AUTOMATED DESIGN OF WASTEWATER COLLECTION SYSTEMS USING GENETIC ALGORITHMS

Lou Y. Liang, Russell G. Thompson and David M. Young

Department of Civil and Environmental Engineering,
The University of Melbourne, Victoria, 3010 Australia

Abstract: The size, shape and slope of pipes are major components of the overall cost of wastewater collection systems. In the past, designers have used charts and specialized rules to determine the size, slopes and materials when designing wastewater collection networks. However, genetic algorithms (GA) provide powerful technique for automating the design and minimizing the construction costs of wastewater networks. This paper describes the development and application of a GA using a repair procedure to incorporate the numerous constraints involved designing a large gravity wastewater collection system.

Keywords: genetic algorithms, constraints and wastewater collection system design

1. INTRODUCTION

The traditional method for designing gravity wastewater collection systems is largely based on trial and error which is very time consuming. Designers typically use charts and specialized rules to determine the diameters, slope and materials of sewers when designing wastewater collection networks. Suitable diameters and slope combinations are selected for all pipes between manholes, so that the wastewater can be transported without violating any hydraulic constraints. Since there is a large range of pipe slopes, diameters and coefficients in the hydraulic relationships, designers can usually only evaluate a small number of networks that do not violate any of the constraints. However, since many of the costs and constraints are not linear, there are no simple procedures to find the least cost design for pipeline networks. Linear programming has been applied to minimize the total cost of sewers subject to constraints [3]. Since this optimization method does not incorporate standard diameters, there is no guarantee of optimality for standard commercial pipe diameters.

Genetic algorithms (GA) provide an efficient technique for finding near exact solutions to a wide range of complex optimization problems. GA are based on the mechanics of natural genetics [7]. They can search large solutions spaces quickly and only require an objective function to be specified. Furthermore, because of the processing influence associated with GA, the method they have much more global orientation than many other methods often used in engineering optimization problems [4]. Recently, GA have been used in the optimization of wastewater collection systems [8]. Many real world problems can be viewed as systems having to satisfy a given set of constraints. Constraint satisfaction is a search procedure that operates in a space of constraint sets. A feasible design is any combination of variables that satisfies the design constraints. Typically, GA make use of penalty function methods to treat infeasible solutions that violate one or more constraints. Where a system has a number of constraints, GA can be become inefficient by creating a large number of infeasible solutions that are usually discarded. However, these infeasible solutions may contain some important information that may assist in identifying the optimal solution. A GA has been applied to obtain the optimal stacking sequences for maximizing the buckling load of composite cylinders. Three stacking rule constraints were implemented [9]. The difficulty of handling the combinatorial constraints in genetic optimizations was overcome using a repair procedure. When a chromosome violated a stacking constraint a repair procedure was implemented. This procedure did not alter the genes of the chromosome but only changed the decoding rules to introduce the constraints. This is similar to recessive genes in biology. The genetic algorithm using the repair procedure was shown to provide higher design reliability [9]. This type of procedure provides a means of increasing the efficiency of GA in systems where there are a number of constraints. This paper describes and evaluates the performance of the repair procedure developed for incorporating a number of hydraulic constraints for designing a large gravity wastewater collection system.

2. GENETIC ALGORITHMS

GA begin, like many other optimization algorithms, by defining the optimization parameters and the cost function. They also terminate like other optimization algorithms, by testing for convergence. However, GA are different from other optimization algorithms in a number of ways [6]. Firstly, GA require the natural parameter set of the optimization problem to be coded as a finite string. Each string represents an artificial chromosome with every string consisting of a number of artificial genes. Secondly, GA use a set of strings to form a set (or population) of solutions. This is in contrast to the single-point approach used by traditional optimization methods. GA therefore, use more global search tactics compared with local search approaches. Thirdly, GA do not require any specific mathematical solution procedures to be used. For example, no complicated calculus-based search algorithms are required in GA. Finally, GA use randomized operators instead of gradient information used by traditional methods [5].

2.1 Reproduction

The initial population (set of solutions) is usually randomly generated. From this population, offspring (new solutions) are produced using three distinct operators. Strong chromosomes (good solutions) will probably be selected several times for inclusion in the new generation and weak chromosomes (poor solutions) may die out from the previous generation.

2.2 Crossover

The crossover operation combines two members of the population by cutting their chromosome strings at a randomly chosen position. Genes are exchanged from the parents and the offspring inherit some genes from each parent.

2.3 Mutation

To avoid premature convergence to a local optimum, a mutation or random change of a number of genes is usually performed. After crossover and mutation, the offspring (new solutions) will generate different values of the objective function.

