CSCanada
Management Science and Engineering
Vol. 5, No. 3, 2011, pp. 149-154
DOI:10.3968/j.mse.1913035X20110503.1z244

ISSN 1913-0341[Print]
ISSN 1913-035X[Online]
www.cscanada.net
www.cscanada.org

Logistics Service Level Improvement Research and Demonstration Based on Queuing Theory

SONG Zhilan[1,*]; LIU Yueyi[1]; LI Kexian[1]

[1]School of Business, Yunnan University of Finance and Economics, Kunming, 50221, China
[*]Corresponding author.
Associate professor, mainly engaged logistic management and engineering.
Address: School of Business, Yunnan University of Finance and Economics, Kunming, 50221, China
Email: ksecszhl@vip.sina.com

Received 20 June 2011; accepted 27 July 2011

Abstract
The paper analyzes the problem of reducing Logistics service level caused by long waiting time in warehouse based on queuing theory and combined with cases. Then proposes three solutions: changing way of queuing, M/M/1 model to optimize service speed, M/M/n model to optimize storage channel. The paper also puts forward some measures and suggestions to make up the shortage of lacking experience.

Key words: Warehousing operations; Queuing theory; Service levels

SONG Zhilan, LIU Yueyi, & LI Kexian (2011). Logistics Service Level Improvement Research and Demonstration Based on Queuing Theory. *Management Science and Engineering, 5*(3), 149-154. Available from: URL: http://www.cscanada.net/index.php/mse/article/view/j.mse.1913035X20110503.1z244 DOI: http://dx.doi.org/10.3968/j.mse.1913035X20110503.1z244

INTRODUCTION

It is common to queue in production operation, such as in warehouse, distribution center and other logistics areas. Especially, with the development of logistics outsourcing and building of logistics centers, cargoes scatter into logistics warehouses rapidly. And with expanding business volume, the phenomenon of queuing too long in entering and out of warehouse, of long time to wait and of customer losing has been appeared. No matter which phenomenon happens, it will damage the image of logistics service providers and lead to a decline in company benefits ultimately. So, whether or not providing fast service has become one of important competitive advantages of logistics service providers.

In this paper, the meanings of studying the problem of queuing in logistics center based on the theory of queuing are following:

(1) It helps to coordinate contradictions between logistics service cost and customers waiting cost so that it can improve the level of designing of service system and achieve a reasonable allocation. It is related to: ①The allocation of equipment and personnel. Considering the factors of ways to reach customer, service, equipment and personnel service, it is wise to increase the number of equipment and person to improve service level, reduce queuing time and attract more customer; ②Principles of priority of queuing. Arrange proper way for customer to queue in various situations. And adopt principles of justice to improve customer satisfaction.

(2) From the point of view of customer, solving the problem of queuing is effective for bettering customer service and saving cost. It can both supply better service in a low cost and increase customer loyalty to enhance competitiveness in the market.

1. THEORY ANALYSIS

Queuing theory is also called theory of stochastic service system, which is established by Danish Engineer Erlang (A.K.Erlang) in 1909 when he was studying telephone system. Nowadays, the theory has been using in varieties areas, such as production industry and so on.

1.1 The Problems Queuing Theory Research

(1) Performance problem. It studies probability laws of various queuing system. The main problems to study are

the distribution of customer, customer waiting time and arrival distribution methods, which include two cases, random and fixed.

(2) Improvement problem. Use queuing theory to design logistics service system and better operation efficiency.

(3) Model of queuing system. It is a system to determine whether the model meets system or not.

1.2 The Composition of Queuing System

A general queuing system is composed of customer source, waiting before service, accepting service and leaving after service, est.

1.3 The Constraints of Queuing Theory and Assumptions

(1) Constraints

Logistics service providers must understand the constraints of queuing theory to analysis service demand and manage queuing problems. The constraints are:

①In normal circumstance, customers' distributions of arrival method and arrival time are uneven. In order to research conveniently, the paper assumes that the arrival rate of customer is balance.

②Each channel's service rate in per hour is not the same, which is as the result of different equipment performance and personnel operation. But the paper argues that the service level of each channel is the same.

③Different patient levels and opinions make different results. Someone may leave at once when he knows it takes a long time to wait, and the others may wait. In this paper, it assumes that all customers would like to wait for a long time.

(2) Hypothesis

Through author's survey in a company, put forward hypothesis of storage operation. Following are hypothesis:

①Customer resource in storage operation is unlimited, and each customer arrives independent, which obeys to Poisson distribution.

②Customer has no special preference to storage channel. Each channel provides same service.

