Going the Extra Mile in Ant Colony Optimization Experimental Evaluation

Ricardo M. A. Silva[*†] *Geber L. Ramalho*[†]

[*] Computer Science Department, Federal University of Lavras
CP 37, CEP 37200.000, Lavras, MG - Brazil
Email: rmas@cin.ufpe.br

[†] Informatic Center, Federal University of Pernambuco
Cidade Universitaria, CEP 50740-540, Recife, PE - Brazil
Email: glr@cin.ufpe.br

1 Introduction

Ant colony optimization (ACO) is a novel and promising meta-heuristic for solving hard combinatorial optimization problems, such as travelling salesman[3] [4] [5] [6] [7] [8], quadratic assignment[9] [10] [11] [12] and set covering problem [13]. Unfortunately, according to systems performance evaluation literature, the methods adopted to experimentally validate the current ACO are not enough accurate and are excessively time-consuming. In fact, some works on heuristic optimization, suffer from the same methodological limitations. However, since the research on ant colony optimization is recent, this community has the opportunity of rethinking the evaluation methods from scratch, avoiding misconception on the evaluation the experimental data.

This paper describes a sound and efficient method for controlling the descriptive phase of ant systems validation. The descriptive phase, devoted to discover which are the relevant factors and how they influence the heuristic performance, is a crucial pre-requisite for the heuristic adjustment and comparison with other approaches. To illustrate the proposed method, we apply it to a case study, evaluating how well ant systems can solve a documented instance of the set-covering problem.

The table (1) classifies the informal and statistical methods according to their purpose (e.g. summarizing data, variability, and so on)[1], while the table (2) presents a review of the main ACO works on different problems, such as, Travelling Salesman Problem (TSP), Quadratic Assign Problem (QAP), Schedulling Problem (SP), Vehicle Routing Problem (VRP), Network Routing (NR), Sequential Ordering Problem (SOP), Frequency Assignment (FAP) and Set Covering Problem (SCP).

The table (2) shows how those ACO works attend the *descriptive phase*, devoted to discover what are the factors and how they influence the heuristic performance; the *fine-tuning phase*, intended to find out the best values of these factors; and the *comparative phase*, designed to confront the heuristic performance with other optimization methods. Observe that most of these ACO papers not only do not attend all phases, but also use informal methods (except [5])[2].

[1] It is obvious that table (2) does not summarize all statistical theory, but it is important in so far as shows an array of the possibilities against to the methods frequently used by the ACO community.

[2] It is worthwhile to remark that the empirical testing cannot prove causality, but it indicates the reliability and validity degree of the results. The more empirical testing is significant statistically, the more one can trust in the results. ACO papers review alerts that the results evaluations should be improved.

This paper describes a sound and efficient method for controlling the ACO descriptive phase, evaluating how well ant systems (AS) can solve a documented instance [19] of the set-covering problem (SCP)[3].

1. Design of Experiments
   1.1. Simple ((a)one-factor-a-time, (b)instances versus alternatives);
   1.2. DOE (e.g. (a)full-factorial, (b)fractional-factorial, (c)Latin-squares, (d)Taguchi);
2. Data analysis
   2.1. Summarizing data by a single number (e.g. (a)mean, (b)median, (c)mode, (d)best);
   2.2. Summarizing data by a distribution (e.g. (a)using histogram plot, (b)using quantile-quantile plot);
   2.3. Summarizing variability (e.g. (a)variance, (b)standard deviation, (c)coefficient of variation, (d)ranges, (e)percentiles, (f)confidence interval);
   2.4. Summarizing significance (e.g. (a)R^2, (b)$F - test$, (c)T statistic, (d)P statistic);
   2.5. Summarizing relationships (e.g. (a)key-tradeoff graphics, (b)observation, (c)ranking, (d)range, (e)regression techniques, (f)calibration techniques);

