

The Multilevel and Multiskill Labor Staffing and Scheduling Model for Data Entry Companies

Zheng Tan, Feng Wu, and Zheng-xiang Zhang

Abstract Data entry, one of the basic businesses in service outsourcing, depends much on human resources. The companies need to reasonably choose and arrange staffs in order to deliver data production in time. In consideration of production process and staff skill, a two-step multilevel multiskill labor staffing and scheduling model was built. The results of the proposed model with the data from the company ICP indicated the availability of the model, which can improve the vendors' capability and the significance of promoting development of the service outsourcing industry in China.

Keywords Labor staffing and scheduling • Linear programming • Multilevel • Multiskill

1 Introduction

Service outsourcing, an important part of high level modern service [1], is very important to industrial restructuring and promotion. The scope is widespread: hospital [2, 3], finance [4, 5], insurance [6], logistics [7] and so on. For the reason of fewer requirements to the staffs and low enter doorsill, the business is highly competitive and the profit space is limited [8]. Data entry outsourcing industry has the problem of the low staff utilization rate: the shortage/redundance of staffs in the program group are the common occurrence. The group has to temporarily transfer some staffs form other group or has no extra orders to distribute to the workers.

Z. Tan (✉) • F. Wu • Z. Zhang

School of Management, Xi'an Jiaotong University, Xi'an 710049, China

The Key Lab of the Ministry of Education for Process Control & Efficiency Engineering, Xi'an, China

e-mail: zheng.tan1986@gmail.com

Even worse, the group cannot deliver the qualified orders on schedule. The company has to pay hefty fines and lose customer trust. However, few studies focus on the operation system of data entry companies. Some papers mainly pay attention to clinical data collection and entry into the information system [9–11].

This paper studied the multilevel and multiskill staff scheduling problem in the data processing company. Depending on the mathematical model, the staff schedules can be generated which can help company deliver the data with low production cost in time.

2 Methodology

2.1 Problem Description

Data entry belongs to system operation service: it mainly process and integrates the data (including information) from bank, insurance and so on. The data processing process in the outsourcing company is shown in Fig. 1.

Segmental model derived from set-covering model proposed by Dantzig [12] and Edie [13] is one kind of developed mathematical model adopted by many researchers [14, 15] for labor staffing problem. Its Basic idea is deciding how shifts covered all the time intervals, then solve the staff demand problem in every time intervals to get the staff arrangement.

According to idea of segmental model, a two-step staff scheduling model, considered the whole data processing process in the outsourcing company was built with the unique constrains coming from the reality.

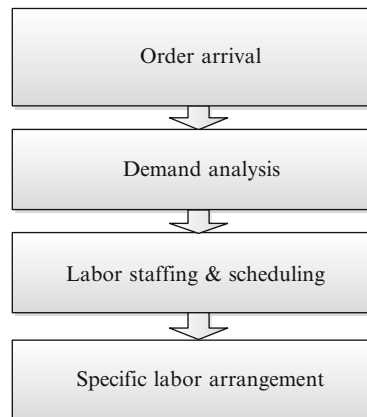


Fig. 1 Data entry process flowchart

2.2 Model Construction

Some terms are explained as follows:

Level of the order processing: the process of the data processing outsourcing business can be divided into three steps: classification, data entry and output. The original order should be transformed to final data product after several treatments. The order processed has transitivity in every level: the order processed in the previous level is the input in the next level.

Order type: the order type can be accepted by the operation system. Only with the same arrival time, delivery time and needed skill can two order be the same order type.

And the parameters used in the model can be seen as follows:

- I: the set of the order type,
- G: the vector of the arrival time of the order.
- M: the vector indicates the number of the accepted order D: the vector indicates the delivering time of the order
- F: the vector indicates the finishing time of the order
- T: the vector indicates the processing time of the order
- L: the vector of the loss when there is back order.
- Z: the set of the staff skill
- N: the vector of the number of the staff in every skill group
- S: the set of skill type
- Q: the set of the working shift
- P: the set of the time interval
- c_{i,q_v} : The cost of the type I order processed in the working shift q_v
- $C_{I \times v}$: The vector indicates order-working shift cost
- e_{i,p_u} : The cost of the type I order processed in time interval p_u :
- $E_{I \times U}$: The vector indicated order- time interval cost
- y_{i,p_u} : The staff number needed to process type I order.
- Y: the matrix indicates the labor request consideration of the order type and time interval
- X: the matrix indicates the staff scheduling consideration of the order type and working shift.

