

Unleashing Constraint Optimisation Problem solving in Big Data environments

Álvaro Valencia-Parra¹, Ángel Jesús Varela-Vaca¹, Luisa Parody², María Teresa Gómez-López¹

¹ {avalencia, ajvarela, maytegomez}@us.es

IDEA Research Group, Universidad de Sevilla, Sevilla

² {mlparody}@uloyola.es

Dpto. Métodos Cuantitativos, Loyola Andalucía, Sevilla

Palabras Clave: Big Data; Optimisation problem; Constraint programming; Distributed data; Heterogeneous data format

Lugar de publicación: Journal of Computational Science (Elsevier B.V) vol. 45 (2020)

Impact factor: JCR IF: 2.644 (Q2, categoría Computer Science, Theory & Methods, posición 28/108, año 2019)

DOI: <https://doi.org/10.1016/j.jocs.2020.101180>

Resumen(Abstract). The use of optimisation problems helps organisations manage their resources, time, and costs of their processes. The increase in the amount of consumable data, as well as the emergence of more sophisticated use cases, lead to a high number of variables and input data. Consequently, the resolution of huge quantities of Constraint Optimisation Problems (hereinafter, COPs), and their complexity, take an extra effort to model and solve them. In this paper, we aim at help stakeholders model their COPs, integrate different data sources with the COP variables and input data, and solving the COPs in a distributed environment by means of Big Data techniques.

In order to face these challenges, we developed FABIOLA (Fast Big Constraint Laboratory). It allows to solve COPs from large datasets in a systematic way, relying on modern Big Data technologies. We provide a user-friendly interface to facilitate the COP modelling, the execution of them, and the analysis of the results.

The approach is applied to an industrial scenario, in which several electricity wholesale companies employ constraint optimisation techniques to optimise the tariff which their customers might hire. By means of asymptotic analysis, we evaluate the performance of our proposal, determining the degree to which the distribution of COPs improves the execution time with respect to the sequential execution as the complexity of the dataset increases. Promising results are obtained.

FABIOLA isolates the resolution of COPs from where the data is located. Our systematic framework facilitates the integration of different data sources, the selection of the inputs of the COPs, the definition of optimisation models, their execution, and querying the results.