Table 1: Experimental evaluation techniques

		Descriptive		**Fine-tuning**		**Comparative**	
Prob.	Alg.	Design	Analysis	Design	Analysis	Design	Analysis
TSP	AS [3][4]	1.1(a)	2.5(a)	1.1(b)	2.5(b)	1.1(b)	2.5(b)
	Ant-Q [5]	NR	NR	?	?	1.1(b)	2.3(d);2.4(c)(d)
	ACS [6]	NR	NR	?	?	1.1(b)	2.5(b)
	MMAS [7]	NR	NR	NR	NR	1.1(b)	2.5(b)
	AS_{rank} [8]	NR	NR	?	?	1.1(b)	2.5(b)
QAP	AS-QAP [9]	NR	NR	NR	NR	1.1(b)	2.5(b)
	HAS-QAP [10]	1.1(a)	2.5(a)	?	?	1.1(b)	2.5(b)
	MMAS-QAP [11]	1.1(a)	2.5(a)	?	?	1.1(b)	2.5(b)
	ANTS-QAP [12]	1.1(a)	2.5(a)	?	?	1.1(b)	2.5(b)
SP	AS-FSP [14]	NR	NR	?	?	1.1(b)	2.5(b)
VRP	AS-VRP [15]	NR	NR	?	?	1.1(b)	2.5(b)
NR	Ant-Net [16]	1.1(a)	2.3(f);2.5(a)	NR	NR	1.1(b)	2.5(b)
SOP	HAS-SOP [17]	NR	NR	1.2(a)	2.5(b)	1.1(b)	2.5(b)
FAP	ANTS-FAP [18]	NR	NR	?	?	1.1(b)	2.5(b)
SCP	AC-SCP [13]	1.1(b)	2.1(a)	NR	NR	1.1(b)	NR

Table 2: Experimental evaluation techniques used in ACO works so far (Numeration in according to table (1). NR = Not Reported and ? = Reported only the final results, without mentioning the evaluation technique).

2 Empirical Testing Illustration

The experimental goal of our case study is discover which factors actually influence, and in what extent, the performance of the ant system applied to SCP. The performance measure adopted in this work as solution quality[4] is the percent error from optimality, $\frac{|x-x^*|}{x^*}$, where x is a heuristic solution and x^* is the optimal solution, or the best known solution, of a problem instance.

[3] The unique application of ACO to SCP [13] is hybrid - an ant colonies and local search combination, it has no report on the experimental descriptive phase and uses simple design plus basic statistics.

[4] The literature contains a variety of criteria for choosing performance measures [2]: solution quality, computational effort, and robustness. However, in order to illustrate an empirical testing, this paper shows only one quality solution

The factors that in principle affect the ant system performance are α (that regulates the influence of pheromone), β (that regulates the influence of heuristic value), ρ (that regulates the evaporation of pheromone), Q (quantity of pheromone), n (number of ants), d (problem density), and s_0 (initial disposition of the ants).

The test instance we have used is called sortcap [19], an SCP instance with 49 demand nodes (48 US capitals plus Washington DC) that are also candidates nodes. The experimental design adopted in this case study was the $2^k r$ factorial design with replication. This design is recommended to reduce the number of levels of each factor (usually large) at the beginning of a performance study. Considering that ant system makes random decisions as it searches for a solution of the problem, there are residuals (errors) that can be quantified by repeating the measurements under the same factor-level combinations. A full factorial design in which each of the k factors is used at two levels requires 2^k experiments. If each of the 2^k experiments is repeated r times, we will have $2^k r$ observations. In our case study, the levels of the factors correspond to reasonable limits (borders) chosen according to the domains[5] of the problem: $\alpha = 0.25$ and 5; $\beta = 0$ and 5; $\rho = 0.1$ and 0.9; $Q = 1$ and 100; $n = 1$ and 90; $d = 0.066$ and 0.234; $s_0 =$ C (all ants start on one demand node determined randomly) and D (all ants are distributed on the demand nodes). In short, considering that we have decided to repeat each experiment 5 times, we performed 640 ($2^7 * 5$) observations.

