Stochastic scheduling

[Updated version of article in *Encyclopedia of Optimization*, C. A. Floudas and P. M. Pardalos, eds., vol. V, pp. 367-372, 2001. Kluwer, Dordrecht]

José Niño-Mora *

July 10, 2002

The field of stochastic scheduling is motivated by the design and operational problems arising in systems where scarce service resources must be allocated over time to jobs with random features vying for their attention. Important examples include manufacturing and computer-communication systems. Consider, e.g., the case of a manufacturing workstation processing different part types, where part arrival and processing times are subject to random variability.

The performance of such systems, as measured by a criterion such as the average time jobs stay in the system (*flowtime*), may be significantly affected by the policy employed to prioritize over time jobs awaiting service (*scheduling policy*). The impact of scheduling policies, together with the high degree of discretionality in the decisions they involve, explain the importance and difficulty of the fundamental problem of stochastic scheduling: to design relatively simple scheduling policies that (nearly) achieve given performance objectives.

The theory of stochastic scheduling addresses this problem in a variety of stochastic service system models. Random features such as job processing times are thus modeled by specifying their probability distributions, which are assumed to be known by the system manager. Model assumptions vary across several dimensions, including the class of scheduling policies considered admissible, job arrival and processing time distributions, type and arrangement of service resources and performance objective to be optimized.

Regarding methods and techniques, it seems fair to say that no unified and practical approach has been developed to design and analyze (nearly) optimal policies across the range of stochastic scheduling models. Although many such models can be cast in the framework of dynamic programming, straightforward application of this technique has not proven very

*Department of Economics and Business, Universitat Pompeu Fabra, Barcelona, Spain; email: jnimora@alum.mit.edu

effective, due to the large (or infinite) size of the resulting formulations (*curse of dimensionality*). Most results have been instead obtained through problem-specific arguments, which often do not extend to seemingly related models. However, a few techniques, such as interchange arguments [26], and some general principles, such as conservation laws [4], have been successfully applied to wider model classes.

Stochastic scheduling models can be classified into three broad categories, which have evolved with a substantial degree of autonomy: models for scheduling a batch of stochastic jobs, multi-armed bandit models, and models for scheduling queueing systems. The development of each of these areas has followed a similar three-stage pattern. The first results elucidated the structure of easily computable policies that solved optimally relatively simple models. An important class of such policies is that of *priority-index rules*: an index is computed for each job type (possibly depending on its current state, but not on that of other jobs), and at each decision epoch jobs of higher index are assigned higher service priority.

A second group of results has sought to identify optimal policies with a simple structure in more general models, often at the expense of introducing additional technical assumptions.

More recent research efforts have addressed harder models, for which the goal of fully characterizing an optimal policy appears out of reach. For these problems researchers aim to design easily implementable heuristic policies with a relatively close to optimal performance. Their degree of suboptimality may be investigated empirically (through simulation) or analytically.

1 Models for scheduling a batch of stochastic jobs

In this class of models a fixed batch of n jobs with random processing times, whose distributions are known, have to be completed by a set of m machines to optimize a given performance objective.

The simplest model in the class represents the problem of sequencing a set of n stochastic jobs on a single machine to minimize the expected weighted flowtime. Job processing times are independent random variables, having a general distribution $G_i(\cdot)$ with mean p_i for job i. Admissible policies must be *nonanticipative* (scheduling decisions are based on the system's history up to and including the present time: they cannot depend on future information, such as the actual processing times of unfinished jobs) and *nonpreemptive* (processing of a job, once started, must proceed uninterruptedly to completion). Let $w_i \geq 0$ denote the cost rate incurred per unit time in the system (waiting or in service) for job i, and let $\tilde{C}_i$ denote its random *completion time* (time at which the job is finished). Let Π denote the class of all admissible policies, and let $E_\pi[\cdot]$ denote expectation under policy $\pi \in \Pi$. The problem can be stated as

$$\min_{\pi \in \Pi} w_1 E_\pi[\tilde{C}_1] + \cdots + w_n E_\pi[\tilde{C}_n].$$

The optimal solution in the special deterministic case is given by the Shortest Weighted Processing Time rule of Smith [36]: sequence jobs in nonincreasing order of the priority index $w_i p_i$. It was shown in [33] that the natural extension of Smith's rule to the above stochastic model (Shortest Weighted Expected Processing Time rule), based on the same index $w_i p_i$, is optimal for it.

