The Road Network Maintenance Problem

Cyrille Dejemeppe & Jean-Baptiste Mairy, N-SIDE 2019 ACP Summer School

Context

The road network of a country allows vehicles to flow from their starting location to their destination. Locations are connected together with a set of roads. The set of locations in a country and the roads between them is referred to as the *road network*.

In order to keep the road network functional, maintenance work has to be performed on road portions. The day a maintenance work is being performed on a given road portion, no vehicle can flow on it. Maintenance works are performed by human workers coming from various work centers. Each maintenance work requires a given amount of human workers from a specified work center. To ensure vehicles are still able to reach their destination, the maintenance works must be carefully planned. To each road section, is associated a traffic perturbation of performing one/several maintenance works on it at a given time. This perturbation is the same whether one or several maintenance works are performed at a given time on the same road section.

A worksheet represents an ordered sequence of maintenance works to be performed. To plan maintenance works, one has to select which worksheet to execute. Either a worksheet is executed completely, meaning all the maintenance works it contains are performed or it is not executed and none of its maintenance work is performed. To each worksheet is associated an *importance* measure that represents the state of decay of the roads contained in the worksheet.

Given a set of worksheets, your goal is to decide which ones will be performed and schedule all the maintenance works they contain throughout the year. Your planning should ensure as most worksheet with high importance measures are performed while minimizing the maximal traffic perturbation associated to those.

The Data

This section identifies the quantities that are used to define an instance of the problem.

Each maintenance work is referred to as an *activity*. To simplify the problem, we consider that each activity is executed in exactly 1 day. Let us define Ω as the set of all the activities. A worksheet $w_k \subseteq \Omega$ is an *ordered sequence* of activities $A_i \in w_k$. Worksheets are disjoint subset of activities: $\forall\, k, l : w_k \cap w_l = \varnothing$. Here are the constants associated to a worksheet w_k:

- u_k, the importance associated to the execution of the worksheet
- b_k, constant telling wether or not the worksheet is forced to be actually executed; 1 if the worksheet must be executed, 0 if it is not forced to be executed
- c_k, the work center from which the worksheet will require its workers

- d_k, the duration of the worksheet, this number can also be deduced from the number of activities the worksheet contains: $d_k = |w_k|$
- est_k, the earliest starting time of the worksheet, meaning that none of its activities can start before this value
- lst_k, the latest starting time of the worksheet, meaning that none of its activities can start after this value

Each worksheet contains one/several activities. Here are the constants associated to an activity A_i :

- h_i, the amount of workers needed to execute the activity
- r_i, the road on which the activity takes place

The set of all work centers is noted C. The total number of workers available in a work center $c_m \in C$ is noted H_m. The total number of days on which the activities have to be scheduled is noted $horizon$. The set of all the road portions is noted R. The traffic perturbation cost associated to performing activities on a portion of road $q \in R$ at time t is noted $p_{q,t}$.

The model

To model this problem, we introduce variables: Here are the variables associated to a worksheet w_k:

- g_k, the actual execution of the worksheet; 1 if the worksheet is executed, 0 otherwise

Here are the variables associated to an activity A_i:

- s_i, the starting time of the activity
- e_i, the ending time of the activity. Note that as activities have a single day duration, we have $e_i = s_i + 1$
- f_i, the actual execution of the activity; 1 if the activity is executed, 0 otherwise

Constraints

Precedences within Worksheet

The activities from a worksheet $w_k \in \Omega$ are all executed in sequence right after each other. Considering the ordered sequence of activities $A_0, A_1, \ldots, A_n$ inside a worksheet w_k, the following constraint holds:

$$\forall\, i \in [0, n-1] : e_i = s_{i+1}$$

Worksheet Earliest Starting Time

Given the earliest starting time est_k of a worksheet $w_k \subseteq \Omega$, the following constraint holds:

$$\forall\, A_i \in w_k : est_k \leq s_i$$

Worksheet Latest Starting Time

Given the latest starting time lst_k of a worksheet $w_k \subseteq \Omega$, the following constraint holds:

$$\forall\, A_i \in w_k : s_i \leq lst_k$$

Complete Worksheets

The activities from a worksheet $w_k \in \Omega$ are either all executed or none is executed:

$$\forall\, w_k \in \Omega,\, \forall\, A_i \in w_k : f_i = b_k$$

Mandatory Worksheets

Given a set of worksheets $W \subseteq \Omega$, all forced worksheets have to be executed:

$$\forall\, w_k \in W : g_k \geq b_k$$

Precedence Between Worksheets

Given two worksheets $w_k, w_l \in \Omega$, if both worksheets are executed and w_k must precede w_l, noted as $w_k \prec w_l$, the following constraint holds:

$$\forall\, A_i \in w_k,\ \forall A_j \in w_l : (b_k \cdot b_l > 0 \wedge w_k \prec w_l) \Rightarrow e_i \leq s_j$$

