

Ant Colony Optimisation for Continuous Domains with Aggregation Pheromones Metaphor

Shigeyoshi Tsutsui

Department of Management Information Sciences, Hannan University
5-4-33 Amamihigashi, Matsubara, Osaka 580-8502, Japan
e-mail: tsutsui@hannan-u.ac.jp

Abstract: *This paper describes an aggregation pheromone system (APS), which is an extension of ACO for continuous domains, using the collective behavior of individuals that communicate using aggregation pheromones. APS is tested on several test functions. Results show APS could solve real-parameter optimization problems fairly well. The sensitivity of control parameters of APS is also studied.*

Keywords: Evolutionary Computation, Ant Colony Optimization, Aggregation Pheromones, Parameter Optimization, Genetic Algorithms

1. Introduction

As a bio-inspired computational paradigm, ant colony optimization (ACO) has been applied with success to a large number of computationally hard problems. ACO simulates the collective behavior of ants, which communicate using pheromone trails. However, ACO is mainly applicable to discrete optimization problems such as the traveling salesman problem (TSP) [2, 5, 6, 19], quadratic assignment problem [14], scheduling problem [4, 7], vehicle routing problem [3], as well as the routing problem in telecommunication networks [21].

Although using pheromone trail metaphor is very effective in solving discrete optimization problems mentioned above, a direct application of the pheromone trail metaphor for solving real-parameter optimization problems is difficult. In this paper, we introduce "aggregation pheromones", also observed in nature, and propose an algorithm "aggregation pheromone system (APS)" for solving real-parameter optimization problems.

When an individual of the same species comes in contact with pheromone, it elicits a response, depending on the type of pheromone. In this way, specific information is conveyed. Pheromones that cause clumping or clustering behavior in a species, which bring individuals into a closer proximity, are referred to as aggregation pheromones [13]. Many functions of aggregation behavior have been observed. These include foraging-site marking and mating [1], finding shelter, and defense. Cockroaches produce a specific pheromone with their excrement when they find safe shelter, which attracts other members of their species [18]. As a result, aggregation pheromones function such that individuals aggregate around a "good position" with positive feedback. Pheromones evaporate with some rate. This prevents other individuals to aggregate to a local position.

The Aggregation Pheromone System (APS) proposed in this paper uses these aggregation pheromones as the basic metaphor of the model. While ACO is mainly applicable in discrete problems, APS can solve real-parameter optimization problems. Although we need more study, results show that it works fairly well in various test functions used in the evolutionary computation community. The remainder of this paper is organized as follows. We introduce the APS in Section 2, experimental analysis is given in Section 3, and Section 4 concludes the paper.

2. The Aggregation Pheromone System

Since the proposed algorithm is a variant of ACO and uses pheromone update rules similar to ACO, ACO is briefly described in Section 2.1. APS is presented in detail in Sections 2.2 and 2.3.

2.1 A Brief Overview of ACO

Here ACO, in this section ACO is briefly overviewed based on [5]. In ACO, search activities are performed with so-called *ants*. A moving ant lays a *pheromone trail* on the ground. An ant encountering a previously laid trail can detect it and decide with high probability to follow it, thus reinforcing the trail with its own pheromone.

In [5], the algorithm is called the *ant system* (AS). For a TSP, AS works as follows. Let m be the total number of ants. Each ant has the following characteristics: (i) it chooses the town to go to with a probability that is a function of the town distance and of the amount of trail present on the connecting edge; (ii) to force the ant to make a legal tour, transitions to already visited towns are disallowed until a tour is completed; (iii) when it completes a tour, it lays a trail on each edge (i, j) visited.

Let $\tau_{ij}(t)$ be the trail intensity on edge (i, j) at iteration t . At $t = 0$, initial values $\tau_{ij}(0)$ for trail intensity are set on edges. When all ants complete their tours, the trail intensity on each edge is updated according to the following formula:

$$\tau_{ij}(t+1) = \rho \cdot \tau_{ij}(t) + \Delta\tau_{ij} \quad (1)$$

where ρ ($0 \leq \rho < 1$) is a coefficient such that $(1 - \rho)$ represents the evaporation of trail between iteration t and $t+1$, $\Delta\tau_{ij} = \sum_{k=1}^m \Delta\tau_{ij}^k$ and $\Delta\tau_{ij}^k$ is the quantity per unit of length of trail laid on the edge (i, j) by the k -th ant between iteration t and $t+1$; it is determined to be inversely proportional to the tour length of k -th ant so that larger values of $\Delta\tau_{ij}^k$ are given if the tour length of the k -th ant is shorter. This process is iterated until the given termination conditions are satisfied.