The model extension where policies are allowed to be *preemptive* (processing of a job may be interrupted at any time, and resumed later), was solved in [34]. The optimal policy is again characterized by priority indices (a job with higher index is processed at each time), which in this case depend on the cumulative amount of processing each job has received.

The optimality of priority-index policies has been extended to restricted classes of models for scheduling a batch of jobs on identical parallel machines. The main performance objectives investigated in this setting are: (1) total expected flowtime minimization,

$$\min_{\pi \in \Pi} E_\pi \left[\sum_{j=1}^{n} \tilde{C}_j \right];$$

and (2) expected *makespan* (finishing time of the last job) minimization,

$$\min_{\pi \in \Pi} E_\pi \left[\max_{1 \le j \le n} \tilde{C}_j \right].$$

The rule that assigns higher priority to jobs with shorter expected processing time (SEPT) has been shown to be optimal for the flowtime objective under the following assumptions: when all the job processing time distributions are exponential [20]; when all the jobs have a common general processing time distribution (having possibly received different amounts of processing prior to start) with a nondecreasing hazard rate function [40]; and, more generally, when job processing time distributions are stochastically ordered [42].

For the expected makespan objective, the rule that assigns higher priority to jobs with longer expected processing times (LEPT) has been shown to be optimal in the following cases: under exponential processing time distributions [10]; and when jobs have a common processing time distribution (with possibly different amounts of processing prior to start) with a nonincreasing hazard rate function [40].

Other models incorporate more general features, such as *uniform* parallel machines, which differ in speed rates. Under relatively strong assumptions, researchers have characterized optimal policies, which exhibit a threshold structure: see [1], [32] for the problem of expected flowtime minimization, and [12] for the problem of expected makespan minimization. An optimal policy for the problem of scheduling a batch of stochastic jobs in a *flow shop* (with m machines in series) is studied in [48].

The optimality of the simple policies mentioned above typically fails to extend to models that violate the required assumptions [13]. Characterizing an optimal policy in such cases appears to be a computationally intractable goal (see [29] for a study on the complexity of decision-making problems under uncertainty, such as stochastic scheduling). This fact has motivated the investigation of suboptimal heuristic policies, often based on the simple rules mentioned above.

For example, it has been shown in [45] that, under mild assumptions, the suboptimality gap for using Smith's rule as a heuristic for parallel machines stochastic scheduling is bounded above by a quantity that is independent of the number of jobs. Therefore, as the latter grows to infinity, the heuristic's relative suboptimality gap vanishes, which represents a desirable kind of asymptotic optimality. A similar asymptotic optimality result, in a model of parallel machines stochastic scheduling with in-tree precedence constraints, was presented previously in [30].

2 Multi-armed bandit models

Models in this class are concerned with the problem of optimally allocating effort over time to a collection of projects, which change state in a random fashion depending on whether they are engaged or not. A classic example is the multi-armed bandit problem which, in its discrete-time version, can be described as follows: there is a collection of N projects, exactly one of which must be engaged at each discrete time epoch $t = 0, 1, \ldots$. Project n can be in a finite number of states $j \in J_n$. If at time t project n, with current state $i \in J_n$, is engaged, two effects result: first, a reward R_i is received, discounted in time by a factor $0 < \beta < 1$; second, the project state changes in a Markovian fashion to j with probability p_{ij}, for $j \in J_n$. States of projects not engaged remain unchanged. The problem consists in finding a nonanticipative scheduling policy, for selecting projects over time, that maximizes the total expected discounted reward over an infinite horizon. Letting Π denote the class of all nonanticipative policies, and letting $j(t)$ denote the state of the project engaged at time t, the problem can be stated as

$$\max_{\pi \in \Pi} E_\pi \left[\sum_{t=0}^{\infty} \beta^t \, R_{j(t)} \right].$$

After being considered intractable for a long time, the problem was solved in a landmark result by Gittins and Jones [19]. The optimal policy is given by Gittins priority-index rule: an index γ_j is computed (in a finite number of steps) for each project state j; then, the rule selects at each time a project with larger current index. The optimality of Gittins rule, for the original model and extensions, has a rich history of proofs, based on different techniques, including interchange arguments [19, 18, 39, 44], dynamic programming [46], intuitive arguments [41], induction arguments [38], and conservation laws/linear programming [4].