Maximal Number of Roads Simultaneously Blocked

Given a subset of roads $\Delta \subseteq R$ and a specified value δ such that $0 < \delta < |\Delta|$, at most m roads can be blocked simultaneously. Considering $\Phi_k = \{A_i \in \Omega \mid r_i = r_k\}$ the set of all activities using a given road r_k, the following constraint holds:

$$\forall\, t \in [0,\, horizon] : \sum_{r_k \in \Delta} \sum_{\substack{A_i \in \Phi_k \\ s_i \leq t < e_i}} (f_i) < \delta$$

Work Center Capacity

The amount of workers from a given work center who are assigned to activities must not exceed the total amount of workers coming from this work center at any point of time. Considering $\Psi_m = \{A_i \in \Omega \mid c_i = c_m\}$ the set of all activities using resources from a given work center c_m, the following constraint holds:

$$\forall\, t \in [0,\, horizon]\,, \forall\, c_m \in C : \sum_{\substack{A_i \in \Psi_m \\ s_i \leq t < e_i}} (f_i \cdot h_i) < H_m$$

Objectives

There are two objectives for the scheduling of the activities: maximizing the total gain and minimizing the maximal traffic perturbation cost reached on a single day.

To simplify the resolution of this problem, we aggregate these two as a single objective as follows:

$$maximize \sum_{w_k \subseteq \Omega} u_k - \max_{t \in [0, horizon]} \left(\sum_{\substack{p \in R \\ \exists f_i = 1 \wedge r_i = q}} p_{q,t} \right)$$

Instance Files

Here is an example of a small instance file:

smallInstance.txt

```
45 5 2 3 7
0 0:5:2 5:8:1 8:13:1 13:20:4 20:22:5 22:37:2 37:43:4 43:45:3
1 0:2:1 2:4:5 4:14:1 14:23:2 23:31:4 31:44:5 44:45:4
2 0:19:4 19:20:3 20:25:3 25:30:4 30:33:5 33:41:5 41:45:4
3 0:0:2 0:3:4 3:6:1 6:14:1 14:21:3 21:27:5 27:33:4 33:39:2 39:45:5
4 0:2:2 2:12:5 12:16:2 16:19:4 19:21:4 21:25:4 25:26:3 26:30:3 30:40:3 40:45:4
0 182
1 123
0 0 0 57 23 43 2 0 0 16 15
1 0 0 72 0 42 3 3 4 0 7 11 12
2 1 0 90 14 43 2 0 0 18 10
M 1 0 4
M 2 1 3 2
P 1 0
```

Each line of the file has a role to determine some characteristics and quantities of the instance. Those lines are explained below.

Instance sizing

The first line of the file defines the size of the instance. It contains 5 numbers that represent — in order — the following quantities:

1. Number of days (**45** in the example)
2. Number of roads (**5** in the example)
3. Number of work centers (**2** in the example)
4. Number of worksheets (**3** in the example)
5. Number of activities (**7** in the example)

Roads

Each of the following *number of roads* lines defines the characteristics of a road. The first number in the line represents the **id** of the road. The remaining of the line contains triplets separated by a white space. Each triplet represents the cost associated to a perturbation on the road during a given day. Each triplet has the following form **start:end:cost**. It means that for each day between **start** (included) and **end** (not included), using the road (i.e., there exists a task using it at this day) will represent a perturbation cost of **cost**.

Work centers

Each of the following *number of work centers* lines defines the characteristics of a work center. Each line contains two number that represent — in order — the following quantities:

1. Road ID
2. Number of available workers

Worksheets

Each of the following *number of worksheets* lines defines the characteristics of a work center. Each line contains a sequence of numbers that represent — in order — the following quantities:

1. ID — The id of the worksheet
2. Work center ID — The id of the work center used by activities in the worksheet
3. Mandatory — **1** if the worksheet is forced to be executed, **0** if it is optional
4. Importance — The importance value associated to the execution of the worksheet
5. EST — The earliest starting time at which the worksheet can begin
6. LST — The latest starting time at which the worksheet can begin
7. Duration — The total number of successive tasks (hence of days) the worksheet is long
8. Road IDs — The next **duration** numbers each represent, in order, the id of the road that is used each day
9. Amounts of Workers — The next **duration** numbers each represent, in order, the number of road workers that is required each day

Maximal Number of Roads Simultaneously Blocked

All the following lines beginning with the letter M represent a subset of roads from which a maximal amount can be blocked simultaneously. Each line contains, following the 'M' letter, a sequence of numbers that represent — in order — the following quantities:

1. Max amount — The maximal amount of simultaneous roads blocked from the specified subset of roads
2. Road IDs — The remaining numbers each represent the id of a road that is part of the subset on which a maximal amount of roads can be blocked simultaneously

Precedences

All the following lines beginning with the letter P represent a precedence between two worksheets. Each line contains two numbers that represent the following quantities:

1. First Worksheet ID — The id of the worksheet that should precede the other one
2. Second Worksheet ID — The id of the worksheet that should be preceded by the other one

Solution File

Here is an example of solution file to the small instance from `smallInstance.txt` shown before:

`smallInstanceSolution.txt`

```
0 35
2 33
```

Solution files should only contain worksheets that are actually executed. Each line in a solution file should contain two numbers representing — in order — the following quantities:

1. Worksheet ID — The id of a worksheet
2. Starting Time — The starting time of the executed worksheet

Model - Easy

This model is an easy variant of the problem that was described earlier. In this version, only a subset of the constraints described earlier are used. Here is the subset of constraints that are part of the model:

- Precedences within Worksheet
- Worksheet Earliest Starting Time
- Worksheet Latest Starting Time
- Complete Worksheets
- Mandatory Worksheets
- Precedence Between Worksheets

The objective is the one defined earlier.