For the same type order, there is relationship between the staff demand in the time interval and the practical staff supply: $Y = X \times R$. When the skill s_j can cover the order; the number of the working staff in the skill group can be expressed as

$$\sum_{i \in J \otimes I} y_{i,p_u} \left(\sum_{i \in J \otimes I} x_{i,q_v} \right).$$

In this staff scheduling problem, the known condition included $O, S, T, N, R, M, D, P, Q, C, E, J \otimes I, q_v \supseteq p_u$ ($1 \leq i \leq I, 1 \leq j \leq J$), we can make Y satisfy some requirements to solve X.

Step 1: labor scheduling problem

$$\min \left[\sum_{i=1}^I \sum_{u=g_i}^V e_{i,p_u} y_{i,p_u} + \sum_{i=1}^{I/2} \sum_{u=d_{i+1/2}}^V \left(l_{i+1/2} \frac{u - d_{i+1/2}}{t_{i+1/2}} y_{i+1/2,p_u} \right) \right] \tag{1}$$

Constrains:

$$y_{i,p_u} = 0, \quad \forall i \in [1, I], u \in [1, g_i] \tag{2}$$

$$\frac{d - g_i}{t_i} \sum_{u=g_i}^d y_{i,p_u} \leq m_i \leq \sum_{u=g_i}^V \left(\frac{u - g_i}{t_i} y_{i,p_u} \right), \quad i \in [1, I], d_i \leq V \tag{3}$$

$$\sum_{\xi=g_{i+1/2}+1}^u \frac{p_{\xi} - p_{\xi-1}}{t_{i+1/2}} y_{i+1/2,p_{\xi}} \leq \sum_{\Delta=g_i+1}^{u-t_i+1/2} \frac{p_{\Delta} - p_{\Delta-1}}{t_i} y_{i,p_{\Delta}}, \quad \forall i \in [1, I/2], u \in [g_i, V] \tag{4}$$

$$\sum_{i \in J' \otimes I} y_{i,p_u} \leq n_{j'}, \quad \forall u \in [g_i, V], n_i \in N, i \in [1, I], j' \in [1, J] \tag{5}$$

$$\sum_{i \in J'' \otimes I} y_{i,p_u} \leq \sum_{j''} n_{j''}, \quad \forall u \in [g_i, V], n_i \in N, i \in [1, I], j'' \in [1, J] \tag{6}$$

$$y_{i,p_u} \geq 0, y_{i,p_u} \in \mathbb{Z} \tag{7}$$

Constraint (2) ensures no staffs are assigned before orders arrival. Constraint (3) guarantees the number of order processed should be between processing ability and the number of the accepted order in the defined period. Constraint (4) indicates the processing sequence from level 1 to level 2: the number of the accumulated order in the level2 should be no more than that in the level 1 at any time; Constraint (5) indicates the number of the workers assigned for the order I should less than the number of the worker in the skill group at any time; Constraint (6) indicates the total number of the workers assigned for the order I should less than the total number of the worker from all skill group. Constraint (7) guarantees the number of the workers should be positive integer.

Step 2: labor staffing problem

$$\min \left[\sum_{i=1}^I \sum_{v=1}^V c_{i,q_v} x_{i,q_v} + \sum_{i=1}^{I/2} \sum_{v=d_{i+1/2}}^V \left(l_{d_{i+1/2}} \frac{v - d_{i+1/2}}{t_{i+1/2}} x_{i+1/2,q_v} \right) \right] \tag{8}$$

Constrains:

$$x_{i,q_v} = 0, \quad \forall i \in [1, I], v \in [1, g_i] \tag{9}$$

$$X \times A \geq Y, \quad \forall u, v \in [1, V], \forall i \in [1, I], \sum_v \frac{v - g_i}{t_i} x_{i,q_v} < m_i \quad (10)$$

$$x_{i,q_{v+1}} = x_{i,q_{v+2}} = \dots = x_{i,q_v} = 0, \quad \forall i \in [1, I], \forall v \in [1, V], \sum_v \frac{v - g_i}{t_i} x_{i,q_v} \geq m_i \quad (11)$$

$$\sum_{\xi=g_i+I/2+1}^u \frac{q_\xi - q_{\xi-1}}{t_{i+I/2}} x_{i+I/2,p_\xi} \leq \sum_{\Delta=g_i+1}^{u-I+I/2} \frac{q_\Delta - q_{\Delta-1}}{t_i} x_{i,q_\Delta}, \quad \forall i \in [1, I], v \in [1, V] \quad (12)$$