③The operation complies with the principle of first come first serve, while no customers leave. It is considered that the way of queuing is a single system.

④The service efficiency of each channel is random which obeys Exponential distribution and there is no difference between each channel service.

⑤Take storage process as a whole, then the efficiency of it can be bettered by designing more channels.

1.4 Models of Queuing

Establish queuing models based on characteristics of queuing system's composition. Generally speaking, there have five models, M/M/1, M/M/n, M/D/1, M/M/1/m/∞ and M/M/n/m/∞. In practice, models, M/M/1 and M/M/n are commonly used. In this paper, mainly apply these two models.

1.4.1 Characters of Models

Characters of model M/M/1: It is a single-channel but single phase. The manners customer reach obey to Poisson distribution while service time obeys to negative exponential distribution.

Characters of model M/M/n: It is a multi-channel but single-stage. The manners customer reach obey to Poisson distribution while service time obeys to negative exponential distribution. There is no limitation to the number of customers. If all channels are busy, customers will wait in a queue. Meanwhile, customers have no special preference to each service channel.

1.4.2 Formulas of M/M/1 and M/M/n to Solve

(1) Formula of M/M/1

The model can be used to search solution to a single-channel but single-stage service system. The indexes include service intensity, idle probability of service system, average number of customer waiting time for service, average number of customer, average waiting time per customer took and average time of stay. Follows are specific formulas.

①Service intensity

$$p = \frac{y}{u} \tag{1}$$

P - Service intensity;
y - The speed of arrival in per unit time;
u - The average service rate of each desk.

②Idle probability of service system

$$P_0 = 1 - P \tag{2}$$

P_0 - Probability of no customer in line.

③Average number of customer waiting for service

$$L_q = \frac{y^2}{u(u-y)} \tag{3}$$

L_q - Average number of customers waiting for service.

④Average number of customer

$$L_s = \frac{y}{(u-y)} \tag{4}$$

L_s - The average number of waiting customer service in system, which not only includes the customers waiting in line, but also includes serving customers.

⑤Average waiting time per customer took

$$W_q = W_s - \frac{1}{u} \tag{5}$$

W_q - The average waiting time for per customer.

⑥Average time of stay

$$W_s = \frac{1}{u-y} \tag{6}$$

W_s - The average time of stay of clients.

(2) Formula of M/M/n

The model can be used to search solutions to a single-stage but multi-channel service system. The indexes include service strength, idle probability of service system, average number of customer waiting for service, average number of customer, average waiting time per customer took and average time of stay. Follows are specific formulas.

①Service intensity

$$p = \frac{y}{un} \tag{7}$$

P - System service intensity;
y - The speed of arrival in per unit time;
u - The average service rate of each desk.
n - Quantity of channels.

②Idle probability of service system

$$p_o = \left[\sum_{k=0}^{n-1} \left(\frac{y}{u}\right)^k \frac{1}{k!} + \left(\frac{y}{u}\right)^n \frac{1}{n!(1-\frac{y}{un})} \right]^{-1} \tag{8}$$

P_0 - The probability of no customers in a single-stage but multi-channel service system.

③Average number of customer waiting for service

$$L_q = yu\left(\frac{y}{u}\right)^n \frac{p_o}{(n-1)!(nu-y)^2} \tag{9}$$

L_q - Average number of customers for service;

④Average number of customer

$$L_s = L_q + \frac{y}{u} \tag{10}$$

L_s - The average number of waiting customer service in system not only includes the customers waiting in line, but also includes serving customers.

⑤Average waiting time per customer took

$$W_q = \frac{L_q}{y} \tag{11}$$

W_q - The average waiting time per customer.

⑥Average time of stay

$$W_s = \frac{L_s}{y} \tag{12}$$

W_s -The average time of stay of clients.

1.4.3 Optimization of Models

It is impossible to eliminate queues in reality, because the service levels and service costs are between the games. If we want to eliminate queues, we need to use a large number of service personnel and equipment that would increase the cost of service. However, if the efficiency of service equipment and service staff cannot meet requirements of customers, it will produce queues, too. Therefore, it is wise to use queuing theory to determine a reasonable level of service to minimize the total cost. Service cost is an increasing function of level of service, while the function of waiting for service level cost is decreasing. The specific functions are shown in Figure 1.

