2011 International Conference on Innovation, Management and Service
IPEDR vol.14(2011) © (2011) IACSIT Press, Singapore

Organisation Staffing Optimisation Using Deterministic Reneging Queuing Model

Trong B. Tran[1], Steven R. Davis and David G. Carmichael

[1,2,3] School of Civil and Environmental Engineering, The University of New South Wales

Abstract. The effect of staffing levels on organisation performance is examined through applying the M/M/c-D reneging queuing model. The paper presents a review of organisation staffing and outlines experiments used to validate the theory. Staffing levels are examined from an output or production orientation in situations where a moderate number of staff functions as a pool of workers that perform similar operations. Both the experimental results and the model indicate that there is an optimum level of staffing which maximises the 'profit' of an organisation.

Keywords: overstaffing, understaffing, human resource management, queuing model, reneging

1. Introduction

Selecting the appropriate number of staff is an important issue for organisations. This problem has occupied the minds of researchers for a long time. However, the issue of what is the appropriate level of staffing is still being debated in the literature. Some people advocate moderate understaffing because they believe that workers in this setting produce more. Others claim that moderate overstaffing has positive effects on an organisation's outcomes. It is hypothesised in this paper that overstaffing leads to improved organisation performance compared to understaffing. Organisation performance is here defined as organisation 'profit', which is taken as the difference between the outputs and inputs of an organisation. The present study tests this hypothesis by conducting experiments on different worker group sizes and applying the M/M/c-D reneging queuing model to analyse group performance.

The paper reviews the literature on organisation staffing in Section 2. In Section 3, the basic concepts and equations necessary to evaluate organisation performance through the M/M/c-D model are listed. The experiment is presented in Section 4, while Section 5 gives the comparison between the experimental results and the theory. In Section 6, optimisation is examined. Conclusions are given in Section 7.

2. A Review of Organisation Staffing

The choice of level of staffing in organisations has occupied the minds of researchers for a long time. However, there is still debate over the effect of overstaffing and understaffing on organisation performance. On the one hand, some people advocate a moderate level of understaffing, arguing that this leads to increased organisation output per head (Barker and Gump, 1964; Bechtel, 1974; Greenberg, 1979; Vecchio and Sussmann, 1981; Oxley and Barrera, 1984). Others argue that slight overstaffing leads to improved organisation performance (Galbraith, 1973; Bourgeois, 1981; Cheng and Kesner, 1997; Tan, 2003; George, 2005).

Supporters of understaffing argue that workers are likely to work more efficiently (Barker and Gump, 1964) and experience higher motivation (Bechtel, 1974; Greenberg, 1979; Vecchio and Sussmann, 1981); hence they produce more. In understaffed groups, Barker and Gump (1964) point out that each person engages in a wider variety of tasks, expends greater effort to achieve organisation goals and takes on more

[1] Corresponding author. Trong B. Tran Tel: +61 2 9385 4290; fax: +612 9385 6139;
Email: tran.trong@student.unsw.edu.au

responsibility. Oxley and Barrera (1984) and Greenberg (1979) suggest that worker motivation may be improved with slight understaffing. Echoing this point, Bechtel (1974) and Vecchio and Sussmann (1981) conduct empirical studies and show that employee levels and worker motivation have a curvilinear relationship. Their studies conclude that moderate understaffing and worker satisfaction are related.

Supporters of overstaffing emphasise that extra staff can be used to support an organisation in absorbing any environmental turbulence (Bourgeois, 1981; Cheng and Kesner, 1997), facilitate organisation strategic directions (Hambrick and Snow, 1977; Thompson, 1967), and enhance organisation performance (Rust and Katz, 2002). The existence of extra staff in an organisation allows the organisation to experience new postures in relation to a changing environment. Hence the organisation is more likely to support special projects (Bourgeois, 1981). Cheng and Kesner (1997) state that organisations with additional staff are more likely to respond aggressively to shifting environmental demands. Having human resource slack in an organisation is argued to enable creative behaviour and experimentation with new strategies, such as introducing new products or entering new markets (Thompson, 1967; Hambrick and Snow, 1977). Knight (1967) supports this point and shows that new processes, new products, and new ideas follow when adding staff to an organisation. Researchers acknowledge that increasing staff levels for an organisation means more cost. Hence they suggest that an organisation should not overstaff greatly (Bourgeois, 1981; Tan, 2003; George, 2005). George (2005) and Tan (2003) conduct studies and confirm the existence of an optimal level of human resource slack in an organisation. In summary, additional staff can be a source of competitive advantage, but too many staff may reduce organisation performance (Tan, 2003). The present paper looks at staffing levels solely from an output or production orientation. It assumes that a moderate number of staff acts as a pool of workers that perform essentially the same operations.

