Discrete-event simulation using estimation of distribution algorithm

Ricardo Pérez, Alberto Ochoa-Zezzatti, Jöns Sánchez and Arturo Hernández

Simulación de eventos discretos usando un algoritmo de estimación de distribuciones

Recibido: noviembre 2, 2011
Aceptado: diciembre 13, 2011

Palabras clave: Modelos de simulación; sistemas de manufactura; investigación de operaciones

Abstract:
The science of decision making or operations research (OR) is present in all levels and in all industries. The Industrial Optimization can carry out the decision-making through analysis of the operation of any system, formulation and use of models to achieve the proposed goals and targets, with proper utilization of available resources. Its scope is very broad, applying to problems of manufacturing, transportation, construction, telecommunications, planning, financial management, health sciences, and services, among others. In this preliminary work, we present a hybrid proposal between discrete event simulation and a distribution estimation algorithm; our goal is to contribute to knowledge in the area. This work is being developed in a manufacturing plant where steel doors and frames are parallel activities.

Keywords:simulation models; manufacturing systems; operations research

IN general, decision making can be applied in all problems related to the management, planning and design [2]. The companies designed to comply lots of product orders or high volume production processes in services have highly repetitive tasks [3]. In these cases the task scheduling is often used to determine the sequence of production batches of a product or service for compliance with delivery order date [3]. Also in a manufacturing context, task scheduling can be used to balance the workload among operators [3].

One of the problems that occur naturally in some companies is the flow-line in manufacturing systems [1]. The problem is how to sequence tasks in machines as efficiently as possible. If not done wisely, this can present an opportunity cost (economic loss) for the company. Therefore, it is important to find the order in which it must process all the tasks in the line to minimize the time of completion [1]. This research aims to analyze and find a better way of implementing the planning of operations in a manufacturing line of steel doors, and by doing a proper scheduling of manufacturing operations, the following objectives are expected to be achieved:

1. Reducing inventories in process within the manufacturing line: to be able to develop different production scenarios with different orders to produce and have the

best possible sequence for these scenarios, it seeks to reduce the inventory of materials and components required in manufacturing, ie to have intelligently only required material and not in excess.

2. Improve occupancy time and the equipment operators: to determine the best sequences for different production scenarios also seeks to find appropriate and justified amount of resources, resulting in a more slender, ie, with greater occupancy rate of personnel and equipment, reduced downtime, reduced extra time spent in production and associated costs, thereby achieving greater productivity.

3. Reduce the average time process flow for each product: clearly having software solutions to manufacturing operations, we will be able to determine the stage that significantly reduces the average flow time expected orders to be processed so as to be produced.

4. Generate efficient alternative policies for the proper sequencing and dispatch of orders: to demonstrate that there are better alternatives for sequencing and delivery of current orders, we want to get a solid knowledge a priori of the impacts that may arise if not considered possible programming options.

5. To systematize the programming decisions at various scenarios, to improve productivity and capacity: establishing clear and well defined guidelines for programming operations.

With a proposed hybrid between discrete event simulation and an estimation algorithm distributions, our goal is to contribute to knowledge in the area:

1. Since we do not have enough current references in the literature on estimation of distribution algorithms in the field of programming operations with discrete event simulation hybrid proposal.

2. The search for solutions to current questions of scheduling operations in manufacturing doors for the capacity, productivity and changes in scenarios.

Basic concepts

There are often conflicts between jobs that use the same resources because it can process only one job at a time [3]. Solving the task scheduling is to decide, for each conflict between tasks and resources, which of the tasks must be executed before [3]. The task scheduling problem is on the allocation of limited resources to certain areas or operations through a certain period of time [9]. Usually the purpose is established by defining the order of tasks per resource allocation that minimizes the completion time of all jobs, also known as makespan [3]. Other possible targets are:

- Minimize the delay of work.
- Minimize time in the system.
- Minimize work in process inventory.
- Minimizing the maximum delay in delivery.
- Minimize the average time of completion of work, among others.

Task scheduling is used to decide which of the tasks must be executed first

The nature of the task scheduling problems is combinatorial and therefore the time required to find the optimal solution grows exponentially with the number of tasks considered [3]. In fact, the problem posed by Muth & Thompson in 1959 in which the objective is to minimize the makespan for 10 jobs 10 machines could not be solved optimally up to 20 years [3]. Research in the field of sequencing and scheduling (sequencing and scheduling) has been extensive [2]. Early efforts have been developed by Johnson [9]. Results of computational complexity, exact optimization schemes and approximation algorithms have been tested in several areas [4].

