ISSN No: 2582-8312

Lattice

THE MACHINE LEARNING JOURNAL

VOLUME-4, **ISSUE**-1 (January - March 2023)

AN INTERNATIONAL PEER REVIEWED JOURNAL
IN DATA SCIENCE & MACHINE LEARNING

ADaSci | THE ASSOCIATION OF DATA SCIENTISTS

www.adasci.org

Decision support to a Retailer's staffing strategy using Mixed Integer Linear Programming

Anand Pratap Singh
Amplify Analytix
anand.singh
@amplifyanalytix.com

Shashank Srinivasan
Amplify Analytix
shashank.srinivasan
@amplifyanalytix.com

Moulik Sthapak
Amplify Analytix
moulik.sthapak
@amplifyanalytix.com

Bharathan Shamasundar
Amplify Analytix
bharathan.rs
@amplifyanalytix.com

***Abstract*— Preparing an efficient and effective staffing roster is a challenging task for retailers. An efficient staffing schedule optimizes payroll cost, while meeting desired customer service levels. Doing this, while complying with labour laws and organizational policies makes the schedule effective. While compliance with legal requirements is mandatory, evaluating the impact of alternate organizational policies helps formulate a suitable staffing strategy. These policies present themselves in a combination of factors around shift duration, standardized shift start times, number & length of breaks during the day, off-days, etc. In this paper, we describe our evaluation of alternate organizational policies on payroll cost of a leading luxury retailer in the UK. Each candidate policy was evaluated using Mixed Integer Linear Programming (MILP), following a two-step approach. In the first step, we forecast the number of customers visiting the store at half-hourly intervals and translate it into expected staff count required on the shopfloor. In the second step, we infuse this, together with legal and organizational policy constraints, into the MILP. Our experiments revealed that flexibility in shift start times, and the mandated minimum staff strength for a store, were the primary drivers of payroll cost. More than three shift start times a day, allowed the flexibility of having a higher mix of part-timers. This translated into a 10-12% reduction in paid hours. This benefit can then be considered in the context of the overhead of managing a large part-time population, and a potential shift in organizational culture. Further, the driver of minimum staff strength highlighted the need for reviewing the store's physical layout.**

I. Introduction:

Creating a daily employee roster to address customer demand for a retailer is challenging because of the uncertainty associated with the number of customers visiting the store. Holidays bring-in a lot of customers in a short duration and maintaining the quality of service becomes crucial. This fluctuating demand directly impacts the operations and management teams who must plan the supply chain, inventory, and staffing schedule. Further, besides developing a specific staffing schedule, the management is presented with a need to commit to a staffing strategy that is best suited to its requirements. Common to all staffing strategies is compliance with labour laws. Some examples of compliance requirements include – limits on total work hours in a week, minimum number of rest days in a seven-day period, etc. Where strategies differ, however, is on considerations that are a matter of organizational policy. One example of this is the flexibility in the number of unique shifts the organization wants to allow. Another example is a model that specifies longer daily working hours, balanced with fewer number of days in a workweek. All this presents a need to evaluate several possible strategies in the first place, before selecting one. For evaluating a particular strategy, we follow a two-step methodology. While the first step is to forecast customer footfalls at every half-hour, the second step is to plug-in the forecasts and build other constraints into a Linear Program to optimally staff the shopfloor. Greater flexibility in shift starts, allows the linear program to allocate more part-timers to the workforce. This lends itself very well to meeting peaks in demand.

A. Existing policy and Assumptions:

We first created a staff schedule based on the existing staffing policy of one of the leading retailers in the UK. The existing policy has a minimum shift duration of 5 hours and a maximum of 12 hours. An employee can start his or her day at only three possible start times, namely early morning, morning or afternoon. The cost per hour for each employee is taken as a constant. So, the cost of a given shift is directly proportional to the total number of working hours in it, notwithstanding the mix in terms of skill or experience. Tea and lunch breaks, as per the breaks policy, are a function of the length of the shift. Shorter duration shifts get single or no tea-breaks. Shift durations beyond a particular threshold get lunch breaks in addition. Further, while the tea breaks are paid, the employee in question is considered unavailable to serve customers. Lunch breaks are unpaid. The shift duration includes the duration of breaks. An important input from business was to consider the average turnaround time to meet customer service requests to be 25 minutes per customer.