$$\sum_{i \in J \otimes I} x_{i,q_v} \leq n_j, \quad \forall q_v \in [1, V], n_i \in N, i \in [1, I], j' \in [1, J] \quad (13)$$

$$\sum_{i \in J \otimes I} x_{i,q_v} \leq \sum_{j''} n_{j''}, \quad \forall u \in [g_i, V], n_i \in N, i \in [1, I], j'' \in [1, J] \quad (14)$$

$$x_{i,q_v} \geq 0, x_{i,q_v} \in \mathbb{Z} \quad (15)$$

Equation (9) ensures no workers are arranged before order arrival. Constraint (10) indicates the worker assigned in every working shift should satisfy the labor demand solved in step 1. Constraint (11) expresses that no workers should be assigned to the working shift once the number of accumulated processed order has equaled to the total order demand (workers assigned in every working shift is more than that in the time intervals belonged to the working shift. The practical progress is faster than which described in the “labor scheduling model”. So the labor staffing in the working shift needs to be stopped compulsively.). Constraint (12) indicates that the processing sequence should from level 1 to level 2, and the number of the accumulated order in level 2 should no more than that in level1 at any time; the formulation from (13), (14), (15) has the similar meaning as the formulation from (5), (6), (7).

3 Results

3.1 Model Input

The data collected for the model is from the company ICP founded in 2002. According to the field research, we choose the number of the arriving order in program A, B, C and VR in 1 day and the demand can be defined. The entry step can be simplified into two step; all the workers can be divided into four skill group: $Z = \{z_1, z_2, z_3, z_4\}$, four group respectively have 15, 15, 12 and 10 workers (Fig. 2).

The specific information of four programs can be seen in Tables 1 and 2.

Fig. 2 Correspondence of skill group and order type

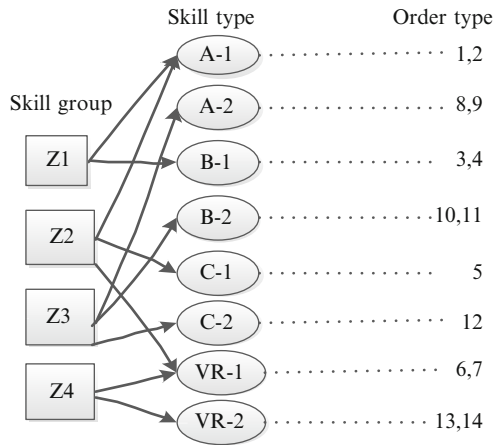


Table 1 Arrival time, delivery time and order amount

	Arrival time	Delivery time	Order received
A	8:30	18:00	286
	13:30	18:00	251
B	9:30	14:30	60
	14:30	9:30 next day	87
C	9:30	18:00	93
VR	8:30	11:30	50
	13:30	16:30	49

Table 2 Process time, labor and back order cost

Program	A		B		C		VR	
Level	1	2	1	2	1	2	1	2
Process time (min)	16	6	12	5	20	10	15	5
Normal labor cost (Yuan)	0.9	0.4	0.7	0.3	1	0.6	0.8	0.4
Overtime labor cost (Yuan)	1.8	0.8	1.4	0.6	2	1.2	1.6	0.8
Back order cost (Yuan)	2.5		2.8		2		3.5	

3.2 Mathematical Results

We solve the model by IBM ILOG CPLEX Optimization Studio12.3 and get the objective function value $f(x) = 287.9$ and the specific order-interval scheduling plan for the company ICP (see Table 3).

The above results are plugged into the 2nd step model and get the order-shift scheduling plan (see Table 4) with the total cost of 285.5.

Table 3 Order-interval staffing result

Interval \ Order	Interval									
	1	2	3	4	5	6	7	8	9	10
1	4	4	4	16	16	11	11	11	0	0
2	0	0	0	0	14	18	18	18	0	0
3	0	1	1	5	5	0	0	0	0	0
4	0	0	0	0	0	2	2	2	6	6
5	0	0	0	7	7	6	6	6	0	0
6	3	3	3	3	3	0	0	0	0	0
7	0	0	0	0	5	3	3	3	0	0
8	1	1	1	6	6	4	4	4	1	1
9	0	0	0	0	3	7	7	7	1	1
10	0	0	0	1	1	1	1	1	0	0
11	0	0	0	0	0	0	0	0	0	0
12	0	0	0	1	1	4	4	4	1	1
13	1	1	1	1	1	0	0	0	0	0
14	0	0	0	0	2	1	1	1	0	0

