



FIRST ESTONIAN STUDY GROUP WITH INDUSTRY

4 – 8 FEBRUARY, 2019, TARTU, ESTONIA

---

## IP for Scheduling Problem

---

*Group Members:*

Miracle Amadi(LUT, Finland)  
Abdullah Makkeh(UT, Estonia)  
Kaveh Khoshkhah(UT, Estonia)  
Donatas Kavaliauskas (VU,  
Lithuania)

*Problem no. - 1*

Sambeet Mishra (TalTech,  
Estonia)  
Agatha Msola (LUT, Finland)  
Dipal Shah (LUT, Finland)

## Contents

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Introduction</b>                         | <b>3</b> |
| <b>2</b> | <b>Problem definition</b>                   | <b>3</b> |
| <b>3</b> | <b>Data description</b>                     | <b>4</b> |
| 3.1      | Average total task . . . . .                | 5        |
| 3.2      | Each hour task distribution . . . . .       | 5        |
| <b>4</b> | <b>Approach</b>                             | <b>7</b> |
| 4.1      | Schematic diagram . . . . .                 | 7        |
| 4.2      | Optimization: Integer Programming . . . . . | 7        |
| <b>5</b> | <b>Results</b>                              | <b>9</b> |
| <b>6</b> | <b>Conclusions</b>                          | <b>9</b> |

# 1 Introduction

Scheduling the working hours for the employees is a challenging task at the public and private sectors. It causes all types of administrative crisis, such as efficient design and plan of working hours to provide the best working environment for the employees and reduce work delays so that customers do not have to wait long for service. Furthermore, while planning an itinerary manager should consider all the real-world constraints for instants, each teams incoming task flow is continuous and the task should be finished within working hours, 40 working hours per week for each employee, national holidays, unannounced sick days, vacation seasons, and many more. Hence, scheduling the working hours for each team and team member can be treated as the oldest optimization problem in the field of computer science and operations research. Our optimization algorithm aims to provide a minimum number of team members and working hours for each team in accordance with the forecast workflow.

The schedule optimization problems can be solved using a number of methods such as,

- Exact algorithm such as Branch and Bound
- Meta-heuristics algorithms
- Heuristics algorithms
- Integer Programming (IP)
- Dynamic Programming
- And others

# 2 Problem definition

Swedbank is one of the most important and active banking firm in the Baltic region. It provides a real-time support service to their customers. However, keeping employees and customer happy is a tedious job for the current administration. Hence, the company aims at scheduling minimum working hours for each team members for a smooth workflow without the unexpected jolts and clash. Swedbank has data for the past two years of workload for the support team. Constructing the team and team members regular schedule has no reasonable meaning since the job flow depends on external or internal conditions

and workflow. Nevertheless, The manager should design the schedule considering all the stochastic irregularities such as team working hours should not increase more than 40 hours per week, and workflow should cover by the colleagues when the employee is on lunch break. The designed schedule should be planned to give answers like:

- How many employees should work in a team,
- Optimal working hours for a team considering historical workload for the last month,
- Batch time arrival information,
- Suitable lunch break time for each employee.

**Objective:** In this work, our main objective is to design a real-time service and support team monthly schedule considering all the possible uncertainty scenarios. The given constraints are as follows,

- Minimum and maximum working hours for the teams and team members.
- Scheduling the total numbers of working teams at a time.
- Team arrival and leaving time.
- Minimize work delays and maximizing each employees work efficiency.
- Avoid unexpected bumps and schedule crashes.

### 3 Data description

The company provided two years of workflow data for 3 different teams with varying working hours. Team 1 and 2 works during weekdays from 8.00 to 17.00 for a total of 8 hours and a few hours (9.00-13.00) on Saturdays. However, team 3 works for all 7 days from 6.00 to 23.00 for a total of 18 hours in a batch. The team 1 and 2 working hours are similar hence, we will further continue our analysis for two major team class 1 and 3.