2.2 Basic Model of the Aggregation Pheromone System

There is a big difference between ACO and APS in how the pheromone functions in the search space. Pheromone density in ACO is defined as a trail on an edge between nodes of a given sequencing problem. In APS, the aggregation pheromone density is defined by a density function in search space X in R^n . In the real world, aggregation pheromones are used by a species to communicate with members of their community to share information about the location of food, safe shelter, potential mates, or enemies.

As in ACO, APS takes a cycle model. In each APS cycle, borrowed from the natural model, m individuals are attracted to positions by the aggregation pheromone in search space X . They are attracted more to the positions where the pheromone density is higher, and less to the positions where the density is lower. Thus, in our system, we consider individuals to choose their positions depending on the aggregation pheromone density, or more specifically, with probability proportional to the aggregation density function in search space X as follows.

Let x represent the variable in search space X , i.e. $x \in X$, and $\tau(t, x)$ be the density function of aggregation pheromone at APS cycle t . In initial APS cycle ($t=0$), the aggregation pheromone distributes uniformly, i.e., $\tau(0, x) = c$, where, c is a constant. The probability density function of aggregation pheromone at APS cycle t , $p_t(t, x)$, is defined as

$$p_t(t, x) = \frac{\tau(t, x)}{\int_X \tau(t, x) dx} \quad (2)$$

Depending on the problem we solve, a fitness function $f(x)$ is assigned for $x \in X$. Each of m individuals emits aggregation pheromone around its position x in X depending on the fitness function $f(x)$. There may be questions regarding which approach is more similar to that used in the natural world, to use an absolute value of $f(x)$, a relative value, or a ranking system. However, in our system we use rank r ($r = 1, 2, \dots, m$) of each individual to make the fitness difference among individuals more distinguishable. The highest ranked individual has a value of m and the lowest ranked individual has a value of 1. An individual which has rank r emits the aggregation pheromone around $x_{t,r}$ with density function represented by $\Delta\tau'(t, r, x_{t,r}, x)$, where $x_{t,r}$ is the position of the individual with rank r . Then the new aggregation pheromone density emitted in APS cycle t by m individuals is

$$\Delta\tau(t, x) = \sum_{r=1}^m \Delta\tau'(t, r, x_{t,r}, x) \quad (3)$$

Here, we assume the total aggregation pheromone volume emitted by m individuals in each cycle t is equal to C , i.e.,

$$\int_X \Delta\tau(t, x) dx = C. \quad (4)$$

This assumption is important when we consider the sampling method in Section 2.3. When all m individuals complete choosing their positions, the total aggregation pheromone density in X is updated according to the following formula, as in Eq. 1:

$$\tau(t+1, x) = \rho \cdot \tau(t, x) + \Delta\tau(t, x). \quad (5)$$

After the pheromone updating is performed, individuals are reset and the next APS cycle starts. This process is iterated. Thus, the pheromone density at promising points in the space increases as the APS cycle is iterated and the APS is expected to converge to a promising solution.

Here, we discuss the form of $\Delta\tau'(t, r, x_{t,r}, x)$ of Eq. 3. We assume that (i) $\Delta\tau'(t, r, x_{t,r}, x)$ is centered around position $x_{t,r}$ (position $x_{t,r}$ has the highest density), (ii) an individual with higher rank emits more pheromone than an individual with lower rank. Further, (iii) we introduce a co-operative phenomenon in which $\Delta\tau'(t, r, x_{t,r}, x)$ is affected by the distribution of other individuals and becomes elongated in the direction of the distribution. In this research, we use the following Gaussian functional form for $\Delta\tau'(t, r, x_{t,r}, x)$ ($r=1, 2, \dots, m$) for Eq. 3.

$$\Delta\tau'(t, r, x_{t,r}, x) = \frac{C}{\sum_{k=1}^m k^\alpha} r^\alpha N(x_{t,r}, \beta^2 \Sigma_t). \quad (6)$$

Here α (>0) is a parameter to adjust the relative importance of rank, Σ_t is the covariance matrix estimated from distribution of m individuals in search space X for cycle t , β ($\beta > 0$) is a parameter to control the width of distribution of the pheromone, $N(x_{t,i}, \beta^2 \Sigma_t)$ is a multivariate normal distribution.

