Control Engineering Practice 21 (2013) 1204–1225

Contents lists available at SciVerse ScienceDirect

Control Engineering Practice

journal homepage: www.elsevier.com/locate/conengprac

Benchmarking flexible job-shop scheduling and control systems

Damien Trentesaux [a,b], Cyrille Pach [a,b,*], Abdelghani Bekrar [a,b], Yves Sallez [a,b], Thierry Berger [a,b], Thérèse Bonte [a,b], Paulo Leitão [c,d], José Barbosa [a,b,c]

[a] University Lille Nord de France, F-59000 Lille, France
[b] UVHC, TEMPO Research Center, F-59313 Valenciennes, France
[c] Polytechnic Institute of Bragança, Campus Sta Apolonia, Apartado 1134, 5301-857 Bragança, Portugal
[d] LIACC—Artificial Intelligence and Computer Science Laboratory, R. Campo Alegre 102, 4169-007 Porto, Portugal

ARTICLE INFO

Article history:
Received 13 July 2012
Accepted 10 May 2013
Available online 19 June 2013

Keywords:
Benchmarking
Flexible job-shop
Scheduling
Optimization
Simulation
Control

ABSTRACT

Benchmarking is comparing the output of different systems for a given set of input data in order to improve the system's performance. Faced with the lack of realistic and operational benchmarks that can be used for testing optimization methods and control systems in flexible systems, this paper proposes a benchmark system based on a real production cell. A three-step method is presented: data preparation, experimentation, and reporting. This benchmark allows the evaluation of static optimization performances using traditional operation research tools and the evaluation of control system's robustness faced with unexpected events.

1. Introduction

Research activities in manufacturing and production control are constantly growing, leading to an increasing variety of scheduling and control solutions, each of them with specific assumptions and possible advantages. Despite this, a very small number attain the stage of industrial implementation or even tests in real conditions for several reasons. One of these reasons is the difficulty to provide robust, reliable performance evaluation of the control systems proposed that would convince industrials to take the risk to implement it. A first step towards a robust, reliable performance evaluation was made several years ago by the operational research (OR) community, which has proposed several benchmarks allowing the algorithms that try to solve static NP-hard optimization problems for production (e.g., routing, scheduling) to be compared.

Benchmarking is comparing the output of different systems for a given set of input data in order to improve the system's performance. In the OR literature, several benchmarks are often cited and widely used: Taillard (1993), Beasley (1990), Reinelt (1991), Kolisch and Sprecher (1996), Demirkol, Mehta, and Uzsoy (1998) and Bixby, Ceria, McZeal, and Savelsbergh (1998). The advantages of all these benchmarks are well known: very large databases using instance generators, and/or updating mechanisms for the community for improved bounds or optimal solutions. The problems are related to traditional OR problems (e.g., traveling salesman) and formalized problems (e.g., MILP); a large number of these benchmarks deal with scheduling problems (e.g., Hybrid Flow Shop Scheduling, Job Shop Scheduling, Hoist Scheduling Problem, Resource Constrained Project Scheduling). Thus, these benchmarks are useful to evaluate the quality of a scheduling method with a structured set of data – with all the data being complete, exact and available at the initial date – with no use of feedback control. As a result, the data handled in these benchmarks are quantitative and static, which allows a clear comparison of performances, in terms of makespan or the number of late jobs, for example.

From a control point of view, these benchmarks respond to part of the problem: the design of a scheduling plan in a static environment sometimes with *a priori* robustness analysis of results. However, it does not allow the dynamic behavior to be evaluated from a control perspective (i.e., a control feedback approach), updating real-time decisions based on observations of real-time events and unstructured data. Furthermore, these OR benchmarks were designed mainly from a theoretical point of view, with little attention paid neither to several constraints imposed by the reality of production systems such as limited production/storage/transport capacity, maintenance/inventory/tool/spacing constraints nor to dynamic events or data such as breakdowns or urgent/canceled orders. Moreover, these benchmarks cannot be adapted to emerging control architectures (e.g.,

* Corresponding author at: UVHC, TEMPO Research Center, F-59313 Valenciennes, France. Tel.: +33 327511322.
E-mail address: cyrille.pach@univ-valenciennes.fr (C. Pach).

0967-0661/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.conengprac.2013.05.004

distributed or coordinated control architectures), other than centralized in which all information is gathered and used by a unique central controller, which leads to incoherent comparisons or results if applied in these emerging architectures.

Despite this, an increasing production control activity is presently being led, focusing on alternative control architectures and their ability to behave in a dynamic environment, such as the proposal by Fattahi and Fallahi (2010). This is mainly due to the evolving industrial need, which can be summed up as follows: from traditional static optimized scheduling towards more reactive, sustainable or agile control. This evolving need leads to the need for more complex performance evaluation, not only expressed traditionally in terms of production delays for a given set of tasks, but also in terms of sustainability or the ability to evolve in a constantly changing world (e.g., energy consumption, carbon footprints).