2.4 Fitness and convergence

A fitness function based on the objective function is generally used to select chromosomes for reproduction. Convergence is often measured using the concept of bias, which is defined as a measure of agreement among the population. If the selection procedure allocates a high probability to good solutions (ie. it is heavy handed), then the population will converge quickly [1].

3. HYDRAULIC FORMULATION

Many practical engineering problems are multi-constrained problems such as designing pipeline networks. In general, constraints are used to represent the numerous hydraulic requirements necessary for Large Gravity Wastewater Collection Systems (LGWCS).

3.1 Hydraulic constraints

1. Diameter progression constraints:
   $D_{i+1} \geq D_i$, i = 1, 2 ,..., n
   Where D_i is the diameter of pipeline i^{th} link
2. Minimum velocity constraints
   $V \geq V_{min}$
   The minimum velocity required for self-cleaning
3. Maximum velocity constraints
   $V_{max} \geq V$
   Excessive velocity can cause erosion of pipe materials by grating sewage
4. Minimum and maximum cover constraints
   Minimum and maximum ground depth are required to cover sewers
5. Invert level constraints
   $INV_i \geq INV_{i+1}$, i = 1, 2,..., n
   The downstream invert flow level must be lower or equal to the upstream flow level

3.2 Design equation for LGWCS

Gravity wastewater collection pipelines are usually designed using the Colebrook-White equation (1). Colebrook published the equation for turbulent flow in circular full flowing pipes in 1939. It was derived from the smooth and rough turbulent logarithmic resistance laws for circular tubes. These were evaluated experimentally by Nikuradse, after theoretical work by Prandtl and Karman [10].

$$V = -2\sqrt{2SgD} \times \log\left(\frac{k_s}{3.71D} + \frac{1.775\nu}{D^{1.5}\sqrt{Sg}}\right) \quad (1)$$

where V, D , S, g, ν and k_s are mean in flow velocity, pipe diameter, pipeline slope, acceleration due to gravity, kinematic viscosity and roughness size respectively.

4. SEARCH OPTIMIZATION

4.1 GA with penalty function

Consider a design that consists of a large pipe with a flat slope and a small pipe with a steep slope, such that they may both be optimal in construction. Finding the set of LGWCS designs using GA with a penalty function is desirable because the trade-off among various diameters and slopes can be observed in our previous work [8]. Normally, two parents are randomly selected from the population.

Based on the strings length, the crossover point is randomly generated to select a segment in one parent between the crossover site. The mutation operator makes random changes to one or more bits of the string. A feasible solution is one that does not violate any of the constraints. However, any random operations that are generated in an evolutionary process cannot be the best possible sequence for all pipes diameters or for all the design criteria. Traditionally, GA use a penalty function for selection, therefore the fitness function must reflect the objective function and any constraint violations. However, the penalty function must be established using a try-and-error process. If the string is infeasible, its total cost is given a very high value, so that it will have a poor chance of being selected in future generations. With hard constraints such as diameter progression constraints, the use of crossover and mutation operators resulted in approximately 20% of new strings violating the progression constraints. There is also a possibility of strings becoming infeasible by violating the hydraulic constraints.

4.2 GA using the repair procedure

To produce feasible offspring, the ascending diameter rule is used to modify the order of bits after the crossover and mutation operators. The aim of the repair procedure is to generate a feasible diameter progression that satisfies the hydraulic constraints. The procedure was applied in the design of the Changbin Industrial Park Project [2]. For illustrative purposes, the selection of pipe diameters are coded as 1, 2, 3,..., 7 corresponding to ϕ300mm, ϕ350mm, ϕ425mm, …, ϕ500mm. In this method, the parents are initially selected from the feasible region. Based on the strings length, the crossover site *site* is randomly generated to select a segment from one parent. The offspring are generated by organizing the elements of the selected segment in one parent according to the repair rule where they appear in the other parent with the rule that the remaining elements are the same as in the first parent. The mutation operator makes a random change to one or more elements of the string. However, there is a possibility of string becoming infeasible by violating the diameter progression. If the string is infeasible, the string will be modified using the repair procedure. The repair procedure for this type of constraint is described as follows.

A finite set D of n variables $D_1, D_2, \ldots D_n$ the domains of parents will represent the set of the domains of all diameter variables related to the parents.