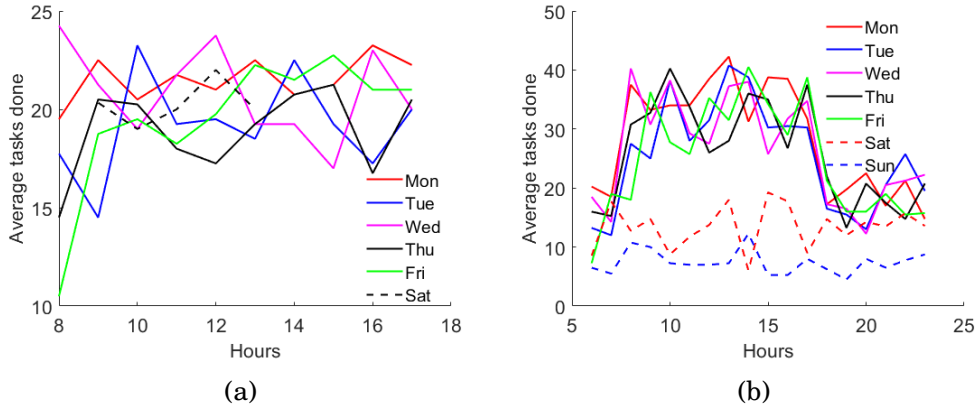


Figure 1: Total finished task for team 1 and 3

### 3.1 Average total task

Figure 1 represents the time series behavior for each day and hour for the team 1 (fig. 1(a)) and 3 (fig. 1(b)). It can be seen from figures that the official working days and hours are quite a busy period for both the team. Team 3 which is working all 7 days in a with for 18 hours have a peak hours during the official working days and hours. The given total task completed at each hour is further subdivided into various group depending upon their completion time such as task finished in 1, 5, 10, 15, 20, and  $\geq 20$  minutes for team 1. However, the tasks completion time for team 3 are much shorter than the previous team. The completion time for team 3 is further subdivided into the 1, 2, 3, 4, 5, and  $\geq 5$  minutes.

### 3.2 Each hour task distribution

Figures 2(a) and 2(b) illustrates team 1 and 3 early workload as per task type. It can be seen from the figure 2(a) that for team 1 majority of task completion is either in fall into 5 or 20 minutes, whereas for team 3 (fig. 2(b)) it is 2 minute particularly for the last few months of the year.

**Data statistics:** We analyzed the incoming task data statistics such as maximum, minimum, average, standard deviation and team efficiency. Table 1 and 2 represents the statistics of team 1 and 3 respectively.

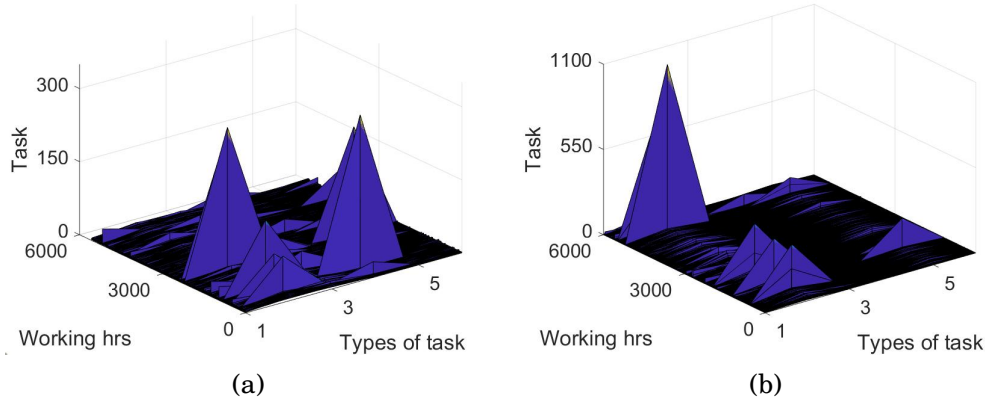


Figure 2: Task distribution for each hour in team 1 and 3

| Total task      | 0-1  | 1-5   | 6-10 | 11-15 | 16-20 | >20  |
|-----------------|------|-------|------|-------|-------|------|
| Maximum         | 34   | 297   | 50   | 318   | 221   | 59   |
| Minimum         | 0    | 0     | 0    | 0     | 0     | 0    |
| Average         | 2.23 | 7.93  | 5.72 | 1.34  | 0.28  | 0.69 |
| St. deviation   | 1.6  | 6.0   | 2.14 | 4.48  | 3.05  | 1.30 |
| Team efficiency | 2.23 | 39.65 | 57.2 | 19.8  | 5.6   | 20.7 |

Table 1: Workflow statistics of team 1.

