Designing Production, new model to Optimal Buffer Procedure of Chinese SMEs

Wen Huaid
Hangzhou Normal University, Hangzhou, 310012, China

Abstract
Small-medium sized enterprises (SMEs) contribute 60% of China’s industrial output and create 80% of China’s jobs. But for the year just past, Chinese SMEs have been experiencing hardships: some have been at risk of collapse. Shortages of electricity, capital and labour have led them to this predicament, and soaring costs have made things worse. Faced with this indictment of faltering growth, the Chinese government’s 12th Five-Year Plan contains a key strategy specifically in support of SMEs. According to the plan, the total number of China’s SMEs will grow steadily over the next five years and achieve an average annual growth rate of 8%. Five primary missions underpin the plan: to improve the capacity of establishing business and creating jobs; to optimise the structure of SMEs; to boost the development of “new, distinctive, specialised and sophisticated” industries and industrial clusters; to upgrade enterprise management; and to refine SME support systems. So what opportunities could this prospective economic about-turn hold for investors and foreign businesses looking to penetrate the Chinese SME market?

Keywords: Optimization, Buffer size, M/M/1/k system and blocking probabilities

Introduction
The analysis and design of production lines is one of the oldest problems in manufacturing (Sevastyanov 1962; Gershwin 1987; Buzacott and Shanthikumar 1993; Heavey 1993; Papadopoulos et. al. 1993; Papadopoulos and Karagiannis 2001). Focus of research is often how to improve efficiency, productivity and/or reduce idle time (Gershwin and Schor 2000; Jeong and Kim 2000; Han and Park 2002; Paik et. al. 2002; Aksoy and Gupta 2005; Sabuncuoglu et. al. 2006; Shaaban and Hudson 2009; Amiri and Mohtashami 2012). Considering the large cost associated with the production lines, a slight improvement in efficiency can lead to significant savings over the life of line (Gershwin and Schor 2000; Khamisabadi et al. 2013; Sabuncuoglu et. al. 2006; Massim 2010; Amiri and Mohtashami 2012).

There are three major decision variables that are considered while designing the production lines. First one is the size of the buffers placed between successive workstations. Second major decision variable is the number of servers allocated to each workstation. Last decision variable is regarding the amount of workload allocated to each workstation. The optimization problem corresponding to size of the buffers placed between successive workstations of the lines is called as the Buffer Allocation Problem (BAP). This is major problem that the designer faces in a tandem production line. What should be the size of buffer and where it should be placed within the line is an important question because buffers can have a great impact on the efficiency of the production line.

This paper is focused on finding an optimal buffer size to be placed between two workstation of a production line. Increasing the buffer size decreases blocking and starving probability which increases throughput and hence profit of system. But in realistic manufacturing system, buffer size can’t be increased beyond floor size available to store work in process inventory. Also increase in buffer size also increases inventory and its associated cost. These costs decrease profit. Considering the tradeoff between costs and increased throughput solution procedure to this buffer optimization problem is formulated in next sections.
Literature Review

Many researchers have done extensive research in the area of design of production lines. The objective in the design of production lines is the optimization of its performance variables. Major performance variable are throughput, associated profit, average flow time, average waiting time, average inventory and system utilization. Performance is constrained by demand and production line characteristics. However size and location of buffer, number of servers per workstation and workload allocated per workstation can be designed to meet system performance objective within given constraints. Numerous models and solutions are developed in the literature for the unique combination of demand and production line characteristics. Here we briefly summarize these studies and further describe our objective in this paper.

Many studies shows that the buffer is key to governing the system performance. Shaaban and Hudson (2009) investigated the productivity problem under the condition of low idle time, average buffer level, and full factorial design. The results show that the buffer capacity is the key factor in determining the idle time and whole line’s average buffer level. Amiri and Mohtashami (2012) stated the productivity can be improved by adding the buffers to limit the propagation of distributions but simultaneously increasing the costs. Massim Y et. al (2010) proposed that the appropriate size of buffer should be determined in order to improve the productivity as well as minimize the costs of manufacturing. Shaaban and Hudson (2009) found that the increased number of buffers will lead to the longer idle time, as well as, the improved buffer capacity will increase the average buffer level but decrease the idle time. Demir et. al (2012) explored the buffer allocation problems in unreliable and non-homogeneous production lines in order to maximize the throughput. S.G. Powell (1994) considers the problem of allocating buffer storage to unbalanced serial production line with three workstations to maximize steady state throughput. Singh and Smith (1997) analyzed problem of buffer allocation for an integer nonlinear network design problem. They utilized an efficient and effective search methodology to generate sub-optimal solutions to the problem for series, merge, and split topologies with M/M/C/K queuing stations. Though buffer design received the most attention in the literature (Buzacott 1967; Seong et. al. 1995; Lutz et. al. 1998; Seong et. al. 2000; Helber 2001; Nahas et. al. 2006; Dolgui et. al. 2007; Nahaset al. 2009; Vergara and Kim 2009; McNamara et. al. 2013), many studies have been done in the area of server allocation (Magazine and Stecke 1996) and workload allocation (Shanthikumar et.al.1991; Yamazaki, Yamazaki et. al. 1992; Hillier and Hillier 2006).