3. Modelling

Reneging queuing models are used in evaluating systems with impatient customers (Barrer, 1957; Bocquet, 2005). This paper applies the M/M/c-D reneging queuing model, where customers lose patience after a deterministic time τ, and leave the queue without being served (Worthington, 2009). Such customers are lost to the system. For a general discussion on applying queuing models, see Carmichael (1987).

Symbols used in following are:

- $1/\lambda$ average interarrival time of customers; λ is the average arrival rate.
- $1/\mu$ average service time per customer; μ is the average service rate.
- ρ servicing factor; $\rho = \lambda/\mu$.
- c number of servers.
- τ maximum time between arrival and start of service that a customer will wait before reneging.
- P_0 probability that there are zero customers in the system at a particular time.
- P_L probability that a customer will not be offered service before time τ and hence renege.
- P_S probability of service; $P_S = 1 - P_L$
- Θ system output (customers per unit time).
- B_p benefit to the organisation of serving one customer
- C_w cost of one worker
- N_w number of workers
- F 'profit'

Several parameters may be considered when measuring system performance: probability of no customers in the system P0; probability of losing a customer PL; probability of service PS; output of the system Θ (Bocquet, 2005). The relevant formulae for these follow.

- The probability of no customers in the system, P_0, is obtained from (Barrer, 1957, p. 654)

$$P_0^{-1} = \begin{cases} \sum_{k=0}^{c} \frac{\rho^k}{k!} + \frac{\rho^{c+1}}{(\rho-c)c!}\left(e^{\mu\tau(\rho-c)}-1\right) & (\rho \neq c) \\ \sum_{k=0}^{c} \frac{\rho^k}{k!} + \frac{\mu\tau\rho^{c+1}}{c!} & (\rho = c) \end{cases} \qquad (1)$$

- The probability of losing customers after time τ (Barrer, 1957, p. 654; Bocquet, 2005, p. 3)

$$P_L = \begin{cases} \dfrac{\frac{\rho^c}{c!} e^{-\mu\tau(c-\rho)}}{\sum_{k=0}^{c-1} \frac{\rho^k}{k!} + \frac{\rho^c}{c!} \frac{\rho e^{-\mu\tau(c-\rho)} - c}{\rho - c}} & (\rho \neq c) \\ \dfrac{\frac{c^c}{c!}}{\sum_{k=0}^{c-1} \frac{c^k}{k!} + \frac{c^c}{c!}(1 + c\mu\tau)} & (\rho = c) \end{cases} \quad (2)$$

- The probability of service

$$P_S = 1 - P_L \quad (3)$$

- The output of the system

$$\Theta = \lambda P_S \quad (4)$$

- Organisation 'profit' is considered to be made up of two components, namely cost associated with the workers and benefit associated with the output (Carmichael, 1987).

$$F = B_p\Theta + C_w N_w \quad (5)$$

Given Bp, Θ, and Cw, the optimisation problem becomes one of finding the number of workers Nw that maximises the 'profit' F.

4. Experiments

4.1. Experiment Design

Experiments were conducted on work groups where workers independently disassembled electronic components from circuit boards. The number of workers in experimental groups ranged from 8 to 12.

Interarrival times of incoming work were regulated to follow an exponential probability distribution, so as to simulate variable work input. Only one piece of work arrived at a time. Each piece of arriving work was served by a single randomly selected worker. In the case that all workers were busy at the time that work arrived three different scenarios were tested. In the first experiment the work left the system immediately. In the second experiment it waited to be served for up to 1 second before leaving. In the third experiment it waited for up to 2 seconds before leaving. Work leaving was regarded as lost work.

4.2. Data collection and data validation

Data collected from the experiments included arrival times, waiting times (if applicable), proposed starting times of service, real starting times of service, service times, and finishing times of service. In the experiments for 8-, 9- and 10-worker groups, the experiments stopped when the number of lost customers reached 100. The 11- and 12-worker group experiments stopped when this number reached 25. After being stopped, the experiments restarted from scratch.

4.3. Results and analysis

The average interarrival time (1/λ) and the average service time (1/μ) respectively in the three experiments were (in seconds): 5.04, 26.92; 5.07, 27.71; and 5.12, 28.08. As a result the servicing factors, $\rho = \lambda/\mu$, are respectively for the three experiments: 5.34; 5.47; 5.48. This indicates that the optimum number of workers would be 5 or 6 if there was no variability in work input.

The experimental results illustrate what might be anticipated, namely that the probability of service and the output increase when the number of workers increases. Adding an extra worker to a smaller group has a bigger effect than adding an extra worker to an already large group. For example, in the third experiment when the group was enlarged from 8 to 9 workers, the probability of service increased by nearly 6% and the output increased by more than 50 per hour. However, when the number of workers changed from 10 to 11, the increase in probability of service was approximately 2% and the output increased by about 10 per hour. The probability of service and hourly output are plotted in Figures 1 and 2.

