2010 IEEE International Conference on Robotics and Automation
Anchorage Convention District
May 3-8, 2010, Anchorage, Alaska, USA

Performance Approximation and Bottleneck Identification in Re-entrant Lines

Yang Liu, Jingshan Li and Shu-Yin Chiang

Abstract— **In this paper, we study a re-entrant line with unreliable exponential machines and finite buffers, operating under last buffer first serve scheduling policy. First, an approximation method is presented to estimate the throughput of the re-entrant line. Then, a system approach to identify bottleneck based on blockage and starvation information is proposed. It has been shown that the approximation method results in acceptable accuracy, and the bottleneck identification method could correctly detect the bottleneck in most cases.**

Keywords: Re-entrant lines, throughput, c-bottleneck.

I. INTRODUCTION

In many manufacturing systems, in order to reduce the overall fixed asset or property, plant and equipment (PP&E) investment, and to increase equipment utilization, re-entrant lines are often used. In a re-entrant line, parts need to visit these expensive machines more than once. For example, in semiconductor manufacturing industry, multiple layers of material need to be imprinted on the wafer layer by layer [1]. In automotive manufacturing, the ignition components need to be kept clean during the production. Thus, they will be cleaned in the washers multiple times, and such operations can be carried out in centralized washers to save space and reduce cost, which results in re-entrant operations [2]. Due to the widespread applications in many manufacturing sectors, in particular, the semiconductor and electronics manufacturing, the modeling, analysis and design of re-entrant lines are of significant importance.

Although throughput analysis of production systems has been studied extensively (see recent review [3]), the study on performance analysis of re-entrant lines is still limited. Most of the research effort devoted to the re-entrant lines has focused on scheduling and control policies, less work is devoted to performance analysis of re-entrant lines, in particular, with unreliable machines and finite buffers.

Queueing network model, fluid model and mean value analysis have been widely used in re-entrant line study [4]-[6]. Stochastic petri-net provides another method for analytical study of re-entrant lines [7]-[9]. However, such models typically either ignore the blocking nature of the system by assuming infinite buffers, or is limited to smaller systems due to computation intensity. Although discrete event simulation can provide accurate analysis [10]-[12], the long model development time and long simulation time restrict its application. Moreover, it is difficult to obtain insight through simulations.

In another direction, systems with similar features to re-entrant lines, such as multiple part types, rework loops, etc., have been studied in recent years [13]-[18]. However, to extend the results to make it applicable to re-entrant lines, substantial effort is still required. Therefore, there is still a need to develop an accurate and effective method to analyze the performance of re-entrant lines with unreliable machines and finite buffers [3], [19]. Paper [2] introduces an approximation method for re-entrant lines with Bernoulli reliability machines. Studying more general cases, e.g., exponential machine reliability models, is necessary.

Bottleneck identification and mitigation are the most effective ways for continuous improvement in manufacturing. Quick and accurate identification of bottlenecks based on collected data is of significant importance for re-entrant line operations. However, such an important issue has not been addressed in the current literature. A system-theoretic method to identify bottlenecks in serial and assembly systems by measuring and comparing blockages and starvations has been developed and successfully applied on the factory floor [20]-[24]. In this paper, we seek a similar method to identify bottlenecks in re-entrant lines.

The goal of this study is to develop an approximation method to evaluate the performance of re-entrant lines with unreliable exponential machines and finite buffers, operating under last buffer first serve (LBFS) scheduling policy, and propose an effective method for bottleneck identification in such lines using the measured data on the factory floor.

The remaining of the paper is structured as follows: Section II formulates the problem. The performance analysis and bottleneck identification methods are introduced in Sections III and IV, respectively. Conclusions are given in Section V.

II. PROBLEM FORMULATION

The structure of a re-entrant line under consideration is illustrated in Figure 1, where the circles represent the

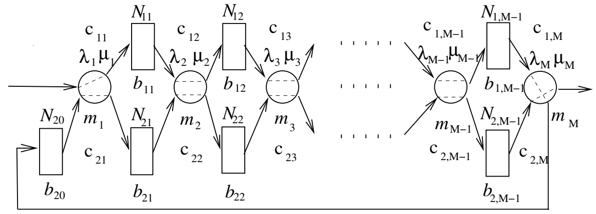

Fig. 1. Asynchronous re-entrant line

Y. Liu and J. Li are with the Department of Electrical and Computer Engineering and Center for Manufacturing, University of Kentucky, Lexington, KY 40506, USA. yang.liu@uky.edu, jingshan@engr.uky.edu

S.-Y. Chiang is with the Department of Information and Telecommunication Engineering, Ming Chuan University, Taoyuan 333, Taiwan, ROC. sychiang@mail.mcu.edu.tw

This work is partially supported by NSF Grant No. CMMI-0727691.