The OR community has changed to consider this evolution. For example, one interesting action, directed by the French Operational Research and Decision Support Society (ROADEF), has led to the organization of several challenges since 1999. A challenge is a set of complex problems to be solved by the community, and the research team that proposed the best results is rewarded.[1] In our opinion, these challenges can be considered as benchmarks that were proposed to the whole community. In the beginning of these challenges, problems were purely static. However, more recently, the problem definition may contain some dynamic data, leading to re-assignment decisions to be made within fixed time window (e.g., the 2009 challenge). Meanwhile, even though production and scheduling were sometimes studied in these challenges, flexible production system's manufacturing and scheduling, and their specific constraints, have never been addressed.

The production control community has also proposed benchmarks intending to allow the coherent comparison of production control architectures and systems, taking the dynamics of the environment into account. For example, Valckenaers et al. (2006) proposed a benchmark that is methodologically oriented, dealing with the way to construct a benchmark for control evaluation. Brennan and O.W. (2002) proposed a benchmark designed to integrate dynamic data. Cavalieri, Macchi, and Valckenaers (2003) proposed a web simulation testbed for the manufacturing control community. Pannequin, Morel, and Thomas (2009) proposed an emulation-based benchmark case study devoted to a product-driven system, and Mönch (2007) proposed a simulation benchmarking system.

These benchmarks are interesting since they try to deal with the dynamic behavior of the system to be controlled, which is harder to formalize in a simple and exclusively quantitative way, like benchmarks from the OR community. If dynamic data, real-time considerations and unpredictable events must be managed and their impact evaluated, this drastically increases the complexity to develop a usable, clearly designed benchmark. In our opinion, this increasing complexity forces the production control benchmarks to focus on specific aspects of benchmarking (e.g., methodological or simulation aspects), restrict the control architecture too much (e.g., product-driven, distributed), or compel the researchers to use specific tools (e.g., simulators). In addition, none of these benchmarks offers operational, fully informed data sets for coherent tests and comparisons. Therefore, despite some very interesting trials and the huge effort, these benchmarks are not often used as the OR community benchmarks.

In the constantly evolving research environment, with a unceasingly increasing importance paid to quality of results, researchers from the production control community, and a growing number of researchers from the OR community, are still seeking for a benchmark that can help them to characterize the static and dynamic behaviors of their control system, taking realistic production constraints into account.

Drawn from the experience of the authors, the conclusion of this literature review is that it is interesting to define a benchmark, allying the advantages of the benchmarks proposed by both communities, usable by both communities, and based upon a physical, real-world system to stimulate benchmarking activities to be grounded in reality. To propose such a benchmark to researchers is the aim of this paper. This conclusion is also consistent with the current determination of the IFAC TC 5.1, which tries to design, use and disseminate of manufacturing control benchmarks.

The rest of the paper is organized as follows. First, Section 2 introduces the proposed benchmark process. Sections 3, 4, and 5 detail the three steps of this benchmarking process: data preparation, experimentation, and reporting. Section 6 presents three applications of the benchmark for illustration purpose. Finally, Section 7 draws the conclusions and presents the prospective for future research.

2. The benchmarking process

Three consecutive steps compose the proposed benchmarking process, which are presented in Fig. 1.

The first step, called *data preparation*, concerns the sizing and the parameterization of the case study. Given a generic model of the target system to schedule and control, the *first benchmarking decision* is to choose the "data set". A data set includes usually an instance of a model of the target system on which the benchmark is applied, accompanied with the input data needed to make this model work. Once this data set is chosen, *the second decision* concerns the definition of the objective function. The couple (data set, objective function) defines the **reference scenario**, called scenario #0. The *third decision* to make in this step is to decide whether or not dynamic behavior should be tested. If yes, then the *fourth and last decision* is that the researchers must decide which dynamic scenarios they are willing to test in a list of dynamic scenarios.

Once defined, scenario #0 contains only static data (i.e., all the data is known at the initial date), which allows researchers to test deterministic optimization mechanisms for a given set of inputs (e.g., OR approaches, simulation or emergent approaches, multi-agents approaches), especially if only few constraints are relaxed. In this stage, performance measures are purely quantitative. Scenario #0 can be used to test different optimization approaches, to evaluate the improvement of certain criteria (e.g., C_{max} values), or to check the basic behavior of a control system in real time where all data are known initially.

The second step, called *experimentation*, is composed of two kinds of experiments:

1. The *static stage*, which concerns the treatment of the reference scenario (i.e., scenario #0), and
2. The *dynamic stage*, which concerns the treatment of the dynamic scenarios.

If the researchers had selected the second option in the previous step, they will execute two types of dynamic scenarios, introducing perturbations (1) on the target system and (2) on the control system itself. In this stage, researchers can test control approaches and algorithms, using the scenario #0 into which some dynamic events are inserted, which defines several other scenarios with increasing complexity.

[1] http://challenge.roadef.org/2012/en/index.php.