MOAQ an Ant-Q Algorithm for Multiple Objective Optimization Problems

Carlos E. Mariano
Instituto Mexicano de Tecnología del Agua
Paseo Cuauhnahuac 8532
Jiutepec, Morelos, 62550, MEXICO
cmariano@tlaloc.imta.mx

Eduardo Morales M.
ITESM - Campus Morelos
Paseo de la Reforma 182-A
Temixco, Morelos, 62580, MEXICO
emorales@campus.mor.itesm.mx

Abstract

This paper describes an Ant-Q algorithm for the solution of multiple objective optimization problems (MOAQ). The algorithm uses families to represent the objectives in the problem. The problem definition considers that each family finds solutions that depend on solutions found by the rest of the families, creating a negotiation mechanism and finding compromise solutions for all the objectives involved. The compromise solutions are domination evaluated in the Pareto sense, assigning rewards to the non-dominated solutions fitting all problem constraints, and punishments to the solutions violating any of them. MOAQ incrementally learns during the optimization process how to generate the compromise solutions. It is shown how MOAQ is used to design water distribution networks which are non-linear multiple objective optimization problems.

1 INTRODUCTION

It can be said that true optimization must be multiobjective; *real problems* usually have more than one function to be optimized, and often these functions cannot be combined into a single one. Historically, the application of classical mathematical methods in the solution of this kind of problems involves an objective function combination to form a scalar objective function, usually through a linear combination (weighted sum) of the multiple objectives, or by turning objectives into constraints. Aside from genetic algorithms, little research has been done, with the use of alternative optimization techniques to solve multiple objective problems. The use of genetic algorithms is often restricted by the need to represent problem constraints, often a difficult procedure. Reinforcement Learning, especially Ant-Q, has shown good results in combinatorial optimization problems [10, 11, 12]. In reinforcement learning [13], an agent learns from its own experiences, receiving rewards and punishments for its actions. In order to increase the exploration capabilities and to reduce convergence times, more than one agent can be used in the solution of a single problem, transforming the basic algorithm to a distributed reinforcement learning algorithm (Ant-Q). Ant-Q, however, has been used only in the solution of single objective optimization problems. Its distributed characteristics, theoretical background, and representation capabilities make it a good candidate for the solution of multiple objective optimization problems.

In this work, an Ant-Q based algorithm, called MOAQ (Multiple Objective Ant-Q), capable of solving multiple objective optimization problems, is presented. In particular, it is shown how MOAQ is applied to the design of irrigation networks. This problem has been attacked considering several assumptions to reduce it to a single objective optimization problem [1, 3, 4, 6, 7, 8, 9].

An integral solution for the design of irrigation water distribution networks, however, requires: (i) the selection of crops and cropped areas to optimize productivity and water usage, (ii) the design of the network layout, (iii) the establishment of the water requirements and minimum water pressures at each hydraulic node, and (iv) the selection of the pipe diameter for all links between nodes. In particular, compromise solutions of crop selection, to maximize profits can produce conflicting results if water availability and cropped surface constraints are not violated. The network design must also satisfy water volume and pressure requirements for each node while minimizing costs. Overall, the problem can be stated as a multiple objective optimization problem.

Section 2 describes in greater detail the objective functions and constraints involved in the design of irrigation networks. In Section 3, the proposed solution using MOAQ is described. Section 4 shows some experimental results and finally, conclusions and future research directions are given in Section 5.

2 DESIGN OF IRRIGATION WATER DISTRIBUTION NETWORKS

Suppose that we have a set of plots for cultivation and that for every plot there is a list of possible crops. In addition, each plot has a hydraulic node supplying water in the required quantities to the cultivated area. Water is supplied to the region by a supply node located somewhere in the region. Irrigation network design requires the definition of the network layout; this means to define the connections between nodes that can guarantee that all hydraulic nodes are connected to the supply node, and the lengths (links can have more that one pipe segment) and diameters of each pipe segment minimize the total pipe investment based on water demands and minimum pressure requirements. The water demand depends on the crop and the cropped area. Crop selection for each plot depends on irrigation water availability in the whole region. The water requirements are obtained through production functions for each crop where crop productivity is represented as function of the required water [2].