The domains of offspring *m with n variables* are defined as (D_m^n) and selected segment from one parent,

$$D_m^n = \begin{cases} D_m^{n+1} & \text{if } D_m^n < D_m^{n+1}, site \le \frac{1}{2}(stringlength) \\ \\ D_m^{n-1} & \text{if } D_m^{n-1} > D_m^n, site > \frac{1}{2}(stringlength) \end{cases}$$

for example,

parent 1: *2,3,3,3,4,5,5,6,7,7,7*
parent 2: *1,2,2,3,4,4,4,5,5,6,6*

consider random crossover site as *6*

offspring 1: *2,3,3,3,4,5,4,5,5,6,6*
offspring 2: *1,2,2,3,4,4,5,6,7,7,7*

Offspring 1 is infeasible with 5, 4 in sites 6, 7. The repair procedure is introduced to modify this string to make it feasible: *2,3,3,3,4,5,5,5,5,6,6*. If the random site for mutation is *9*, this gene is changed from *5* to *2*. Again, this offspring becomes infeasible since it violates the constraint. Using the repair procedure, the new offspring becomes feasible: *2,3,3,3,4,5,5,5,6,6,6*. After the string is repaired and hence feasible, the string is decoded into design variables and material types. The second step is determined from the hydraulic analysis of the network to determine if any inactive hydraulic constraints are violated. If the inactive constraints are violated, the GA has to generate a new population.

5. RESULTS AND DISCUSSION

Most sewers are laid under roads to avoid interference with private property when connections and repairs have to be made. Sewers are commonly laid in straight lines, with manholes provided where any change of diameter, gradient or direction occurs. Where the sewers are laid at shallow depths, especially under roads or where they are very deep, concrete surroundings are needed to provide further strength. The principal purpose of this study is to identify the least cost design of a large gravity wastewater collection system. The cost function for this model can be written as follows:

$$f(c) = Min \sum_{i=1}^{n} \left(Sc_i + Ec_i + Bc_i + Dc_i + Pc_i \right)$$

where,
n - total number of sewer links
Sc_i - sewer material cost of link i
Ec_i - soil excavation cost of link i
Bc_i - backfill of soil cost of link i
Dc_i - soil dump cost for link i
Pc_i - penalty cost for link i

The cost of the sewer network can be obtained using information concerning soil excavation, backfill, dump, sewer type, slope, burying, etc. Table 1 shows the quantity of soil earthworks for section 2401 using the traditional (reasonable) design method and both GA appraoches. These results illustrate that GA significantly reduces construction costs and there was little difference in cost performance between the 2 GA procedures. Compared with the traditional design method, GA achieved a 9% reduction. Excavations occur in pipeline LGWCS, and this involve breaking out and removal of soils to the line and levels shown in Figure 1. Figure 2 compares the convergence speed of the GA with penalty function and the GA with repair procedure for 2000 runs. These results show that the GA using the repair procedure converges faster, with the GA using the penalty function being more unstable. For each sewer the maximum flow has to be estimated and a diameter chosen to suit the gradient available. The diameter is chosen so that the estimated flow will be carried with the sewers running, that is the pipe is assumed to be half-full but not surcharged and the hydraulic gradient is the same as the invert gradient. Due to the properties of circular sections, a sewer will actually flow at about 80% of full depth with the design discharge and there will be a margin available above the design discharge before the sewer becomes surcharged. Sewers can often flow in a surcharged condition without causing flooding. Table 2 shows the hydraulic gradient and sewer diameter obtained using the GA with penalty function. In this study, if the diameter of sewer is greater or equal to 400mm, it will be made from resin reinforced concrete. Otherwise, vitrified clay pipes are chosen. Table 3 shows the best design using the GA with repair procedure has larger diameters and milder hydraulic gradients than the design obtained using the GA with a penalty function.

GA have been shown to achieve good performance for a wide range of optimization problems. The design methods involved in this system were mainly conventional, despite some unusual structural features above the pipeline and other industrial infrastructure below the ground section of the layout of pipeline. The aim of this paper is evaluate GA for LGWCS design. However, GA algorithms are generally not suited to optimization problems where there are a number of constraints. To overcome this weakness rules can be developed to control the GA operators to control the search for solutions.

Table 1. Earthwork quantities

Design method	Excav. (m^3)	Backfill (m^3)	Dump (m^3)	Cost *f(c)* (NT$m)
Traditional	4863	4273	590	3.38
GA (penal.)	4304	3840	464	3.08
GA (repair)	4401	3930	471	3.07

