

Chapter 8

VisTHAA: A Statistical Tool for Comparison of Heuristics

Laura Cruz-Reyes

*National Mexican Institute of Technology,
Mexico*

Claudia Gómez

*National Mexican Institute of Technology,
Mexico*

Mercedes Pérez Villafuerte

*National Mexican Institute of Technology,
Mexico*

Nelson Rangel Valdez

*National Mexican Institute of Technology,
Mexico*

Marcela Quiroz-Castellanos

*National Mexican Institute of Technology,
Mexico*

César Medina Trejo

*National Mexican Institute of Technology,
Mexico*

Enith Martínez-Cruz

National Mexican Institute of Technology, Mexico

ABSTRACT

In this chapter, a scientific tool designed to facilitate fair comparisons of heuristics is introduced. Making a fair comparison of the performance of different algorithms is a general problem for the heuristic community. Most of the works on experimental analysis of heuristic algorithms have been focused on tabular comparisons of experimental results over standard sets of benchmark instances. However, from a statistical point of view, and according to the experimental design theory, a minimum requirement to compare heuristic algorithms is the use of non-parametric tests. Non-parametric tests can be used for comparing algorithms whose results represent average values, in spite of the inexistence of relationships between them, and explicit conditions of normality, among others. The proposed tool, referred to as VisTHAA, incorporates four non-parametric statistical tests to facilitate the comparative analysis of heuristics. As a case study, VisTHAA is applied to analyze the published results for the best state-of-the-art algorithms that solve the one-dimensional Bin Packing Problem.

DOI: 10.4018/978-1-4666-9779-9.ch008

INTRODUCTION

Many real-world problems are NP-hard, i.e. it is believed that there is no exact algorithm that solves them, and whose running time does not increase exponentially with the problem size (Garey & Johnson, 1979). There are two ways to attack the NP-hard problems. The first one is using *exact methods* that require exponential computational time. The second way, the most used in practice for large problems, corresponds to the use of *heuristic algorithms*, commonly called *heuristics*. Heuristics are procedures for which it is not possible to prove that the optimal solution will be obtained in finite time. Hence, a heuristic can be considered as a rule of thumb for guiding the solution search, achieving a good performance in terms of runtime and solution quality.

Heuristic algorithms include strategies based on common sense to provide approximate solutions in a reasonable time, but do not guarantee the optimal solution. Throughout the search for the best possible solutions for NP-hard problems, a wide variety of heuristic algorithms have been proposed. However, the “No free lunch” theorem (Wolpert & Macready, 1997) states that there is no efficient algorithm being better in behavior for all possible situations.

Metaheuristic algorithms are a special class of heuristics that include general algorithmic frameworks that can be applied to an extensive range of problems for producing heuristics of specific purpose. Generally, the principle of guided search of a metaheuristic is inspired by some natural phenomena. For example, in the case of evolutionary algorithms, they are based on genetic inheritance, and the Darwinian metaphor of “Natural Selection”.

The behavior of a metaheuristic is controlled by a set of components and parameters that are set to induce a desirable performance on a given optimization problem. When a metaheuristic is applied to a particular NP-hard problem, some modules need to be developed and some parameters need to be tuned. These two problems are called *structural* and *parametric* setting, respectively; the combination of these two problems is called *tuning*. The tuning process is crucial for optimizing a metaheuristic; the choice of values for the search parameters has a substantial impact on the performance of the procedure (Halim, 2009).

Heuristic algorithms are too complex to be studied analytically, but become accessible through an experimental analysis. In many cases, the design and tuning of metaheuristics are made manually and without theoretical support, hindering the understanding of their good or bad performance (Snodgrass, 2010). From a scientific perspective, a new algorithmic proposal must demonstrate an improvement in performance when comparing with the state-of-the-art algorithms, making a series of computational experiments and demonstrating statistical validity (Adenso-Diaz & Corral, 2007). Several situations have hampered the fair comparison between algorithms, including: the lack of standard test cases; the lack of standards for reporting computational experiments; incomplete and heterogeneous experimental results addressing one problem, presenting a single result, or using different equipment with unstandardized processing time, among others; and the ignorance of the most appropriate statistical test to validate results. Currently there are many attempts to rectify this situation. Guides to properly report computational experiments are presented in (Crowder et al., 1979; McGeoch, 2007; Preuss, 2007). A simple methodology for scaling different computer equipments is reported in (Quiroz-Castellanos et al., 2015). Some guides to identify the most appropriate statistical test can be found in (Adenso-Diaz & Corral, 2007; Derrac et al., 2011).

25 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/visthaa/145624

Related Content

Customer Value Perceptions: Testing of a Conceptual Model in the Frame of Own-Country Geographic Indication Foods

Toula Perrea, Katerina Melfou, Spiros Mamalis and Panoraia Papanagiotou (2016). *International Journal of Food and Beverage Manufacturing and Business Models* (pp. 1-11).

www.irma-international.org/article/customer-value-perceptions/145321

Developing High Affect-Based Trust in U.S.-China Business Negotiations

Maria Lai-Ling Lam (2021). *International Journal of Applied Management Theory and Research* (pp. 1-13).

www.irma-international.org/article/developing-high-affect-based-trust-in-us-china-business-negotiations/268896

The Role of Sustainable Performance Measurement System in Global Supply Chain

Kijpokin Kasemsap (2016). *Strategic Management of Sustainable Manufacturing Operations* (pp. 331-348).

www.irma-international.org/chapter/the-role-of-sustainable-performance-measurement-system-in-global-supply-chain/152402

Simulation Output Analysis and Risk Management

E. Jack Chen (2016). *Analyzing Risk through Probabilistic Modeling in Operations Research* (pp. 200-220).

www.irma-international.org/chapter/simulation-output-analysis-and-risk-management/140426

Measuring Tourist Attitudes and Consumer Behavior

Hanin Hosni, Henk Kievit, Periklis D. Drakos, George Angelakis and Gert van Dijk (2018). *International Journal of Food and Beverage Manufacturing and Business Models* (pp. 29-41).

www.irma-international.org/article/measuring-tourist-attitudes-and-consumer-behavior/210636