# Scheduling on Unrelated Parallel Machines

Piotr Wojciechowski1

<sup>1</sup>Lane Department of Computer Science and Electrical Engineering West Virginia University

April 17, 2014

# Outline



Scheduling on Unrelated Parallel Machines

- Definition
- IP formuation
- Algorithm

# Outline



# Scheduling on Unrelated Parallel Machines

- Definition
- IP formuation
- Algorithm

UPM

Definition IP Algorithm

# Definition

We are given a set of *n* jobs  $J = \{j_1, \ldots, j_n\},\$ 

UPM

Definition IP Algorithm

# Definition

We are given a set of *n* jobs  $J = \{j_1, \ldots, j_n\}$ , a set of *m* machines  $M\{m_1, \ldots, m_n\}$ ,



We are given a set of *n* jobs  $J = \{j_1, \ldots, j_n\}$ , a set of *m* machines  $M\{m_1, \ldots, m_n\}$ , and a time matrix **P** where  $p_{ij}$  is the time it takes the *i*th machine to run the *j*th job.

We are given a set of *n* jobs  $J = \{j_1, ..., j_n\}$ , a set of *m* machines  $M\{m_1, ..., m_n\}$ , and a time matrix **P** where  $p_{ij}$  is the time it takes the *i*th machine to run the *j*th job. Our goal is to schedule the jobs to that the makespan is minimized.

We are given a set of *n* jobs  $J = \{j_1, ..., j_n\}$ , a set of *m* machines  $M\{m_1, ..., m_n\}$ , and a time matrix **P** where  $p_{ij}$  is the time it takes the *i*th machine to run the *j*th job. Our goal is to schedule the jobs to that the makespan is minimized.

### Note

This is a generalization of minimum makespan scheduling in which the running time for each job depends on the machine to which it is assigned.

We are given a set of *n* jobs  $J = \{j_1, ..., j_n\}$ , a set of *m* machines  $M\{m_1, ..., m_n\}$ , and a time matrix **P** where  $p_{ij}$  is the time it takes the *i*th machine to run the *j*th job. Our goal is to schedule the jobs to that the makespan is minimized.

### Note

This is a generalization of minimum makespan scheduling in which the running time for each job depends on the machine to which it is assigned. A opposed to the PTAS obtained for the original problem we will demonstrate a 2-approximation.

# Outline



# Scheduling on Unrelated Parallel Machines

- Definition
- IP formuation
- Algorithm

The integer program for scheduling on unrelated parallel machines need to encode the following properties:

The integer program for scheduling on unrelated parallel machines need to encode the following properties:

• Each job can be scheduled on at most one machine.

The integer program for scheduling on unrelated parallel machines need to encode the following properties:

- Each job can be scheduled on at most one machine.
- If job  $j_j$  is scheduled on the  $m_i$  machine the running time is  $p_{ij}$ .

The integer program for scheduling on unrelated parallel machines need to encode the following properties:

- Each job can be scheduled on at most one machine.
- If job  $j_i$  is scheduled on the  $m_i$  machine the running time is  $p_{ij}$ .
- **(2)** We want to minimize the longest total running time over all machines.

# IP formulation

# IP formulation

This problem can be represented by the following integer program where the variable  $x_{ij}$  is 1 if job  $j_i$  is assigned to machine  $m_i$ :

min t

UPM

#### IP Algorithm

## IP formulation

$$\min t$$

$$\sum_{i \in M} x_{ij} = 1 \qquad j \in J$$

Definition IP Algorithm

### IP formulation

$$\min_{i \in M} t$$

$$\sum_{i \in M} x_{ij} = 1 \qquad j \in J$$

$$\sum_{i \in J} x_{ij} \cdot p_{ij} \leq t \qquad i \in M$$

Definition IP Algorithm

### IP formulation

$$\begin{array}{ll} \min t \\ \sum\limits_{i \in M} x_{ij} = 1 & j \in J \\ \sum\limits_{i \in J} x_{ij} \cdot p_{ij} \leq t & i \in M \\ x_{ij} \in \{0,1\} & i \in M, j \in J \end{array}$$

Definition IP Algorithm

### IP formulation

This problem can be represented by the following integer program where the variable  $x_{ij}$  is 1 if job  $j_i$  is assigned to machine  $m_i$ :

$$\min t$$

$$\sum_{i \in M} x_{ij} = 1 \qquad j \in J$$

$$\sum_{i \in J} x_{ij} \cdot p_{ij} \leq t \qquad i \in M$$

$$x_{ij} \in \{0, 1\} \qquad i \in M, j \in J$$

# Integrality Gap

The main issue with this formulation is that it has an unbounded integrality gap.