For more complex model extensions, the Gittins rule is no longer optimal. The incorporation of costs/delays for switching between projects is studied in [2], where a partial characterization of an optimal policy is given. This reduces the computational burden for calculating an optimal policy, which however remains impractical for large problems (as it grows exponentially with the model size).

The model extension where projects not engaged continue to evolve, possibly with different transition probabilities, and a fixed number $m \geq 1$ of projects must be engaged at each time, was addressed by Whittle in [47]. In the setting of a time-average version of such *restless bandit problem*, he proposed a heuristic priority-index rule based on the optimal solution to a *relaxed* linear programming formulation of the problem: the requirement that *exactly* m projects be selected at each time is relaxed by requiring instead that m projects be selected only *on average*. Whittle's rule reduces to that of Gittins when specialized to the classical model. His conjecture regarding the asymptotic optimality of the index rule was established, under appropriate conditions, in [43]. A different priority-index heuristic, obtained from Whittle's linear programming relaxation, together with improved performance bounds, has been developed and tested computationally in [7]. A polyhedral framework for analysis and computation of the Whittle index and extensions, based on the notion of *partial conservation laws*, has been recently developed in [28, 27].

3 Queueing scheduling control models

Models in this class are concerned with the problem of designing optimal service disciplines in queueing systems, where the set of jobs to be completed, instead of being given at the start, arrives over time at random epochs. The main class of models in this setting is that of *multiclass queueing networks* (MQNs), widely applied as versatile models of computer-communications and manufacturing systems.

The simplest types of MQNs involve scheduling a number of job classes in a single server. Similarly as in the two model categories discussed previously, simple priority-index rules have been shown to be optimal for a variety of such models. Consider the case of a multiclass $M/G/1$ queue, where N job classes vie for the attention of a single server: Jobs of class j arrive at the system at epochs given by a Poisson process with rate α_j, and their service times are drawn independently from a common distribution $G_j(\cdot)$ with mean $1/\mu_j$. Class j jobs incur linear holding costs at a rate $c_j \geq 0$ per unit time in the system (waiting or in service). The goal is to find a nonanticipative and nonpreemptive scheduling policy, for deciding which job to serve at each decision epoch, that minimizes the expected steady-state holding cost rate. Let Π denote the class of all such admissible policies, and let $E_\pi[L_j]$ denote the steady-state expected number of class j jobs in the system under policy $\pi \in \Pi$. The problem can be stated as

$$\min_{\pi \in \Pi} c_1 E_\pi[L_1] + \cdots + c_N E_\pi[L_N].$$

Its solution is given by the classical $c\mu$-rule [15]: select for service at each decision epoch a job with larger priority-index $c_j \mu_j$. The $c\mu$-rule is also optimal among preemptive policies when service times are exponential.

The optimality of priority-index policies for the model extension that incorporates Markovian job feedback (when a class i job completes service it changes class to j with probability p_{ij}, $1 \leq j \leq N$, and leaves the system with probability $1 - \sum_{j=1}^{N} p_{ij}$) was established by Klimov in [24]. The optimal priority indices are efficiently computed by N-step *Klimov's algorithm*. The result was shown to extend to a time-discounted objective in [37].

An account of these results based on the *achievable region method*, which seeks to characterize the region of achievable system performance (e.g., mean queue lengths) by means of linear (or convex) programming constraints, has been given in [14, 17, 35, 4] (in an increasing level of generality). The performance of Klimov's rule, when used as a heuristic for the model extension that incorporates identical parallel servers, has been analyzed using this approach in [22]: a *relaxed* linear programming formulation of the performance region is shown to yield closed-form suboptimality bounds, which imply asymptotic optimality in the heavy-traffic limit.

More general MQN models involve features such as changeover times for changing service from one job class to another [25], or multiple processing stations, which provide service to corresponding nonoverlapping subsets of job classes. Due to the intractability of such models, researchers have aimed to design relatively simple heuristic policies which achieve a performance close to optimal. This goal has been hindered by formidable technical challenges, including the *stability problem* for multiclass queueing networks with multiple stations [9]: in general it is not known what conditions on model parameters ensure that a given policy is *stable* (the time-average number of jobs in the system is finite). As a result, computer simulation remains the most widely used tool in applications of these models. Theoretical approaches currently under active development include: study of heuristic scheduling policies based on diffusion approximations of the original system under heavy-traffic conditions [23], [31]); study of policies based on fluid approximations [11, 3] and development of performance bounds and policies by means of the achievable region method [8, 5, 6, 21, 22, 16, 27].

