The small world of pheromone trails

Paola Pellegrini, Andrea Ellero

Department of Applied Mathematics
University Ca' Foscari of Venice
Dorsoduro 3825/E, I–30123 Venice, Italy
paolap@pellegrini.it, ellero@unive.it

Abstract. In this paper we consider $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System and Ant Colony System. They are generally recognized to be the best performing algorithms of the Ant Colony Optimization family. They are characterized by a quite different way for dealing with the pheromone trail. We propose an experimental analysis for observing whether this difference impacts significantly on the characteristics of the pheromone distributions produced during the runs. The results obtained are analyzed by using some concepts derived by the literature on small-world networks. It comes out that ants actually build small-world pheromone graphs during their runs. This behavior is interpreted here as a sort of decomposition of the instances tackled.

Keywords: $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System, Ant Colony System, small-world graphs, problem decomposition.

1 Introduction

In this paper we consider two Ant Colony Optimization (ACO) algorithms: $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System ($\mathcal{MM}$AS) and ant colony system (ACS). They are recognized as the two best performing procedures of the Ant Colony Optimization family [1]. They are characterized by a pretty different behavior with respect to pheromone distribution and exploitation, as described in the following.

The aim of this paper is analyzing the effect of such a difference in the way the two algorithms deal with the distribution of the pheromone trail on solution components. In order to study this distribution, we use two elements coming from the literature on the small-world phenomenon. This concept originally belongs to the field of social sciences: A social network exhibits the small-world phenomenon if, roughly speaking, any two individuals in the network are likely to be connected through a short sequence of intermediate acquaintances [2–4]. Shifting this definition to mathematics, a small-world network is a graph in which nodes are neighbors of few others, but most nodes can be reached from every other in few steps. Moreover, the network is highly clustered, in the sense that if two vertices are neighbors of a third one, with high probability they will be connected. The small-world phenomenon is of particular interest here, since its definition allows us to interpret the evolution of a run (in terms of pheromone distribution) as some kind of splitting an instance in sub-instances. In this sense, each cluster represents a sub-instance: when it is reached, various ways are available for visiting the neighborhood, while when it is left only few paths are likely to be chosen.

In order to make this concept more explicit, and to observe the difference in the pheromone structure imposed by $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System and Ant Colony System, the rest of the paper is organized as follows: in Section 2 the two algorithms are described, while in Section 3 some elements on small-world graphs are summarized. Sections 4 and 5 report the experimental setup and results, and Section 6 concludes the paper.

2 $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System and Ant Colony System

Let us consider the generic combinatorial optimization problem to be solved as mapped on an edge-weighted graph $G = (N, E)$, with $N = \{1, 2, \cdots, n\}$ set of nodes, and $E = \{(i, j) : i, j \in N\}$ set of edges. The graph is such that each solution to the combinatorial optimization problem corresponds to at least one path on the graph itself. The weights associated to the edges are such that the cost of a path (V_s) equals the cost of the associated solution s.

In $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System [5], the state transition rule is named *random-proportional rule.* The probability of choosing the generic edge (i, j) is biased by the heuristic information η_{ij} and by the pheromone trail τ_{ij}. The relative weight of these values are controlled by two parameters of the algorithms, namely α and β.

After the activity of m ants, the pheromone evaporates on all the edges of the graph. The amount of this evaporation depends on the value of parameter ρ, $0 < \rho < 1$. Moreover, some pheromone is deposited on those used in a specific solution: One solution is used for the reinforcement. It is either the *iteration best* one, i.e. the best among the last m solutions built, or the *best so far* one, i.e. the best solution constructed. Finally, the pheromone is constrained between a lower and an upper bound τ_{min} and τ_{MAX}.

In ACS [6], the *pseudorandom-proportional rule* is applied: With probability q_0, the step to perform is the one with the highest combination of the values of the heuristic information η_{ij} and by the pheromone trail τ_{ij}. Only the former is weighted by using an exponent β. With probability $1 - q_0$ the random proportional rule is applied (with $\alpha = 1$). The pheromone update is performed both after each iteration (*global update*) and after each ant has added one component to the solution under construction (*local update*). In the global pheromone update, the pheromone changes only on the edges belonging to the best so far solution. Its evaporation and deposit are controlled by parameter ρ. In the local pheromone update, after ant k has included edge (i, j) in the path, the pheromone level on (i, j) itself is updated, and the role of ρ is taken by another parameter ξ.