### IP formulation

This problem can be represented by the following integer program where the variable  $x_{ij}$  is 1 if job  $j_i$  is assigned to machine  $m_i$ :

$$\min t$$

$$\sum_{i \in M} x_{ij} = 1 \qquad j \in J$$

$$\sum_{i \in J} x_{ij} \cdot p_{ij} \leq t \qquad i \in M$$

$$x_{ij} \in \{0, 1\} \qquad i \in M, j \in J$$

# Integrality Gap

The main issue with this formulation is that it has an unbounded integrality gap. For example consider trying to schedule one job on *m* machines where  $p_{i1} = m$  for all  $i \in M$ .

We can avoid this situation by making the LP unable to schedule jobs longer than the total makespan.

We can avoid this situation by making the LP unable to schedule jobs longer than the total makespan. This cannot be done in a single LP but can be expressed as a family of Linear Programs.

We can avoid this situation by making the LP unable to schedule jobs longer than the total makespan. This cannot be done in a single LP but can be expressed as a family of Linear Programs.

# Pruned LPs

The new linear programs take the following form:

We can avoid this situation by making the LP unable to schedule jobs longer than the total makespan. This cannot be done in a single LP but can be expressed as a family of Linear Programs.

### Pruned LPs

The new linear programs take the following form:

• For  $T \in \mathbb{Z}^+$  create the set of machine-job pairs  $S_T = \{(i, j) : x_{ij} \leq T\}$ .

We can avoid this situation by making the LP unable to schedule jobs longer than the total makespan. This cannot be done in a single LP but can be expressed as a family of Linear Programs.

### Pruned LPs

The new linear programs take the following form:

- For  $T \in \mathbb{Z}^+$  create the set of machine-job pairs  $S_T = \{(i, j) : x_{ij} \leq T\}$ .
- 2 Create the linear program LP(T) defined as

$$\sum_{i:(i,j)\in S_T} x_{ij} = 1 \qquad j \in J$$

We can avoid this situation by making the LP unable to schedule jobs longer than the total makespan. This cannot be done in a single LP but can be expressed as a family of Linear Programs.

### Pruned LPs

The new linear programs take the following form:

• For  $T \in \mathbb{Z}^+$  create the set of machine-job pairs  $S_T = \{(i, j) : x_{ij} \leq T\}$ .

2 Create the linear program LP(T) defined as

$$\sum_{\substack{i:(i,j)\in S_T}} x_{ij} = 1 \qquad j \in J$$
$$\sum_{i:(i,j)\in S_T} x_{ij} \cdot \rho_{ij} \le t \qquad i \in M$$

We can avoid this situation by making the LP unable to schedule jobs longer than the total makespan. This cannot be done in a single LP but can be expressed as a family of Linear Programs.

### Pruned LPs

The new linear programs take the following form:

• For  $T \in \mathbb{Z}^+$  create the set of machine-job pairs  $S_T = \{(i, j) : x_{ij} \leq T\}$ .

2 Create the linear program LP(T) defined as

$$\sum_{i:(i,j)\in S_T} x_{ij} = 1 \qquad j \in J$$
$$\sum_{i:(i,j)\in S_T} x_{ij} \cdot p_{ij} \leq t \qquad i \in M$$
$$x_{ij} \geq 0 \qquad (i,j) \in S_T$$

We can avoid this situation by making the LP unable to schedule jobs longer than the total makespan. This cannot be done in a single LP but can be expressed as a family of Linear Programs.

#### Pruned LPs

The new linear programs take the following form:

• For  $T \in \mathbb{Z}^+$  create the set of machine-job pairs  $S_T = \{(i, j) : x_{ij} \leq T\}$ .