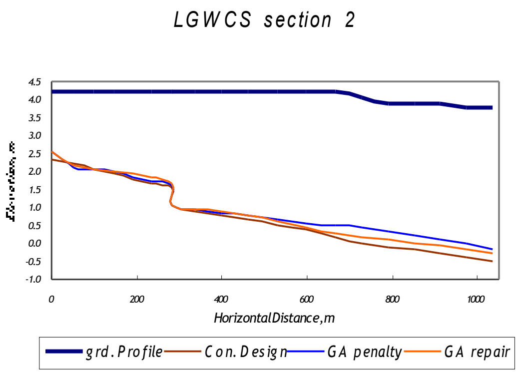

Fig 1. Optimal and conventional design profile

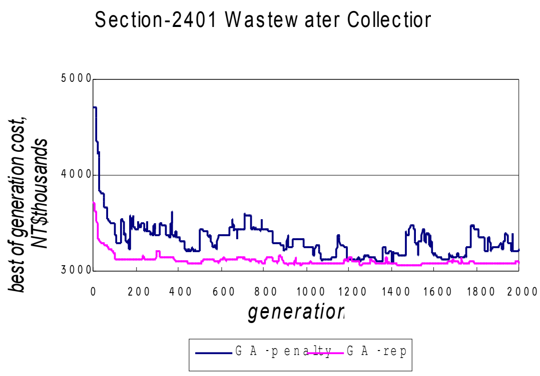

Fig 2. Performance of GA methods

CONCLUSIONS

The traditional method of designing large gravity wastewater collection systems is based on trial and error. If a design does not meet the constraints, the process has to be repeated again from upstream to downstream until each constraint

Table 2. The least cost LGWCS design using the GA with a penalty function

Link No.	Diameter, m	Link Slope
SE1001	0.300	0.0018
SE1002	0.300	0.0012
SE1003	0.300	0.0022
SE1004	0.350	0.0034
SE1005	0.350	0.0024
SE1006	0.350	0.0028
SE1007	0.375	0.0010
SE1008	0.375	0.0020
SE1009	0.375	0.0010
SE1010	0.400	0.0014
SE1011	0.400	0.0016
SE1012	0.450	0.0008
SE1013	0.450	0.0014
SE1014	0.450	0.0014
SE1015	0.450	0.0016
SE1016	0.450	0.0026
SE1017	0.450	0.0016
SE1018	0.450	0.0026

Table 3. The least cost LGWCS design using the GA based on repair procedure

Link No.	Diameter, m	Link Slope
SE1001	0.300	0.0014
SE1002	0.300	0.0024
SE1003	0.300	0.0012
SE1004	0.300	0.0012
SE1005	0.300	0.0024
SE1006	0.300	0.0038
SE1007	0.350	0.0012
SE1008	0.350	0.0012
SE1009	0.350	0.0016
SE1010	0.375	0.0026
SE1011	0.400	0.0024
SE1012	0.400	0.0024
SE1013	0.475	0.0014
SE1014	0.475	0.0014
SE1015	0.475	0.0014
SE1016	0.475	0.0016
SE1017	0.475	0.0016
SE1018	0.475	0.0016

is not violated for all sewers. This paper presented the application of GA to minimise the construction cost of wastewater collection systems. A repair procedure was developed for handling hydraulic constraints. The performance of this procedure was compared with a GA using a penalty function as well as the traditional design approach. Both GA produced lower cost designs than the traditional design method for a network with a large number of sewer links. The efficiency of the GA using the penalty function was found to be largely dependent on the complexity of the network. The performance of the GA using the repair procedure was observed to be more stable compared with that of the GA using a penalty function.

REFERENCES

[1] Bauer, R. J., *Genetic Algorithms and Investments Strategies*, Wiley Finance Education, New York, 1994

[2] *Changbin Coastal Industrial Park Hsien-Si Zone East I Wastewater Collection System Engineering Design*, in Chinese, SEC, 1992.

[3] Elimam, A. A., Charalambous, C. and Ghobrial, F. H., "Optimum design of large sewer networks", *J. of Envir. Engrg.*, ASCE, 115(6), 1171-1190,1989

[4] Goldberg, D. E., and Kuo, C. H., "Genetic Algorithms in Pipeline optimization", *J. of Comp. In Civ. Engrg.,* ASCE, Vol. 1, No. 2, pp. 128-141, 1987.

[5] Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Publishing Co., Inc., Reading, Mass., 1989.

[6] Haupt, R. L., Haupt S. E., *Practical Genetic Algorithms*, A Wiley-Interscience Publication, John Wiley & Sons, Inc. New York, 1998.

[7] Holland, J. H., *Adaptation in Natural and Artificial Systems,* University of Michigan Press, Ann Arbor, Michigan, USA, 1975.

[8] Liang, L. Y., Thompson, R. G., and Young, D. M.,"Designing Wastewater Collection System Using Genetic Algorithms", accepted by *Lecture Notes in Artificial Intelligence*, Springer-Verlag.

[9] Todoroki, A., Sasai, M.,"Stacking sequence optimization for buckling of laminated cylinders using a genetic algorithm with a recessive gene like repair strategy", *ASME, Pressure Vessels and Piping Division PVP* v370, p43-50, 1998.

[10] Wallingford, H., Barr, D.I.H., *Tables for hydraulic design of pipes, sewers and channels*, Thomas Telford, London, 1998.