The two algorithms are not hybridized with any local search procedure. This choice is owed to the fact that we are interested in the nature of the algorithms themselves rather than in the absolute quality of the results. For a more detailed description of the procedures, we refer the reader to the specific literature.

3 Small-world graphs

For observing the characteristics and the evolution of the pheromone trails on solution components, we analyze the structure of the graph we obtain considering all the nodes of the original one $G = (N, E)$, and the subset of edges on which the pheromone level is above the average: $G' = (N, E'), E' \subseteq E : \tau_{i,j} \geq |E|^{-1} \sum_{(i,j) \in E} \tau_{i,j}, \forall (i,j) \in E'$. G' will be referred to as pheromone graph. We study the structure of the pheromone graphs built during a run of the algorithm. This is done by using two measures that are typical of the literature on small-world networks [7], namely the characteristic path length (L) and the clustering coefficient (C). The former is defined as the average number of edges that must be traversed in the shortest path between pairs of nodes in the graph. The latter, instead, describes the neighborhood structure of the graph. In particular, if a node v is directly connected to k_v vertices, then this neighborhood defines a subgraph in which at most $k_v(k_v - 1)/2$ edges can exist. The clustering coefficient of v (C_v) is the ratio between this maximum and the number of edges that actually compose the subgraph. The average clustering coefficient is the average of this measure over all the nodes of the graph.

In order to state whether a graph G with n vertices exhibits the small-world property, its characteristic path length and clustering coefficient are compared to those of a random graph with the same number of nodes. In the latter $kn/2$ out of all possible $n(n-1)/2$ edges are chosen at random with equal probability (k indicates the average degree of the vertices of G). In such a random graph, an asymptotic approximation of the two measures we are considering are: $L_{random} \sim \ln(n)/\ln(k)$, $C_{random} \sim k/n$ [8]. If $L \sim L_{random}$ and $C \gg C_{random}$, the graph under analysis is a small-world network. In words, it represents a clustered structure and, in average, a quite short path is sufficient to move from one node to another. In the following we will consider the ratios L/L_{random} (L ratio) and C/C_{random} (C ratio) in order to analyze the presence of the small-world property. Usually a graph is recognized to have this property if L ratio ~ 2.5 or lower and C ratio ~ 5 or higher [9–12].

To our aim, a small-world pheromone graph may be interpreted as the algorithm splitting an instance in sub-instances: computational resources are used for focusing on the clusters. Once one node of a cluster is reached, the possibility for visiting its neighbors are various, while, when a cluster is left, only few paths have high probability of being selected.

In the analysis proposed here, we use these concept for observing whether the pheromone graphs created during a run by $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System and Ant Colony System allow to think to this kind of problem decomposition.

Table 1. Values tested for the parameters.

$\mathcal{MAX}$–$\mathcal{MIN}$ Ant System

parameter	values tested
m	50, 100, 150, 200
β	2, 3, 4, 5, 6
ρ	0.02, 0.07, 0.12, 0.17, 0.22, 0.27, 0.32
α	1, 2, 3, 4

Ant colony system

parameter	values tested
m	5, 10, 20, 50
β	2, 3, 4, 5, 6
ρ	0.16, 0.25, 2.33, 0.42, 0.5
q_0	0.75, 0.8, 0.85, 0.9
ξ	0.1, 0.2

Table 2. Values chosen for the parameters after the tuning procedure.

Fairly explorative parameters $\mathcal{MM}$AS

n	m	ρ	β	α	n	m	ρ	β	α
200	50	0.17	4	1	400	50	0.22	4	1
300	50	0.22	4	1	500	50	0.32	4	1

Fairly explorative parameters ACS

n	m	ρ	β	q_0	ξ	n	m	ρ	β	q_0	ξ
200	20	0.33	4	0.75	0.1	400	10	0.16	4	0.75	0.1
300	20	0.25	4	0.75	0.1	500	20	0.25	4	0.75	0.1

Explorative parameters $\mathcal{MM}$AS

n	m	ρ	β	α	n	m	ρ	β	α
200	150	0.07	3	1	400	150	0.12	3	1
300	150	0.07	3	1	500	150	0.12	4	1

Explorative parameters ACS

n	m	ρ	β	q_0	ξ	n	m	ρ	β	q_0	ξ
200	50	0.42	4	0.75	0.1	400	20	0.25	4	0.75	0.1
300	20	0.33	4	0.8	0.1	500	20	0.42	5	0.75	0.1

4 Experimental setup

The classical combinatorial optimization problem called traveling salesman problem (TSP) is considered in the experiments. Four sets of instances are tackled, with 200, 300, 400 and 500 nodes respectively.