In the descriptive context, data analysis using regression techniques is composed by the following steps: (1) a regression model proposal, to estimate or predict a random variable as a function of several other variables; (2) the regression model validation, to measure the precision of this model; (3) the data interpretation, to infer conclusions about the algorithm behaviour; and (4) an optional practical approach to find the combination of factor levels that produces a good performance.

The Regression Model Proposal

The regression model fitted to the 128 observations with 5 repetitions each, was the following multiple linear, second-order regression model:

$$\begin{aligned}\sqrt{\text{perc. error}} = {} & 0,502 - 0,0544\alpha - 0,0954\beta + 0,269\rho - 0,000875n + 2,15d + \\ & +0,00824\alpha\beta - 0,000176\alpha Q + 0,153\alpha d - 0,0339\beta\rho + 0,000093\beta Q + \\ & +0,000573\beta n - 0,000669\rho n - 0,671\rho d - 0,0156nd \end{aligned} \qquad (1)$$

The Regression Model Validation

For each term of the regression equation (1), the table (3) shows the coefficients (Coef), their standard deviations (SECoef), T and P statistics. The T and P are statistical significance measures. In our case study for instance, at the 95% confidence, $|T|$ must be more than 1.76, and P less than 0.05. There are also overall "good fit" indicators of the regression model, such as, R^2 and F statistics (see table (3)). While the bigger R^2 better, F must be greater than 1.71 at the 95% confidence. Hence, the regression model (1) has got excellent precision indicators in view of its statistics (see table 3): $R^2 = 81.3\%$, $F = 194.35$, $T \leq -2.47$ or $T \geq 2.88$, and $P[0.000, 0.014]$.

In addition, also there are visual tests for verifying the regression assumptions. For instance, if one uses a linear model (like Equation (1)) the assumptions, among others, are: (1) independent residuals, the scatter plot of residual versus the fitted values cannot presents trends (see Figure 1.(a)); and (2) normally distributed error, the residual histogram should be normally distributed ((Figure 1.(b)), and the normal quantile-quantile plot of errors should be linear (Figure 1.(c)).

The Data Interpretation

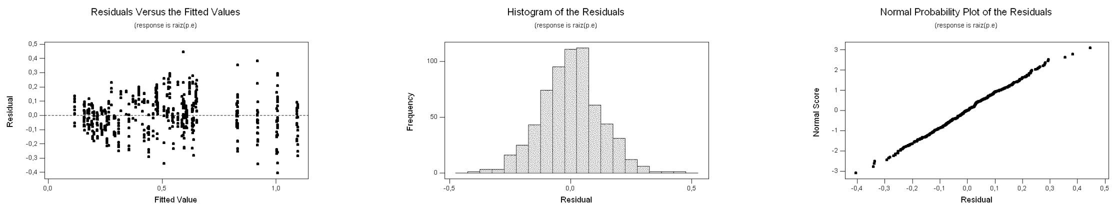
According to the equation (1), the mean percentual error of the ant system for the set covering problem

performance measure.

[5] For instance, considering that $\rho \in [0, 1]$, we chose $\rho = 0.1$ and $\rho = 0.9$. And, if $\alpha \geq 0$, one reasonable choice will be $\alpha = 0.25$ and $\alpha = 5$.

Factor	Coef	SE Coef	T	P	Seq SS	Imp. (%)
Constant	0,50157	0,02493	20,12	0,000		
α	-0,054367	0,004908	-11,08	0,000	1,3976	3,45%
β	-0,095355	0,004545	-20,98	0,000	14,3490	35,40%
ρ	0,26873	0,03012	8,92	0,000	0,2832	0,67%
n	-0,0008748	0,0002818	-3,10	0,002	5,7511	14,20%
d	2,1488	0,1260	17,06	0,000	10,2255	25,20%
$\alpha\beta$	0,0082417	0,0008125	10,14	0,000	1,5326	3,78%
αQ	-0,00017583	0,00003233	-5,44	0,000	0,3166	0,78%
αd	0,15275	0,02414	6,33	0,000	0,5965	1,47%
$\beta\rho$	-0,033898	0,004824	-7,03	0,000	0,7354	1,81%
βQ	0,00009338	0,00003237	2,88	0,004	0,1239	0,306%
βn	0,00057276	0,00004336	13,21	0,000	2,5986	6,41%
ρn	-0,0006692	0,0002710	-2,47	0,014	0,0908	0,224%
ρd	-0,6714	0,1433	-4,68	0,000	0,3269	0,81%
nd	-0,015648	0,001288	-12,15	0,000	2,1976	5,42%
$R^2 = 81.3\%$, $F = 194.35$, $\sum_{i=1}^{14} SeqSS_i = 40{,}5252$						