Please send all correspondence to Prof. Jingshan Li.

machines and the rectangles are the buffers. The machines, buffers, and their interactions are addressed by the following assumptions.

1) The system consists of M machines and $2M-1$ buffers. The first time jobs are processed at machines m_i, $i=1,\ldots,M$, and stored in buffers b_{1i}, $i=1,\ldots,M-1$. After first time processing at m_M, all jobs are sent to buffer b_{20}, waiting for second time processing. Then the jobs are reprocessed at m_1 through m_M, but stored in b_{2i}, $i=1,\ldots,M-1$. Jobs leave the system after being processed at m_M for the second time.
2) The capacity (or speed) of machine m_i is $c_{1,i}$ for the first time jobs and $c_{2,i}$ for the second time jobs, $i=1,\ldots,M$.
3) The up- and downtime of machine m_i are exponentially distributed random variables with parameters λ_i and μ_i respectively, $i=1,\ldots,M$. In other words, λ_i and μ_i are the failure and repair rates, and $1/\lambda_i$ and $1/\mu_i$ are the average up- and downtime of machine m_i, respectively.
4) Each buffer b_k, $k=11$, 12, …, $(1,M-1)$, 20, 21, 22, …, $(2,M-1)$, has finite capacity N_k, $0<N_k<\infty$.
5) Machine m_i, $i=1,\ldots,M-1$, is blocked by the first (respectively, second) time job at time t if it is up, buffer b_{1i} (respectively, b_{2i}) is full and machine m_{i+1} does not take a job from it at time t. Machine m_M is blocked by the first time job at time t if it is up, buffer b_{20} is full and machine m_1 does not take a job from b_{20} at time t. Machine m_M is never blocked by the second time job.
6) Machine m_i, $i=2,\ldots,M$, is starved at time t if it is up, both buffers $b_{1,i-1}$ and $b_{2,i-1}$ are empty, and their upstream machines fail to send any job into them at time t. Machine m_1 is never starved by the first time job.
7) The second time jobs have higher priorities than the first time ones, i.e., when it is up, machine m_i, $i=1,\ldots,M-1$, always takes job from buffer $b_{2,i-1}$ if it is not empty and m_i is not blocked by b_{2i}, otherwise it will take a part from buffer $b_{1,i-1}$ if it is not empty and m_i is not blocked by b_{1i}. Machine m_M will take a part from $b_{2,M-1}$ if it is not empty, otherwise m_M loads from $b_{1,M-1}$ if it is not empty and m_M is not blocked by buffer b_{20}. Such a priority loading policy is typically referred to as last buffer first serve policy [1].

Let TP be the throughput, i.e., the average number of finished (or second time) jobs by the last machine m_M per unit of time, of the re-entrant line. The problem addressed in this paper is formulated as follows:

Given the re-entrant production system 1)-7), develop a method to estimate the system throughput as a function of the system parameters, investigate system structural properties, and introduce a bottleneck identification method based on the measured data.

Solutions to the problem are presented in Sections III and IV below.

III. PERFORMANCE ANALYSIS

A. Idea of the Approach

The main difficulty of analyzing re-entrant line is that the jobs visit each machine more than once. Thus, the machine capacity needs to be allocated to multiple job processing by following the priority-based loading policy. In addition, it still has the complexity caused by blockage and starvation, which is commonly seen in other production lines. All these make the exact analysis of re-entrant lines impossible. Therefore, an iterative approximation method is developed to estimate the system throughput. The main idea of this method is to represent an M-machine re-entrant line in Figure 1 by a $2M$-machine serial line (see Figure 2 for illustration). The first M machines (denoted as m_1' to m_M') are dedicated for the first time jobs and the second M machines (denoted as m_1'' to m_M'') are for the second time jobs. Such an equivalence represents the part's flow path within the re-entrant line, i.e., a job begins with visiting m_1 through m_M for first time processing, then via buffer b_{20}, it returns to $m_1,\ldots,m_M$ for second time operations. The parameters of the machines (m_i' and m_i'', $i=1,\ldots,M$) need to be modified to take into account the nature of the shared processing of both the first and second time jobs.