2 Create the linear program LP(T) defined as

$$\sum_{i:(i,j)\in S_T} x_{ij} = 1 \qquad j \in J$$
$$\sum_{i:(i,j)\in S_T} x_{ij} \cdot p_{ij} \leq t \qquad i \in M$$
$$x_{ij} \geq 0 \qquad (i,j) \in S_T$$

Our goal is to find the smallest T such that LP(T) is feasible, call this value  $T^*$  (note that  $T^*$  is still a lower bound for *OPT*).



We can avoid this situation by making the LP unable to schedule jobs longer than the total makespan. This cannot be done in a single LP but can be expressed as a family of Linear Programs.

### Pruned LPs

The new linear programs take the following form:

• For  $T \in \mathbb{Z}^+$  create the set of machine-job pairs  $S_T = \{(i, j) : x_{ij} \leq T\}$ .

2 Create the linear program LP(T) defined as

$$\sum_{i:(i,j)\in S_T} x_{ij} = 1 \qquad j \in J$$
$$\sum_{j:(i,j)\in S_T} x_{ij} \cdot p_{ij} \leq t \qquad i \in M$$
$$x_{ij} \geq 0 \qquad (i,j) \in S_T$$

Our goal is to find the smallest T such that LP(T) is feasible, call this value  $T^*$  (note that  $T^*$  is still a lower bound for *OPT*). We will then take an extreme point solution to  $LP(T^*)$  and round it to obtain an integer solution with makespan no more than  $2 \cdot T^*$ .

### IP Algorithm

# Properties of Extreme Point Solutions

We first need to show that an extreme point solution of LP(T) satisfies certain properties.

We first need to show that an extreme point solution of LP(T) satisfies certain properties.

### Theorem

Any extreme point solution to LP(T) has at most n + m nonzero variables.

We first need to show that an extreme point solution of LP(T) satisfies certain properties.

#### Theorem

Any extreme point solution to LP(T) has at most n + m nonzero variables.

### Proof.

• Let  $r = |S_t|$  be the number of variables on which LP(T) is defined.

We first need to show that an extreme point solution of LP(T) satisfies certain properties.

#### Theorem

Any extreme point solution to LP(T) has at most n + m nonzero variables.

#### Proof.

- Let  $r = |S_t|$  be the number of variables on which LP(T) is defined.
- An extreme point solution must satisfy at least r constraints with equality.

We first need to show that an extreme point solution of LP(T) satisfies certain properties.

#### Theorem

Any extreme point solution to LP(T) has at most n + m nonzero variables.

#### Proof.

- Let  $r = |S_t|$  be the number of variables on which LP(T) is defined.
- An extreme point solution must satisfy at least r constraints with equality.
- At least r (n + m) of these must be non-negativity constraints.

We first need to show that an extreme point solution of LP(T) satisfies certain properties.

#### Theorem

Any extreme point solution to LP(T) has at most n + m nonzero variables.

#### Proof.

- Let  $r = |S_t|$  be the number of variables on which LP(T) is defined.
- An extreme point solution must satisfy at least r constraints with equality.
- At least r (n + m) of these must be non-negativity constraints.
- At least r (n + m) variables are set to 0.
# Properties of Extreme Point Solutions

We first need to show that an extreme point solution of LP(T) satisfies certain properties.

#### Theorem

Any extreme point solution to LP(T) has at most n + m nonzero variables.

- Let  $r = |S_t|$  be the number of variables on which LP(T) is defined.
- An extreme point solution must satisfy at least r constraints with equality.
- At least r (n + m) of these must be non-negativity constraints.
- At least r (n + m) variables are set to 0.
- At most n + m variables are nonzero.

Any extreme point solution to LP(T) must set at least n - m jobs integrally.

Any extreme point solution to LP(T) must set at least n - m jobs integrally.

# Proof.

• Let **x** be an extreme point solution to LP(T).

Any extreme point solution to LP(T) must set at least n - m jobs integrally.

- Let **x** be an extreme point solution to LP(T).
- Let  $\alpha$  and  $\beta$  be the number of jobs set integrally and fractionally respectively.

Any extreme point solution to LP(T) must set at least n - m jobs integrally.