The function defined by Eq. 6 satisfies all the three conditions mentioned above. It has the highest value at $x_{t,r}$, higher rank individuals have larger functional values, and covariance matrix $\beta^2 \Sigma_t$ reflects a distribution of individuals. With larger values of α , the ratio of the total amount of pheromone produced by a higher ranked individual increases. With larger values of β , the aggregation pheromone spreads more widely in the search space X , and with smaller values of β , aggregation pheromone spreads less. Then, $\Delta\tau(t, x)$, the total pheromone density emitted by m individuals in cycle t , can be obtained from Eq. 3. We can see also that $\Delta\tau(t, x)$ satisfies Eq. 4, i.e., the total amount of aggregation pheromone at each cycle equals C .

2.3 Sampling Technique

In this subsection, we discuss how to sample individuals from aggregation pheromone density function $\tau(t+1, x)$ obtained in Eq. 5. As described in 2.2, we sample individuals with probability proportional to the aggregation density $\tau(t+1, x)$. To perform the proportional sampling, we need to obtain probability density function $p_\tau(t+1, x)$ from $\tau(t+1, x)$. $\tau(t+1, x)$ in Eq. 5 can be rewritten as

$$\tau(t+1, x) = \rho^{t+1} \tau(0, x) + \sum_{h=0}^t \rho^h \Delta\tau(t-h, x). \quad (7)$$

From Eqs. 2, 4, and 7, $p_\tau(t+1, x)$ is obtained as

$$p_\tau(t+1, x) = \frac{\rho^{t+1}}{\sum_{k=0}^{t+1} \rho^k} \cdot \frac{\tau(0, x)}{C} + \sum_{h=0}^t \frac{\rho^h}{\sum_{k=0}^{t+1} \rho^k} \cdot \frac{\Delta\tau(t-h, x)}{C}. \quad (8)$$

In general, if a probability density function (pdf) $f(x)$ can be decomposed into sub-pdfs $f_k(x)$ as $f(x) = p_1 f_1(x) + p_2 f_2(x) + \dots + p_S f_S(x)$, then we can sample $f(x)$ as follows: first we choose $f_s(x)$ ($s=1, 2, \dots, S$) according to probability (p_1, p_2, \dots, p_S) , then we sample $f_s(x)$, where, (p_1, p_2, \dots, p_S) is a probability distribution. Using this method, we can sample $p_\tau(t+1, x)$ of Eq. 8 in the following manner. In Eq. 8, the terms $\tau(0, x)/C$ and $\Delta\tau(t-h, x)/C$ can be pdfs from Eq. 4. Thus, first we choose $\tau(0, x)/C$ or $\Delta\tau(t-h, x)/C$ ($h=0, 1, \dots, t$) according to probability $\rho^h / \sum_{k=0}^{t+1} \rho^k$ ($h=0, 1, \dots, t, t+1$). We call this *cycle sampling*. Next, we sample $\tau(0, x)/C$ or $\Delta\tau(t-h, x)/C$. Sampling of $\tau(0, x)/C$ is simple because it is a uniform sampling. Sampling of $\Delta\tau(t-h, x)/C$ is similar to cycle sampling. Generally, $\Delta\tau(t-h, x)/C$ can be obtained from Eq. 6 as

$$\frac{\Delta\tau(t-h, x)}{C} = \sum_{r=1}^m \frac{r^\alpha}{L} N(x_{t-h,r}, \beta^2 \Sigma_{t-h}), \quad (9)$$

where $L = \sum_{k=1}^m k^\alpha$. As with cycle sampling, we first choose rank r according to the probability r^α / L ($r=1, 2, \dots, m$). We call this sampling *rank sampling*. After rank sampling, we sample $N(x_{t-h,r}, \beta^2 \Sigma_{t-h})$, using Cholesky decomposition [20]. To perform this sampling, based on Eq. 8, we

need a large memory to store $x_{t-h,r}$ vector values and covariance matrix $\beta^2 \Sigma_{t-h}$ when APS cycle t becomes large. In this case $\rho^h \rightarrow 0$ for large h since $0 < \rho < 1$. Thus, we can limit the maximum number of cycles to keep data up to a constant H . Then, Eq 8 for $t \geq H$ can be represented as

$$p_\tau(t+1, x) = \sum_{h=0}^{H-1} \frac{\rho^h}{\sum_{l=0}^{H-1} \rho^l} \cdot \frac{\Delta\tau(t-h, x)}{C}. \quad (10)$$