Table 1 has a total of 145.18 task completion team efficiency for one hour. The amount of workload can be easily handled by 6 employees, however, one should consider all the other possible stochastic conditions as described earlier. Hence, from the statistical analysis, it seems 3 to 4 member team can survive a day.

| Total task      | 0-1  | 1-5   | 6-10 | 11-15 | 16-20 | >20  |
|-----------------|------|-------|------|-------|-------|------|
| Maximum         | 97   | 1089  | 70   | 73    | 203   | 19   |
| Minimum         | 0    | 0     | 0    | 0     | 0     | 0    |
| Average         | 1.09 | 12.66 | 0.22 | 0.011 | 5.84  | 0.05 |
| St. deviation   | 1.91 | 17.10 | 1.12 | 0.65  | 5.15  | 0.43 |
| Team efficiency | 1.09 | 63.3  | 2.2  | 0.165 | 116.8 | 0.95 |

Table 2: Workflow statistics of team 3.

Table 2 represents the team 3 statistics. A team has a total task completion efficiency of 184.51 per hour. The given quantity of work can be performed by a minimum 4 members team and a maximum of 5 members. However, there

should be at least 13 employees should work in a day irrespective the size of the teams to have a smooth workflow without delays.

## 4 Approach

The given problem can be modeled as an optimization problem. It can be solved with Mixed Integer Programming.

$$\begin{aligned} &\text{minimize} && a^t x \\ &\text{subject to} && Cx = b \\ & && Gx \leq h \\ & && x \in \mathbb{R} \cup \{0, 1\}. \end{aligned}$$

In our case, the objective function chosen is (min.  $a^t \delta$  where  $a$  is the vector of priorities for time delays and  $\delta$  is the vector of time delays). The mixed IP model simulation flow is described in Figure 3.

### 4.1 Schematic diagram

Our mixed IP model optimization starts with statistical data analysis. We derive the model tuning parameters information from the team statistics. We get a rough overview of how many employees are needed per day/month to handle the average workload. We transfer this initial information to the optimizer to derive an optimized schedule. In addition, we also consider the expected workload crisis and how it will affect employees overtime. We further analyze our optimized working schedules results from different angles and generate new tuning parameters to make a new schedule. The process was continued until the satisfied results are obtained.

### 4.2 Optimization: Integer Programming

This subsection presents the optimization method used. The problem is modeled as Mixed Integer Programming. First, the sets, parameters and variables will be explained.