A method as those amply used in other contexts of production and mathematical programming is not appropriate in this area, since their ILP mixed formulation results in high running times, for actual size problems, which makes it little advisable [3]. Other procedures are heuristics. These are based on logical rules extracted from the user experience to generate valid solutions in reasonable times [3].

For four decades, have proposed heuristics that find approximate solutions almost immediately [6]. For scheduling problems Baker [7] published compendia of heuristic methods, while authors such as Campbell et al. [8] popularized their specialized heuristics. A useful example is the special case where problems are considered as 2 machines, it has been solved by Johnson [1]. Johnson's rule is an exact algorithm, which gives the optimal sequence for 2 machines problem with n tasks [1]. As is well known in the area, when the problem is greater than or equal to 3 machines, there is no simple rule that provides an optimal solution to the problem [1]. Most problems in this area are combinatorial optimization problems and a large proportion of them belong to the class of NP-hard problems. NP-hard problems are a subset of the class NP (problems which can not have a solution in polynomial time for all instances) with the characteristic that all problems of this kind can be reduced to NP [9] . The problems which can find a solution algorithm in polynomial time form the P type, which is a subset of the class NP [9]. NP-hard problem arises when an algorithm that tries to fix it, increases its running time, in the worst case, exponentially to the size of the problem [9].

Heuristic algorithms are easily generalized to several types of problems, such as genetic algorithms, simulated annealing or tabu search; these are called metaheuristics algorithms [3].Most of them emulate processes that occur in nature. Thus, genetic algorithms mimic the process of natural selection over successive generations of individuals representing possible considered solutions. The new individuals are created across couples who are selected giving preference to those solutions with better objective function value (fittest to the environment) [3]. Given the importance of the scheduling problem from the beginning, genetic algorithms, with different variants have been successfully applied to this problem [5]. Some authors have gone to the area of evolutionary computing, seeking efficient storage structures of chromosomes and variations in breeding and mutation operations to obtain better results in such problems [5]. Genetic algorithms are efficient metaheuristics problems, as they perform a good balance between search space exploration and exploitation of subschemes which are encoded within the solutions. For this reason, variation mechanisms are used that are sensitive to the selected representations and often require knowledge of its behavior for fixing the values of its parameters [10]. However, the estimation of distribution algorithms EDAs (Estimation of Distribution Algorithms) are a class of algorithms based on evolutionary computing paradigm mechanisms that replace variation (crossover and mutation) traditionally used by evolutionary algorithms. The population of new solutions are generated by simulating a probability estimate produced by the information of the solutions generated in previous iterations [10]. They are derived from evolutionary computation approaches, unlike those that simulate a probability distribution and variation mechanisms require. Estimate a probability distribution requires a process of learning from data models that are produced by feedback information. The performance observed in FDTs approaches, it is very promising, both from the standpoint of the quality of the solutions generated, and the speed of convergence of the algorithm (number of tests required). However, learning the probabilistic model is usually the most costly step in terms of computation and the need to promote new ways to reduce the computational cost, which represents the main weakness of this novel method [10]. Furthermore, to adequately handle the actual dynamic environments, where there are discrete events that affect system performance, as is the case for the manufacture of steel doors, simulation modeling is used. The discrete event simulation is an important branch that has contributed to the successful definition of programs of operations due to the possibility of integrating the conditions of the manufacturing environment not only for the machines or processes but also the human factor involved in the operations of different companies [11].

The discrete event simulation significantly supports those decisions where it is difficult to predict with certainty the scenarios that would result if some detail was not considered or omitted at the time of program operations, making the simulation not only an additional tool to the decision maker, but also a technique which reduces the economic impact of these complex decisions. [11] demonstrates the benefits achieved globally. The construction of complex models in manufacturing becomes accessible through current technology in simulation and accelerates the ability to identify opportunities to schedule operations and detect impacts that generates the same schedule. Experimentation and analysis done through simulation can integrate the experience of making decision makers optimize what has already been discussed above, only with a wider focus, where multiple constraints and objectives can be met through simulation.

The simulation makes it possible to analytically evaluate complex models of manufacturing, reducing the risk of poor planning and predicting impacts well in advance not only in programming operations, but in actual production.

[11,12,13,14] set out simulation modeling that has been implemented by corporations around the world, the objective was to improve the design and operation of their systems as stochastic process, getting complex analytical expressions for the variables of input / output. Furthermore, when the system interdependence and variability increase, not only decreases the performance, but also the ability to predict human resource performance [15].

