Integer Programming Models for Multi-Processor Scheduling of Parallelizable Jobs with Resource-Dependent Durations

Yu. Zakharova, M. Sakhno

September 27, 2022

Omsk Department of Sobolev Institute of Mathematics SB RAS
The research is supported by RSF grant 22-71-10015.

Report structure

A mixed integer programming (MIP) approach for:

- single-processor jobs with regular criteria,
- parallelizable (multiprocessor and parallel) jobs,
- parallelizable jobs with resource-dependent durations.

We use continuous-time formulations with event points (positions).

Scheduling for single-processor jobs

$J = \{1, ..., n\}$ is the set of jobs.

$I = \{1, ..., m\}$ is the set of machines (processors).

p_j is the duration of job j.

Preemptions are disallowed.

Regular criteria: makespan, total completion times, maximum lateness, total tardiness, etc.

Problems are NP-hard.

Additional constraints: partial order/release dates/deadlines (due dates)/setup types.

MIP model for single-processor jobs[1]

K is the set of event points (positions), $|K| \leq n$.

Variables:

$$x_{jik} = \begin{cases} 1, & \text{if job } j \text{ is performed on processor } i \text{ in event point } k, \\ 0 & \text{otherwise.} \end{cases}$$

C_{ik} is the completion time of a job in event point k of processor i.

Constraints:

$$\sum_{j \in J} x_{jik} \leq 1, i \in I, k \in K, \quad (1)$$

$$\sum_{k \in K} \sum_{i \in I} x_{jik} = 1, j \in J, \quad (2)$$

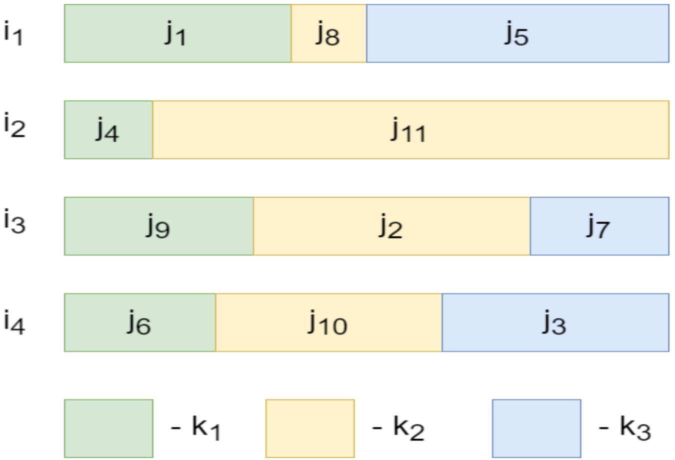
$$\sum_{j \in J} x_{j,i,k-1} \geq \sum_{j \in J} x_{j,i,k}, i \in I, k \in K, \quad (3)$$

$$C_{ik} \geq C_{i,k-1} + \sum_{j \in J} p_j x_{jik}, i \in I, k \in K. \quad (4)$$

[1] J. Blazewicz, M. Dror, J. Weglarz, Mathematical programming formulations for machine scheduling: A survey, 1991.

Scheduling for single-processor jobs

By event point[2] we mean a subset of variables in MIP model, which characterize a selection of a certain set of jobs and their starting and completion times.



[2]We use the positions based approach in contrast to the adjacency approach.

Some regular criteria

The makespan

$$C_{\max} \geq C_j, \; j \in J.$$

The total completion time

$$C_{\sum} = \sum_{j \in J} C_j.$$

The maximum lateness

$$L_{\max} \geq C_j - d_j, \; j \in J, .$$

The total tardiness

$$T_{\sum} = \sum_{j \in J} T_j,$$

$$T_j \geq 0, \; T_j \geq C_j - d_j, \; j \in J.$$

Here $C_j \geq C_{ik} - T_{max}(1 - x_{jik}), j \in J, i \in I, k \in K.$

Parallelizable jobs

Parallel jobs

- **Rigid jobs**: the number of required processors is given and fixed ($size_j$).
- **Moldable jobs**: the number of required processors is chosen by the scheduler before starting a job, and is not changed until the job termination (δ_j).
- **Malleable jobs**: the number of required processors is chosen by the scheduler, and can be changed at runtime (δ_j).

Multiprocessor jobs

- **Single mode jobs**: the set of required processors is given and fixed (fix_j).
- **Multimode jobs**: alternative sets of processors may be used (set_j).