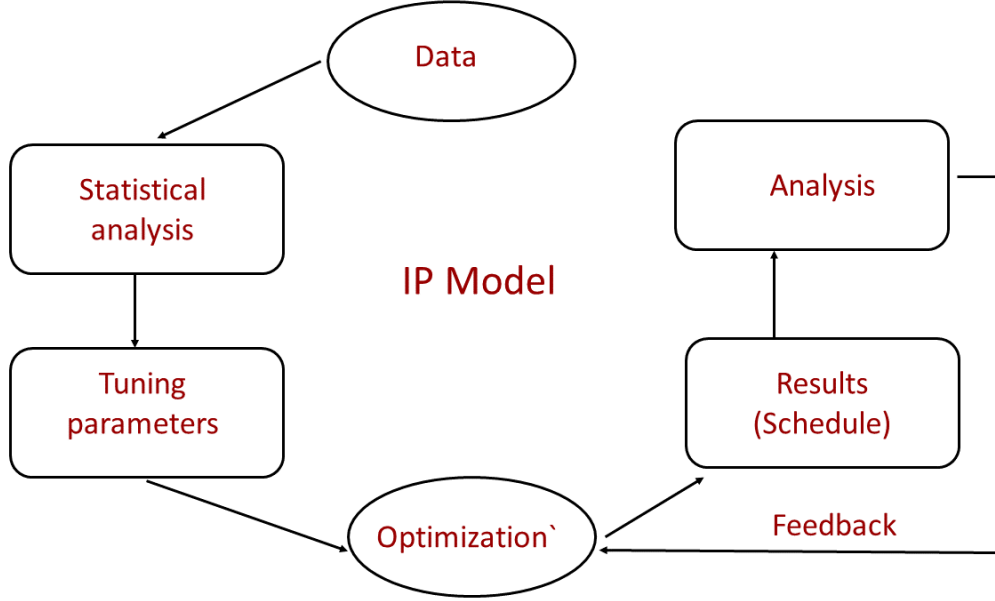


Figure 3: Schematic diagram for the IP workflow optimization

| Sets | Description        |
|------|--------------------|
| $S$  | set of time slots  |
| $D$  | set of days        |
| $T$  | set of tasks types |
| $V$  | set of employees   |

Table 3: Model Sets

**Model:**

$$\begin{aligned}
 &\text{minimize} && \sum_{i,j \in S \times D} a_{(i,j)} \delta_{(i,j)} \\
 &\text{subject to} && \sum_{i,j \in S \times D} (\beta_{(i,j),v} - \alpha_{(i,j),v}) + |D| \leq \theta_0 - \theta_2 && \forall v \in V \\
 &&& \sum_{i,j \in S \times D} (\beta_{(i,j),v} - \alpha_{(i,j),v}) + |D| \geq \theta_1 - \theta_2 && \forall v \in V \\
 &&& \sum_{i \in S} \alpha_{(i,j),v} = 1 && \forall v \in V, j \in D \\
 &&& \sum_{i \in S} \beta_{(i,j),v} = 1 && \forall v \in V, j \in D \\
 &&& \alpha_{(i,j),v} = 0 && \forall v \in V, j \in D, i \in S \setminus \{\theta_3, \theta_4\} \\
 &&& \sum_{i \in S} i (\beta_{(i,j),v} - \alpha_{(i,j),v}) \geq -1 && \forall v \in V, j \in D \\
 &&& \tau_{(i,j),v} \leq \sum_{k=1}^i \alpha_{(k,j),v} && \forall v \in V, i \in S, j \in D \\
 &&& \tau_{(i,j),v} \leq 1 - \sum_{k=1}^{i-1} \beta_{(k,j),v} && \forall v \in V, i \in S, j \in D
 \end{aligned}$$



| Parameters  | Description  |
|-------------|--|
| $c_k$       | duration of tasks of type $k$                                    |
| $a_{i,j}$   | delay priorities at time $i$ on day $j$                          |
| $w_{i,j,k}$ | number of tasks of type type $k$ received at time $i$ on day $j$ |
| $\theta_0$  | upper bound on the total working hours                           |
| $\theta_1$  | lower bound on the total working hours                           |
| $\theta_2$  | preparation time   |
| $\theta_3$  | starting time of batch 1   |
| $\theta_4$  | starting time of batch 2   |
| $\theta_5$  | the slot index of the starting time of the lunch break           |
| $\theta_6$  | the slot index of the ending time of the lunch break             |
| $\theta_7$  | the duration of the lunchtime                                    |

Table 4: Model Parameters

| Variables        | Description   |
|------------------|---|
| $\alpha_{i,j,v}$ | $\alpha_{i,j,v} = 1$ , if the starting slot time of work for employee $v$ is $i$ on day $j$ |
| $\beta_{i,j,v}$  | $\alpha_{i,j,v} = 1$ , if the ending slot time of work for employee $v$ is $i$ on day $j$   |
| $\tau_{i,j,v}$   | $\tau_{i,j,v} = 1$ , if the employee $v$ is working during time $i$ on day $j$              |
| $\delta_{i,j,v}$ | the amount of delay during time slot $i$ on day $j$   |