Simulation allows to analytically evaluate complex models of manufacturing, reducing the risk of poor planning and predicting impacts well in advance not only in programming operations, but in actual production

The simulation process illustrates what would happen to a system under a particular configuration. This configuration includes the program for the system, the physical arrangement of the machines, the rules of information flow and work as well as failures and expected shifts. The criteria can be included in the route to give higher priority to certain events over others, including failure to follow through on the other most important operation, or you can create files to show the expected time of commencement and termination for each job. Thus, the simulation can perform sequencing once the basic information and orders process defined in the model. In fact, it can feed alternative production programs via spreadsheets and create different scenarios to consider several alternatives. The result of each stage can be analyzed simultaneously, so that the best sequence is chosen. However, the simulation is not a pure programming system, as in an MRP II or a package of production scheduling. For the simulation to be credible, it must consider the occurrences of real life and the randomness and interdependence between resources and parts, operators, among others, while in a production program, that cannot be considered [13]. This paper settles a proposed hybrid able to find good solutions in reasonable smaller times.

Solution methodology

This paper poses the problem of finding a sequence of n tasks in a flow line (flow shop) of m machines that minimizes the number of tasks that are delivered late. Mathematically the problem is to find a permutation of the tasks that allows to give the least number of late tasks which is a typical problem of sequencing tasks. It is important that the tasks meet specified delivery times, because in most cases, the violation of these days involves a penalty [2]. In a flow line, it has a set of n tasks that must be processed in a set of m machines each one. Each task has the same order of routing through the machines, ie each of the tasks should be processed first in the machine 1, then the machine 2, and so on until the machine m. It is assumed that each task is initially available and it has a delivery time, also the sequence of tasks on each machine is the same and each machine must process one task at a time [2]. The processing time of task j on machine i is denoted by p_{ij}. The delivery time task j is denoted by d_j. The time when a task exists in the system (ie, time of task completion in the last machine which needs to be processed) is a variable that is denoted by C_j [2]. Penalty unit (variable) of the task j is defined in the equation (1) as:

$$U_j = \begin{cases} 1, \text{if } C_j > d_j \\ 0, \text{if } C_j \leq d_j. \end{cases} \tag{1}$$

The goal is to find a sequence of n tasks to minimize the number of tasks that are delivered late. The function to be minimized is called the objective function [2]. Then the objective function (see Equation 2) is represented as:

$$\min \sum_{j=1}^{n} U_j. \tag{2}$$

To minimize this number, it should be found a sequence that optimizes this objective function (ie, delivering the lowest number of late tasks) [2]. We start from the position of developing a hybrid proposal between discrete event simulation and an estimation algorithm that offers better solutions distributions in scheduling operations. Therefore the hybridized proposal (see Figure 1) is generally described as follows:

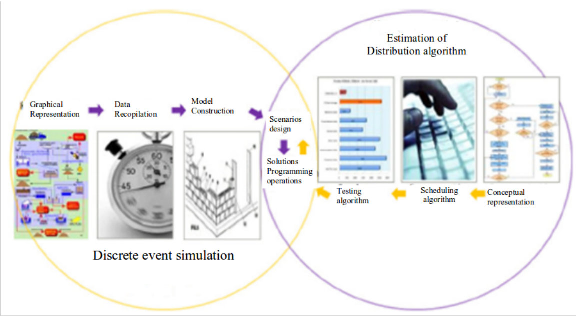

Figura 1. Hybrid proposal system to resolve Discret-event simulation.

We start from the position of developing a hybrid proposal between discrete event simulation and an estimation algorithm

---

For the case of discrete event simulation we have the following phases:

1. Generate a visual and conceptual representation of the manufacturing process, where the variables show the greatest impact on the process, define outputs or specific numerical values under study, the inputs such as initial operating conditions, a priori known values of operation, and uncertain values with a probability distribution.

2. Data collection processes, operators, equipment and measurements for each of the different product models offered by the company. Then analyze the data to determine probability distributions that fit those processes in order to correctly represent the real situation of manufacturing.

3. Building the model through specialized software that is flexible and practical in its construction, which would modify the operating conditions and dynamically generate real operating scenarios that can later be validated with a level of certainty to ensure that the constructed model makes it possible to analytically evaluate complex models of manufacturing of doors, reducing the risk of poor planning and predicting impacts well in advance not only in programming operations, but in actual production.

4. Develop scenarios to detect impacts, it generates operations scheduling, or some detail that was not considered or omitted, illustrating what would happen under particular operating conditions.