II. Problem Description

To prepare a workforce schedule at half hourly level that optimizes payroll cost, adapts to the peak as well as off-peak demands while maintaining the quality of service to customers. To ensure the right allocation of workforce, it is important to estimate the customer service demand. Since workforce scheduling is at half-hourly levels, demand too should be estimated at the same level. Labor laws and company's break policy also need to be considered in the process.

III. Methodology

Here we are adopting a two-step methodology with the first step to estimate the customer coming in the store at half hourly level and second step to arrive at optimal workforce

schedules which ensures all business and legal constraints are followed.

A. Estimating customer demands

To estimate half hourly customer demands, we used the estimated weekly sales projection for the financial year 2022 provided to us by operational stakeholders of the retailer. The weekly sales projection for the year 2022 was compared to the aggregated weekly historical sales data of the year 2019. The year 2019 proved to be the ideal choice year as this was the most recent year which did not suffer from customer demand fluctuations caused by the COVID effect. The percentage of the customers arriving at the store for each working day at hourly level was identified, and this percentage was applied to the weekly sales projection to arrive at a daily sales projection for 2022. In the next step, the half hourly sales projection for each day of the week was calculated using the percentage of proportion of customers arriving at each working day at a half hourly level in the year 2019. This percentage of proportions was then applied to the daily sales projection for the year 2022 as calculated from the previous step. The resulting estimates at a half hourly level for the year 2022 are now ready to serve as an input to a mixed integer linear programming. The estimates were made preserving the structure of the shift hours of the stores as seen in 2019.

B. MILP (Mixed Integer Linear Programming)

MILP is one of the techniques out of all other linear programming techniques that solves for variables in which few of them can take integer values. This methodology is widely used in industries as most of the time it is needed when the nature of the required solution must be expressed in integer numbers thus making it a relevant proposition. It uses techniques like root relaxation and branch and bound algorithms which are considered as state-of-the-art algorithms, to arrive at the solution in significantly reduced time.

C. Decision variables

In optimization we often define decision variables in a way that best serves to represent the business problem. They are used to define objectives and constraints in a mathematical representation or notation. In this paper, we define three decision variables X, Y and P with three, two and two dimensions respectively. The dimensions for X variable are days (with length of the axis being the number of days in a calendar month for example 0-30), employee (with length of the dimension being total number of employees for example 0-40) and shift pattern (a 24-bit long pattern of 1's and 0's because store opens from 9AM to 9PM making it 12 hour and for half hourly level it is 24). Similarly, Y is a 2-dimensional variable with days and employee axes.

$$X_{d,e,s}=\begin{cases}1 \text{ If employee 'e' on day 'd' takes 's' shift pattern,}\\ 0 \text{ Otherwise}\end{cases}$$

$$Y_{d,e}=\begin{cases}1 \text{ If employee 'e' on day 'd' is working,}\\ 0 \text{ Otherwise}\end{cases}$$

$$P_{e,sd}=\begin{cases}1 \text{ if employee 'e' is allocated a shift of duration 'sd',}\\ 0 \text{ Otherwise}\end{cases}$$

P is another two-dimensional variable with axes being (number of employees 'e' and possible shift duration 'sd'). Every possible shift pattern 's' will have a certain shift duration denoted with 'sd'. Since the minimum shift duration is 5 hours and maximum 12 hours with an incremental half hour, this makes a total of 15 possible shift durations. Mathematically we can write'sd' ∈ (5,5.5,6,6.5, ...12). Although solving for X only would be enough to get an optimal roster plan, the challenge lies in formulating a few businesses constraints which cannot be represented with X decision variable alone. Since X and Y are two decision variables having two axes in common, they need to be linked for them to convey the same piece of information by adding additional nonbusiness constraints. For example, if for a given e, d, s variable X takes the value 1 then for the same e, d variable Y cannot be 0 meaning if on a given day an employee is working on a certain hour, the same employee cannot be on leave for the day. Similarly, we need to link decision variable P with decision variable X explained in the later section. The essence of variable P is in the context of solving constriant1 described below. Which basically tracks what shift duration an employee has been allocated to regardless of the day of the month.