Table 5: Model Variables

$$\begin{aligned}
\sum_{k \in \mathbf{T}} c_k w_{(i,j),k} &\leq \sum_{v \in \mathbf{V}} \tau_{(i,j),v} + \delta_{(i,j)} & \forall i \in \mathbf{S}, j \in \mathbf{D} \\
\sum_{i=\theta_5}^{\theta_6} \sum_{k \in \mathbf{T}} c_k w_{(i,j),k} &\leq \sum_{i=\theta_5}^{\theta_6} \left( \sum_{v \in \mathbf{V}} \tau_{(i,j),v} + \delta_{i,j} \right) - \theta_7 \sum_{v \in \mathbf{V}} \tau_{(\theta_5,j),v} & \forall j \in \mathbf{D} \\
\sum_{\ell=i} \sum_{k \in \mathbf{T}} c_k w_{(\ell,j),k} &\leq \sum_{\ell=i} \sum_{v \in \mathbf{V}} \tau_{(\ell,j),v} + \sum_{v \in \mathbf{V}} \tau_{(|\mathbf{S}|,j),v} & \forall i \in \mathbf{S}, j \in \mathbf{D}.
\end{aligned}$$

## 5 Results

The following are the basic experiments we made:

## 6 Conclusions

The python environment is used for developing the scripts. Pyomo package is used for the development of the algebraic model. The *GLPK* solver (open-

| Employee | Day | Start Time | End Time |
|----------|-----|------------|----------|
| 1        | 1   | 1          | 13       |
| 1        | 2   | 1          | 17       |
| 1        | 3   | 1          | 1        |
| 1        | 4   | 1          | 1        |
| 1        | 5   | 1          | 5        |
| 1        | 6   | 1          | 1        |
| 1        | 7   | 1          | 1        |
| 1        | 8   | 1          | 12       |
| 1        | 9   | 1          | 1        |
| 1        | 10  | 1          | 1        |
| 1        | 11  | 1          | 1        |
| 1        | 12  | 1          | 1        |
| 1        | 13  | 1          | 1        |
| 1        | 14  | 1          | 1        |
| 1        | 15  | 1          | 1        |
| 1        | 16  | 1          | 8        |
| 1        | 17  | 1          | 1        |
| 1        | 18  | 1          | 5        |
| 1        | 19  | 1          | 1        |
| 1        | 20  | 1          | 6        |
| 1        | 21  | 6          | 6        |
| 1        | 22  | 1          | 18       |
| 1        | 23  | 6          | 12       |
| 1        | 24  | 1          | 4        |
| 1        | 25  | 1          | 8        |
| 1        | 26  | 1          | 9        |
| 1        | 27  | 1          | 11       |
| 1        | 28  | 1          | 6        |
| 2        | 1   | 1          | 2        |
| 2        | 2   | 6          | 5        |
| 2        | 3   | 1          | 11       |
| 2        | 4   | 1          | 8        |
| 2        | 5   | 1          | 1        |
| 2        | 6   | 6          | 17       |

Figure 4: Example of the schedule produced

| # Employees | Solution | Time     | Delays |
|-------------|----------|----------|--------|
| 100         | y        | 11 sec   | none   |
| 90          | y        | 9.21 sec | none   |
| 80          | y        | 8.53 sec | none   |
| 70          | y        | 7.4 sec  | none   |
| 60          | y        | 7.7 sec  | none   |
| 50          | y        | 7.8 sec  | none   |
| 40          | y        | 5.5 sec  | none   |
| 30          | y        | 6 sec    | none   |
| 20          | y        | 1859 sec | 6hrs   |
| 10          | n        | 0.2 sec  | n/a    |