For the case of the estimation algorithm distributions, we have the following phases:

1. Generate a conceptual graph and the steps to be executed by the algorithm for estimating distributions.

2. Scheduling algorithm for estimating distributions in a powerful and versatile language capable of receiving and transferring information to the simulator.

3. Performance testing and performance of the algorithm is programmed to ensure that is feasible and responsive to the demands for which it was programmed, according to the current literature.

4. Transfer the algorithm simulator scenarios to process and determine the best scheduling solution operations to determine the optimal sequence for each scenario.

Once you have software solutions, you will proceed to validate these solutions in terms of operation, to ensure that such sequences can be implemented without hindrance.

Conclusion

This work is being developed in a manufacturing plant where steel doors and frames are parallel activities, and it is expected to make this hybrid proposal successful. In the Brewing Industry its hybrid proposal was successful. To the extent that the scheduling of operations is to produce better alternative solutions in the manufacture of steel doors for programming operations, there will be a contribution of knowledge in this field.

Bibliografía

[1] M.L. Valenzuela y R.Z. Rios. “Comparando métodos heurísticos para secuenciar tareas en líneas de flujo”. Ingenierías, UANL, 7(25), pp. 35-39, Oct - Dic, 2004.

[2] M.A. Salazar y R.Z. Rios. "Minimización heurística del número de tareas tardías al secuenciar líneas de flujo". Ingenierías, UANL, 7(23), pp. 52-57, Abr - Jun, 2004.

[3] P. Sanchez y S. Lopez de Haro. "Programación de tareas, un reto diario en la empresa".Anuales de Mecánica y Electricidad, UPCO, pp. 24-30. May – Jun 2005.

[4] R.Z. Rios y J.F. Bard. "Secuenciando óptimamente líneas de flujo en sistemas de manufactura". Ingenierías. UANL. 4(10). pp. 38-45. Ene – Mzo. 2001.

[5] B. Pérez y M.A. Osorio. "Análisis comparativo de heurísticas para el problema de calendarización de trabajos con transferencia cero". 1er Congreso Mexicano en Computación Evolutiva (COMCEV), Centro de Investigación en Matemáticas (CIMAT), pp. 43-54.

[6] S.H. Zanakis y J.R. Evans. "Heuristic Optimization: why, when, and how to use it". Interfaces. 11(5). October 1981.

[7] K. Baker. Introduction to sequencing and scheduling, John Wiley & Sons, New York 1974.

[8] H.G. Campbell. R.A. Duder y M.L. Smith. "A heuristic algorithm for the n-job, m-machine sequencing problem". Management Science. 45(4). pp. 630-637. 1970.

[9] E. Delgado. C.J. Cortés. O. Duarte. "Aplicación de algoritmos genéticos para la programación de tareas en una celda de manufactura". Ingeniería e Investigación. Universidad Nacional de Colombia. 25(2). pp. 24-31. 2005.

[10] D. Pandolfi. A. Villagra. G. Leguizamón. "Algoritmo de Estimación de Distribución aplicado al problema de secuenciamiento de Flow Shop". XII Reunión de Trabajo en Procesamiento de la Información y Control, pp. 1-6. 2007.

[11] SJE Taylor; S. Robinson. So where the next? A survey of the future for discrete-event simulation. Journal of Simulation 0, pp. 1-6. 2006.

[12] Diamond, R.; Harrell, R.C.; Henriksen, O.J, Nordgren, B.W.; Pegden, C.D., Rohrer W.M.; Waller, P.A.; Law, M.A. The current and future status of simulation software (Panel). Proceedings of the 2002 Winter Simulation Conference, pp. 1633–1640. 2002.

[13] Harrell, R. C.; Ghosh, B. K.; y Bowden, R. O. Simulation Using Promodel. 2nd edition. McGraw-Hill, New York. 2004.

[14] Rosen, S. L.; Harmonosky, C. M.; and Traband, M. T. Optimization of systems with multiple performance measures via simulation: Survey and recommendations. Computers & Industrial Engineering 54(2) pp. 327-339. 2008.

[15] Lloyd, S. y Melton, K. Using Statistical Process Control to Obtain More Precise Distribution Fitting Using Distribution Fitting Software. Simulators International XIV 29(3), pp. 193-98. 1997.

Acerca del autor o autores

Ricardo Pérez and Jöns Sánchez are with CIATEC, México. Alberto Ochoa-Zezzatti is with the Juárez City University México, and Arturo Hernández is with CIMAT, México. alberto.ochoa@uacj.mx