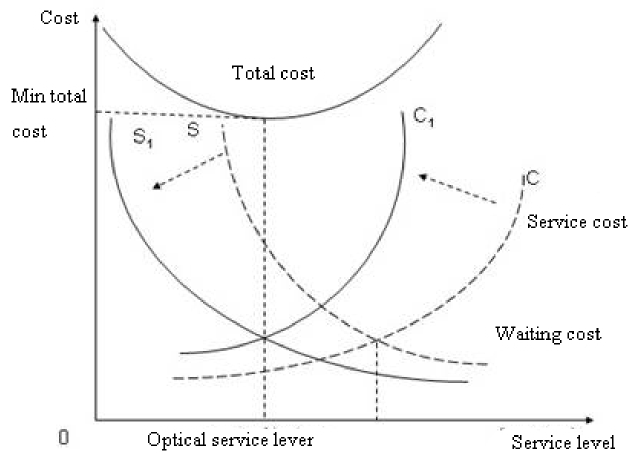

Figure1
The Relationship Between Cost and Service Level

Time becomes more important for customers. Figure1 shows that the customer waiting cost increases from to in reality. If improve service level or redesign the system, the waiting time will shorten and service costs will decrease from to . In order to reach the best level of service, it is available to optimize the service system from the following two aspects.

(1) Optimization of service speed for M/M/1 model

M/M/1 model can improve the speed of equipment and personnel to solve the problem of waiting too long in queuing. When we use M/M/1 model to optimize the speed of service, we need to know service cost for per unit time of each channel and each waiting cost when is one. It is smart to use total cost to optimize service system.

The formula of total cost is:

$$f(u) = C_2u + SL_s(u) \tag{13}$$

$f(u)$ - Total cost in per unit time;
C_2 -service cost of each service channel in per unit of time when u is equal to one;
S - Each customer’s waiting cost in the per unit time;
$L_s(u)$ - The average number of customer in service system.

The formula (4) is:

$$L_s = \frac{y}{(u-y)}$$

So the function of total cost in per unit time $f(u)$ is:

$$f(u) = C_2u + S\frac{y}{(u-y)} \tag{14}$$

To find the minimum, suppose that $\frac{df(u)}{du} = 0$. The derivation may:

$$\frac{df(u)}{du} = C_2 - \frac{Sy}{(u-y)^2} = 0$$

After simplified, it becomes:

$$u^* = y + \sqrt{\frac{S}{C_2 y}} \quad (15)$$

As $\frac{d^2 f(u)}{du^2} = \frac{2Sy}{(u-y)^3}$, and when $u > y$, it is possible to optimize the formula, $\frac{2Sy}{(u-y)^3} > 0$.

According to analysis, when the level of service is proper, the total system cost is lowest.

(2) Optimization of service channel for M/M/n model

It is possible to change the quantity of channel to better the service level and decrease waiting time. When use M/M/n model to optimize the service channel, it needs to know the service cost of each system in per unit time, waiting cost of each customer and average number of customer. Then apply M/M/n model to calculate the total cost.

The formula is:

$$F(n) = Cn + SL_s(n) \quad (16)$$

$F(n)$ - Total cost in per unit time;
C - Service costs in per unit time of each service system;
S - Each customer's waiting cost in the per unit time;
$L_s(n)$ - The average number of customers.

Then use the marginal analysis method to minimize the total cost of service system.

$$F(n) \le F(n-1)$$
$$F(n) \le F(n+1) \quad (17)$$

Union formula (16) and (17)

$$Cn + SL_s(n) \le C(n-1) + SL_s(n-1)$$
$$Cn + SL_s(n) \le C(n+1) + SL_s(n+1) \quad (18)$$

Be simplified:

$$L_s(n) - L_s(n+1) \le \frac{C}{S} \le L_s(n-1) - L_s(n) \quad (19)$$

Finally, it is easy to get the number of channel, n(n=1, 2, 3…) based on knowing the difference between $L_s(n)$ and $L_s(n+1)$ as well as the arrangement $\frac{C}{S}$.

On theory, it can shorten customer waiting time by determining the best service levels and the quantity of channels. However, the method of queuing must be change in practice.

2. EXAMPLE OF WAREHOUSE STORAGE

This study sets a coffee company as an example. Storage has a greater impact on the company.

2.1 The Status of Warehouse Storage Operations

(1) Process of storage

The process includes carrying, inspection, procedures for storage and inventory.

(2)Situation of customer or goods arrival

Time to harvest coffee is generally form December to June, but the heavy time of entering storage is in March and April. The company has three accesses and operations of inventory are provided from 8 a.m to 11 a.m so that customers always arrive at the same time. Table 1 shows collected data.