- Let **x** be an extreme point solution to LP(T).
- Let  $\alpha$  and  $\beta$  be the number of jobs set integrally and fractionally respectively.
- We have that  $\alpha + \beta = n$ .

Any extreme point solution to LP(T) must set at least n - m jobs integrally.

- Let **x** be an extreme point solution to LP(T).
- Let  $\alpha$  and  $\beta$  be the number of jobs set integrally and fractionally respectively.
- We have that  $\alpha + \beta = n$ .
- Assigning a job fractionally results in at least two nonzero variables so  $\alpha + 2 \cdot \beta \le n + m$ .

Any extreme point solution to LP(T) must set at least n - m jobs integrally.

- Let **x** be an extreme point solution to LP(T).
- Let  $\alpha$  and  $\beta$  be the number of jobs set integrally and fractionally respectively.
- We have that  $\alpha + \beta = n$ .
- Assigning a job fractionally results in at least two nonzero variables so  $\alpha + 2 \cdot \beta \le n + m$ .
- Thus,  $\beta \leq m$  and  $\alpha \geq n-m$ .

We now need to specify how fractionally assigned jobs are handled.

• Construct the bipartite graph G = (M, J, E) where  $(m_i, j_j) \in E$  if  $x_{ij} > 0$ .

- Construct the bipartite graph G = (M, J, E) where  $(m_i, j_j) \in E$  if  $x_{ij} > 0$ .
- Let  $F \subseteq J$  be the set of fractionally assigned jobs.

- Construct the bipartite graph G = (M, J, E) where  $(m_i, j_j) \in E$  if  $x_{ij} > 0$ .
- Let  $F \subseteq J$  be the set of fractionally assigned jobs.
- Define *H* to be the subgraph of *G* induced by the vertex set  $M \cup F$ .

- Construct the bipartite graph G = (M, J, E) where  $(m_i, j_j) \in E$  if  $x_{ij} > 0$ .
- Let  $F \subseteq J$  be the set of fractionally assigned jobs.
- Define *H* to be the subgraph of *G* induced by the vertex set  $M \cup F$ .
- Our goal is to find a matching in *H* so that every job in *F* is assigned to a unique machine in *M*.

IP Algorithm

### Handling Fractionally Assigned Jobs

- Construct the bipartite graph G = (M, J, E) where  $(m_i, j_j) \in E$  if  $x_{ij} > 0$ .
- Let  $F \subseteq J$  be the set of fractionally assigned jobs.
- Define *H* to be the subgraph of *G* induced by the vertex set  $M \cup F$ .
- Our goal is to find a matching in *H* so that every job in *F* is assigned to a unique machine in *M*.







Definition IP Algorithm

# Theorem

G is a pseudo-forest.

G is a pseudo-forest.

# Proof.

• We need to show that each connected component of *G* is a pseudo-tree.



G is a pseudo-forest.

- We need to show that each connected component of *G* is a pseudo-tree.
- Let  $G_c$  be a component of G.

G is a pseudo-forest.

- We need to show that each connected component of *G* is a pseudo-tree.
- Let  $G_c$  be a component of G.
- We can restrict **x** and LP(T) to the the machines and jobs in  $G_c$ .

G is a pseudo-forest.

- We need to show that each connected component of *G* is a pseudo-tree.
- Let G<sub>c</sub> be a component of G.
- We can restrict **x** and LP(T) to the the machines and jobs in  $G_c$ .
- Call the results  $\mathbf{x}_{\mathbf{c}}$  and  $LP_{c}(T)$ .

G is a pseudo-forest.

- We need to show that each connected component of G is a pseudo-tree.
- Let G<sub>c</sub> be a component of G.
- We can restrict **x** and LP(T) to the the machines and jobs in  $G_c$ .
- Call the results  $\mathbf{x}_{\mathbf{c}}$  and  $LP_{c}(T)$ .
- Since x is an extreme point solution of LP(T) then we must have that x<sub>c</sub> is an extreme point solution to LP<sub>c</sub>(T).

G is a pseudo-forest.