MIP model for rigid jobs

$$w_{jk} = \begin{cases} 1, & \text{if job } j \text{ is executed in event point } k, \\ 0 & \text{otherwise.} \end{cases}$$

T_{jk}^{st} and T_{jk}^{f} is the start and completion time of job j in event point k.

$$\sum_{j \in J} x_{jik} \leq 1, i \in I, k \in K, \tag{5}$$

$$\sum_{k \in K} w_{jk} = 1, j \in J, \tag{6}$$

$$\sum_{i \in I} x_{jik} = size_j w_{jk}, k \in K, j \in J, \tag{7}$$

$$T_{jk}^{f} \geq T_{jk}^{st}, k \in K, j \in J, \tag{8}$$

$$T_{jk}^{f} - T_{jk}^{st} = w_{jk} p_j, k \in K, j \in J, \tag{9}$$

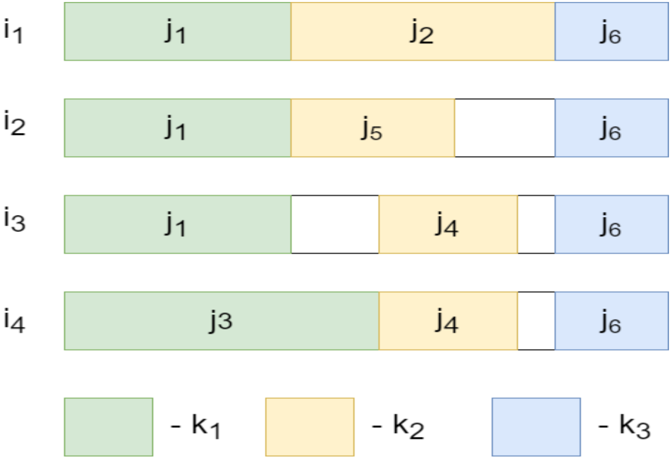
$$\sum_{k \in K} (T_{jk}^{f} - T_{jk}^{st}) \geq p_j, j \in J, \tag{10}$$

$$T_{jk}^{st} \geq T_{j'k'}^{f} - T_{max}(2 - x_{jik} - x_{j'ik'}), \tag{11}$$

$$j \neq j' \in J, i \in I, k' < k \in K, k \neq 1.$$

Parallelizable jobs

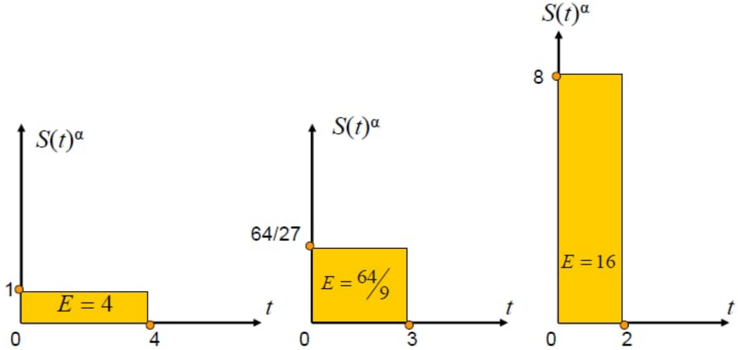
By event point we mean a subset of variables in MIP model, which characterize a selection of a certain set of jobs and their starting and completion times[3].



[3] C.A. Floudas, X. Lin, Mixed Integer Linear Programming in Process Scheduling: Modeling, Algorithms, and Applications, 2005

Resources

Energy $E = \int_{t_0}^{t_1} s^{\alpha}(t)\mathrm{d}t$, where $s(t)$ is the speed of a processor at time t and $\alpha > 1$ is the constant.



Convex model for the case of energy constraint

Energy consumption:

$$s_j \sum_{k \in K} (T_{jk}^{f} - T_{jk}^{st}) \geq W_j, j \in J. \tag{12}$$

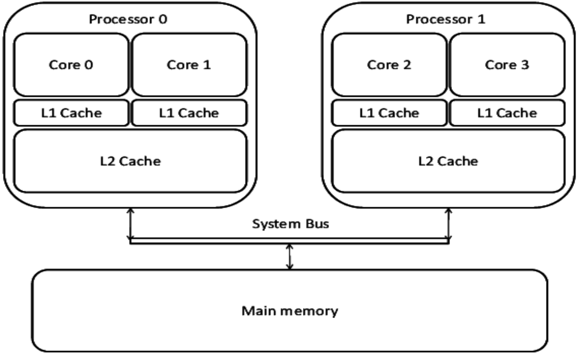
$$\sum_{j \in J} W_j (s_j)^{\alpha - 1} \leq E. \tag{13}$$

Variable $s_j \geq 0$ represent the speed of job $j \in J$.

Parameter W_j is the processing volume of job $j \in J$ (the execution time on one processor with unit speed).

Resources

A data bus is a part of the system bus that is used to transfer data between computer components.



