Measuring Variables Effect to Statistically Model the Multi-Robot Patrolling Problem by Means of ANOVA

David Portugal and Rui P. Rocha

Institute of Systems and Robotics, Department of Electrical and Computer Engineering, University of Coimbra, 3030-290 Coimbra, Portugal
{davidbsp,rprocha}@isr.uc.pt

Abstract. This paper focuses on analyzing extensive results generated from running diverse multi-robot patrolling algorithms with different configurations towards measuring the influence of the variables of the general problem. In order to do this, a statistical technique by the name of Analysis of Variance (ANOVA) is employed to compare the parameters and identify the ones which give raise to the total dispersion of the data set, at the same time accessing their contribution to the obtained results. It is shown that by applying such technique, it is possible to compute a data-related model for the Multi-Robot Patrolling Problem (MRPP).

Keywords: Multi-Robot Patrolling, ANOVA and Statistical Analysis.

1 Introduction

This work addresses cooperative multi-robot systems in area patrolling missions. In such tasks, agents must coordinately visit all regions of an area, at regular intervals, in order to protect or supervise them. Since it involves several variables, the multi-robot patrolling problem is highly complex and no optimal patrolling strategy has been proven to exist for all cases.

In a recent work [1], we implemented and compared five representative strategies to solve the problem using Stage/ROS simulator [2]. Aiming to promote transparent comparisons between all algorithms, it was necessary to test them in environments with different connectivity and teamsizes. Naturally, the combination of these factors produced a large set of results, which were only qualitatively discussed therein.

2 Contribution to Value Creation

Analyzing the state of the art in this field, the majority of works contain limitations like the absence of studies on scalability, performance and on other parameters of the problem, as well as many simplifications at the experimental work level. In this work, the main objective is to go beyond empirical analysis by conducting a quantitative analysis, identifying the contribution of each factor/parameter {Algorithm, Teamsize, Connectivity} to the dispersion of the results using ANOVA and building an initiatory data-related model that explains the results obtained in the context of the Multi-Robot Patrolling Problem (MRPP).

L.M. Camarinha-Matos et al. (Eds.): DoCEIS 2012, IFIP AICT 372, pp. 199–206, 2012.
© IFIP International Federation for Information Processing 2012

3 Related Work

During the last decade, several authors have proposed different algorithms for solving the MRPP, which mainly differ in terms of planning strategy, communication paradigm, coordination scheme, complexity, agent perception, decision-making and performance evaluation.

Pioneer work on this field was presented in [3] and [4], where very simple approaches with reactive and cognitive agents were proposed as well as a first evaluation criterion based on idleness. Following these works, more advanced approaches based on graph theory became also popular like [5] and [6], which use graph partitioning to assign efficient patrol routes to different robots. Based on a similar idea, in [7] a technique to reduce the patrolling graph to multiple spanning trees, assigning a spanning tour around each tree was recently presented. In contrast, many authors have proposed new algorithms for multi-robot patrolling based on other concepts, like task allocation [8], negotiation mechanisms [9], swarm intelligence [10], neural networks [11] or game theory [12]. For a more complete survey of multi-robot patrolling strategies, the interested reader is referred to [13].

On the other hand, ANOVA is a powerful statistical technique which enables the comparison between parameters of more than two populations [14]. From the analysis of the total dispersion present in a data set, it allows us to identify the source of the variations that led to that dispersion and evaluate the contribution of each factor, determining whether a significant relation exists between variables. The experiments presented in this work consider three factors (algorithm, teamsize and connectivity), thus the analysis of variance used to study their effects is called a three-way ANOVA.

ANOVA has been used as a tool in a wide range of scientific literature. For example, the relationships between group size and efficiency in robotic ant-like foraging tasks are studied in [15]. Different teams of robots were compared in 3 sets of experiments with different food distributions. To test the effect of group size, one-way ANOVAs were performed. Relationships between sets of experiments and group size were also inspected using two-way ANOVAs. Generally, it was concluded that group size has a significant effect on colony foraging efficiency. Results also show that in environments where robots are able to recruit other robots (call others to where the "food" is) the foraging efficiency significantly increases when compared to situations where robots cannot communicate.