The irrigation water distribution network design problem can be stated as follows: *find the best set of crops for the plots that maximize regional productivity and water usage considering that the water to each plot must be transported the shortest distance to each hydraulic node over the minimal cost network that satisfies the pressure and flow restrictions.*

The problem can be formalized with the following objectives:

1. Minimize the cost of the network:

$$min\ z_1 = \sum_{l} \sum_{m} C(l,m)X(l,m)$$

where, $C(l,m)$ and $X(l,m)$ are, respectively, the cost per unit of pipe segment and length of the mTh.. segment of link l.

2. Maximize the profit:

$$max\ z_2 = \sum_{p=1}^{Np} [A_c(p)(\$_c P_c - C(p) - W(p))] \quad (1)$$

where, Np is the number of plots, $A_c(p)$ is the area in plot p to cultivate with crop c, $\$_c$ is the profitability of crop c, P_c is the production per unit of area of crop c, $C(p)$ are the input costs, $W(p)$ is the cost of the required water for plot p cultivated with crop c.

The solution is subject to the following constraints: (a) The variation in pressure ($H(l,m)$) in the mth segment of length l is equal to the hydraulic gradient in a pipe segment ($J(l,m)$) times its length:

$$\triangle H(l,m) = J(l,m)X(l,m) \quad (2)$$

(b) The required pressure at a particular node s connected to node n must be greater than the minimum required pressure at n (H_{min_n}) and sufficient to satisfy the resulting head losses in the water transport from s to n, but not greater than the maximum possible pressure at node s (H_{max}):

$$H_{min_n} \leq H_s \pm \sum_{l} \sum_{m} J_{l,m} X_{l,m} \leq H_{max_s} \quad (3)$$

(c) Water requirements must not be greater than the amount of water available for irrigation. ϕ is the amount of water from the total volume V_t that is available for irrigation, considering V_t is unrestricted in the region:

$$\sum_{p=1}^{Np} V_c(p) \leq \phi V_t$$

(d) The sum of cropped areas ($A_c(p)$) in the region must not be greater than the available area (A_t):

$$\sum_{p=1}^{Np} A_c(p) \leq A_t$$

(e) The volume of water to supply plot p ($V_c(p)$) planted with crop c must not be greater than that required per unit of area cultivated with the best crop option (D_p) times the maximum cultivated area ($A_{max,c}(p)$):

$$V_c(p) \leq D_p A_{max,c}(p)$$

(f) The supply of water for the selected crop for each plot ($D_c(p)$), considering the plot area ($A_c(p)$) cultivated, must not be greater than the capacity of the pipe segment conveyance plot p ($Q(p)$):

$$\sum_{p=1}^{Np} D_c(p) A_c(p) \leq Q(p)$$

(g) The profit from each plot must be greater than that obtained considering a reduction in the cropped area, in all the plots, proportional to the water reduction ratio:

$$A_c(p)\,(\$_c P_c - C(p) - W(p)) \geq$$
$$[A_{max,c}(p)\,(\$_c P_{max,c} - C(p) - W_{max}(p))]\,\phi$$

(h) The profit reduction produced by a crop selected for plots ($\xi(c,p)$), must be similar to that of the whole region:

$$\xi(c,p) = \frac{A_c(p)\,(\$_c P_c - C(p) - W(p))}{A_{max,c}(p)\,(\$_c P_c - C(p) - W_{max}(p))}$$

(i) The length ($L(l)$) of link l is equal to the sum of lengths ($X(l,m)$) of its m segments:

$$\sum_m X(l,m) = L(l)$$

(j) The segment lengths and the cropped areas must be positive:

$$X(l,m) \geq 0, A_c(p) \geq 0$$

The hydraulic gradient $J(l,m)$ is assumed to obey the Hazen-Williams equation; namely, $J(l,m)$ satisfies:

$$J(l,m) = \alpha \left(\frac{Q(l)}{C}\right)^{1.852} D(l,m)^{-4.87}$$

where α is a coefficient whose value depends on the units used, C is the Hazen-Williams coefficient, $Q(l)$ is the flow in link l, and $D(l,m)$ is the diameter of the $m-$th pipe segment in link l.