For both $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System and Ant Colony System, we consider two configurations of parameters for each set of instances: a fairly explorative configuration and an explorative one. In order to fix them, we apply the tuning procedure F-Race [13]. The stopping criteria adopted is the total number of objective function evaluations (which is proportional to computational time). The reasoning is based on the fact that ACO algorithms, in order to get good performances, need to devote resources to both exploration and exploitation. If the available time is short, good parameters should imply a level of exploration such that some time is left to exploitation. On the other hand, if the computational time is longer, the exploration level can be higher, still allowing for exploitation [14]. Therefore, the maximum number of objective function evaluations allowed for a run is varied: we consider first a short run (25000 evaluations) in order to get a fairly explorative configuration, and then a long one (175000 evaluations) in order to get a more explorative one.

All the combinations of the values reported in Table 1 are considered for the tuning procedure. These values are chosen on the basis of the literature on the two algorithms [1]. Five hundred instances are used for the tuning phase. The values selected after the tuning procedure are reported in Table 2.

5 Experimental results

In the experiments proposed, we apply the two algorithms presented in Section 2 to the traveling salesman problem. Ten TSP instances are used for each set, with 200, 300, 400 and 500 nodes respectively. The number of objective function evaluations considered as stopping criterion is 175000. Beside observing the quality

Table 3. Number of objective function evaluations after which the pheromone graph loses the small-world property (results are recorded each 10 iterations).

Fairly explorative parameters

inst	n=200		n=300		n=400		n=500	
	$\mathcal{MM}$AS	ACS	$\mathcal{MM}$AS	ACS	$\mathcal{MM}$AS	ACS	$\mathcal{MM}$AS	ACS
1	3500	13800	3500	12000	4000	25600	2500	88800
2	3000	3200	2500	1200	3500	26800	2500	31200
3	4000	10200	2500	2600	7500	71400	3000	15800
4	3500	4300	3000	2600	3500	101400	3500	45600
5	3000	4400	3000	19600	4000	44200	3000	33600
6	4000	3000	2500	14000	4000	23200	2500	27500
7	3500	1600	3500	24600	4500	44800	3000	16700
8	4000	1200	4000	18600	5000	13200	3000	15200
9	4000	7800	3000	18400	4000	92600	2500	25000
10	4000	6000	4000	20000	4500	22000	3500	30000

Explorative parameters

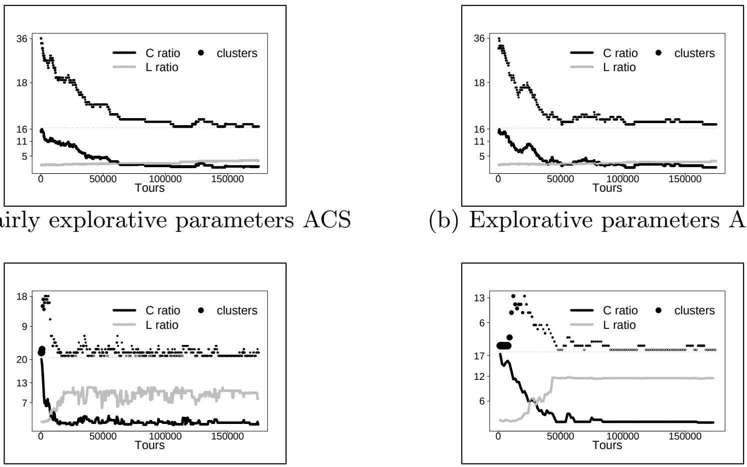
inst	n=200		n=300		n=400		n=500	
	$\mathcal{MM}$AS	ACS	$\mathcal{MM}$AS	ACS	$\mathcal{MM}$AS	ACS	$\mathcal{MM}$AS	ACS
1	24000	43200	55500	19400	24000	56800	40500	20600
2	27000	2000	31500	30900	18000	35200	22500	15200
3	30000	27800	45000	22500	21000	32000	30000	14600
4	34500	45300	37500	32400	22500	42200	24000	21700
5	36000	1200	46500	35100	30000	15600	21000	15400
6	36000	200	52500	43600	46500	35400	27000	9700
7	36000	1200	51000	25000	33000	43400	25500	13900
8	37500	2000	55500	30200	33000	43000	22500	21400
9	28500	2700	49500	24000	27000	24100	22500	55800
10	51000	12200	36000	26000	27000	16700	37500	8000

of the results returned for each run, we consider the dynamics of the distribution of pheromone on solution components during the run itself. More in detail, for each run we focus on the pheromone graphs as defined in Section 3: Every ten iterations, we observe the graph having all the nodes of the original one, and only the edges on which the pheromone trail is above the average. For each of these pheromone graphs, we compute the C and L ratios (Section 3), and we verify whether we are in presence of a small-world network.