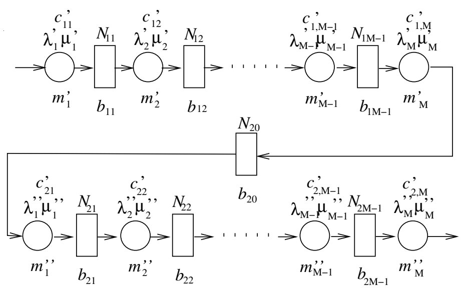

Fig. 2. Equivalent serial line

To obtained the parameters of the virtual machines m_i' and m_i'', $i=1,\ldots,M$, capacity allocation of machines m_i to the first and second time jobs needs to be estimated. Due to the high priority of second time jobs, it is equivalent to claim that m_i is ready to process second time jobs as long as it is up. A capacity is allocated to first time jobs only when it is up and cannot work on second time jobs. Thus, machine m_i'' can keep the original machine parameters, while m_i' needs to be modified to exclude processing on second time jobs. Thus, we reduce the production capacity of $c_{1,i}, i=1,\ldots,M-1$ by multiplying the probability that buffer $b_{2,i-1}$ is empty or buffer $b_{2,i}$ is full. Since the probabilities of buffers being empty or full are unknown, we introduce iterations. At the beginning of the iteration, these probabilities are initialized. Using a serial line analysis method, we analyze the $2M$-machine serial line with modified machine parameters and obtain the probabilities that buffers $b_{2,i-1}$ are empty, $i=1,\ldots,M$ and b_{2i} are full, $i=1,\ldots,M-1$. Then we introduce modifications on m_i', and carry out next iteration, and continue until the procedure is convergent. Finally, we obtain the estimated machine parameters and the system throughput.

B. Recursive Procedure

Let $TP_M(\cdot)$ denote the calculation of throughput for a serial line, $\widehat{ST}_i$ and $\widehat{BL}_i$ be the approximation of starvation and blockage probabilities, respectively. (due to space limitation, we omit the calculation of $TP_M(\cdot)$, a detailed description of it can be found in [24].) The estimation of $\widehat{ST}_i$ and $\widehat{BL}_i$ is obtained with the following argument: Consider machine m_i. If it is not starved, then a serial line consisting of m_i and all its downstream machines and buffers can be constructed and its throughput can be calculated. The main difference between this throughput and the throughput of the whole equivalent serial line ($2M$ machines) is mostly due to the starvation of m_i to buffer $b_{2,i-1}$. Similar rationale is applied to blockage estimation. A recursive procedure to approximate throughput of a re-entrant line is introduced as follows:

Procedure 1:

$$\begin{aligned}
\widehat{ST}_{2,i}(s+1) &= 1-\widehat{TP}(s)/TP_M(c_{2,i+1},\lambda_{i+1},\mu_{i+1},\\
&\quad \ldots,c_{2,M},\lambda_M,\mu_M,N_{2,i+1},\ldots,N_{2,M-1}),\\
&\quad i=0,\ldots,M-2,\\
\widehat{ST}_{2,M-1}(s+1) &= 1-\widehat{TP}(s)/(c_{2,M}\mu_M/(\lambda_M+\mu_M)),\\
\widehat{BL}_{2,i}(s+1) &= 1-\widehat{TP}(s)/TP_M(c'_{1,1}(s),\lambda_1,\mu_1,\\
&\quad \ldots,c'_{1,M}(s),\lambda_M,\mu_M,c_{2,1},\lambda_1,\mu_1,\\
&\quad \ldots,c_{2,i},\lambda_i,\mu_i,N_{1,1},\ldots,N_{1,M-1},\\
&\quad N_{2,0},N_{2,1},\ldots,N_{2,i-1}), \qquad (1)\\
&\quad i=1,\ldots,M-1,\\
c'_{1,i}(s+1) &= c_{1,i}[\widehat{ST}_{i,2}(s+1)+\widehat{BL}_{i,2}(s+1)\\
&\quad -\widehat{ST}_{i,2}(s+1)\widehat{BL}_{i,2}(s+1)],\\
&\quad i=1,\ldots,M,\\
c'_{1,M}(s+1) &= c_{1,M}\widehat{ST}_{2,M-1}(s+1),\\
\widehat{TP}(s+1) &= TP_M(c'_{1,1}(s+1),\lambda_1,\mu_1,\\
&\quad \ldots,c'_{1,M}(s+1),\lambda_M,\mu_M,\\
&\quad c_{2,1},\lambda_1,\mu_1,\ldots,c_{2,M},\lambda_M,\mu_M,),
\end{aligned}$$