As can be seen from the above equations, the problem is non-linear due to the gradient head restrictions in Eq. 2 and Eq. 3 and to the objective function for the productivity, Eq. 1. Traditional mathematical programming techniques have been used, but only after making several rigorous assumptions. Some of them include considering the network layout independently of the overall network design, the linearization of the objective function in segments, and the assumption that all the network hydraulic parameters are given [1, 3, 4, 6, 7, 8, 9].

3 MOAQ MULTIPLE OBJECTIVE ANT-Q ALGORITHM

In this section, a multiple objective Ant-Q algorithm is presented. In particular, the algorithm is applied to the solution of irrigation network design. However, the approach can be extended to more objectives and to other optimization problems.

The basic idea behind MOAQ is to have a family of agents for each objective. Each family tries to optimize an objective considering the solutions found for the other objectives. In our algorithm, all the families must have the same number of agents, say m. As in Ant-Q, all the agents in one family try to find a solution at the same time. The m solutions found for one objective in one cycle influence the m successive starting points of the following family (objective). This process continues with all the families (objectives). Once all the families have been tried, a reward is given to the non-dominated solutions (to all the agents in each family responsible for that solution) that satisfied all the constraints. Results violating constraints are punished and the rest of the solutions are ignored. The whole process is repeated several times until only non dominated solutions (Pareto set) satisfying all the constraints are found or a predetermined number of iterations is satisfied (see table 1).

The main important points of the algorithm can be summarized as follows:

- The number of objectives is the same as the number of families
- The number of agents is the same in all the families
- The solution of one ant in one family (objective) affects (is used in) the solutions of corresponding agents in other families (objectives)
- Each family run once and rewards and punishments are given after all the families have been tried. Process is repeated until only non-dominated solutions.

Table 1: MOAQ algorithm

Given a list of n families (objectives) and
m agents for each family
Let N be the maximum number of iterations
Let $IN = 0$ (initialize iteration number)
Initialize all the Q values for all the families
Until only non-dominant solutions or $IN > N$
    let $IN := IN + 1$
    for $i = 1$ to n
        initialize parameters of family(i)
        for $j = 1$ to m
            map to ant j the solution found
            with ant j in objective $i - 1$
            find a solution for objective i
    evaluate solutions found
    for $j = 1$ to m
        if solution(j) violates any constraint
            apply punishment to all its components
        else if solution(j) is non dominated
            apply reward to all its components
            introduce solution(j) into Pareto set
            remove all dominated solution
            from Pareto set
        else if solution(j) is dominated
            neither apply reward nor punishment

The order given to the objectives depends on the particular problem. In some cases, the order can be arbitrary. In irrigation networks, it was necessary to determine first an initial solution for the crops in each plot to establish the water demands and then determine the network layout with this solution.

3.1 DESIGN OF IRRIGATION NETWORKS

For the solution of irrigation water distribution network design, two cooperating families of agents are proposed. The first family (*family1*) searches for the least cost network layout, while the second family (*family2*) selects crops and defines the optimal area to cultivate for each plot in the region. In particular, *family1* looks for the minimum distance network layout connecting all the hydraulic nodes that supply water to each plot in an agricultural region. The proposed network must guarantee that largest volume of water travels the shortest distance to minimize costs. Agents in *family2* select a crop for each plot region. This selection determines the water required for each plot and its productivity. This second family tries to maximize the global productivity without exceeding the irrigation water availability and without violating the restrictions of the problem. To generate compromise solutions a relation between the distance that water must travel from the supply to the node and water volume required in each plot is used to evaluate the connection possibilities. The most demanding nodes should be closer to the supply reducing with this the water transport rate in the links. The criteria where to move an ant in a family follows the approach given by [12].

3.2 OPTIMAL NETWORK LAYOUT

For *family1*, a heuristic value is used to evaluate the advisability benefit of connecting node x to node y:

$$H(x,y) = \frac{Q_T}{\$_N \cdot L_N}$$

where Q_T is total demanded flux calculated for the case when node y is connected to the network, $\$_N$ is the network cost when node y is connected to the network, and L_N is the network length when node y is connected to the network. The heuristic value favors connections that increase the total demanded network flux with the lowest cost and shortest length increments. In addition, a real value $Q(x,y)$ associated with each connection is used. The combination of these two parameters is used to select the next node to be connected whose selection is made under the following rule.