Table 1
Statistical of the Number of Customers to Reach in March and April

Data	Quantity	Data	Quantity	Data	Quantity	Data	Quantity	Data	Quantity	Week
3-1	9	3-15	0	3-29	14	4-12	3	4-26	9	Mon.
3-2	8	3-16	10	3-30	15	4-13	1	4-27	7	Tues.
3-3	10	3-17	12	3-31	15	4-14	5	4-28	6	Wed.
3-4	6	3-18	11	4-1	8	4-15	14	4-29	12	Thur.
3-5	2	3-19	14	4-2	4	4-16	13	4-30	9	Fri.
3-6	0	3-20	18	4-3	0	4-17	0			Sat.
3-7	0	3-21	0	4-4	0	4-18	0			Sun.
3-8	4	3-22	4	4-5	9	4-19	12			Mon.
3-9	5	3-23	2	4-6	11	4-20	8			Tues.
3-10	9	3-24	9	4-7	12	4-21	8			Wed.
3-11	13	3-25	12	4-8	10	4-22	11			Thur.
3-12	14	3-26	11	4-9	14	4-23	10			Fri.
3-13	0	3-27	0	4-10	0	4-24	0			Sat.
3-14	0	3-28	0	4-11	0	4-25	0			Sun.
Subtotal	80		103		112		85		43	
Total						423				

It is easy to observe that the number of customer is opposite completely in Saturday and Sunday. In order to study the question easily, we do not consider the above situation.

Using formula (20) can calculate the average number of customer arrival.

$$a = A_1 / B_1 \quad (20)$$

a - The average number of customer arrival in one hour;

A_1 - The number of all customers (not including weekends);
B_1 - The total time of service.
a = (423-18)/(45×3)=3

(3)Service rate of entering storage channel
Most of operations of entering storage are manual, whose operation efficiency is low.

Table 2
Statistic of Service Time of Entering Storage in March and April

No.	Time (min.)	No.	Time (min.)	No.	Time (min.)	No.	Time (min.)
1	50	11	90	21	38	31	35
2	30	12	30	22	44	32	52
3	40	13	40	23	48	33	46
4	55	14	36	24	52	34	40
5	35	15	50	25	32	35	40
6	46	16	80	26	28	36	51
7	60	17	42	27	54	37	39
8	45	18	35	28	58	38	48
9	48	19	25	29	60	39	49
10	54	20	50	30	60	40	30
Subtotal	463	Subtotal	478	Subtotal	474	Subtotal	430
Total					1845		

Based on the data, use formula (21) can calculate the average number of clients of each channel.

$$b = A_2 / B_2 \tag{20}$$

b - The number of serviced customer in per hour;
A_2 - One hour;
B_2 - The average time of each customer to be provided.
b=60/ (1845/40) =1.3

2.2 The Solution to Shorten Entering Storage Waiting Time

For these problems, take use of queuing theory to improve storage service.

According to the actual analysis of collected data, customer reaches in every 3 hour but there are only three channels in warehouse. If assume that customers stay in three channels and do not be allowed to change queue, the arrival speed of customer for each channel is one in per hour, and the average rate of channel service is 1.3 in per hour. So we can get following conclusions: service cost of each channel is ¥200; Service cost of each customer in per hour is ¥153.85 and waiting cost in per hour is ¥400.

To shorten the queuing time of storage, following data are used.

C =200 ¥/ h; S = 400 ¥ / h; C_2 = 153.85 ¥/h; y = 3.

In three M/M/1 queuing models, y = 1 and u = 1.3.

(1) Better ways of queuing

The company has three M/M/1 queuing model currently which can be changed into one. M/M/3 queuing model is formed by 3 channels and do not be allowed to change team. If the rate of customer arrival is one in per hour, it needs three M/M/1 channel. But in M/M/3 queuing model, there only needs one queue, which has three services windows. Calculate solutions to the corresponding index based on model M/M/1 and model M/ M / n. The results are shown in Table 3.

Table 3
Comparison Between M/M/3 Model and M/M/1 in Queuing Way

Indexes	M/M/3 model	Three M/M/1 models
Rate of idle P_0	0.067	0.23(for one system)
Average number of customers waiting for service L_q	1.988	2.56×3
The average number of waiting customer L_s	4.298	3.33×3
Average time to stay W_s	85.86 minutes	199.8 minutes
Average time to wait W_q	39.76 minutes	153.6 minutes

According to Table 3, when three M/M/1 queuing models are changed into M/M/3 queuing model, every parameters reduce by several times which introduce that model M/M/3 can shorten the queuing time as well as improve services level.

(2)Improve speed of service

Without enlarging the storage channels, it is possible to improve service speed by increasing the speed of equipment and staff. Through the optimization formula (15) we can calculate the optimal service rate u^*.

$$u^* = y + \sqrt{\frac{S}{C_2 y}} = 1 + \sqrt{\frac{400}{153.85}} = 2.61$$

Service speed before and after are compared in Table 4.