3. Experimental Study

3.1 Experimental Methodology

In the APS cycle model, we reserve the best e ($e = m \times E_{\text{rate}}$) individuals at each cycle. They are transferred to the next cycle. This acts as an elitist strategy in GAs [8]. Parameter E_{rate} is used to control the number of elites in each cycle. We use the following three test functions: the Ellipsoidal function ($F_{\text{Ellipsoidal}}$), the Ridge function (F_{Ridge}), and the Rosenbrock function ($F_{\text{Rosenbrock}}$). F_{Ridge} has weak linkage among variables. $F_{\text{Rosenbrock}}$ has strong linkage among variables. $F_{\text{Ellipsoidal}}$ has no linkage among variables. Problem size $n = 20$ is used for all test functions.

$$F_{\text{Ellipsoidal}} = \sum_{i=1}^n ix_i^2, \quad (-3.12 \leq x_i < 7.12), \quad (11)$$

$$F_{\text{Ridge}} = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2, \quad (-44 \leq x_i < 84), \quad (12)$$

$$F_{\text{Rosenbrock}} = \sum_{i=2}^n (100(x_1 - x_i^2)^2 + (x_i - 1)^2), \quad (-2.048 \leq x_i < 2.048), \quad (13)$$

The default value of the parameters is as follows: $m = 100$, $\rho = 0.92$, $\alpha = 4$, $\beta = 0.6$, $E_{\text{rate}} = 0.1$, and $H=200$. These default values were chosen to tune APS to work well when tested with the $F_{\text{Rosenbrock}}$ function. We evaluated APS by measuring their #OPT (number of runs in which APS succeeded in finding the global optimum) and MNE (mean number of function evaluations to find the global optimum in those runs where it did find the optimum). We considered the detection of the solution successful if all parameters (x_1, \dots, x_n) of the best individual were within the range $[(o_j - 0.0001), (o_j + 0.0001)]$ for all j , where (o_1, \dots, o_n) is the optimal point. 20 runs are performed. Each run continued until the global optimum is found or a maximum of 500,000 evaluations is reached.

3.2 Results

(1) Results with default parameter values

Table 1 shows the results of APS using the default parameter values described in Section 3.1. The results are compared with the results of SPX, a typical state-of-art crossover operator in real-coded GAs [9]. APS showed better results than SPX on $F_{\text{Ellipsoidal}}$ and F_{Ridge} , although it is possible to further tune APS as described below (see Fig. 3). On $F_{\text{Rosenbrock}}$, APS showed much better results than SPX. Fig. 2 shows the convergence processes of both APS and SPX on $F_{\text{Rosenbrock}}$. We can see that APS converges more stably and rapidly than SPX.

Table 1 Results of APS with default values

Function	$F_{\text{Ellipsoidal}}$	F_{Ridge}	$F_{\text{Rosenbrock}}$	
APS	#OPT	20	20	
	MNE	80300.0	101660.0	121760.0
	STD*	1469.7	1765.4	8882.3
SPX	#OPT	20	20	20
	MNE	138900.0	179078.3	265766.3
	STD*	1362.2	2337.1	27152.5

* STD: Standard Deviation

(2) Sensitivity of parameters

To see the effect of the evaporation coefficient ρ in APS, we varied the value of ρ from the default value in the range $[0.8, 0.98]$ (Fig. 3). When the value of ρ becomes larger, pheromone emitted in previous cycles remains for a longer period of time. With $F_{\text{Ellipsoidal}}$ and F_{Ridge} , APS with ρ values in $[0.84, 0.86]$ showed #OPT = 20 and the better (smaller) MNE values. With $F_{\text{Rosenbrock}}$, which has strong linkage among variables, APS with ρ value 0.92 showed smaller MNE with #OPT = 20.