with initial conditions

$$\widehat{ST}_{i,2}(0)\in(0,1),\quad \widehat{BL}_{i,2}(0)\in(0,1),\quad i=1,\ldots,M,$$

and s is the iteration number, s=0,1,2,....

Based on extensive numerical experiments, we formulate the following numerical fact:

Numerical Fact 1: Procedure 1 results in two convergent sequences for even and odd number of iterations, i.e.,

$$\lim_{s\to\infty}\widehat{TP}(s)=\begin{cases}\widehat{TP}_{even} & \text{if } s \text{ is even,}\\ \widehat{TP}_{odd} & \text{if } s \text{ is odd,}\end{cases} \qquad (2)$$

where $\widehat{TP}_{even}$ and $\widehat{TP}_{odd}$ are the throughput estimates for even and odd number of iterations, respectively.

Then, we select their average as the throughput estimation of the re-entrant line, denoted as $\widehat{TP}$,

$$\widehat{TP}=\frac{\widehat{TP}_{even}+\widehat{TP}_{odd}}{2}. \qquad (3)$$

In this way, an estimate of the throughput of a re-entrant line in steady state is obtained.

C. Accuracy

The accuracy of the approximation is investigated numerically. Specifically, we consider more than 100 re-entrant lines by randomly and equiprobably selecting machine and buffer parameters from the following sets:

$$\begin{aligned}
M &\in \{2,3,5,10,20,50\},\\
e_i &\in [0.75,0.95],\\
c_{j,i} &\in [1,2],\quad j=1,2, i=1,\ldots,M\\
T_{i,down} &= \frac{1}{\mu_i}\in[1,20],\quad i=1,\ldots,M, \qquad (4)\\
N_{ji} &\in \lfloor k\cdot\max\{T_{i,down},T_{i+1,down}\}\rfloor,\\
&\quad j=1,2, i=1,\ldots,M-1,\\
N_{20} &\in \lfloor k\cdot\max\{T_{M,down},T_{1,down}\}\rfloor,\\
k &\in [1,3],
\end{aligned}$$

where $\lfloor x\rfloor$ denote the largest integer less than or equal to x. The machine average uptime, $T_{i,up}=\frac{1}{\lambda_i}$, is calculated through average downtime, $T_{i,down}$, and isolated efficiency, $e_i=\frac{\mu_i}{\lambda_i+\mu_i}$, $i=1,\ldots,M$.

For each of these lines, both analytical method using Procedure 1 and simulation approach using *SIMUL8* ([25]) are pursued to evaluate system throughput. In each simulation, 5,000 cycles of warmup time are assumed, which is sufficient to guarantee that the steady state is reached, and the next 200,000 cycles are used for collecting steady state statistics. 20 replications are carried out to obtain the average throughput, with 95% confidence intervals typically less than ±0.001. Then, the differences between analytical and simulation results are evaluated as

$$\epsilon=\frac{\widehat{TP}-TP}{TP}\cdot 100\%, \qquad (5)$$

where TP and $\widehat{TP}$ are the throughputs obtained by simulation and recursive procedure, respectively.

The results of this investigation are illustrated in Figure 3. Table I provides the average and maximum errors in all experiments, and the percentages where errors are within 5 or 10%. It is shown that in average the discrepancy is roughly 2%. In more than 87% of cases, the difference is within 5%, with a few exceptions (about 2% of cases) going up to more than 10% (maximum less than 12%). Thus, we conclude that in all the cases we studied, Procedure 1 results in an acceptable accuracy for throughput estimation, and can be used for design and analysis of re-entrant lines.