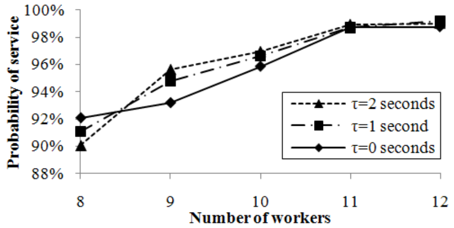

Figure 1. Probability of service of experimental groups.

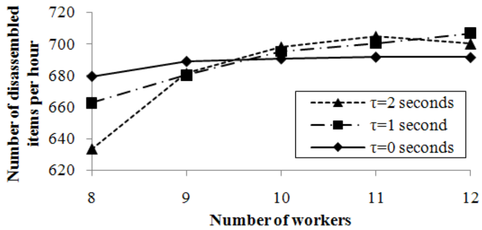

Figure 2. Hourly output of experimental groups.

5. Comparison of the Experiment and the Model

Given the arrival rate of work λ and the service rate μ, Equations (1) to (4) can be used to calculate P0, PL, PS, and Θ. Figures 3 and 4 show the comparison between the experimental results and the queuing model results for PS and Θ for the $\tau = 1$ second experiment. The results from the other experiments show similar trends.

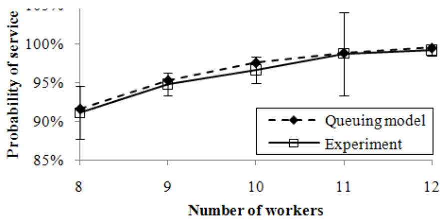

Figure 3: Comparing probability of service between the experiment and the model for groups in the $\tau = 1$ second experiment. Error bars refer to experimental results.

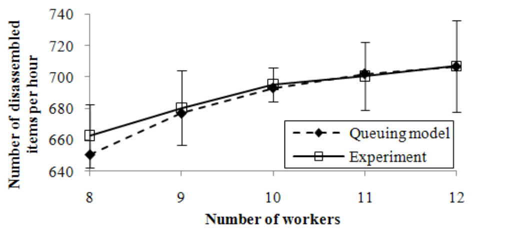

Figure 4: Comparing hourly output between the experiment and the model for groups in the $\tau = 1$ second experiment. Error bars refer to experimental results.

Correlation coefficients were used to measure how well the queuing model fits the experimental results for both the probability of service and hourly output. Statistical tests for all three experiments produced correlation coefficients in the range 0.91 to 0.997. The p-values were all less than 3% for both parameters implying good agreement with the experimental results at the 95% confidence level.

6. Optimising Staffing Level

When adding workers to a group, the probability of service and the output of the group increase. However, the cost of adding workers may be in excess of the benefits that the additional workers bring to the organisation. The optimisation study in this section aims at finding the number of workers in a group that bring the maximum 'profit' to the group.

The output Θ can be calculated from Equation (4), while Bp and Cw can be estimated from the market. Figure 4 presents results where the number of workers in the groups range from 1 to 20, $\lambda = 0.1972$, $\mu = 0.0361$, $\tau = 1$, Bp= \$0.8 and Cw= –\$25 (the negative sign here indicates that the group needs to outlay money to hire workers). λ, μ and τ are the same values as were used in the second experiment above. Cw was assessed from the labour market for appropriate semi-skilled workers. Bp was estimated as the average price of the electrical components on the second hand market. Figure 5 confirms that there is a size for which the group gains the maximum benefit and beyond that point, the 'profit' will decrease. In this case the optimum number of workers is 8. This is quite different to the value of 5 or 6 that would be derived from the servicing factor, as discussed in Section 4c, for the case of no variability in work input.

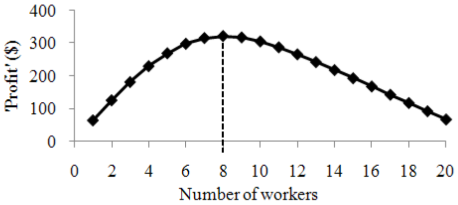

Figure 5: 'Profit' of groups using base case values

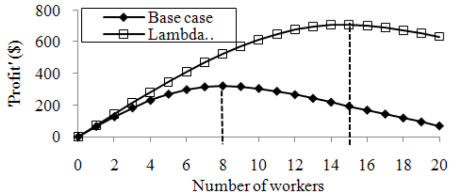

Figure 6: Sensitivity of 'profit' of groups to doubling λ.

A sensitivity analysis was carried out by varying each of the variables λ, μ, τ, Bp and Cw separately to observe the effect on the optimum number of workers. The example in the previous paragraph is used as the base case. Figure 6 shows, for example, the effect of doubling λ. In this case the optimum number of workers increases from 8 to 15 and the maximum 'profit' correspondingly nearly doubles.