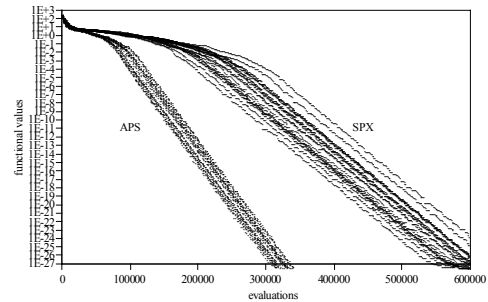


Fig. 2 Convergence process on $F_{\text{Rosenbrock}}$

For this problem the value of ρ is larger than those resulting from $F_{\text{Ellipsoidal}}$ and F_{Ridge} . With larger values of ρ , APS uses information of the pheromone density in the past cycles more than with smaller values of ρ and can find solution efficiently on functions that have strong linkage among variables.

To see the effect of the parameter α , which adjusts the relative importance of rank when each individual emits pheromone, we tested for α in the range of [2, 8] with step size 2 (Fig. 4). With larger values of α , higher ranked individuals emit pheromone at

increasing rates. With larger values of α , the performance of APS increased, as seen in the results with $F_{\text{Ellipsoidal}}$ and F_{Ridge} . However, as seen in the results with $F_{\text{Rosenbrock}}$, which has strong linkage among variables, larger α values cause side effects in convergence.

Parameter β controls the width of distribution of the pheromone emitted by individuals. With increasingly larger values of β , individuals emit pheromone more. We tested β in the range of [0.5, 0.7] with step size 0.02 (Fig. 5). With function $F_{\text{Ellipsoidal}}$ and F_{Ridge} , having respectively no and weak linkage among variables, smaller values of β have better performance. With function $F_{\text{Rosenbrock}}$, which has a strong linkage among variable, both smaller and larger values of β have poor performance, showing over-exploring and over-exploiting behavior.

4. Conclusions

In this paper, we have described the *aggregation pheromone system (APS)*, which uses *aggregation pheromones* as the basic metaphor of the model, for solving real-parameter optimization problems. We studied APS with several test problems. The results showed that APS could solve real-parameter optimization problems fairly well. We also explored the sensitivity of the control parameters of APS.

There are many opportunities for further research related to APS. These include study of the relationship between the parameters and the resulting performance of APS, the scalability analysis of the algorithm, comparative study with evolutionary algorithms such as UNDX [15], SPX [9], and EDAs (estimation of distribution algorithms) [11, 16].

Acknowledgements

The author would like to thank Associate Professor Gordon Wilson, Hannan Univ. for his valuable comments on this paper. This research is partially supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan under Grant-in-Aid for Scientific Research number 16500143.