The data of the instances for the easy model makes the problem easier to solve. Indeed, in easy instances, the following data specificities are found:

- The perturbation cost associated to blocking a road remains constant through all days
- The worksheets all contain the same number of activities
- There is only one work center and it has a capacity larger than the sum of all workers needed by all activities
- All worksheets are mandatory

Here is an example of a small *EASY* instance file:

smallInstanceEASY.txt

```
45 5 1 3 9
0 0:45:3
1 0:45:5
2 0:45:1
3 0:45:4
4 0:45:1
0 10000
0 0 1 94 0 42 3 0 3 3 1 1 1
1 0 1 55 0 42 3 1 2 1 1 1 1
2 0 1 15 0 42 3 2 1 0 1 1 1
P 1 2
```

And here is a potential solution for this instance:

`smallSolutionEASY.txt`

```
0 6
1 0
2 3
```

Model - Medium

This model is an medium variant of the problem that was described earlier. In this version, only a subset of the constraints described earlier are used. Here is the subset of constraints that are part of the model:

- Precedences within Worksheet
- Worksheet Earliest Starting Time
- Worksheet Latest Starting Time
- Complete Worksheets
- Mandatory Worksheets
- Precedence Between Worksheets
- Work Center Capacity

The objective is the one defined earlier.

The data of the instances for the medium model makes the problem easier to solve. Indeed, in medium instances, the following data specificities are found:

- The perturbation cost associated to blocking a road remains constant through all days
- The worksheets all contain the same number of activities

Here is an example of a small *MEDIUM* instance file:

smallInstanceMEDIUM.txt

```
45 5 2 3 9
0 0:45:3
1 0:45:5
2 0:45:1
3 0:45:4
4 0:45:1
0 117
1 224
0 1 0 94 0 42 3 0 3 3 8 8 17
1 1 0 55 0 42 3 1 2 1 19 13 20
2 1 0 15 0 42 3 2 1 0 6 7 16
P 2 1
```

And here is a potential solution for this instance:

smallSolutionMEDIUM.txt

```
0 37
1 31
2 0
```

Model - Hard

This model is the full model of the problem that was described earlier. In this version, all the constraints described earlier are used. Here is the set of constraints that are part of the model:

- Precedences within Worksheet
- Worksheet Earliest Starting Time
- Worksheet Latest Starting Time
- Complete Worksheets
- Mandatory Worksheets
- Precedence Between Worksheets
- Work Center Capacity
- Maximal Number of Roads Simultaneously Blocked

The objective is the one defined earlier.

Here is an example of a small *HARD* instance file:

smallInstanceHARD.txt

```
45 5 2 3 7
0 0:45:3
1 0:37:4 37:43:1 43:45:4
2 0:35:1 35:45:4
3 0:21:4 21:45:3
4 0:2:5 2:31:5 31:45:4
0 254
1 166
0 0 0 57 23 43 2 0 0 11 15
1 0 0 72 0 42 3 3 4 0 15 18 9
2 1 0 90 14 43 2 0 0 8 20
P 1 2
M 1 0 1 4
M 1 0 3 4
```

And here is a potential solution for this instance:

smallSolutionHARD.txt

```
0 35
1 23
2 39
```

Solution Checker

We provide you with a runnable solution checker. This solution checker will check that the solution that you propose for a specific instance is valid. If it is not valid, a small list of the constraints violated will be displayed. If the solution is valid, the score of the solution will be displayed.

You can launch the solution checker in a terminal using the following command:

Terminal

```
% For Unix
> ./bin/check-solution DIFFICULTY PATH_TO_INSTANCE PATH_TO_SOLUTION

% For Windows
> \bin\check-solution.bat DIFFICULTY PATH_TO_INSTANCE PATH_TO_SOLUTION
```

The arguments of the solution checker are as follows:

1. `DIFFICULTY` is the difficulty of the instance you are considering either EASY, MEDIUM or HARD.
2. `PATH_TO_INSTANCE` is the path to the instance file defining the instance considered.
3. `PATH_TO_SOLUTION` is the path to the solution file for the instance considered.

Here are some example of uses:

Terminal

```
% For Unix EASY Instance
> ./bin/check-solution EASY EASY_5000_1500.txt SOL_EASY_5000_1500.txt

% For Unix HARD Instance
> ./bin/check-solution HARD HARD_5_3.txt SOL_HARD_5_3.txt

% For Windows MEDIUM Instance
> \bin\check-solution.bat MEDIUM MEDIUM_1000_100.txt SOL_MEDIUM_1000_100.txt

% For Windows EASY Instance
> \bin\check-solution.bat EASY EASY_2500_1000.txt SOL_EASY_2500_1000.txt
```