4 Preliminaries

Every patrolling algorithm implicated in the results set, assume a topological environment, which is a common assumption in the MRPP [13], and the decision-making process is related to the computation of paths in the navigation graph so that robots effectively patrol a given environment, visiting all vertices of the graph.

Since they are governed by different patrolling rules, it is necessary to define an evaluation metric to compare the performance of these algorithms. In this work the average graph idleness Idl_G is used, defined as the average idle time of all vertices of the graph G during the multi-robot patrolling mission [1]. Results are therefore expressed by Idl_G in seconds. The lower the Idl_G, the better the performance is.

The dataset considered is the result of the extensive simulation process described in [1]. Note that these simulations are fairly realistic, given that a community recognized robotic framework was used (ROS [17]), considering the robot's dynamics; and results are based on the actual time instead of simulation iterations, unlike previous works.

Five state-of-the art patrolling algorithms were implemented, namely: Conscientious Reactive (CR) [3], Heuristic Conscientious Reactive (HCR) [4], Heuristic Pathfinder Conscientious Cognitive (HPCC) [4], Cyclic Algorithm for Generic Graphs (CGG) [5] and Generalized MSP Algorithm (MSP) [5]. These algorithms were combined with six different teamsizes: 1, 2, 4, 6, 8 and 12; and three environments with different connectivity properties: low connectivity (A), medium connectivity (B) and high connectivity (C). All trials were repeated three times.

Tables 1, 2 and 3 present a summary of the performance results obtained in the experiments. In the next section three-way ANOVA is applied to quantitatively measure the influence of the parameters tested in the experiments.

5 Applying ANOVA to the Dataset

The purpose of this analysis is to access the influence of three important parameters in the results: patrolling algorithm or strategy, teamsize and environment connectivity.

Linear models are considered, assuming that the probability distribution of the response is normal, mutually independent and homoscedastic (i.e., the variance of the data inside the groups is equivalent). In the general case, the ANOVA fundamental test, which makes use of the F-statistics distribution usually with 95% of confidence bounds, verifies the significance of the factors by checking the group's variable effect (α_i) [14]:

$$\begin{aligned} &H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_I = 0 \\ &H_1 : \exists \alpha_i \neq 0, \end{aligned} \tag{1}$$

where H_0 is the null hypothesis and H_1 is the alternative hypothesis. If the probability of the null hypothesis is near 0, a main effect is present due to the associated factor, meaning that the result is statistically significant. Arbitrarily high n-way ANOVA divides the total variation, given by the deviation of all observations from the global mean, into variations given by different factors and residual variation (or error). Also, if the model has a non-additive effect, variations given by the interaction of factors are also considered. These variations are calculated through sums of squares.

In this work, a model with first-order interaction effects between pairs of factors was adopted. This model explains 99.75% of the variation of the results.

Table 1. Numerical Results for Map A (*Idl_G* values in seconds)

Teamsize	CR	HCR	HPCC	CGG	MSP	$\overline{x}_{.jA}$
1	1734.09	1962.42	1740.37	1717.36	1704.36	1771.72
2	843.93	1146.27	791.20	845.49	930.04	911.39
4	433.38	652.84	434.11	451.70	476.92	489.79
6	367.11	506.90	377.73	348.46	381.97	396.44
8	271.70	442.39	361.62	288.72	253.19	323.53
12	287.14	412.65	352.79	265.47	183.74	300.36
$\overline{x}_{i.A}$	656.23	853.91	676.30	652.87	655.04	$\overline{x}_{..A} = 698,87$

Table 2. Numerical Results for Map B (Idl_G values in seconds)