TABLE I
ACCURACY FOR ASYNCHRONOUS RE-ENTRANT LINE

Average $\lvert\epsilon\rvert$	max $\lvert\epsilon\rvert$	$\lvert\epsilon\rvert<5\%$	$\lvert\epsilon\rvert<10\%$
2.3%	11.5%	87%	98%

D. Structural Properties

Using the method developed above, we investigate some system properties.

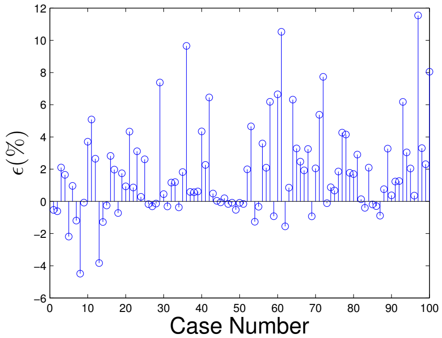

Fig. 3. Accuracy of Procedure 1

a) Monotonicity: It has been shown in [24] that monotonicity holds in serial lines and assembly systems, i.e., system throughput can be improved by increasing machine reliability and/or buffer capacity. Similar properties are observed for $\widehat{TP}$ in re-entrant lines as well.

Property 1: The system throughput, $\widehat{TP}$, defined by Procedure 1 and estimate (3), is monotonically increasing with respect to μ_i, $i = 1, \ldots, M$, and N_{ij}, $ij = 11, \ldots, (1, M-1), 20, 21, \ldots, (2, M-1)$, and decreasing with respect to λ_i, $i = 1, \ldots, M$.

b) Reversibility: Reversibility is also observed in serial production lines [24]. For re-entrant lines, let the line in Figure 4 be the reversed line of the one in Figure 1. Higher priority is still assigned to buffer b_{2i}, $i = 1, \ldots, M-1$. Let $\widehat{TP}$ and $\widehat{TP}^{rev}$ denote the throughputs obtained through Procedure 1 and estimate (3), respectively. Then identical throughputs are obtained.

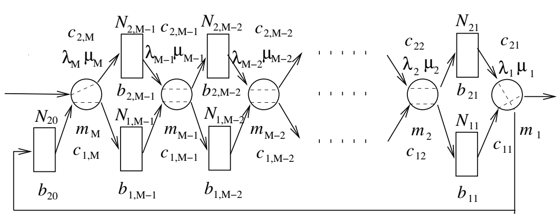

Fig. 4. Reversed re-entrant line

Property 2: The system throughput defined by Procedure 1 and estimate (3) satisfies

$$\widehat{TP} = \widehat{TP}^{rev}. \tag{6}$$

The above properties are observed based on investigations of $\widehat{TP}$ using recursive procedure 1, and have been verified by simulations.

IV. BOTTLENECK IDENTIFICATION

A bottleneck machine is the machine that impedes the system performance in the strongest manner. In other words, its improvement will lead to the largest improvement in system throughput comparing with improvement on other machines. Therefore, a bottleneck can be defined as the machine with the largest partial derivative with respect to one of its parameters. In particular, the c-bottleneck (c-BN), i.e., improvement with respect to machine capacity, is one of the most useful means in continuous improvement. Therefore, in this paper, we study c-BN in re-entrant lines.

Definition 1: Machine m_i, $i = 1, \ldots, M$, is the c-BN machine if

$$\frac{\partial TP}{\partial c_i} \geq \frac{\partial TP}{\partial c_j}, \quad \forall j \neq i. \tag{7}$$

Remark 1: Based on Definition **??**, a machine with the smallest c_i (or $c_i e_i$) may not be the c-BN. In some cases, even the best machine which has the largest c_i and $c_i e_i$can be the c-BN, whose improvement can lead to the largest improvement in system throughput (i.e., largest $\Delta TP/\Delta c_i$). Thus, a systematic method to identify c-BN is needed.

However, this definition could not lead to direct identification of c-BN. First, the partial derivatives of the throughput cannot be measured on the factory floor during normal operations; Second, in most cases these partial derivatives cannot be calculated analytically since even the throughput cannot be evaluated exactly. Therefore, indirect method is pursued to identify c-BN. Similar to the case of serial and assembly lines, we seek to identify the c-BN using blockage and starvation information. Since such measurement can be available on the factory floor, and does not require any knowledge of machine and buffer parameters, such a method can be easily applied in operations.