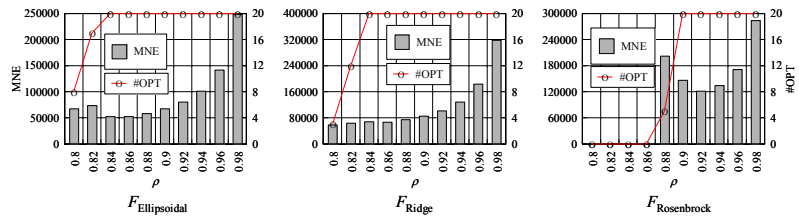


Fig. 3 Effect of the coefficient of pheromone evaporation ρ

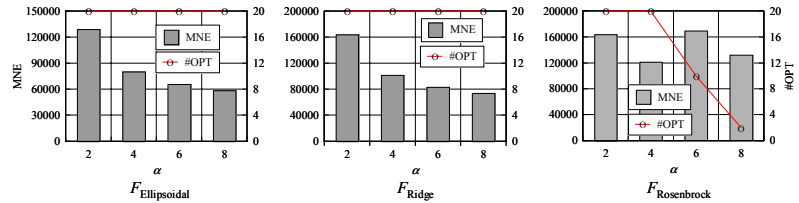


Fig. 4 Effect of parameter α

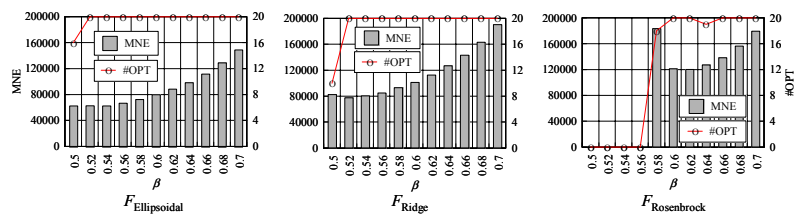


Fig. 5 Effect of parameter β

References

- [1] Bell, W. J.: Chemo-orientation in Walking Insects, *Chemical Ecology of Insects*, Bell, W. J. and Carde, R. T. Eds, 93-109 (1984).
- [2] Bullnheimer, B., Hartl, R. F., and Strauss, C.: A New Rank Based Version of the Ant System: A Computational Study, *Central European Journal for Operations Research and Economics*, 7(1):25-38 (1999).
- [3] Bullnheimer, B., Hartl, R. F., and Strauss, C.: Applying the Ant System to the Vehicle Routing Problem, *Meta-Heuristics: Advances and Trends in Local Search Paradigms for Optimization*, Voss, S., et al (Eds.), Kluwer (1999).
- [4] Costa, D. and Hertz, A.: Ants Can Colour Graphs. *Journal of the Operational Research Society*, 48:295-305 (1997).
- [5] Dorigo M., Maniezzo, V., and Colorni, A.: The Ant System: Optimization by a Colony of Cooperating Agents, *IEEE Trans. on SMC-Part B*, 26(1):29-41(1996).
- [6] Dorigo M. & L.M. Gambardella (1997). Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem. *IEEE TEC*, 1(1):53-66 (1997).
- [7] Forsyth P. and Wren, A.: An Ant System for Bus Driver Scheduling. *Proc. of the 7th Int. Workshop on Computer-Aided Scheduling of Public Transport* (1997).
- [8] Goldberg, D. E.: *Genetic algorithms in search, optimization and machine learning*. Reading, MA: Addison-Wesley publishing company (1989).
- [9] Higuchi, T., Tsutsui, S., and Yamamura, M.: Theoretical analysis of simplex crossover for real-coded Genetic Algorithms, *Proc. of the PPSN VI*, 365-374 (2000).
- [10] Kuntz P., Layzell, P., and Snyers, D.: A Colony of Ant-like Agents for Partitioning in VLSI Technology. *Proc. of the 4th European Conf. on Artificial Life*, 417-424, MIT Press (1997).
- [11] Larranaga, P. and Lozano, J. A. (eds): *Estimation of distribution algorithms*, Kluwer Academic Publishers (2002).
- [12] Law, J. H. and Regnier, F. E.: *Annual Review of Biochemistry*, 40:533-548 (1971).
- [13] Lorenzo Figueiras, A.N. and Lazzari, C. R.: Aggregation behavior and interspecific responses in three species of Triatomine, *Memórias do Instituto Oswaldo Cruz* . 93(1):133-137 (1998).
- [14] Maniezzo, V.: Exact and approximate nondeterministic tree-search procedures for the quadratic assignment problem, *CSR 98-1*, Scienze dell'Informazione, Università di Bologna (1998).
- [15] Ono, I. and Kobayashi, S.: A Real-Coded Genetic Algorithm for Function Optimization Using the Unimodal Normal Distribution Crossover, *Proc. of the 7th ICGA*, 246-253, (1997).
- [16] Pelikan, M., Goldberg, D. E., and Lobo, F.: A survey of optimization by building and using probabilistic models. *Computational Optimization and Applications*, 21(1), 5-20 (2002). Also Technical Report IlliGAL Report 99018, University of Illinois at Urbana-Champaign (1999).
- [17] Pena, J. M., Lozano, J. A., and Larranaga, P.: Benefit of Data Clustering in Multimodal Function Optimization via EDAs, *Estimation of Distribution Algorithms*, Larranaga, P. and Lozano, J. A. (Eds), Kluwer Academic Publishers, Chap. 10:211-229 (2002).
- [18] Sakuma, M. and Fukami, H.: Aggregation arrestant pheromone of the German cockroach, *Blattella germanica* (L.) (Dictyoptera: Blattellidae): isolation and structure elucidation of blastellastanoside-A and -B, *Journal of Chem. Ecol.* 19:2521-2541 (1993).
- [19] Stützle, T. and H. Hoos, H.: The MAX-MIN Ant System and local Search for Combinatorial Optimization Problems: Towards Adaptive Tools for Global Optimization, *Proc. of the 2nd Metaheuristics Int. Conf. (MIC-97)* (1997).
- [20] Schatzman, M., Taylor, J., and Schatzman, M.: *Numerical Analysis: A Mathematical Introduction*, Oxford Univ Press (2002).
- [21] Schoonderwoerd, R., Holland, O., Bruten, J., and L. Rothkrantz, L.: Ant-based Load Balancing in Telecommunications Networks. *Adaptive Behavior*, 5(2):169-207 (1997).