Teamsize	CR	HCR	HPCC	CGG	MSP	$\bar{x}_{.jB}$
1	1315.79	1283.59	1235.67	1347.30	1401.80	1316.83
2	675.44	654.61	670.44	675.64	749.42	685.11
4	363.46	373.45	298.77	335.45	375.15	349.25
6	238.57	273.60	254.96	234.18	248.92	250.05
8	198.90	217.38	225.44	172.39	185.28	199.88
12	172.4	255.62	212.3	143.94	-	$\bar{x}_{.12B}$
$\bar{x}_{i.B}$	494.09	509.71	482.93	484.82	$\bar{x}_{MSP.B}$	$\bar{x}_{..B}$

Table 3. Numerical Results for Map C (Idl_G values in seconds)

Teamsize	CR	HCR	HPCC	CGG	MSP	$\bar{x}_{.jC}$
1	715.30	714.23	737.93	767.25	766.41	740.23
2	353.06	351.15	358.45	385.09	423.60	374.27
4	193.30	186.59	188.03	200.53	209.82	195.65
6	141.68	138.64	135.74	142.94	148.09	141.42
8	104.00	108.45	118.75	113.71	95.22	108.03
12	101.82	105.64	118.36	94.35	-	$\bar{x}_{.12C}$
$\bar{x}_{i.C}$	268.19	267.45	276.21	283.98	$\bar{x}_{MSP.C}$	$\bar{x}_{..C}$

The ANOVA table, presented in Figure 1, illustrates the model used. The only factor that has no relevant significance is the algorithm-teamsize interaction. In this case, the null hypothesis is accepted:

$$F-test_{Alg*TS} = \frac{MS_{Alg*TS}}{MSE} = \frac{1840.7}{1112} = 1.66 < F_{20,38}(\propto = 0.05) < 1.85 \quad (2)$$

Analysis of Variance

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Algorithm	64953.1	4	16238.3	14.6	0
Teamsize	12763618.2	5	2552723.6	2295.67	0
Connectivity	2574860.1	2	1287430.1	1157.79	0
Algorithm*Teamsize	36814	20	1840.7	1.66	0.0891
Algorithm*Connectivity	120574.6	8	15071.8	13.55	0
Teamsize*Connectivity	1279680.4	10	127968	115.08	0
Error	42255	38	1112		
Total	17234127	87			

Constrained (Type III) sums of squares.

Fig. 1. ANOVA Table considering a model of three factors and first-order interactions

In fact, this is expectable as seen on the Idl_G values of each of the tables, which do not differ much when each of the associated columns are compared as a whole. The significance of the factor algorithm is also reduced; however, the null hypothesis cannot be accepted in this case. Idl_G values vary deeply when teamsize changes which is the factor with more significance in the MRPP. After teamsize, the most significant factor is the connectivity of the environment as expected. Additionally, the interaction of these two factors is the most significant interaction. Figure 2 presents the approximately additive effect of interaction teamsize-algorithm and the non-additive effect of the interaction teamsize-connectivity is presented in Figure 3.

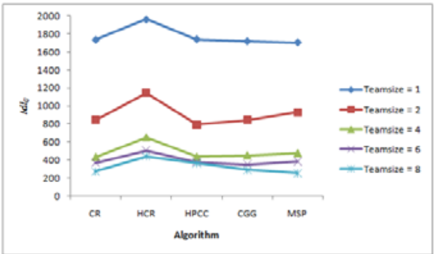

Fig. 2. The factors algorithm and teamsize present an approximate additive effect, which is clear by the near parallel curves. In this chart, the environment used was map A.

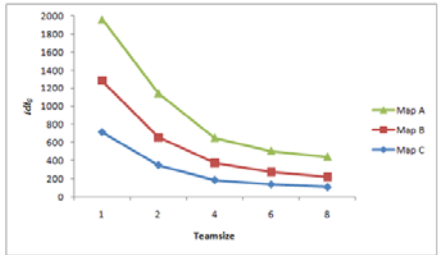

Fig. 3. The factors teamsize and connectivity present a clear non-additive effect. In this chart, the algorithm used was HCR.