As for what the quality of the result is concerned, for all the runs, at the beginning Ant Colony System performs better than $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System, while the latter ends up being the best at the end of the process. By the way, the number of objective function evaluations after which the value of the best solution returned by $\mathcal{MM}$AS is smaller than the one returned by ACS with the explorative parameters, is often higher than the value chosen as stopping criterion for these configurations.

After this observations, let us focus on the presence of small-world pheromone graphs during the runs. Interestingly, the presence of such networks can be observed in all the runs performed. Table 3 reports the number of objective function evaluations after which the small-world phenomenon weakens, *i.e.* after which the relevant values are no more above (or below, respectively) the thresholds reported in Section 3.

For visualizing the small-world presence in the pheromone graphs, we propose the results obtained for a representative instance with 400 nodes (instance number 2 in Table 3). Figure 1 reports the C and L ratios computed on the pheromone graph for the two algorithms and for the different sets of parameters. Moreover, in the upper part of the graphics (above the dashed line), the number and size of the clusters are represented: The size of the bullets is pro-

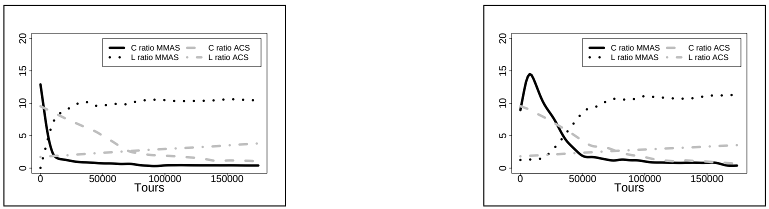


(a) Fairly explorative parameters ACS (b) Explorative parameters ACS

(c) Fairly explorative parameters $\mathcal{MM}$AS (d) Explorative parameters $\mathcal{MM}$AS

Fig. 1. C and L ratios, and number of clusters in one instance with 400 nodes.

portional to the average size of the clusters, and their y-coordinates correspond to the number of clusters that are present in the graphs. As it can be observed, clusterization is somehow present during the whole runs, but it is clearly higher at the beginning. A difference can be detected in the behavior of $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System and Ant Colony System, regardless the set of parameters considered: at the very beginning the former works with one or two large clusters, which are then decomposed in smaller ones. The latter, instead, immediately creates several small groups and only progressively decreases their number. In some sense, then, $\mathcal{MM}$AS chooses more carefully which nodes are to be grouped together, at the cost that, in some sense, it works for some time without any indication given by the pheromone. ACS, on the other hand, exploits from the first iteration the indirect communication that characterizes ACO algorithms, at the cost of possibly choosing the clusters in an imprecise way. According to the results obtained, ACS's strategy may be more powerful when the computational resources are very limited, while it is in general outperformed by $\mathcal{MM}$AS's one when the time available is longer. For summarizing the trends followed, Figure 2 represents the Friedman super smoother [15] computed on the observations of the C and L ratios in the set of instances with 400 nodes, for ACS and $\mathcal{MM}$AS. The pictures concerning the other sets are qualitatively similar: The difference implied by the sets of parameters chosen is much stronger for $\mathcal{MM}$AS than for ACS. This can be read as a sort of greater ability of $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System to adapt to the conditions in which it has to work, and, on the other hand, as a higher stability of Ant Colony System. Besides the difference in the duration of small-world effect, it is interesting to notice that both algorithms present the same type of behavior: At the beginning of a run, when the need for exploration



(a) Fairly explorative(n=400)

(b) Explorative (n=400)

Fig. 2. Friedman super smoother computed on the observations of the C and L ratios in the set of instances with 400 nodes, for ACS and $\mathcal{MM}$AS.

is higher, the pheromone is distributed in a very particular way. The nodes are grouped in clusters. As a consequence, when an ant reaches one vertex of a cluster, several possibilities for choosing the following vertex have a probability significantly higher than zero. Intuitively, then, the order in which the nodes of a cluster appear in the different solutions will vary. When a cluster has been completely visited, instead, only very few edges are rich of pheromone.