D. Objective Function

The objective is represented in terms of decision variables is the following:

$$\text{Minimize}\left(\sum X_{d,e,s}\cdot \text{Cost}_s\right)$$

Where 's' represents an instance of the shift pattern and Cost[s] represents the cost of the shift pattern. The hourly wage rate for every employee irrespective of the type (full time or part time) is the same, therefore we can compute cost directly from the shift patterns and associate it with every instance of the shift pattern and this assigned cost would be the same regardless of the nature and skills of the employee. The costs associated with a shift pattern are inclusive of breaks. We have two classes of breaks, one being a "Tea break", the other being "Lunch break". Tea breaks are half an hour in duration and are paid but Lunch breaks are 1 hour in duration and are unpaid as per company policy. In the manner that every shift pattern has a cost associated with it, a similar cost is also associated to the duration of every shift pattern namely *duration[s]* or $duration_s$. This tells about the duration of the shift pattern after taking into consideration whether the breaks are part of the weekly working hour limits as per labor law. So, every shift pattern 's' has two information associated with it, its cost and its duration denoted by $Cost_s$ or cost[s] and $duration_s$ or duration[s]. Mathematically we are trying to minimize the sum of the product of decision variable and its corresponding cost. Since every shift from the enumerated list of shift patterns (one shift pattern being a 24-bit long pattern on 1's and 0's telling if at a particular half hour, an employee on a certain day should work or not) comes with a cost depending on the duration of the shift and inclusive of breaks which may be paid or unpaid as per company's break policy. Multiplying the decision variable with it will force the optimization model to pick a shift pattern out of all shift enumeration which is the most cost effective. To handle leave on a certain day we added a 24-

bit long sequence of 0's in the shift patterns and if model wants to allocate a holiday to any employee, it can select and assign this shift pattern (all 0's) on a given day. The cost associated with this pattern would also be 0.

E. Constraint

The following is the constraints list that were the part of basic constraints among all the scenarios or alternate policies.

1. Employees should be assigned the same shift duration throughout the monthly schedule.
2. Employees are required to start their shift 30 minutes before or end their shift 30 minutes after the start/end times of business.
3. Employees should get at least 1 weekend off in a month. 1 weekend off refers to a consecutive off on Saturday and Sunday.
4. Employees can only start at 3 possible timeslots in a day.
5. There must be at least a 90-minute gap after the end of the first break and the start of a second break for any employee.
6. The first break should be given no later than 6 hours in a shift.
7. One-hour break to be given no less than 2.5 hours after the start of shift time and must be given between the hours 12pm and 7.30pm.
8. The last break should finish no later than 1 hour before the end time of a shift.
9. The number of employees at any time interval should satisfy the estimated customer demand.
10. A minimum number of employees should always work at any given time interval based on the desired factor (currently, the factor is determined based on the standard deviation of the demand on any day)
11. Site operational hours are 10 AM to 9 PM from Monday to Saturday.
12. Site operational hours are 11:30 AM to 6 PM on Sunday.
13. Employees can work between 5 and 12 hours in a day.
14. The employees should have a rest period of 11 consecutive hours in every 24-hour period
15. The employees can work for a maximum of 48 hours in a 7-day period
16. An employee can only work 1 shift a day
17. Tea breaks and lunch breaks should be applied as per the policy
18. An employee should get minimum 2 days rest in a 7-day period

F. Mathematical representations

A few of the constraint from the above section are represented as described below:

$$\text{Constraint 16 } \sum_{s=0}^{S} X_{d,e,s} = 1$$

where 'S' denotes total count of possible shifts (1)