A more complete model could eventually be obtained by considering second-order interaction of the three factors. However, this interaction is only responsible for the remaining 0.25% variation of the results and, therefore, it is ignored.

6 Extracting a Data-Related Model

As seen before, the ANOVA table in Figure 1 shows that the interaction between algorithm and teamsize has no relevant significance, which is an important property associated to the results obtained. Therefore we can build a simplified model which does not consider that interaction (as well as second-order interactions), accounting for 99.55% of the variation of the results. To describe such situations, the general expression of the model is [16]:

$$x_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_{jk} + \varepsilon_{ik} + E_{ijk}, \qquad (3)$$

with the following restrictions:

$$\sum_i \alpha_i = 0; \ \sum_j \beta_j = 0; \ \sum_k \gamma_k = 0; \ \sum_j \delta_{jk} = 0; \ \sum_i \varepsilon_{ik} = 0, \qquad (4)$$

and:

$$i = \{CR, HCR, HPCC, CGG, MSP\}, \quad j = \{1, 2, 4, 6, 8, 12\}, \quad k = \{A, B, C\}. \qquad (5)$$

μ represents a global parameter, while the effects of algorithm, teamsize and connectivity are described by α_i, β_j and γ_k respectively. δ_{jk} represents the effect of the teamsize-connectivity interaction and ε_{ik} describes the effect of the algorithm-connectivity interaction. E_{ijk} is the error and $E_{ijk} \sim N(0,\sigma^2)$. The parameters of the model can be estimated efficiently via the least squares methods. Thus, we obtain:

$$
\begin{aligned}
\hat{\mu} &= \overline{x_{...}} = \frac{1}{IJK}\sum_i\sum_j\sum_k x_{ijk} \\
\hat{\alpha}_i &= \overline{x_{i..}} - \hat{\mu} = \frac{1}{JK}\sum_j\sum_k x_{ijk} - \overline{x_{...}} \\
\hat{\beta}_j &= \overline{x_{.j.}} - \hat{\mu} = \frac{1}{IK}\sum_i\sum_k x_{ijk} - \overline{x_{...}} \\
\hat{\gamma}_k &= \overline{x_{..k}} - \hat{\mu} = \frac{1}{IJ}\sum_i\sum_j x_{ijk} - \overline{x_{...}} \\
\hat{\delta}_{jk} &= \overline{x_{.jk}} - \overline{x_{.j.}} - \overline{x_{..k}} + \overline{x_{...}} \\
\hat{\varepsilon}_{ik} &= \overline{x_{i.k}} - \overline{x_{i..}} - \overline{x_{..k}} + \overline{x_{...}}.
\end{aligned}
\tag{6}
$$

As seen in Tables 2 and 3, the Idl_G values for a team of 12 robots using the MSP algorithm are not available due to the nature of the algorithm, which is not able to partition those particular environments in 12 regions. In order to have a coherent model and calculate the average values ($\overline{x}_{MSP.B}$, $\overline{x}_{MSP.C}$, $\overline{x}_{.12B}$ and $\overline{x}_{.12C}$) and large averages ($\overline{x}_{..B}$ and $\overline{x}_{..C}$) left open in Tables 2 and 3, a curve fitting method by means of a general power function of the type $a{\cdot}x^b$ was used, as shown in Figures 4 and 5, to estimate the result for a team of 12 robots in those environments based on the data obtained with the same algorithm with lower teamsizes. The general power function model was applied because of the trend shown by the data as the teamsize grows when using the algorithm in the given environment, as well as its non-negative constraint, which completely suits the problem. The estimations obtained were:

$x_{MSP,12,B} = 133.45$ (with R^2: 0.9995)
$x_{MSP,12,C} = 77.26$ (with R^2: 0.9994).