Considering a stochastic value $t \in [0,1]$ calculated on each iteration, and an initial t_s value equal to t_{max} that decreases with each iteration by a value determined with the formula $t_s(t) = \frac{t_s(t-1)*\lambda}{t_{max}}$, where $\lambda < 1$ we have:

$if\ t \geq t_s$

$$\max_{y \in no_net_k, x \in net_k} \{Q(x,y)^\beta H(x,y)^\delta\} \tag{4}$$

Otherwise, the next node is selected considering the probability of each node to be connected, calculated with the following equation:

$if\ x \in net_k \wedge y \in no_net_k$

$$\frac{Q(x,y)^\beta H(x,y)^\delta}{\sum_{z \in no_net_k} [Q(x,z)^\beta H(x,z)^\delta]} \tag{5}$$

where, k is the kth ant of *family1*, *net* and *no_net* are, respectively, lists of connected and disconnected nodes

for a particular agent. During the first iterations, exploration (with uniform distribution) is favored over exploitation, and this is gradually reversed on later iterations. β and δ are exponents representing the relative importance of their bases.

Learning Q-values are calculated using the following update rule:

$$Q(x,y) \leftarrow (1-\alpha)\cdot Q(x,y) + \alpha \cdot \left[r(x,y) + \gamma \cdot \max_{z \in no_net_k, p \in net_k} Q(p,z) \right]$$

where α is the learning step, γ is the discount factor, $r(x,y)$ is the reward given to the best solution found in each iteration calculated using the following rule:

$$r(x,y) = \begin{cases} \frac{1\times 10^6}{L_{n_dom} \cdot \$_{n_dom}} & (x,y) \in net_{n_dom} \\ 0 & otherwise \end{cases}$$

where L_{n_dom} is the length of the network layout evaluated as non dominated, $\$_{n_dom}$is the network cost. Rewards are given after all the agents in both families have been tried.

3.3 OPTIMIZATION OF THE IRRIGATED AREA

For each plot there is a list of possible crops. Each crop has its own productivity function that depends on the supplied water. As in the network layout optimization approach, a $Q(c,p)$ value is considered for all the crops in the list for each plot. A heuristic value, $\xi(C,P)$ is used to evaluate the superiority of crop with respect to the others. It evaluates the profit distribution, which must be similar in all the plots, avoiding the existence of predominant crops.

$$\xi(c,p) = \frac{A_c(p)\,(\$_c P_c - C(p) - W(p))}{A_{max,c}(p)\,(\$_c P_c - C(p) - W_{max}(p))} \forall c \in list$$

where, $list$ is a list of crop options for each plot p, that must be considered by ant k in *family2* during the optimization process. A numerical value for $\xi(c,p)$ can be determined in many ways. In this paper, $\xi(c,p)$ is initially assigned a value close to one, which means that there is no water restriction and the crop that generates the greatest economical benefits is selected for each plot. Normally, this crop consumes large quantities of water producing infeasible results, so the $\xi(c,p)$ value is gradually reduced until feasible solutions are found.

The action choice rule is similar to Eqs..4 and 5 . In this case it has the following form:

$if\ t \geq t_s$

$$\max_{c \in list_p} [Q(c,p) + X(c,p)] \tag{6}$$

Otherwise the next node is selected considering the probability of connecting the remaining nodes calculated by the following equation:

$if\ c \in list_p \wedge i \in list_p$

$$\frac{[Q(c,p) + X(c,p)]}{\sum_{i \in list_p} [Q(i,p) + X(i,p)]} \tag{7}$$

The Q values are updated using a similar procedure as in the optimal network layout.

$$Q(c,p) \leftarrow (1-\alpha)\cdot Q(c,p) + \alpha \cdot \left[r(c,p) + \gamma \cdot \max_{i \in list_p} Q(i,p) \right]$$

and the reward function is defined by:

$$r(c) = \begin{cases} profit \cdot \sum_j A_c(p) & c \in k_{n_dom} \\ 0 & otherwise \end{cases}$$

where *profit* is given by z_2 in Eq.. 1. In the previous reward function, crop c was selected for each plot by the agents evaluated as non dominated (k_{n_dom}).