For Constraint 18 first we need to link decision variables X and Y explained on equation (2). So, Constraint 18 can be written as combination of following two constraints:

$$\sum_{s=0}^{S-1} X_{d,e,s} = Y_{d,e}$$

where S-1' denotes all shifts except the holiday shift (2)

$$\sum_{w=0}^{6} Y_{w,e} \leq 5,$$

where w is a sliding window of 7 days length (3)

For Constraint 15

$$\sum_{w=0}^{6} \sum_{s=0}^{S} X_{w,e,s} \cdot \text{duration}_s \leq 48,$$

where w is a sliding window of 7 days length (4)

Constraint 9&10 are related to the minimum employee cover, and these should scale as per demand. Since we have demand estimated at half hourly level from the forecast for customer service request, Let's say we have *demand[t]* available at half hourly level for the entire calendar month and given that it takes around 25 mins on average to complete one customer service request. The equation for these constraints is formulated as shown below:

$$\sum_{e=0}^{E} \sum_{s=0}^{S} X_{d,e,s} \cdot s_t \geq \text{Min}\left(\text{Minimum cover}, \left(\frac{25}{30}\right) \cdot \text{demand}_t\right),$$

$\forall$ t in working hours (5)

Operating hours of the store may vary depending on the day of the week as mentioned in constraints 10&11 above. To address Constraint 1, we already have *duration[s]* for every shift pattern 's', therefore we can leverage this information to formulate this constraint and while using another decision variable 'P' as mentioned above. This is a 2-dimensional decision variable with employee and shift duration as axes. This constraint comes with a degenerated constraint, so we can further split this into constraints since an employee can take only one type of shift durations in entire month we can write:

$$\sum P_{e,sd} = 1 \ \forall\ e \in E, \text{ where 'sd'} \in \text{durations} \quad (6)$$

And the second part can be written as:

$$\sum_{s=0}^{S} X_{d,e,s} \leq P_{e,sd} \ \forall\ s \in \{\text{duration}_s = sd\}$$

where 'sd' is duration$_s$ for given 's' (7)

Every other constraint has been taken care of while generating possible shift patterns itself.

IV. Candidate Policies or Scenarios:

The following table presents the various scenarios that were explored after considering business stakeholder's input and the MILP model were built. To explain the scenarios, additional constraints were added per scenarios compared to the base constraint explained above and given to MILP models and then compare what is best suited for the current job market from the employment standpoint that minimizes overall labor costs and still attracts employees. In all the scenarios shift start times were flexible although the shift duration was changed according to the nature of the employee being part or full timer.

LIST OF SCENARIOS EXPLORED

S No.	Scenario Details
1	3 consecutive days working followed by at least 3 days off Full Time Employee – Daily Shift Duration – 5 to 12 hours Part Time Employee – Daily Shift Duration – 5 to 7 Hours
2	5 days working and at least 2 days off Full Time Employee – Daily Shift Duration – 5 to 12 hours Part Time Employee – Daily Shift Duration – 5 to 7 hours
3	Employees working 6-hour shifts for 5 days per week Daily Shift Duration – 6 hours Days off – 2 days
4	Full Time Employees working 5 days x 7.5-hour shifts Daily Shift Duration – 7.5 hours Days off – 2 days
5	4 consecutive days working and 4 days off Daily Shift Duration – 5 to 12 hours Days off – 4 days
6	5 days working and exactly 2 days off Full Time Employee – Daily Shift Duration – 8/8.5/9 hours (working 5 days) Part Time Employee – Daily Shift Duration – 6.5 / 8.5 hours (working 2 or 3 days in a week)
7	3 consecutive days working and 3 days off Full Time Employee – Daily Shift Duration – 5 to 12 hours Part Time Employee – Daily Shift Duration – 5 to 7 hours

V. Challenges

Solving optimization problems is often challenging. The challenges involve mathematical representations of business constraints. A more effective and tighter constraint can cut down the solution space more quickly and efficiently thus reducing the run time complexity. We present a few challenges during the development below.

