On the Use of Multivariate Regression for Performance Characterisation of JavaSpaces

Frederic Hancke (frederic.hancke@ua.ac.be)
Tom Dhaene (tom.dhaene@ua.ac.be)
Jan Broeckhove (jan.broeckhove@ua.ac.be)
Department of Mathematics and Computer Science
Middelheimlaan 1, 2020 Antwerp, Belgium

Abstract

This paper discusses research results on the performance characterisation of JavaSpaces. Whereas Design of Experiments provides a technique to obtain response surface models of second order with as few experiments as possible, the construction of higher order models makes more precise predictions possible. Analysis of Variance is iteratively used in combination with Linear Regression to obtain information about the performance characteristics of JavaSpaces, resulting in an accurate response surface model as a function of the input factors.

1. Introduction

This contribution reflects the first results of ongoing research in the field of performance characterisation of distributed platforms in general, in this paper JavaSpaces [3], a Java implementation of the Linda Tuple Spaces concept [1, 4]. In recent research [6, 2, 5] a well known statistical technique, called *Design of Experiments (DoE)* [7], was used for modeling the performance with as few experiments as possible. Using this technique implies the construction of response surface models (RSM) of maximum order 2. This is a serious limitation when using it to cover wide input ranges.

In this paper all experiments lie on a uniform grid in the input space, instead of using a predefined DoE design. This implies that the output of the experiments can be fitted to multivariate polynomial models of order higher than 2. The data is fitted in an adaptive way by incrementally augmenting the order of the RSM until a good fit is obtained.

2. Higher Order Modeling

Analysis of Variance (ANOVA) [8] allows to detect the significance of the influences for each input factor of the input space, where input factors can be powers of the main input factors (e.g. $A, B, B^2, \ldots$) as well as interactions between the main input factors (e.g. $AB, AB^2, \ldots$). Before using *Linear Regression* to obtain a higher order RSM, only the significant input factors are selected based on the results of the ANOVA.

Fitting the experimental data by incrementally augmenting the order of the RSM, is here applied to the JavaSpaces platform. Each experiment consists of **tTa** tasks of runtime **tMe** to be distributed to **tWo** workers in a *Master-Slave* environment. These three controllable input factors form the input space. The only outputfactor is **pWCT**, the wallclock time. From this, the *efficiency* **pEff** can be calculated as:

$$\text{pEff} = \frac{\text{tTa} * \text{tMe}}{\text{pWCT} * \text{tWo}} \tag{1}$$

tTa was found to have no influence on pEff, whereas both tMe and tWo had influence of order 4. The visualisation of the RSM can be seen in Figure 1. The data points used for contructing the model are shown with $+$. Additional verification points are shown with $\times$.

3. Conclusions

This contribution discussed intermediate but relevant results of research on the performance characterisation of JavaSpaces. The construction of higher order models allows to predict the overall performance accurately. These models will later be used in the scheduling process of computationally complex problems.

References

[1] Yale linda group. http://www.cs.yale.edu/ Linda/ linda.html.

[2] H. De Neve, F. Hancke, T. Dhaene, J. Broeckhove, and F. Arickx. On the use of DoE for characterization of JavaSpaces. In *Proceedings ESM*, pages 24–29. Eurosis, 2003.

Proceedings of the IEEE Computer Society's 12th Annual International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems (MASCOTS'04)
1526-7539/05 $20.00 © 2005 **IEEE**

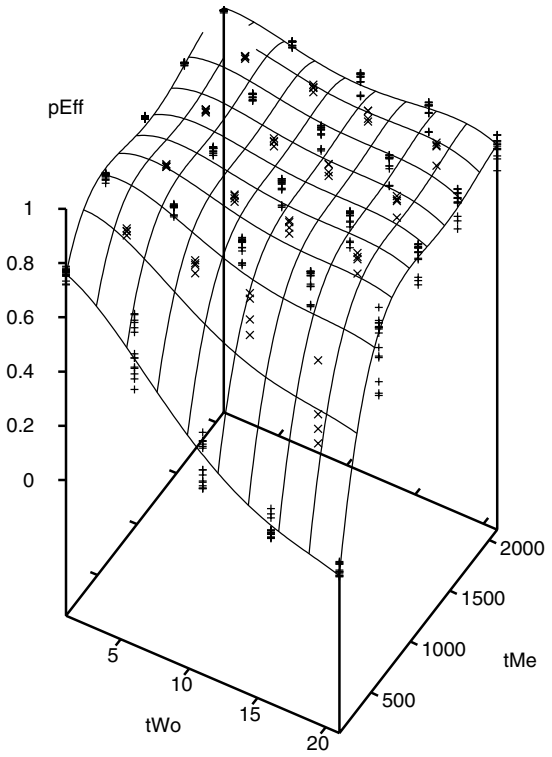

Figure 1. 3D plot of pEff

[3] E. Freeman, S. Hupfer, and K. Arnold. *JavaSpaces Principles, Patterns, and Practice*. Addison Wesley, 1999.

[4] D. Gelernter. Generative communication in Linda. *ACM TOPLAS*, 7(1):80–112, January 1985.

[5] F. Hancke, T. Dhaene, and J. Broeckhove. Performance characterisation of Tuple Spaces with Design of Experiments. In *Proceedings VECPAR, Volume I*, pages 175–187, 2004.

[6] F. Hancke, G. Stuer, D. Dewolfs, J. Broeckhove, F. Arickx, and T. Dhaene. Modelling overhead in JavaSpaces. In *Proceedings Euromedia*, pages 77–81. Eurosis, 2003.

[7] D. C. Montgomery. *Design and Analysis of Experiments*. John Wiley and Sons Inc., New York, 2001.

[8] J. Neter, M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. *Applied Linear Statistical Models*. WCB/McGraw-Hill, fourth edition, 1996.