The algorithm starts with the same number of agents for each family. Agents in *family2* generate feasible crop configurations and each one is assign them to the agents in *family1*. Given the crop configuration water demands can be calculated, network layouts produced and pipe diameters selected for each link in the network. At this stage, network costs can be determined and the network - crop configuration solutions evaluated. The evaluation is made considering non domination criteria in the Pareto sense. A reward, calculated with its corresponding equation, is given to the links and crops selected in all the non dominated solutions found on each iteration. Two punishment were considered. If constraints of a single element (plot or link) are violated, only the element is punished, but if generalized constraints are violated (constraints considering summation of values), all the elements will be punished. The punishment is calculated considering the normalized difference between the value for the proposed solution and the boundary value for the constraint. After the rewards and punishments are given, the whole process is repeated until only non-dominated solutions are found or a given number of iterations are occurred.

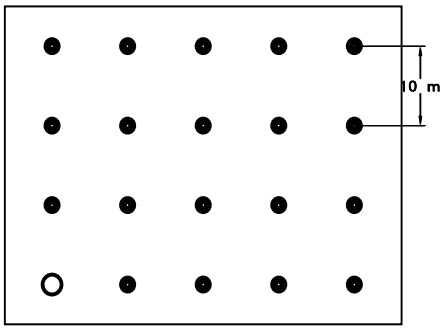

Figure 1: Hydraulic nodes confuguration, • are hydraulic nodes and ∘ is the supply node.

4 RESULTS

The algorithm was tested in the design of a 20 hydraulic node network configured as shown in Figure 1. The data for the plots under consideration are listed in table 2. Table 3 contains the data for the crops used in the multiple objective optimization problem. The parameters used in this work were: $\lambda = 0.9$, $\gamma = 0.3$, $\alpha = 0.1$, $\xi_{initial} = 0.9$, $\beta = 1.0$, $\delta = 2.0, T_{max} = 100$ for network optimization, $T_{max} = 200$ for crop selection and profit optimization, considering 10 agents for both families. Initial Q values for links in the network optimization where initialized using $Q = \sum_{i=1}^{N_l} (L(i) * \phi V_t)/(N_t^2 - N_t)$ for every link between nodes. Q values for crops in plots were initialized considering the profit divided by the total irrigated area times the water demand for the whole region given no water constraints ($Q = \sum_{i=1}^{N_p} P(i)/(\sum_{i=1}^{N_p} A(i) * \sum_{i=1}^{N_p} V(i))$) for each crop in the plots.

Figure 2 shows the final Pareto set in the objectives space. A better understanding of the algorithm behavior can be seen in Figure 3, where the fraction of non-dominated solutions in the set versus the iteration number is shown. In this case, the plot was constructed considering that each individual in the set of solutions obtained was compared against all the other solutions, including the solutions in the Pareto set. The fraction of non-dominated solutions has a continuous growing behavior. As the problem difficulty increases, more evaluations are required to get a large fraction of non-dominated solutions in the Pareto set. Once the maximum is obtained, it is maintained over the time, until the algorithm finishes. Thus, the algorithm behavior is stable for some number of iterations.

Table 2: Plot data

Plot No	x,y,z	Area (ha)
1	10,10,0	Supply
2	20,10,0	20
3	30,10,0	50
4	40,10,0	100
5	50,10,0	30
6	10,20,0	8
7	20,20,0	25
8	30,20,0	40
9	40,20,0	10
10	50,20,0	11
11	10,30,0	10.3
12	20,30,0	15.4
13	30,30,0	20.5
14	40,30,0	21
15	50,30,0	201
16	10,40,0	21
17	20,40,0	201
18	30,40,0	55
19	40,40,0	33
20	50,40,0	21

Table 3: Crop data (Taken from FAO 33).

