, no. 6, pp. 556–568, 2008.

[5] O. Piller, M. Propato, and M. E. Tryby, *Linear Algebra Analysis for Contaminant Source Identification in Water Distribution Systems*, 2007, ch. 513, pp. 1–10.

[6] G. Dorini, P. Jonkergouw, Z. Kapelan, F. di Pierro, S. Khu, and D. Savic, *An Efficient Algorithm for Sensor Placement in Water Distribution Systems*, 2006, ch. 100, pp. 1–13.

[7] J. Guan, M. Aral, M. Maslia, and W. Grayman, "Identification of contaminant sources in water distribution systems using simulationoptimization method: Case study," *Journal of Water Resources Planning and Management*, vol. 132, no. 4, pp. 252–262, 2006.

[8] M. Li, S. Liu, L. Zhang, H. Wang, F. Meng, and L. Bai, "Non-dominated sorting genetic algorithms-iibased on multi-objective optimization model in the water distribution system," *Procedia Engineering*, vol. 37, no. 0, pp. 309 – 313, 2012, the Second {SREE} Conference on Engineering Modelling and Simulation (CEMS 2012). [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877705812018942

[9] K. Y. Yip, P. Patel, P. M. Kim, D. M. Engelman, D. McDermott, and M. Gerstein, "An integrated system for studying residue coevolution in proteins," *Bioinformatics*, vol. 24, no. 2, pp. 290–292, 2008. [Online]. Available: http://bioinformatics.oxfordjournals.org/content/24/2/290.abstract

[10] M. A. Potter and K. A. de˜jong, "A cooperative coevolutionary approach to function optimization," *Lecture Notes in Computer Science*, vol. 866, 1994.

[11] Z. Yang, K. Tang, and X. Yao, "Large scale evolutionary optimization using cooperative coevolution," *Information Sciences*, vol. 178, no. 15, pp. 2985 – 2999, 2008, nature Inspired Problem-Solving. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S002002550800073X