Preprocessing

A configuration is the feasible subset of jobs that can be executed simultaneously, joined with their speeds and number of utilized processors. C is the set of all configurations.

The following parameters are calculated in the preprocessing:

W_j is the processing volume of job $j \in J$ (the execution time on one processor with unit speed).

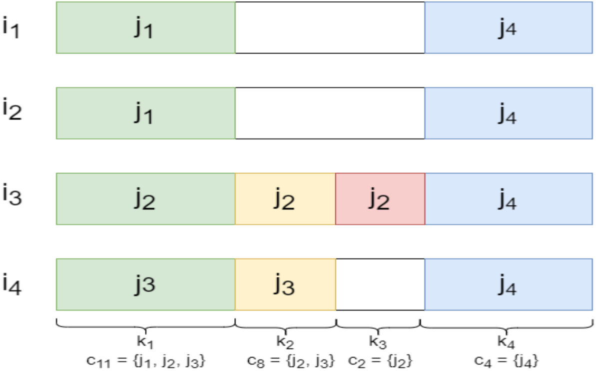
v_{jc} is the speed of job j in configuration c.

$$q_{jc} = \begin{cases} 1, & \text{if job } j \text{ is performed in configuration } c, \\ 0 & \text{otherwise,} \end{cases}$$

E_c is the instantaneous power for configuration c.

Event points and configurations

Exactly one configuration is executed in each event point.



MIP model for jobs with resource-dependent durations

Allocation of configurations in event point.

$$\sum_{c \in C} x_{kc} \leq 1, \; k \in K, \tag{14}$$

$$t_{kc} \geq 0, \; k \in K, \; c \in C, \tag{15}$$

$$t_{kc} \leq x_{kc} T_{max}, k \in K, \; c \in C, \tag{16}$$

MIP model for jobs with resource-dependent durations

Full execution and non-preemtion of jobs.

$$\sum_{k\in K}\sum_{c\in C} t_{kc} v_{jc} = W_j,\ j\in J, \tag{17}$$

$$\sum_{k\in K} y_{jk} = 1,\ j\in J, \tag{18}$$

$$\sum_{c\in C} x_{kc} q_{jc} - \sum_{c\in C} x_{k-1,c} q_{jc} \le y_{jk},\ j\in J,\ k\in K, \tag{19}$$

MIP model for jobs with resource-dependent durations

Completion times

$$T_k^f \geq T_{k-1}^f + t_{kc},\ k \in K,\ c \in C, \tag{20}$$

$$C_j^f \geq T_k^f - T_{max}(1 - x_{kc}),\ j \in J,\ k \in K,\ c \in C, q_{jc} > 0, \tag{21}$$

Additional constraints

$$\sum_{k \in K} \sum_{c \in C} E_c t_{kc} \leq E, \tag{22}$$

$$\sum_{k \in K} (k+1) x_{kc_1} \leq (1 - \sum_{k \in K} x_{kc_1}) K_{max} + \sum_{k \in K} k x_{kc_2}, \tag{23}$$

$$c_1, c_2 \in C, a_{c_1,c_2} > 0.$$

Resource is energy

We discretize the possible speed values and consider only a finite number of speeds at which the processors can run.

Theorem.[4] Given $\varepsilon > 0$. A schedule of energy consumption at most $E+\varepsilon$ and objective at most OPT can be found using the presented model. The number of speed values is polynomial in $\frac{1}{\varepsilon}$ and exponential in the size of the input.

[4] A. Kononov, Yu. Zakharova Speed scaling scheduling of multiprocessor jobs with energy constraint and makespan criterion, Journal of Global Optimization, 2021.

Experimental evaluations

The CPLEX package was used to solve the MIP model for jobs with resource-dependent durations.

For each number of jobs 192 test instances with different number of processors (2, 3, 4) and partial order types (bitree, random, one-to-many-to-one, trivial) were generated.

N_{opt} is the percentage of found optimal solutions.

R_{avg} is the average relative gap.

T_{avg} is the average time of CPLEX package for finding optimal solution. The CPLEX package was limited by 20 minutes.

	N_{opt}	R_{avg}	T_{avg}	**max |C|**
4 jobs	100%	0%	482 ms	16
6 jobs	99%	0.02%	39 s	57
7 jobs	88%	0.1%	104 s	99
8 jobs	61%	0.4%	152 s	163
10 jobs	28%	1.7%	134 s	386

Conclusion and further research

- We provide analysis of the event point approach for different types of jobs and resources, influencing on job durations.
- The proposed model is tested on real data.
- Experimental comparison of convex and linear model for energy constraint problems.
- Computing lower and upper bounds for metaheuristics and heuristics.

Thank you for your attention!