Table 4 Order-shift scheduling result

Order \ Shifts	Shifts			
	1	2	3	4
1	4	16	11	0
2	0	14	18	0
3	6	0	0	0
4	0	0	2	6
5	0	14	1	0
6	3	3	0	0
7	0	14	0	0
8	1	6	4	0
9	0	3	7	0
10	1	1	0	0
11	0	0	0	0
12	0	1	4	1
13	1	1	0	0
14	0	2	1	0

4 Discussion

Order-shift scheduling plan is generated on the basis of order-interval plan. There are just some dedicate differences of the total cost between 1st step model and 2nd step model; order-shift scheduling plan emphasizes on the processing deadline so that there is no need to arrange staff when one kind of the order have been delivered to the customer. While some staff assigned to the shift cannot process the order before the orders arrive and they just make production preparation, for example debugging the computer or the work platform and rapid response to the orders arrived earlier. From the aspect of the company, the orders can also be postponed to the next shift with more staffs to finish them. We should consider this problem to decrease staffs' idle time and deliver the orders in time.

5 Conclusion

The multilevel, multiskill labor staffing and scheduling model is proposed from the aspect of the data processing company. The effectiveness of the model is confirmed by the empirical data from the company ICP. The main contributions of the paper are

- A. The efficient operation system with low cost is very important for data processing companies. The core part is labor staffing and scheduling. Few studies focus on this kind of problem which the companies are eager to solve. This paper just makes up for the blank.
- B. The two-step model is proposed under the condition of deterministic demand with multilevel and multiskill. It can be reduced to linear programming model which is finally solved by IBM ILOG CPLEX Optimization Studio12.3
- C. The practical data collected from the company is used for the model. The result proves the validity of the scheduling model.

However, the model in this paper just considers two level processes. Actually the processes of data processing are more than two steps. Further research is needed to extend the model for more levels and balance the labor assignment between intervals and shifts.

Acknowledgment This research is supported by the National Natural Science Foundation of China (Grant no. 71071126). The authors are grateful for the referees whose suggestions let to an improvement of an earlier version of this paper.

References

1. Yin Guopeng, Yang Bo (2010) The capabilities assessment model for offshore IT outsourcing vendors in China: an empirical investigation at Beijing. *Stud Sci Sci* 28(3):388–396
2. Hsiao CT, Pai JY, Chiu H (2009) The study on the outsourcing of Taiwan’s hospitals: a questionnaire survey research. *BMC Health Serv Res* 9:78
3. Diana ML (2009) Exploring information system outsourcing in US hospital-based health care delivery systems. *Health Care Manage Sci* 12(4):434–450
4. Yang Longfeng, Wang Tieshan, Wu Yanqin (2009) The decision model and strategy choice of financial service outsourcing for commercial bank. *Sci Technol Manage Res* 6:366–369
5. Clark GL, Monk AHB (2013) The scope of financial institutions: in-sourcing, outsourcing and off-shoring. *J Econ Geogr* 13(2):279–298
6. Zhao Guohui (2008) The study of the insurance service outsourcing problem. *China Insur* 9:49–52
7. Gotzamani K, Longinidis P, Vouzas F (2010) The logistics services outsourcing dilemma: quality management and financial performance perspectives. *Supply Chain Manage Int J* 15(6):438–453
8. McIvor R, Humphreys P, Mckittrick A, Wall T (2009) Performance management and the outsourcing process: lessons from a financial services organization. *Int J Oper Prod Manage* 29(10):1025–1048
9. Beretta L, Aldrovandi V, Grandi E et al (2007) Improving the quality of data entry in a low-budget head injury database. *Acta Neurochir (Wien)* 149:903–909
10. Davide M (2008) Double-blind control of the data manager doesn’t have any impact on data entry reliability and should be considered as an avoidable cost. *Control Clin Trials* 18(66):432–450
11. Tom W (2009) Automated versus manual audit in the emergency department. *Control Clin Trials* 15:351–366
12. Dantzig GB (1954) A comment on Edie’s ‘Traffic delays at toll booths’. *J Oper Res Soc Am* 2(3):339–341
13. Edie LC (1954) Traffic delays at toll booths. *Oper Res* 2(2):107–137
14. Cezik MT, Ecuver PL’ (2008) Staffing multiskill call centers via linear programming and simulation. *Manag Sci* 54(2):107–137
15. Avranmidis AN, Chan W, Gendreau M, Ecuver PL’, Pisacane O (2010) Optimizing daily agent scheduling in multiskill call center. *Eur J Oper Res* 200(2):822–832