- We need to show that each connected component of G is a pseudo-tree.
- Let G<sub>c</sub> be a component of G.
- We can restrict **x** and LP(T) to the the machines and jobs in  $G_c$ .
- Call the results  $\mathbf{x}_{\mathbf{c}}$  and  $LP_{c}(T)$ .
- Since **x** is an extreme point solution of LP(T) then we must have that  $\mathbf{x}_{c}$  is an extreme point solution to  $LP_{c}(T)$ .
- We have that the number of non-zero variables in x<sub>c</sub> cannot exceed the number of machines plus the number of jobs in LP<sub>c</sub>(T).

G is a pseudo-forest.

- We need to show that each connected component of G is a pseudo-tree.
- Let G<sub>c</sub> be a component of G.
- We can restrict **x** and LP(T) to the the machines and jobs in  $G_c$ .
- Call the results  $\mathbf{x}_{\mathbf{c}}$  and  $LP_{c}(T)$ .
- Since **x** is an extreme point solution of LP(T) then we must have that  $\mathbf{x}_{c}$  is an extreme point solution to  $LP_{c}(T)$ .
- We have that the number of non-zero variables in x<sub>c</sub> cannot exceed the number of machines plus the number of jobs in LP<sub>c</sub>(T).
- Thus the number of edges in G<sub>c</sub> cannot exceed the number of verticies.



# Theorem

H has a matching in which each job in F is assigned to a machine in M.

# Theorem

H has a matching in which each job in F is assigned to a machine in M.

# Proof.

• In G, every integrally set job has exactly one edge incident to it.

## Theorem

H has a matching in which each job in F is assigned to a machine in M.

- In G, every integrally set job has exactly one edge incident to it.
- When we remove these jobs and edges the remaining graph is precisely H.

H has a matching in which each job in F is assigned to a machine in M.

- In G, every integrally set job has exactly one edge incident to it.
- When we remove these jobs and edges the remaining graph is precisely H.
- Thus, H is also a pseudo-forest.

## Theorem

H has a matching in which each job in F is assigned to a machine in M.

- In G, every integrally set job has exactly one edge incident to it.
- When we remove these jobs and edges the remaining graph is precisely H.
- Thus, H is also a pseudo-forest.
- In *H* each job vertex has degree at least 2 and so the leaves are all machine verticies.

## Theorem

H has a matching in which each job in F is assigned to a machine in M.

- In G, every integrally set job has exactly one edge incident to it.
- When we remove these jobs and edges the remaining graph is precisely H.
- Thus, H is also a pseudo-forest.
- In *H* each job vertex has degree at least 2 and so the leaves are all machine verticies.
- If a machine node *i* is a leaf we can assign to it its parent job node and remove both from the graph.

## Theorem

H has a matching in which each job in F is assigned to a machine in M.

- In G, every integrally set job has exactly one edge incident to it.
- When we remove these jobs and edges the remaining graph is precisely H.
- Thus, H is also a pseudo-forest.
- In *H* each job vertex has degree at least 2 and so the leaves are all machine verticies.
- If a machine node *i* is a leaf we can assign to it its parent job node and remove both from the graph.
- This process continues until the only edges in the graph form even length cycles.

# Theorem

H has a matching in which each job in F is assigned to a machine in M.

- In G, every integrally set job has exactly one edge incident to it.
- When we remove these jobs and edges the remaining graph is precisely H.
- Thus, H is also a pseudo-forest.
- In H each job vertex has degree at least 2 and so the leaves are all machine verticies.
- If a machine node *i* is a leaf we can assign to it its parent job node and remove both from the graph.
- This process continues until the only edges in the graph form even length cycles.
- Alternating edges in these cycles form a matching in which each remaining job vertex is assigned to a unique remaining machine vertex.

# Theorem

H has a matching in which each job in F is assigned to a machine in M.

- In G, every integrally set job has exactly one edge incident to it.
- When we remove these jobs and edges the remaining graph is precisely H.
- Thus, H is also a pseudo-forest.
- In H each job vertex has degree at least 2 and so the leaves are all machine verticies.
- If a machine node *i* is a leaf we can assign to it its parent job node and remove both from the graph.
- This process continues until the only edges in the graph form even length cycles.
- Alternating edges in these cycles form a matching in which each remaining job vertex is assigned to a unique remaining machine vertex.
- Thus at the end of this procedure each job in *F* has been uniquely assigned to one of the machines in *M*.