The optimization problem corresponding to buffer size design is called as the Buffer Allocation Problem (BAP). Buffers affect efficiency of production line. Efficiency loss in production line occurs when station is either blocked or starved. A workstation is blocked when service of an item is complete at the workstation but it cannot be moved in the buffer because it is full. In this situation blocked workstation has to wait till next workstation pick up an item from buffer creating place for item on blocked station to proceed further. On the other workstation is said to starving when it has completed processing an item and is moved to buffer next to it but there is no item in buffer before it to pick and start serving. Such workstation remains idle till workstation before it competes processing an item and move it to buffer between them. Focus of this paper is on BAP in manufacturing line consisting of two workstations separated by finite buffer space. Processing times of workstations are exponentially distributed and arrivals are considered as Poisson. This problem can be modeled as M/M/1/k queuing networks with blocking. Except in special cases, e.g. queuing networks with reversible routing under repetitive blocking (Spinellis and Papadopoulos 2000), closed queuing networks with blocking (Perros 1994) queuing networks with blocking do not have product form solution. Therefore exact solution can only be obtained by numerical techniques. This approach has been criticized given the time and efforts involved along with limitation on size of network. As a result considerable amount effort has been devoted for approximate solutions for queuing networks with blocking (Dallery and Frein, 1993; Perros, 1994). Given the recent development in computer and information technology drawbacks of numerical techniques can easily overcome. Thus an effective procedure to find optimal solution to BAP is developed in the next section.
Model

Consider the model depicted in Fig. 1. There are two workstation separated by finite buffer of size \( k \). \( a \) is arrival rate of items to upstream workstation (Workstation 1). \( b \) and \( c \) are the processing rate of upstream and downstream workstation (Workstation 2) respectively. Processing rates follow exponential distribution and arrivals follow passion distribution. There is a single server at each workstation. It is assumed that workstations are reliable, i.e., never undergoes breakdown. Downstream workstation is never blocked. There are no limitations on arrivals. This system can be modeled as \( M/M/1/k \) Markovian system.

Problem is to determine size of buffer \( k \) placed in between two workstations. \( k \) represents a finite positive number, as it is more realistic for manufacturing system because of inventory cost associated and floor space required for work in process inventory. This finite nature of buffer size \( k \) gives rise to blocking and starving of workstations. Calculating probabilities of blocking and starving enables to us to analyze its impact on efficiency of production line hence throughput. Throughput of manufacturing system governs net profit of and thus there is direct link between increasing size of buffer size and increased profit. On the other hand buffer storage is expensive due to its direct cost, and due to the increase of the work-in-process (WIP) inventories. Thus optimal buffer size should be determined to maximize profit.

![Diagram](image)

Methodology

Consider a simple case in above model where there is no buffer \( (k=0) \) in between two workstations. If an item at workstation 2 is undergoing processing and processing on an item at workstation 1 is completed then later must wait until processing is completed at workstation 2. Workstation 1 is said to be blocked and arrivals coming at workstation 1 are turned away. In addition to this situation, arrivals are also turned away if Workstation 1 is busy. This system undergoes five different probability states named as \( P_1, P_2, \ldots, P_5 \). Description of these states and diagrammatic representation is shown in Table 1 and Fig. 2 respectively.