Consider a re-entrant line shown in Figure 5. The two rows under each machine indicate the starvation and blockage probabilities. (Note that here the blockage and starvation imply that the machine is blocked or starved by both first and second time jobs, respectively.) Then, by comparing the starvation probability with the upstream blockage probability, we assign arrows as follows:

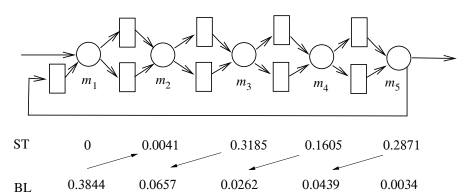

Fig. 5. c-BN identification in re-entrant line

Arrow Assignment Rule: If $BL_i > ST_{i+1}$, assign the arrow pointing to m_{i+1} from m_i. If $BL_i < ST_{i+1}$, assign the arrow pointing to m_i from m_{i+1}.

Then, there will exist a machine with no emanating arrow, this machine is the c-BN machine.

When there are more than one machine having no emanating arrows, we define the severity index S as follows:

$$\begin{aligned} S_j &= \mid ST_{j+1} - BL_j \mid + \mid BL_{j-1} - ST_j \mid, \\ &\qquad j = 2, \ldots, M-1, \\ S_1 &= \mid ST_2 - BL_1 \mid, \\ S_M &= \mid BL_{M-1} - ST_M \mid. \end{aligned} \tag{8}$$

Then the machine with the largest severity index is the primary c-BN.

Thus, a c-BN indicator can be formulated as follows:

c-BN Indicator: If there is a machine with no emanating arrow, it is the c-BN of the re-entrant line. If there are several machines with no emanating arrows, the primary c-BN is the one with the largest severity index.

An illustration of such identification is shown in Figure 6. The actual c-BN is calculated by evaluating the partial derivative ($\Delta TP/\Delta c$) using simulations. As one can see, machine m_2 has the largest partial derivative, and the bottleneck indicator also identifies it as a c-BN.

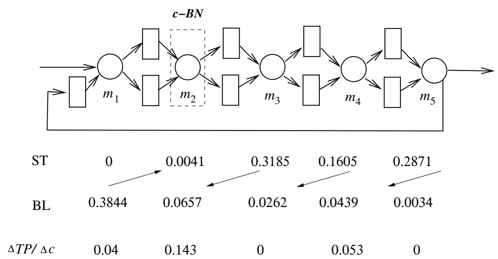

Fig. 6. c-BN identification: single bottleneck case

In the case of design phase or no available blockage and starvation information, recursive procedure 1 (assuming all buffer and machine information is available) can be used to calculate blockage and starvation probabilities of the equivalent serial line. Then a serial line bottleneck identification method ([20], [24], same as the one introduced here) is used to identify the c-BN. As shown in Figure 7, it also identifies m_2 as the c-BN.

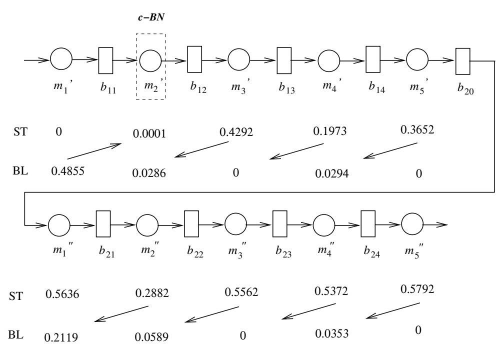

Fig. 7. c-BN identification: single bottleneck case using analytical approach

For multiple bottlenecks case, Figure 8 presents an example where both machines m_1 and m_3 having no emanating arrows. m_1 has severity index $S_1 = 0.1042$, and m_3 has $S_3 = 0.0869$. Thus, m_1 is the primary c-BN. Such an identification is also verified by calculating $\widehat{BL}_i$ and $\widehat{ST}_i$ of the equivalent serial line using procedure 1 and then identifying c-BN of the serial line (see Figure 9).