References

[1] A. K. Agrawala, E. G. Coffman, Jr., M. R. Garey, and S.K. Tripathi. A stochastic optimization algorithm minimizing expected flow times on uniform processors. *IEEE Transactions on Computers*, c-33:351–356, 1984.

[2] M. Asawa and D. Teneketzis. Multi-armed bandits with switching penalties. *IEEE Transactions on Automatic Control*, 41:328–348, 1996.

[3] D. Atkins and H. Chen. Performance evaluation of scheduling control of queueing networks: fluid model heuristics. *Queueing Systems. Theory and Applications*, 21:391–413, 1995.

[4] D. Bertsimas and J. Niño-Mora. Conservation laws, extended polymatroids and multiarmed bandit problems: polyhedral approach to indexable systems. *Mathematics of Operations Research*, 21(2):257–306, 1996.

[5] D. Bertsimas and J. Niño-Mora. Optimization of multiclass queueing networks via the achievable region method: part I, the single-station case. *Mathematics of Operations Research*, 24(2):306–330, 1999.

[6] D. Bertsimas and J. Niño-Mora. Optimization of multiclass queueing networks via the achievable region method: part II, the multi-station case. *Mathematics of Operations Research*, 24(2):331–361, 1999.

[7] D. Bertsimas and J. Niño-Mora. Restless bandits, linear programming relaxations and a primal-dual index heuristic. *Operations Research*, 48(1):80–90, 2000.

[8] D. Bertsimas, I. C. Paschalidis, and J. N. Tsitsiklis. Optimization of multiclass queueing networks: Polyhedral and nonlinear characterizations of achievable performance. *The Annals of Applied Probability*, 4:43–75, 1994.

[9] M. Bramson. Instability of FIFO queueing networks. *The Annals of Applied Probability*, 4:414–431, 1994.

[10] J. Bruno, P. Downey, and G. N. Frederickson. Sequencing tasks with exponential service times to minimize the expected flow time or makespan. *Journal of the ACM*, 28:100–113, 1981.

[11] H. Chen and D. D. Yao. Dynamic scheduling of a multiclass fluid network. *Operations Research*, 41:1104–1115, 1993.

[12] E. G. Coffman, Jr., L. Flatto, M. R. Garey, and R. R. Weber. Minimizing expected makespans on uniform processor systems. *Advances in Applied Probability*, 19:177–201, 1987.

[13] E. G. Coffman, Jr., M. Hofri, and G. Weiss. Scheduling stochastic jobs with a two-point distribution on two parallel machines. *Probability in the Engineering and Informational Sciences*, 3:89–116, 1989.

[14] E. G. Coffman, Jr. and I. Mitrani. A characterization of waiting time performance realizable by single server queues. *Operations Research*, 28:810–821, 1980.

[15] D. R. Cox and W. L. Smith. *Queues*. Methuen, London, 1961.

[16] M. Dacre, K. Glazebrook, and J. Niño-Mora. The achievable region approach to the optimal control of stochastic systems. *Journal of the Royal Statistical Society Series B, Methodological*, 61(4):747–791, 1999.

[17] A. Federgruen and H. Groenevelt. Characterization and optimization of achievable performance in general queueing systems. *Operations Research*, 36:733–741, 1988.

[18] J. C. Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society Series B, Methodological*, 41:148–177, 1979.

[19] J. C. Gittins and D. M. Jones. A dynamic allocation index for the sequential design of experiments. In *Progress in statistics (European Meeting of Statisticians, Budapest, 1972)*, pages 241–266. Colloq. Math. Soc. János Bolyai, Vol. 9. North-Holland, Amsterdam, 1974.

[20] K. D. Glazebrook. Scheduling tasks with exponential service times on parallel processors. *Journal of Applied Probability*, 16:685–689, 1979.

[21] K. D. Glazebrook and J. Niño-Mora. A linear programming approach to stability, optimisation and performance analysis for Markovian multiclass queueing networks. *Annals of Operations Research*, 92:1–18, 1999. Models and Algorithms for Planning and Scheduling Problems: Cambridge Workshop (1997).

[22] K. D. Glazebrook and J. Niño-Mora. Parallel scheduling of multiclass $M/M/m$ queues: approximate and heavy-traffic optimization of achievable performance. *Operations Research*, 49(4):609–623, 2001.