<table>
<thead>
<tr>
<th>Probability States</th>
<th>( n_1, n_2 )</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_1 )</td>
<td>0,0</td>
<td>Empty system</td>
</tr>
<tr>
<td>( P_2 )</td>
<td>1,0</td>
<td>Machine 1 busy and machine 2 empty</td>
</tr>
<tr>
<td>( P_3 )</td>
<td>0,1</td>
<td>Machine 1 empty and machine 2 busy</td>
</tr>
<tr>
<td>( P_4 )</td>
<td>1,1</td>
<td>Machine 1 and 2 busy</td>
</tr>
<tr>
<td>( P_5 )</td>
<td>b,1</td>
<td>Machine 2 busy, machine 1 finished processing, item is waiting for machine 2 to become available, that is system is blocked</td>
</tr>
</tbody>
</table>
Figure 2. Probability States For Two Sequential Workstations Without Buffer

For this multidimensional Markov chain, following steady state equations can be written (Gross and Harris, 1985).

\[-a^*P_1+c^*P_3=0\]
\[-b^*P_3+a^*P_4+c^*P_5=0\]
\[-(a+c)^*P_3+b^*P_2+c^*P_5=0\]
\[-(b+c)^*P_2+a^*P_3=0\]
\[-c^*P_5+b^*P_4=0\]

Using boundary conditions we get \(P_1+P_2+P_3+P_4+P_5=1\)

When a buffer is introduced in between first and second workstation the number of probability states increases. At \(k = 1\) probability states are increased to 7. For this system seven steady state equations along with boundary condition can be written. With each unit increase in buffer size, two probability states are increased so are steady state equations. Probability \(P_i\) to \(P_{2k+5}\) for system with buffer size \(k\) are shown in Fig. 3.
Figure 3. Probability States For Two Sequential Workstations Without Buffer

Following $2k+5$ set of equations can be written for system with buffer size $k$.

\[-a*P_1+c*P_2=0\]
\[-b*P_2+a*P_1+c*P_3=0\]
\[-(a+c)*P_3+b*P_2+c*P_4=0\]
\[-(b+c)*P_4+a*P_3+c*P_5=0\]
\[-(a+c)*P_5+b*P_4+c*P_6=0\]
\[-(b+c)*P_6+a*P_5+c*P_7=0\]
\[-(a+c)*P_7+b*P_6+c*P_8=0\]
\[-(b+c)*P_8+a*P_7+c*P_9=0\]
\[-(a+c)*P_9+b*P_8+c*P_{10}=0\]

\[\vdots\]

\[-(b+c)*P_{2k+4}+a*P_{2k+5}=0\]
\[-(a+c)*P_{2k+5}+b*P_{2k+6}=0\]

Using boundary conditions we get $P_1+P_2+P_3+P_4+\ldots+P_{2k+4}+P_{2k+5}=1$

Above set of equation has to be solved to calculate probability of each state from $P_1$ to $P_{2k+5}$. This solution will be in terms of $k$ which is variable this analysis. Major complexity is generated due to the fact that number of equations to be solved and number of variables (probability states) to be determined are dependent on $k$. For such problems it is difficult to come up with one close form equation but analyzing the solutions obtained for probability states for varying values of $k$ may help
us to develop underlying pattern in the solution. Steady state equations for buffer sizes 0 to 4 are solved using Maple and solutions for probability states are given in appendix. Due to space limitations we could not include more solutions. After analyzing these solutions we can say that solution for any probability state for any value of $k$, can be represented in three parts—multiply coefficient term, remaining coefficient term and common denominator.

$$\text{Probability} = \text{multiplying coefficient for } a * \text{multiplying coefficient for } b * \text{multiplying coefficient for } c \text{remaining coefficient terms} / \text{common denominator}$$

$$P_j = A[i,j]*B[i,j]*C[i,j]*RC[i,j]/CD$$

Where $A$, $B$, $C$ are multiplying coefficient vector for $a$, $b$, $c$ respectively and $RC$ and $CD$ are the remaining coefficient term and common denominator. In general for buffer size $k$, multiplying coefficient terms are:

- For $a$: $\{1, a, a^2, a^3, a^4, ..., a^{k+2}, a^{k+3}, a^{k+4}\}$
- For $b$: $\{b, 1, b, b^2, b^3, ..., b^k, b^{k+1}, b^{k+2}\}$
- For $c$: $\{c^{k+2}, c^{k+3}, c^{k+4}, ..., c^4, c^3, c^2, c, 1\}$

Remaining coefficient term for $k=0$ consist of five rows. In order from top to bottom these rows are $b+c$, $a+b+c$, $b+c$, $1$ and $1$. For $k=1$, in addition to these 5 rows, two more new rows are added to top. These rows are $ac+(b+c)^2$ and $a^2+(b+c)^2+a(b+2c)$. With increase in unit value of $k$, there is a pattern of two new rows adding to the top. There exists a close relation between first two rows of each set. These first two rows from third set onwards can be computed using general formula as follows.