Table 3: Regression table



(a) Residuals versus Fits (b) Residual Histogram (c) Normplot of Residuals

Figure 1: Verifying Regression Visuality for $\sqrt{\text{perc. error}}$

with size 49x49[6] is 0.25 ($Constant^2 = 0,50157^2 = 0.25157$); the effect of α is 0.00296 ($\alpha^2_{coef} = 0,054367^2 = 0.00296$); β is 0.00909; ρ is 0.07222; n is 0.00000076528; d is 4.62; the interaction between α and β accounts 0.00006793; and so on. Note that Q does not affect by itself the performance measure, except when it is interacting with other factors. Observe also that s_0 is not a relevant factor due to fact that it not even to appear in equation (1).

The importance of a factor is measured by the proportion of the total variation in the response that is explained by the factor. Table (3) shows the variation (SeqSS) and importance of each term of the regression equation (1), according to square root of the performance measure ($\sqrt{\text{perc. error}}$). Note that the three greatest influential factors are; β (35,40%), the density d of the problem (25,20%) and the number of the ants n (14,20%). Observe that the importance of these three factors are quite big in comparison with factors such as α (3,45%) and ρ (0,67%). One of the conclusions that we can infer is that it is not recommended to use a pure ant approach (i.e. without heuristic information about the problem) in a set covering problem with size 49x49.

A Good Combination of the Factor Levels

[6] Although the density of the problem was considered as a factor, the size of the problem was not. Therefore, all conclusions concern only problems with size equal to 49x49 (demand nodes x facility nodes).

As discussed in previous paragraphs, the regression technique serves to understand how the factors influence the heuristic performance. However, the regression model (1) can also be used to find a good combination of these factors values[7], in a manner more easy and fast than when there is only the heuristic to simulate. If the regression model is well behaved, an analytic evaluation can be used. Otherwise, informal methods (e.g. observation, ranking and range) or optimization methods (e.g. enumeration, greedy search) can be used. For instance, using the observation method, a good combination of the factor levels in our case study was: $\alpha = 4$, $\beta = 1.5$, $\rho = 0.2$, $Q = 90$ and $n = 150$ (for $d = 0.066$, with percentual error equals to 0.0325); and $\alpha = 0.25$, $\beta = 5$, $\rho = 0.8$, $Q = 1$ and $n = 150$ (for $d = 0.234$, with percentual error equals to 0.0252).

3 Final Remarks

In this paper, we have explained a descriptive evaluation of the ant systems performance. The statistical foundation of this evaluation establishes a structured framework, which facilitates the design of experiments and shorten the time needed to get relevant results, as well as guarantees greater accuracy of the results. For instance, there are some performance considerations with regard to case-study studied: (1) the importance of α and ρ is quite less than β and the number n of ants, so it is not recommended to use the pure ant system approach; (2) the start position s_0 of the ants is not relevant for the solution quality; (3) the quantity of pheromone Q does not affect the performance, except when it is interacting with β.

As further work, we intend to apply the same method to assess ant systems solutions to other optimization problems, in particular to other facility location problems such as center problems, median problems and fixed charge.