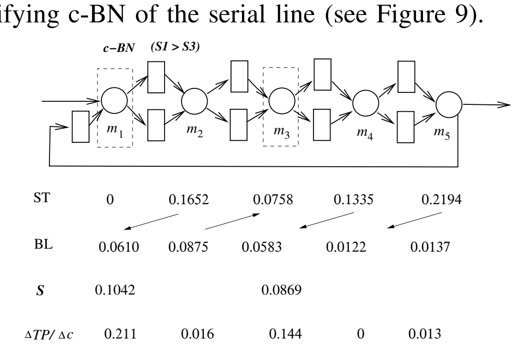

Fig. 8. c-BN identification: multiple bottlenecks case

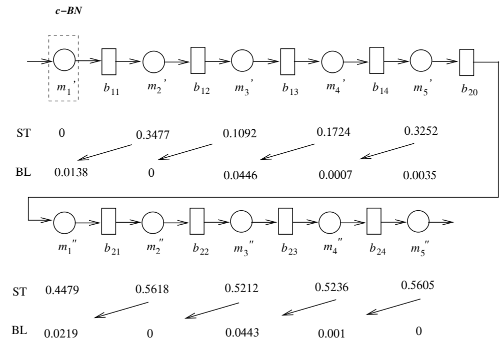

Fig. 9. c-BN identification: multiple bottlenecks case using analytical approach

To verify the correctness of this method, numerical experiments have been carried out. In most cases, we obtain correct identification. However, a few counter examples are discovered. One of them is shown in Figure 10, where the indicator suggests m_4 as the c-BN, but calculation of partial derivative through simulations points to m_5. However, as we can see, m_4 has the second largest partial derivative (which is quite close to that of m_5). This implies that it may lead to identification of the second bottleneck even though it points to an incorrect one. Similar scenario occurs for identification using procedure 1 where the bottleneck is also pointed to machine m_4 (Figure 11).

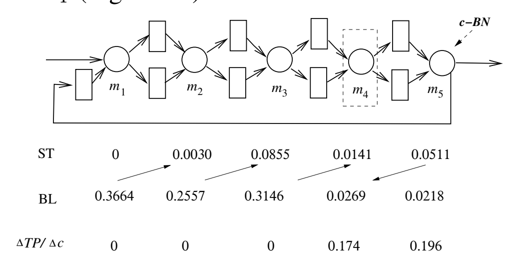

Fig. 10. c-BN identification: counter example in single bottleneck case

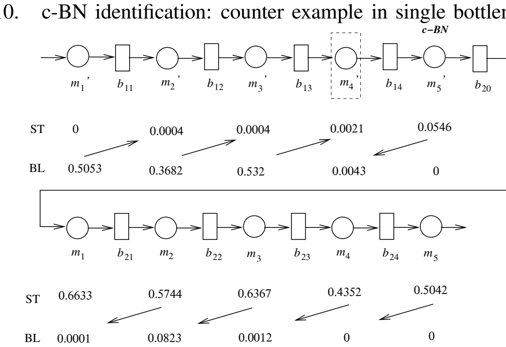

Fig. 11. c-BN identification: counter example in single bottleneck case using analytical approach

In the multiple bottleneck case, Figure 12 indicates that the set of machines without emanating arrows includes the actual c-BN, but the primary bottleneck is wrongly assigned. Both m_1 and m_3 are identified as bottlenecks, but severity index points to m_3 rather than the real bottleneck m_1. This is mainly due to that the severity index (8) is proposed by heuristics. Similar case is observed for identification using procedure 1 (Figure 13).

Among more than 100 lines we have tested, except for a few cases (about 2%), the c-BN indicator can return correct

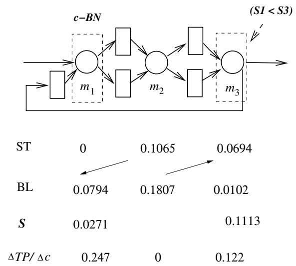

Fig. 12. c-BN identification: counter example in multiple bottlenecks case

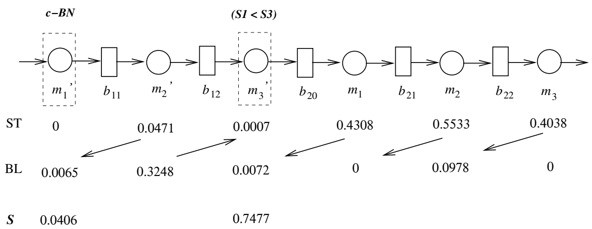

Fig. 13. c-BN identification: counter example in multiple bottlenecks case using analytical approach

identification for single bottleneck case. For all the cases with incorrect identifications, they all point to the second largest derivatives (Table II). In the multiple bottleneck case, the actual bottleneck is always within the bottleneck set identified by the indicator, about 70% of them the primary c-BN is selected correctly (Table III). In all of our experiments, the bottlenecks identified through blockage and starvation in re-entrant lines are always coincident with the bottlenecks suggested by the equivalent serial lines through recursive procedure 1.