- row 1 of set $i = (b+c)*(\text{row 1 of set } i-1) + a*c*(\text{row 2 of set } i-1)$
- row 2 of set $i = (\text{row 1 of set } i-1) + a*(\text{row 2 of set } i-1)$

Since all probabilities sum to one, common denominator is a sum of numerators of all probability states. This denominator can be calculated as:

$$CD= \sum A[i,j]*B[i,j]*C[i,j]*RC[i,j]/\text{for } j = 1 \text{ to } 2k+5$$

Thus probability state for any size of buffer can be calculated using above logic.

**Optimal Size Of Buffer**

In this section equation for net profit of system is developed taking into consideration throughput and associated costs with buffer. The throughput of system increases with decrease in blocking and starvation probabilities. From analytical results computed in earlier section blocking and starvation probabilities decreases with increase in buffers size. But in realistic manufacturing system, buffer size can’t be increased beyond floor size available to store work in process inventory. Also increase in buffer size increases inventory holding cost and storage facility cost. This inventory cost decreases net profit. These costs are subtracted from expected gains obtained from throughput to find net profit of system. Optimal size of buffer is found when net profit of system is highest.

**5.1 Throughput and Expected gain from the system**

Throughput is defined as rate at which unit are manufactured through the system, i.e., rate at which unit is leaving the last server of manufacturing line. To calculate throughput it is sufficient to know actual output of last server. Thus throughput of the system is:

$$\text{Throughput}=\text{Probability that last server is busy} / \text{Mean time to produce single unit by last server}$$

$$\text{Throughput} = (1 - P_1 - P_2) / (1/c) = c * (1 - P_1 - P_2)$$

Expected gain or revenue per unit time from the system is directly proportional to throughput of the system. It is given by $G(k)$, where $m$ is fixed gain in releasing a part from system

$$G(k) = m * c * (1 - P_1 - P_2)$$
5.2 Costs associated with the buffer

There are two types of costs associated with buffer and work in process inventory. First cost is average inventory holding cost which is proportional to amount of inventory present in buffer per unit time (Papadopoulos et al. 1993). This cost for all \( k=0 \) is given by \( C_i(k) \), where \( c_i \) is fix inventory carrying cost per part per unit time

\[
C_i(k) = c_i (P_5 + P_6) + (k-1)(P_7 + P_8) + \ldots + (1)(P_{2k+3}) + P_{2k+4} + P_{2k+5} 
\]

Second type of cost is average storage facility cost which is directly proportional to size of buffer (Papadopoulos et al. 1993). It is given by \( C_i(k) \), where \( c_i \) is fixed cost per part per unit time of maintaining the inventory storage facility

\[
C_i(k) = c_i k
\]

5.3 Expression for Net Profit

Net profit \( NP(k) \), per unit time from the system is expected gain per unit time minus associated cost per unit time (Papadopoulos et al. 1993).

For \( k = 0 \)

\[
NP(k) = G(k) = m^* c^* (1 - P_1 - P_2)
\]

For \( k > 0 \)

\[
NP(k) = G(k) - C_i(k) - C_j(k) \\
NP(k) = m^* c^* (1 - P_1 - P_2) - c_i ((k)(P_5 + P_6) + (k-1)(P_7 + P_8) + \ldots + (1)(P_{2k+3}) + P_{2k+4} + P_{2k+5}) - c_2 k
\]

Above equation can be used to find the optimal size of buffer. Calculate profit of system for \( k=0, k=1 \), and so on. If net profit is decreasing it means best solution is not to have any buffer between two workstations. If net profit is increasing then keep on computing profit for higher values of \( k \) till value of \( k \) with highest profit is reached. If maximum size of \( k \) due to space limitations is reached before hitting maximum profit point then best solution is to have buffer with maximum possible size. This buffer should be placed in front of the production capacity limiting machine.

Conclusions

This paper addresses the problem of optimal buffer allocation in production line with two workstations separated by finite buffer space. Increasing the buffer size decreases blocking and starving probability which increases throughput and hence profit of system. On the other hand increase in buffer size also increases inventory and its associated cost reducing the profit. This tradeoff is considered to find optimal buffer size.

REFERENCES