References

[1] Draper N.R., Smith H. *Applied Regression Analysis* Wiley series in probability and mathematical statistics, 1981

[2] Barr R.S., Golden B.L., Kelly J.P., Resende M.G.C., Stewart W.R. *Designing and Reporting on Computational Experiments with Heuristic Methods.* Technical report, 1995.

[3] Dorigo M., Maniezzo V., Colorni A. *Positive feedback as a search strategy.* Technical report 91-016, Dipartimento di Eletronica, Politecnico di Milano, Italy, 1991.

[4] Dorigo M., Maniezzo V., Colorni A. *The ant system: optimization by a colony of cooperating agents.* IEEE Trans. System Man and Cybernetic B 26(1) (1996) 29-41.

[5] Gambardella L.M., Dorigo M. *Ant-Q: a reinforcement learning approach to the travelling salesman problem.* Proceedings of the 12th International Conference on Machine Learning, Palo Alto, CA, pp. 252-260, 1995.

[6] Dorigo M., Gambardella L.M. *Ant colony system: a cooperative learning approach to the travelling salesman problem.* IEEE Trans. Evol. Comput. 1 (1) (1997) 53-66.

[7] Stutzle T., Hoos H. *Improvements on the Ant System: Introducing MAX-MIN Ant System.* Proceedings of the International Conference on Artificial Neural Networks and Genetic Algorithms, Vienna, pp.245-249, 1998.

[8] Bullnheimer B., Hartl R.F., Strauss C. *A new rank version of the Ant System: a computational study.* Technical Report POM-03/97, Institute of Management Science, University of Vienna, Austria, 1997.

[7] Within of the range determined by extreme levels have chosen during experimental design phase.

[9] Maniezzo V., Colorni A., Dorigo M. *The ant system applied to the quadratic assignment problem.* Technical Report IRIDIA/94-28, IRIDIA, Universite Libre de Bruxelles, Belgium 1994.

[10] Gambardella L.M. Taillard E.D. *Ant colonies for the QAP.* Technical Report IDSIA-4-97, IDSIA, Lugano, Switzerland, 1997.

[11] Stutzle T., Hoos H. *MAX-MIN Ant system and local search for combinatorial optimization problem.* Meta-heuristics: Advances and trends in Local search paradigms for optimization, Kluwer Academic, Boston, MA, pp 313-329, 1999.

[12] Maniezzo V. *Exact and aproximate nondeterministic tree-search procedures for the quadratic assignment problem.* Technical Report CSR 98-1, Scienze dellInformazione, Universita di Bologna, Sede di Cesena, Italy, 1998.

[13] Hadji R., Rahoual M., Talbi E., Bachelet V. *Ant Colonies for the Set Covering Problem* Abstract Proceedings of ANTS'2000 From Ant Colonies to Artificial Ants: Second International Workshop on Ant Algorithms, Brussels, Belgium, pp. 63-66, 2000.

[14] Stutzle T. *An ant approach to the flow shop problem.* Technical Report AIDA-97-07, FG Intellektik, FB Informatik, TH Darmstadt, September, 1997.

[15] Bullnheimer B., Hartl R.F., Strauss C. *An improved Ant System algorithm for thr vehicle routing problem.* Technical Report POM-10/97, Institute of Management Science, University of Vienna, Austria, 1997.

[16] Di Caro G., Dorigo M. *AntNet: a mobile agents approach to adaptive routing.* Technical Report IRIDIA/97-12, IRIDIA, Universite Libre de Bruxelles, Belgium, 1997.

[17] Gambardella L.M., Dorigo M. *HAS-SOP: an hybrid Ant System for the sequential ordering problem.* Technical Report IDSIA-11-97, IDSIA, Lugano, Switzerland, 1997.

[18] Maniezzo V., Carbonaro A. *An ANTS heuristic for the frequency assignment problem.* Technical Report CSR 98-4, Scienze dellInformazione, Universita di Bologna, Italy, 1998.

[19] Daskin M.S. *Network and Discrete Location. Models, Algorithms, and Applications.* Wiley-Interscience Series in Discrete Mathematics and Optimization, 1995.