As already mentioned, this characteristic of the runs can be seen as the algorithms splitting the instances in sub-instances, *i.e.* the clusters. This procedure is very effective, as the results on ACO algorithms applied to the TSP reported in the literature show. Moreover, it reproduces the algorithmic idea behind several procedure used for solving, among a large set of problems, the traveling salesman problem [16]: This idea consists in recursively breaking down a problem into two or more sub-problems of the same type, until these become simple enough to be solved directly.

6 Conclusions

In this paper we propose the application of $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System and Ant Colony System to the traveling salesman problem. The original contribution of the research consists in observing the characteristics of the pheromone distributions, beside the relative performance of the algorithms.

In the experiments proposed, during each run, we observed the graphs obtained by discarding, at different points in time, the edges with a very low level of pheromone which are less likely to be selected as far as alternative arcs exist.

These pheromone graphs ended up in exhibiting the small-world properties: simplifying, at the beginning of each run each of them appeared as a set of clusters linked by few chains. The presence of the small-world properties in the pheromone graphs at the beginning of each run, when the need of exploration is higher, and their vanishing later on, represents a strong link between the behaviors of the two ACO algorithms. This is particularly surprising given the different ways pheromone is treated in the two procedures.

We read this behavior as ants decomposing the instances in sub-problems, without even the implementer being aware of it. The solutions to the sub-problems are then combined to give a solution to the original problem. One of the main problems of *divide et impera* approaches lies in fixing a rule for selecting the groups to be tackled separately. ACO algorithms solve this problem for us, building their small-world on the instances they are given.

Of course, further experiments need to be done in order to state that these observations can actually be extended to the whole ACO family, or at least to a great part of it. Moreover, we need to deal with some other optimization problem. The next step will consist in extending the analysis to a problem in which ACO algorithms build oriented pheromone graphs, namely the Quadratic Assignment Problem.

References

1. Dorigo, M., Stützle, T.: Ant Colony Optimization. MIT Press, Cambridge, MA, USA (2004)
2. Pool, I., Kochen, M.: Contacts and influence. Social Networks **1** (1978) 1–48
3. Milgram, S.: The small world problem. Psychology Today **2** (1967) 60–67
4. Newman M.E.J., Barabasi A.L., WattsD.J.: The Structure and Dynamics of Complex Networks. Princeton University Press, Princeton (2006)
5. Stützle, T., Hoos, H.H.: Improving the Ant System: A detailed report on the $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System. Technical Report AIDA-96-12, FG Intellektik, FB Informatik, Technische Universität Darmstadt, Darmstadt, Germany (1996)
6. Dorigo, M., Gambardella, L.M.: Ant Colony System: A cooperative learning approach to the traveling salesman problem. IEEE Transactions on Evolutionary Computation **1**(1) (1997) 53–66
7. Watts, D.: Networks, dyanmics, and the small-world phenomenon. The American Journal of Sociology **105**(2) (1999) 493–527
8. Bollobas, B.: Random Graphs. Academic Press, London (1985)
9. Watts, D., Strogatz, S.: Collective dynamics of 'small-world' networks. Nature **393** (1998) 440–442
10. Abe, S., Suzuki, N.: Small-world structure of earthquake network. Physica A **337** (2004) 357–362
11. Montoya, J.M., Solé, R.V.: Small wolrd patterns in food webs. Journal of theoretical biology **214**(3) (2002) 405–412
12. Stam, C.J.: Functional connectivity patterns of human magnetoencephalographic recordings: a 'small-world' network? Neuroscience Letters **355** (2004) 25–28
13. Birattari, M.: The Problem of Tuning Metaheuristics as Seen from a Machine Learning Perspective. PhD thesis, Université Libre de Bruxelles, Brussels, Belgium (2004)
14. Pellegrini, P.: ACO: parameters, exploration and quality of solutions. PhD thesis, Department of Applied Mathematics, Università Ca' Foscari, Venice, Italy (2007)
15. Friedman, J.H.: A variable span smoother. Technical Report 5, Department of Statistics, Stanford University, Stanford, CA, USA (1984)
16. Halton, J.H., Terada, R.: A fast algorithm for the euclidean traveling salesman problem, optimal with probability one. SIAM J. Comput. **11**(1) (1982) 28–46