1. Formulating the "same shift duration in a month" constraint which is mentioned as constraint1 above was difficult, we solved this constraint first non-linearly and later linearly, but this constraint was still time consuming and computationally expensive as it picked the most optimal duration that should fit for all the days in a month or a week. Additionally, the complexity of the constraint increases as Sunday store timings add a heavily biased solution to the interpretation of this constraint.

2. Adding full-timers and part timers as a capability inside the model was complex as the constraint were different for full and part timers and part timers as per scenarios. If the model can address the ratio (part time to full time mix) itself without explicitly trying different ratios, could have result in lowering the complexity. This can also be considered as future scope of improvement.

3. Execution time of the model is highly sensitive to the input and unpredictable as there are multiple working parts that interact with each other to an impact on the model and hence determining the run-time is difficult

4. Adding a new constraint to the existing setup is sometime exceedingly difficult as it can change the entire formulation done so far or it may lead to the current design not being able to incorporate the formulation of the new constraint. Solving MILP for various policy from the policy table described above leads to addition of new constraints, for example three consecutive days of working followed by three consecutives off needs to be formulated as a separate constraint.

VI. Computation

We first tried to solve the MILP model with open-source solvers available for free, but the performance of these solvers turned out to be low. As the number and complex constraints are added the open-source solvers started taking days to solve the MILP problem and were not able to prove optimality within stipulated time limits. Hence, we tried GUROBI solver which is a commercial solver available in the market. After converting all the breaks related Non-Linear constraints to Linear constraints by generating all possible shift patterns which are valid and passing these patterns to gurobi solver, we were able to reduce the time complexity to a few hours and prove the optimality. Gurobi solver is designed to exploit the underlying infrastructure and thus making it the fastest solver available.

VII. Business Impact

To evaluate the various candidate policies the total labour cost and workforce schedules were analyzed and compared to the existing policy. The workforce schedules were compared based on how the number of part-time and full-time employees were allocated to various shift start times and the distribution of the shift duration among them. Some of the key insights and impacts are listed below.

1. The exercise enabled the business team to have more visibility on how the part-time employees can be best placed at the different times of the day such that at peak hours customers can be addressed without reducing Quality-of-Service and at the same time reducing the shift durations to lower the cost of labor. This gives an opportunity to select ideal shift start times along with number shift start time slots which are only three (mentioned above) in the existing policy or scenario.
2. Based on the scenarios explained above, MILP model suggested an important input on what should be the optimal mix of part- and full-timers they should hire considering the estimated customer traffic and hence reducing the overall

cost 10-12% (around 50-man hour work) compared to existing policy.

3. Interestingly, in all the scenarios a basic constraint (Constraint 10) which is the minimum employee cover at any given time turned out to be a little higher in the existing policy and should be considered as a scope for further cost reduction.

Acknowledgment

I would like to Thank Amplify Analytix for providing this opportunity and the support throughout the project. I would like to express my gratitude to Bharathan R Shamasundar, Chief data scientist and Managing Partner at Amplify Analytix, for providing us with the necessary business and technical guidance needed for this paper.

References

[1] C. S. Azmat and M. Widmer. A case study of single shift planning and scheduling under annualized hours: A simple three-step approach. "European Journal of Operational Research", 153: 148–175, 2004.

[2] A. Billionnet. Integer programming to schedule a hierarchical workforce with variable demands. "European Journal of Operational Research", 114: 105–114, 1999.

[3] R.N. Burns and G.J. Koop. A modular approach to optimal multiple-shift manpower scheduling. "Operations Research", 35(1): 100–110, 1987.

[4] A. Corominas, A. Lusa and R. Pastor. Annualised hours: A real flexibility tool. "OR Insight", 18: 10–14, 2005.

[5] R. Hung. Single-shift off-day scheduling of a hierarchical workforce with variable demands. "European Journal of Operational Research", 78(1): 49–57, 1994.

ADaSci | THE ASSOCIATION OF DATA SCIENTISTS

BANGALORE, INDIA