# Outline



# Scheduling on Unrelated Parallel Machines

- Definition
- IP formuation
- Algorithm
The approximation algorithm proceeds as follows:

• Let  $\alpha$  be the makespan of the schedule formed by assigning each job to the machine on which it has the shortest running time.

The approximation algorithm proceeds as follows:

- Let  $\alpha$  be the makespan of the schedule formed by assigning each job to the machine on which it has the shortest running time.
- **2** Using a binary search over the interval  $[\frac{\alpha}{m}, \alpha]$  find the smallest  $T \in \mathbb{Z}^+$  such that LP(T) is linearly feasible. Call this value  $T^*$ .

The approximation algorithm proceeds as follows:

- Let  $\alpha$  be the makespan of the schedule formed by assigning each job to the machine on which it has the shortest running time.
- **2** Using a binary search over the interval  $[\frac{\alpha}{m}, \alpha]$  find the smallest  $T \in \mathbb{Z}^+$  such that LP(T) is linearly feasible. Call this value  $T^*$ .

**③** Find **x**, an extreme point solution to  $LP(T^*)$ .

The approximation algorithm proceeds as follows:

- Let  $\alpha$  be the makespan of the schedule formed by assigning each job to the machine on which it has the shortest running time.
- **2** Using a binary search over the interval  $[\frac{\alpha}{m}, \alpha]$  find the smallest  $T \in \mathbb{Z}^+$  such that LP(T) is linearly feasible. Call this value  $T^*$ .
- **③** Find **x**, an extreme point solution to  $LP(T^*)$ .
- Assign each integrally set job to the machine it is assigned in x.

The approximation algorithm proceeds as follows:

- Let *α* be the makespan of the schedule formed by assigning each job to the machine on which it has the shortest running time.
- **2** Using a binary search over the interval  $[\frac{\alpha}{m}, \alpha]$  find the smallest  $T \in \mathbb{Z}^+$  such that LP(T) is linearly feasible. Call this value  $T^*$ .
- Sind **x**, an extreme point solution to  $LP(T^*)$ .
- Assign each integrally set job to the machine it is assigned in **x**.
- Solution Assign each fractionally set job to a unique machine according to the matching found in *H*.

The algorithm has an approximation factor of 2.

# Theorem

The algorithm has an approximation factor of 2.

## Proof.

• We have that  $T^* \leq OPT$ .



# Theorem

The algorithm has an approximation factor of 2.

- We have that  $T^* \leq OPT$ .
- The makespan of the schedule corresponding to  $\mathbf{x}$  is no more than  $T^*$ .

## Theorem

The algorithm has an approximation factor of 2.

- We have that  $T^* \leq OPT$ .
- The makespan of the schedule corresponding to  $\mathbf{x}$  is no more than  $T^*$ .
- Thus when restricted to integrally assigned jobs the makespan of the schedule is still at most T\*.

### Theorem

The algorithm has an approximation factor of 2.

- We have that  $T^* \leq OPT$ .
- The makespan of the schedule corresponding to  $\mathbf{x}$  is no more than  $T^*$ .
- Thus when restricted to integrally assigned jobs the makespan of the schedule is still at most T\*.
- According to the matching in *H* each machine is assigned at most one originally fractionally assigned job.

### Theorem

The algorithm has an approximation factor of 2.

- We have that  $T^* \leq OPT$ .
- The makespan of the schedule corresponding to  $\mathbf{x}$  is no more than  $T^*$ .
- Thus when restricted to integrally assigned jobs the makespan of the schedule is still at most T\*.
- According to the matching in *H* each machine is assigned at most one originally fractionally assigned job.
- This increases the makespan of the schedule by at most T\*.

## Theorem

The algorithm has an approximation factor of 2.

- We have that  $T^* \leq OPT$ .
- The makespan of the schedule corresponding to  $\mathbf{x}$  is no more than  $T^*$ .
- Thus when restricted to integrally assigned jobs the makespan of the schedule is still at most T\*.
- According to the matching in *H* each machine is assigned at most one originally fractionally assigned job.
- This increases the makespan of the schedule by at most T\*.
- Thus the resultant schedule has total makespan at most  $2 \cdot T^* \leq 2 \cdot OPT$ .