Table 6: Results for experiments for Team 3 with two batches. The arriving tasks were drawn from exponential distribution with mean 80 and std 5

| # Employees | Solution | Time     | Total amount of Expected Delays |
|-------------|----------|----------|---------------------------------|
| 100         | y        | 11 sec   | none                            |
| 90          | y        | 10 sec   | none                            |
| 80          | y        | 8 sec    | none                            |
| 70          | y        | 7 sec    | none                            |
| 60          | y        | 8 sec    | none                            |
| 50          | y        | 8 sec    | none                            |
| 40          | y        | 6 sec    | none                            |
| 30          | y        | 6 sec    | none                            |
| 20          | y        | 1929 sec | 6hrs                            |
| 10          | n        | 0.4 sec  | n/a                             |

Table 7: Results for experiments for Team 3 with two batches. The arriving tasks were drawn from exponential distribution with mean 80 and std 5

source) is used to solve the model. However, a range of solvers can be used to solve the model. This is an open-source initiative and it allows any institutions to use it freely. The project is made available on GitHub platform public-ally as "Swedbank employee scheduling".

Scaling up the input data, it is advisable to use commercial solvers as "Gurobi" or "CPLEX". However for smaller instances of the problem can be solved in acceptable time and accuracy with open source solvers. paragraph

In a future development, new constraints such as the employee efficiency can be considered in the model as additional constraint.

| # Employees | Solution | Time    | Total amount of Expected Delays |
|-------------|----------|---------|---------------------------------|
| 100         | y        | 20 sec  | none                            |
| 90          | y        | 15 sec  | none                            |
| 80          | y        | 18 sec  | none                            |
| 70          | y        | 12 sec  | none                            |
| 60          | y        | 17 sec  | none                            |
| 50          | y        | 10 sec  | none                            |
| 40          | y        | 8 sec   | none                            |
| 30          | y        | 13 sec  | none                            |
| 20          | n/a      | n/a     | n/a                             |
| 10          | n        | 0.4 sec | n/a                             |

Table 8: Results for experiments for Team 3 with two batches. The arriving tasks were drawn from Poisson distribution with mean 100 and std 5

| # Employees | Solution | Time    | Delays |
|-------------|----------|---------|--------|
| 100         | y        | 2 sec   | none   |
| 90          | y        | 2 sec   | none   |
| 80          | y        | 1 sec   | none   |
| 70          | y        | 1 sec   | none   |
| 60          | y        | 1 sec   | none   |
| 50          | y        | 0.9 sec | none   |
| 40          | y        | 0.9 sec | none   |
| 30          | y        | 0.6 sec | none   |
| 20          | y        | 0.4 sec | none   |
| 10          | y        | 0.2 sec | n/a    |

Table 9: Results for experiments for Team 1 and 2 with two batches. The arriving tasks were drawn from normal distribution with mean 80 and std 5

| # Employees | Solution | Time    | Delays |
|-------------|----------|---------|--------|
| 100         | y        | 2 sec   | none   |
| 90          | y        | 2 sec   | none   |
| 80          | y        | 1 sec   | none   |
| 70          | y        | 1 sec   | none   |
| 60          | y        | 1 sec   | none   |
| 50          | y        | 0.9 sec | none   |
| 40          | y        | 0.9 sec | none   |
| 30          | y        | 0.6 sec | none   |
| 20          | y        | 0.4 sec | none   |
| 10          | y        | 0.2 sec | 1hr    |

Table 10: Results for experiments for Team 1 and 2 with two batches. The arriving tasks were drawn from exponential distribution with mean 80 and std 5

| # Employees | Solution | Time    | Delays |
|-------------|----------|---------|--------|
| 100         | y        | 4 sec   | none   |
| 90          | y        | 3 sec   | none   |
| 80          | y        | 3 sec   | none   |
| 70          | y        | 2 sec   | none   |
| 60          | y        | 1 sec   | none   |
| 50          | y        | 1 sec   | none   |
| 40          | y        | 1 sec   | none   |
| 30          | y        | 0.9 sec | none   |
| 20          | y        | 0.9 sec | none   |
| 10          | y        | 0.9 sec | 1hr    |

Table 11: Results for experiments for Team 1 and 2 with two batches. The arriving tasks were drawn from Poisson distribution with mean 100 and std 5