Table 4
Service Speed Before and After

Indexes	u =1.3	u =2.61
Rate of idle P_0	0.23(for one system)	0.62(for on system)
Average number of customers waiting for service L_q	2.56×3	0.24×3
The average number of waiting customer L_s	3.33×3	0.62×3
Average time to stay W_s	199.8 minutes	37.2 minutes
Average time to wait W_q	153.6 minutes	14.4minutes

(3) Optimize storage channels

On perspective of cost, it is best design to make the total of customer waiting cost and storage services minimum. Calculate solutions to n by formula (16), (17), (18) of M / M / n model:

$$L_s(n) - L_s(n+1) \le \frac{C}{S} \le L_s(n-1) - L_s(n)$$

$p = \frac{y}{nu} < 1$, namely $p = \frac{3}{1.3n} < 1$. So the solution is n > 2.31. We can also calculate the corresponding value of n combining with M/M/n model's relevant formulas and the results are shown in Table 5.

Table 5
Values to *n*

n	Ls(n)	Ls(n)-Ls(n+1)	Ls(n-1)-Ls(n)	F(n)
3	4.298	1.785		2119.2
4	2.513	0.164	1.785	1805.2
5	2.349	0.035	0.164	1939.2
6	2.314		0.035	2125.6

According to the formula, $L_S(4) - L_S(5) \le \frac{C}{S} \le L_S(3) - L_S(4)$, we can calculate the quantity of channel. The service indexes of 4 channels are indicated in table 6.

Table 6
Service Indexes of 4 Channels

Indexes	Four channels
Rate of idle P_0	0.051
Average number of customers waiting for service L_q	0.203
The average number of waiting customer L_s	2.513
Average time to stay W_s	50.4 minutes
Average time to wait W_q	4.2 minutes

2.3 Analysis and Comparison of Three Solutions

2.3.1 Comparison of Service Levels

The indexes of service level of three different optimization solutions are indicated in Table 7.

Table7
Indexes of Service Level of Three Different Optimization Solutions

Indexes	Four channels	Speed better	M/M/3 medol
Rate of idle P_0	0.051	0.62(for one system)	0.067
Average number of customers waiting for service L_q	0.203	0.24×3	1.988
The average number of waiting customer L_s	2.513	0.62×3	4.298
Average time to stay W_s	50.4 minutes	37.2 minutes×3	85.86minutes
Average time to wait W_q	4.2 minutes	14.4 minutes×3	39.76minutes

After comparing three optimization solutions in Table 7, it is best solution to set 4 channels to improve service speed.

3.3.2 Comparison of Total Cost

The total cost of storage of current single-channel in per unit time

$F(n) = Cn + SL_s(n) = 200 \times 1 + 400 \times 3.33 = 1532$

The total cost of three channels is 4596.

(2)The total cost of improving service speed of a single M/M/1 model in per unit time

$f(u) = C_2u + SL_s(u) = 153.85 \times 2.61 + 400 \times 0.62 = 649.55$

The total cost of three channels is 1948.65.

(3) The total cost of changed M/M/3 and M/M/4 model in per unit time

$F(n) = Cn + SL_s(n)$

$F(3) = 200 \times 3 + 400 \times 4.298 = 2119.2$

$F(4) = 200 \times 4 + 400 \times 2.513 = 1805.2$

Through analysis and comparison, it can be seen that setting four storage channels is the best solution.

CONCLUSION

With the rapid increasing in logistics services, customers have higher demand of time. So whether can provide fast service or not becomes a competitive advantage. Waiting too long will make customer dissatisfied. The paper use queuing theory to study the scheduling problem of warehouse to find the most proper quantity of entering storage channels that can improve service efficiency. At the same time, the paper introduces queuing models o address the problem of storage operations by collected data in survey. The quantitative approach makes up the weakness of judging by experience and put forward some measures and suggestions.

REFERENCES

[1] JIN Fang, FANG Kai, & Wang Jinglin (2004). AGV Dispatching Based on Queuing Theory Studies. *Science Instrument, S1*.

[2] ZHAO Yuan, & DING Wenying (2008). Determine a Reasonable Amount of Crane Based on Queuing Theory. *Logistics Technology, 8*.

[3] SHEN Jian, & WANG Dong (2009). Analyses for RFID Garment Manufacturing on Queuing Theory. *Computer Information, 5*.

[4] FENG Lingge, & WANG Hao (2006). Operation Management. *Press of Capital Economics University, 5*.

[5] HU Liege. (2007). *Operations Research of Logistics*. Beijing: Press of China Transportation.