Figure 7 shows the sensitivity diagram relating the five variables to the optimum number of workers. The optimum number of workers is most sensitive to λ and least sensitive to τ. The optimum number of workers is increased by increasing λ or Bp, or by decreasing μ or Cw. However, it is unaffected by changing τ.

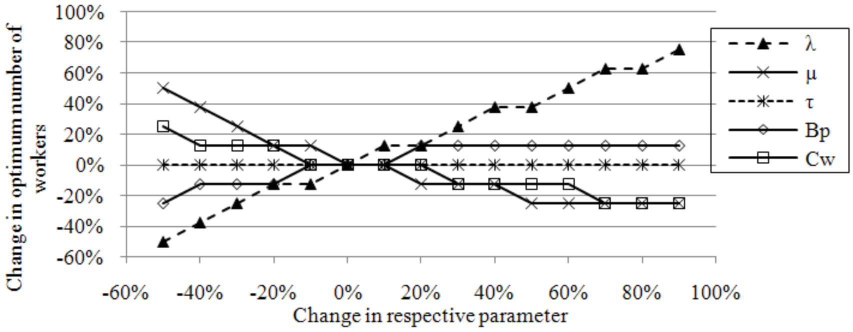

Figure 7. Sensitivity diagram.

7. Findings and Conclusions

The study shows that the M/M/c-D model can be used to describe the probability of service and output of working groups when two conditions occur simultaneously. Firstly, the arrival rate of work obeys an exponential distribution. Secondly, the maximum time that a customer will wait between arrival and start of service before reneging is a fixed value. The model may be used to optimise the staffing levels of work groups.

Results from the research also point out that a group should not be greatly overstaffed. Increasing the number of workers creates more output and hence brings more benefits for the organisation. However, the cost of additional workers is constant, while the benefits that each worker brings decreases as the number of workers increases. An optimisation study is recommended for establishing the appropriate number of employees in an organisation.

8. References

[1] Barker, R. G. and Gump, P. V. (1964). Big School, Small School: High School Size and Student Behavior: Stanford University Press.

[2] Barrer, D. (1957). Queuing with impatient customers and ordered service. Operations Research, 5(5), 650-656.

[3] Bechtel, R. B. (1974). The undermanned environment: A universal theory? In D. H. Carson (Ed.), Man environment interactions (Part VIII). Stroudsburg, Penn.: Dowden, Hutchinson and Ross.

[4] Bocquet, S. (2005). Queuing Theory with Reneging. Victoria: Defence Science and Technology Organisation.

[5] Bourgeois, L. J. (1981). On the Measurement of Organizational Slack. The Academy of Management Review, 6(1), 29-39.

[6] Carmichael, D. G. (1987). Engineering queues in construction and mining. Chichester: Ellis Horwood Limited.

[7] Cheng, J. L. C. and Kesner, I. F. (1997). Organizational Slack and Response to Environmental Shifts: The Impact of Resource Allocation Patterns. Journal of Management, 23(1), 1-18.

[8] Galbraith, J. R. (1973). Designing complex organizations. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

[9] George, G. (2005). Slack Resources and the Performance of Privately Held Firms. Academy of Management Journal, 48(4), 661-676.

[10] Greenberg, C. I. (1979). Toward an integration of ecological psychology and industrial psychology: Undermanning theory, organization size, and job enrichment. Journal of Nonverbal Behavior, 3(4), 228-242.

[11] Hambrick, D. C. and Snow, C. C. (1977). A contextual model of strategic decision making in organizations. Academy of Management Proceedings, 37, 109-112.

[12] Knight, K. E. (1967). A Descriptive Model of the Intra-Firm Innovation Process. The Journal of Business, 40(4), 478-496.

[13] Oxley, D. and Barrera, J. M. (1984). Undermanning Theory and the Workplace - Implication of setting size for job satisfaction and social support. Environment and Behavior, 16(2), 211-234.

[14] Rust, K. G. and Katz, J. P. (2002). Organizational Slack and Performance: The Interactive Role of Workforce Changes. Retrieved from http://www.midwestacademy.org/Proceedings/2002/papers/Rust.doc.

[15] Tan, J. (2003). Curvilinear Relationship Between Organizational Slack and Firm Performance: Evidence from Chinese State Enterprises. European Management Journal, 21(6), 740-749.

[16] Thompson, J. D. (1967). Organisations in action. New York: McGraw-Hill.

[17] Vecchio, R. P. and Sussmann, M. (1981). Staffing sufficiency and job enrichment: Support for an optimal level theory. Journal of Occupational Behaviour, 2(3), 177-187.

[18] Worthington, D. (2009). Reflections on queue modelling from the last 50 years. Journal of the Operational Research Society, 60, S83-S92.