Note that R^2 (or R-square) is a fraction between 0.0 and 1.0, which has no units and quantifies goodness of fit. Higher values indicate that the model fits the data better. Considering these values we obtain the following results for the values that were left open and present the results for the other large averages related to the algorithms and teamsize factors ($\overline{x}_{i..}$ and $\overline{x}_{.j.}$) extracted from the three tables:

$\overline{x}_{MSP.B} = 515.67$	$\overline{x}_{CR..} = 472.84$	$\overline{x}_{.1.} = 1276.26$
$\overline{x}_{MSP.C} = 286.73$	$\overline{x}_{HCR..} = 543.69$	$\overline{x}_{.2.} = 656.92$
$\overline{x}_{.12B} = 183.54$	$\overline{x}_{HPCC..} = 478.48$	$\overline{x}_{.4.} = 344.90$
$\overline{x}_{.12C} = 99.49$	$\overline{x}_{CGG..} = 473.89$	$\overline{x}_{.6.} = 262.63$
$\overline{x}_{..B} = 497.44$	$\overline{x}_{MSP..} = 485.81$	$\overline{x}_{.8.} = 210.48$
$\overline{x}_{..C} = 276.51$		$\overline{x}_{.12.} = 194.46$

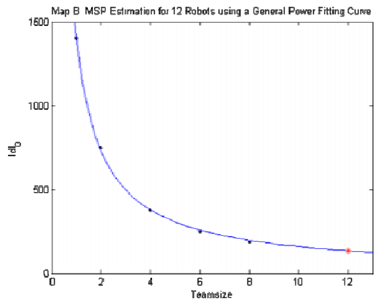

Fig. 4. Estimation of the Idl_G value for a teamsize of 12 robots using MSP algorithm in map B. A General Power Model of type $a \cdot x^b$, with a = 1408 and b = -0.9482 was used to fit the curve to the 5 points previously obtained with the same algorithm and lower teamsize.

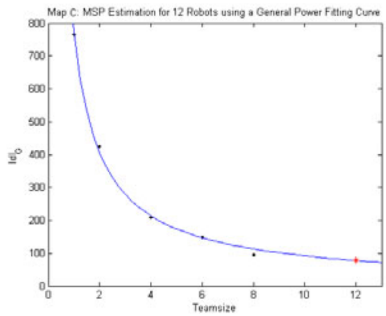

Fig. 5. Estimation of the Idl_G value for a teamsize of 12 robots using MSP algorithm in map C. A General Power Model of type $a \cdot x^b$, with a = 772.2 and b = -0.9264 was used to fit the curve to the 5 points previously obtained with the same algorithm and lower teamsize.

Having all data available, it is now possible to compute the parameters of the model using (6). Due to space limitations, the numerical results obtained are omitted. The model's parameters help us to understand the values of each x_{ijk}, clarifying the sources of variability and identifying the main factors that explain the results obtained.

7 Conclusions and Future Work

In this work, an extensive quantitative analysis of results obtained using multi-robot patrolling strategies with different teamsizes and environments was described. Analysis of Variance (ANOVA) was used to understand the significance of the different factors involved in the problem and a data-related model was calculated to explain the results.

The results presented were somehow unforeseen, given that the patrolling algorithms proved to be much less significant as a factor than teamsize and connectivity of the environment, which represents an important conclusion to this field. Therefore, future work should focus mostly on approaches that ensure scalability and are appropriate, or perhaps can adapt, to all kinds of environment.

Acknowledgments. This work was financially supported by a PhD grant (SFRH/BD/64426/2009) from the Portuguese Foundation for Science and Technology (FCT) and the Institute of Systems and Robotics (ISR-Coimbra) also under regular funding by FCT. The authors gratefully acknowledge António Portugal (Dept. Chemical Engineering, University of Coimbra), for his contribution and feedback.