[23] J. M. Harrison. Brownian models of queueing networks with heterogeneous customer populations. In *Stochastic Differential Systems, Stochastic Control Theory and Their Applications*, volume 10 of *IMA Volumes in Mathematics and Its Applications*, pages 147–186. Springer, New York, 1988.

[24] G. P. Klimov. Time-sharing service systems. I. *Theory of Probability and its Applications*, 19:532–551, 1974.

[25] H. Levy and M. Sidi. Polling systems: applications, modeling, and optimization. *IEEE Transactions in Communications*, 38:1750–1760, 1990.

[26] P. Nain, P. Tsoucas, and J. Walrand. Interchange arguments in stochastic scheduling. *Journal of Applied Probability*, 27:815–826, 1989.

[27] J. Niño-Mora. A new priority index for restless bandits and admission control: a polyhedral approach. *Mathematical Programming Series A*. Forthcoming.

[28] J. Niño-Mora. Restless bandits, partial conservation laws and indexability. *Advances in Applied Probability*, 33(1):77–98, 2001.

[29] C. H. Papadimitriou. Games against nature. *Journal of Computer and System Sciences*, 31(2):288–301, 1985. Special issue: Twenty-fourth annual symposium on the foundations of computer science (Tucson, Ariz., 1983).

[30] C. H. Papadimitriou and J. N. Tsitsiklis. On stochastic scheduling with in-tree precedence constraints. *SIAM Journal on Computing*, 16:1–6, 1987.

[31] M. I. Reiman and L. M. Wein. Dynamic scheduling of a two-class queue with setups. *Operations Research*, 4:532–547, 1998.

[32] R. Righter. Job scheduling to minimize expected weighted flowtime on uniform processors. *System Control Letters*, 10:211–216, 1988.

[33] M. H. Rothkopf. Scheduling with random service times. *Management Science*, 12:707–713, 1966.

[34] K. C. Sevcik. Scheduling for minimum total loss using service time distributions. *Journal of the ACM*, 21:66–75, 1974.

[35] J. G. Shanthikumar and D. D. Yao. Multiclass queueing systems: Polymatroidal structure and optimal scheduling control. *Operations Research*, 40:S293–299, 1992.

[36] W. E. Smith. Various optimizers for single-stage production. *Naval Research Logistics Quarterly*, 3:59–66, 1956.

[37] D. W. Tcha and S. R. Pliska. Optimal control of single-server queueing networks and multiclass $M/G/1$ queues with feedback. *Operations Research*, 25:248–258, 1977.

[38] J. N. Tsitsiklis. A short proof of the Gittins index theorem. *The Annals of Applied Probability*, 4:194–199, 1994.

[39] P. P. Varaiya, J. C. Walrand, and C. Buyukkoc. Extensions of the multiarmed bandit problem: the discounted case. *IEEE Transactions on Automatic Control*, 30:426–439, 1985.

[40] R. R. Weber. Scheduling jobs with stochastic processing requirements on parallel machines to minimize makespan or flowtime. *Journal of Applied Probability*, 19:167–182, 1982.

[41] R. R. Weber. On the Gittins index for multiarmed bandits. *The Annals of Applied Probability*, 2:1024–1033, 1992.

[42] R. R. Weber, P. Varaiya, and J. Walrand. Scheduling jobs with stochastically ordered processing times on parallel machines to minimize expected flowtime. *Journal of Applied Probability*, 23:841–847, 1986.

[43] R. R. Weber and G. Weiss. On an index policy for restless bandits. *Journal of Applied Probability*, 27:637–648, 1990.

[44] G. Weiss. Branching bandit processes. *Probability in the Engineering and Informational Sciences*, 2:269–278, 1988.

[45] G. Weiss. Turnpike optimality of Smith's rule in parallel machines stochastic scheduling. *Mathematics of Operations Research*, 17:255–270, 1992.

[46] P. Whittle. Multiarmed bandits and the Gittins index. *Journal of the Royal Statistical Society Series B, Methodological*, 42:143–149, 1980.

[47] P. Whittle. Restless bandits: Activity allocation in a changing world. In J. Gani, editor, *A Celebration of Applied Probability*, volume 25A of *Journal of Applied Probability*, pages 287–298. Applied Probability Trust, 1988.

[48] S. H. Wie and M. Pinedo. On minimizing the expected makespan and flow time in stochastic flow shops with blocking. *Mathematics of Operations Research*, 11:336–342, 1986.