Crop	CFC (m^3/(s ha))	Yield (ton/ha)	Price ($/ton)
Alfalfa	0.00697	28	10
Banana	0.00958	60	200
Bean	0.00218	2	50
Cabbage	0.00218	35	40
Citrus	0.00523	45	80
Cotton	0.00566	5	100
Grape	0.00348	30	300
Groundnut	0.00305	3	20
Maize	0.00348	9	20
Olive	0.00261	20	300
Onion	0.00239	45	20
Pea	0.00218	3	30
Pepper	0.00392	25	100
Pineapple	0.00436	90	10
Potato	0.00305	35	10
Rice	0.03048	8	5
Safflower	0.00523	4	400
Sorghum	0.00283	5	200
Sugarbeet	0.00327	60	50
Sugarcane	0.01088	100	45
Sunflower	0.00436	3.5	20
Tobacco	0.00261	2.5	20
Tomato	0.00261	65	50
Watermelon	0.00261	35	40
Wheat	0.00283	6	20

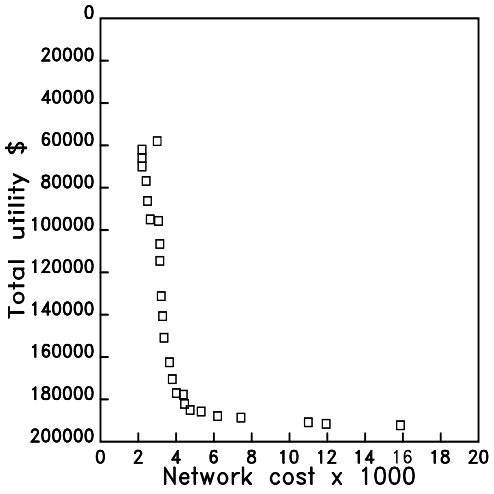

Figure 2: Final Pareto set.

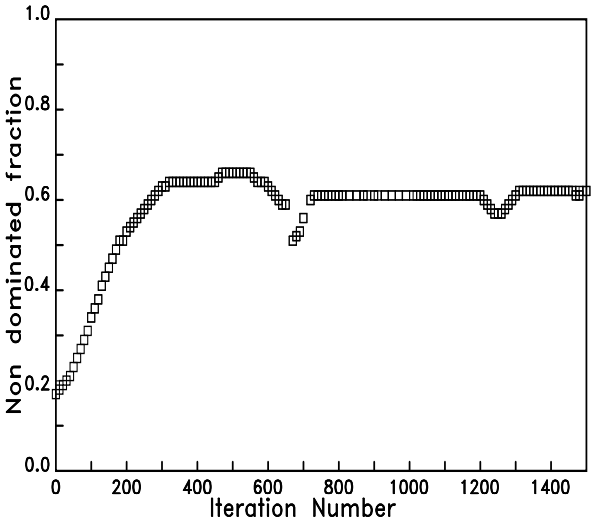

Figure 3: Non dominated solutions per iteration.

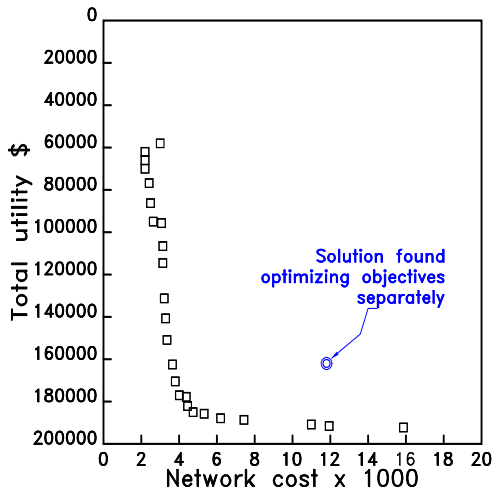

Figure 4: Solution found optimizing objectives separetely.

The design of irrigation water distribution networks is an extremely complex problem and, as mentioned in Section 1, it has been treated and solved considering very rigorous assumptions. With MOAQ, designs for irrigation networks were obtained considering not only the network layout design but also the crop selection and cultivated area. It is worth noting that all the solutions found in the Pareto set by MOAQ satisfied all the constraints, making them feasible in hydraulic and economic terms.

Figure 4 shows the network design solution obtained using distributed reinforcement learning considering the optimization problem separately. The crop selection problem was solved prior to the network optimization and considering the hydraulic parameters obtained with the solution for crops. When the solution in Figure 4 is compared with those contained in the Pareto set, Figure 2, many of the latter are much better for both objective functions. It can be stated that algorithm performance was good for the solution of non linear multiple objective optimization problems, as can be seen with the results reported.