References

1. Portugal, D., Rocha, R.P.: On the Performance and Scalability of Multi-Robot Patrolling Algorithms. In: Proc. of the 2011 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR 2011), Kyoto, Japan, November 1-5, pp. 50–55 (2011)
2. ROS. org: Stage stack, `http://www.ros.org/wiki/stage`
3. Machado, A., Ramalho, G., Zucker, J., Drogoul, A.: Multi-Agent Patrolling: an Empirical Analysis of Alternative Architectures. In: 3rd Int. Workshop on Multi-Agent-Based Simulation, Italy, pp. 155–170 (July 2002)
4. Almeida, A., Ramalho, G., Sanana, H., Tedesco, P., Menezes, T., Corruble, V., Chaveleyre, Y.: Recent Advances on Multi-agent Patrolling. In: Bazzan, A.L.C., Labidi, S. (eds.) SBIA 2004. LNCS (LNAI), vol. 3171, pp. 474–483. Springer, Heidelberg (2004)
5. Portugal, D., Rocha, R.P.: MSP Algorithm: Multi-Robot Patrolling based on Territory Allocation using Balanced Graph Partitioning. In: Proc. of the 25th ACM Symposium on Applied Computing (SAC 2010), Sierre, Switzerland, March 22-26, pp. 1271–1276 (2010)
6. Chevaleyre, Y.: Theoretical Analysis of the Multi-agent Patrolling Problem. In: IEEE/WIC/ACM Int. Conf. Proc. of the Intelligent Agent Technology (IAT 2004), Beijing, China, September 20-24, pp. 302–308 (2004)
7. Fazli, P., Davoodi, A., Pasquier, P., Mackworth, A.K.: Complete and Robust Cooperative Robot Area Coverage with Limited Range. In: Proc. of the IEEE/RSJ Int. Conference on Intelligent Robots and Systems (IROS 2010), Taipei, Taiwan (2010)
8. Sempé, F., Drogoul, A.: Adaptive Patrol for a Group of Robots. In: Proc. of the Int. Conf. on Intelligent Robots and Systems (IROS 2003), Las Vegas, Nevada (October 2003)
9. Menezes, T., Azevedo Tedesco, P., Ramalho, G.L.: Negotiator Agents for the Patrolling Task. In: Sichman, J.S., Coelho, H., Rezende, S.O. (eds.) IBERAMIA 2006 and SBIA 2006. LNCS (LNAI), vol. 4140, pp. 48–57. Springer, Heidelberg (2006)
10. Chu, H., Glad, A., Simonin, O., Sempé, F., Drogoul, A., Charpillet, F.: Swarm Approaches for the Patrolling Problem, Information Propagation vs. Pheromone Evaporation. In: Proc. of the International Conference on Tools with Art. Intelligence, France, pp. 442–449 (2007)
11. Guo, Y., Parker, L., Madhavan, R.: Collaborative Robots for Infrastructure Security Applications. SCI, vol. 50, pp. 185–200. Springer, Heidelberg (2007)
12. N. Basilico, N. Gatti, T. Rossi, S. Ceppi and F. Amigoni, "Extending Algorithms for Mobile Robot Patrolling in the Presence of Adversaries to More Realistic Settings". In Proc. of WI-IAT 2009 IEEE/WIC/ACM *Int. Joint Conference on Web Intelligence and Intelligent Agent Technology*, IEEE Press, Vol. 02, pp. 557-564, Milan, Italy, September 15-18 2009.
13. Portugal, D., Rocha, R.: A Survey on Multi-Robot Patrolling Algorithms. In: Camarinha-Matos, L.M. (ed.) DoCEIS 2011. IFIP AICT, vol. 349, pp. 139–146. Springer, Heidelberg (2011)
14. Scheffé, H.: The Analysis of Variance. John Wiley & Sons, New York (1959)
15. Krieger, M., Billeter, J., Keller, L.: Ant-like task allocation and recruitment in cooperative robots. Nature 406, 992–995 (2000)
16. Guimarães, R.C., Sarsfield Cabral, J.A.: Estatística. McGraw-Hill (1997)
17. Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R., Ng, A.: ROS: an open-source Robot Operating System. In: Proceedings of the 2009 IEEE International Conference on Robotics and Automation (ICRA 2009), Workshop on Open Source Software, Kobe, Japan, May 12-17 (2009)