On the other hand, MOAQ facilitates the representation of multiple objectives and all the restrictions involved without losing expressivity, making it applicable to other multiple objective problems without significant changes.

5 CONCLUSIONS AND FUTURE WORK

Many real world applications require compromise solutions for several conflicting objectives. The use of traditional mathematical methods has been severely restricted and alternative techniques need to be tried. In this paper, an Ant-Q based algorithm called MOAQ for the solution of multiple objective optimization problems has been presented. MOAQ uses a family of agents for each objective and by sharing partial results between agents, global compromise solutions can be found. In particular, MOAQ was used to solve design problems of irrigation water distribution networks.

As future work, we would like to test MOAQ with more objectives. In particular for the design of irrigation networks, we could try to minimize the length of the network with the following objective: $min\ z_3 = \sum_{i,j} L_{(i,j)}$. This is needed because the minimum cost network does not always has the minimum length. We would also like to test MOAQ on other multiple objective optimization problems, some of which are suggested in [5]and[14].

Acknowledgments

Authors gratefully acknowledges support by Consejo Nacional de Ciencia y Tecnología (CONACyT), Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM) Campus Morelos and Instituto Mexicano de Tecnología del Agua (IMTA).

References

[1] Alperovits, E., Shamir, U. (1977). Design of Optimal Water Distribution Systems, *Water Research Resources* 13(6): 885-900

[2] Doorenbos J., Kassam A.H. (1986). Yield Response to Water, *Food and Agriculture Organization of the United Nations, Irrigation and Dreinage paper 33*, Rome Italy, 1986.

[3] Eiger, G. Shamir, U., Ben-Tal, A. (1994). Optimal Design of Water Distribution Networks, *Water Research Resources* 30(9): 2637-2646.

[4] Fujiwara, O., Khang D.B. (1990). A two phase decomposition method for optimal design of looped water distribution networks, *Water Resources Research* 26(4): 539-549.

[5] Horn J., Naftopolis N., and Goldberg D. E. (1994). A Niched Pareto Genetic Algorithm for Multiobjective Optimization Problems. In *First IEEE Conference on Evolutionary Computation, IEEE World Congress on Computational Intelligence,* Vol. 1, 1994 (ICEC'94), Piscataway, NJ: IEEE Service Center, pp. 82-87.

[6] Labye, Y. Olson, M.A. Galand, A., Tsiourtis, N. (1988). Design and Optimization of Irrigation Distributions Networks, *FAO Irrigation and Drainage Paper 44*, Rome.

[7] Loganathan, G.V. Greene, J., Ahn, T.J. (1995). Design Heuristic for Globally Minimum Cost Water-Distribution Systems, *Journal of Water Resources Planning and Management* 121(2): 182-192.

[8] Ormsbee, L.E., Wood, D.J. (1986). Hydraulic Design Algorithms for Pipe networks, *ASCE Journal of Hydraulic Engineering* 112(12): 1195-1207.

[9] Quindry, G.E., Brill, E.D., Liebman, J.C. (1981). Optimization of Looped Distribution Systems, *ASCE Journal of Environmental Engineering* 107(EE4): 665-679.

[10] Dorigo, M., Maniezzo, V., Colorni, A. (1996). The Ant-System: Optimization by a colony of cooperating agents, *IEEE Transactions on Systems, Man and Cybernetics*, 26(1): 1-13 (Part B).

[11] Dorigo, M., Gambardella, L.M. (1997). Ant Colony System: Cooperative Learning Approach to the Traveling Salesman Problem, *IEEE Transactions on Evolutionary Computation*,1(1): 53-65, April 1997.

[12] Gambardella, L.M., Dorigo, M. (1995) Ant-Q: A Reinforcement Learning approach to the traveling salesman problem. In *Proceedings of the 12th. International Machine Learning Conference* (ML-95), pp., 252-260, Morgan Kaufmann.

[13] Sutton, R.S., Barto, A.G. (1998). *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, MA.

[14] Valenzuela-Rendón, M., Y Uresti-Charre, E. (1997). A non generational genetic algorithm for multiobjective optimization. In *Proceedings of the seventh international conference on genetic algorithms*, Morgan Kaufmann Publishers, San Mateo Ca. pp. 658-665.