TABLE II

C-BN IDENTIFICATION ACCURACY - SINGLE BOTTLENECK

Correct	97.7%
Incorrect but points to 2nd c-BN	100%

TABLE III

C-BN IDENTIFICATION ACCURACY - MULTIPLE BOTTLENECKS

Correct	69.2%
Incorrect but provides correct c-BN set	100%

Based on the above, we conclude that the c-BN indicator can be used as an effective tool for continuous improvement of re-entrant lines.

V. CONCLUSIONS

In this paper, we introduce a recursive procedure to estimate the throughput of re-entrant lines with asynchronous exponential machine reliability model. The idea of the approximation is to transform the M-machine re-entrant line into an equivalent $2M$-machine serial line. The estimates are of high accuracy. In addition, we present a bottleneck identification method based on starvation and blockage for re-entrant lines. In most cases, such a method leads to correct identification. These developments will provide production engineers and managers quantitative tools to design and improve systems with re-entrant operations.

REFERENCES

[1] P.R. Kumar, "Re-entrant Lines," *Queueing Sys.*, vol. 13, pp. 87-110, 1993.

[2] C. Wang and J. Li, "Approximate Analysis of Re-entrant Lines with Bernoulli Reliability Models," to appear in *IEEE Trans. on Autom. Sci. and Engineering*, 2010.

[3] J. Li, D.E. Blumenfeld, N. Huang and J.M. Alden, "Throughput Analysis in Production Systems: Recent Advances and Future Topics," *Int. J. of Prod. Res.*, vol. 47, pp. 3823-3851, 2009.

[4] J.G. Dai and G. Weiss, "Stability and Instability of Fluid Models for Certain Re-entrant Lines," *Math. of Oper. Res.*, vol. 21, pp. 115-134, 1996.

[5] Y. Narahari and L.M. Khan, "Performance Analysis of Scheduling Policies in Re-entrant Manufacturing Systems," *Comp. & Oper. Res.*, vol. 23, pp. 37-51, 1996.

[6] S. Kumar and P.R. Kumar, "Queueing Network Models in the Design and Analysis of Semiconductor Wafer Fabs," *IEEE Trans. on Robot. and Autom.*, vol. 17, pp. 548-561, 2001.

[7] M.C. Zhou and M.D. Jeng, "Modeling, Analysis, Simulation, Scheduling, and Control of Semiconductor Manufacturing Systems: A Petri Net Approach," *IEEE Trans. on Semicon. Manuf.*, vol. 11, pp. 333-357, 1998.

[8] M.D. Jeng, X.L. Xie and W.Y. Hung, "Markovian Timed Petri Nets for Performance Analysis of Semiconductor Manufacturing Systems," *IEEE Trans.s on Sys., Man and Cyber. – Part B*, vol. 30, pp. 757-771, 2000.

[9] J.Y. Choi and S.A. Reveliotis, "Relative Value Function Approximation for the Capacitated Re-entrant Line Scheduling Problem," *IEEE Trans. on Autom. Sci. and Eng.*, vol. 2, pp. 285-299, 2005.

[10] L.M. Wein, "Schedling Semiconductor Wafter Fabrication," *IEEE Trans. on Semicon. Manuf.*, vol. 1, pp. 115-130, 1988.

[11] S.H. Lu, D. Ramaswam and P.R. Kumar, "Efficient Scheduling Policies to Reduce Mean and Variance of Cycle-Time in Semiconductor Manufacturing Plants," *IEEE Trans. on Semicon. Manuf.*, vol. 7, pp. 374-388, 1994.

[12] B.-W. Hsieh, C.-H. Chen and S.-C. Chang, "Scheduling Semiconductor Wafter Fabrication by Using Ordinal Optimization-Based Simulation," *IEEE Trans. on Robot. and Autom.*, vol. 17, pp. 599-608, 2001.

[13] S. Helber, "Approximate Analysis of Unreliable Transfer Lines with Splits in the Flow of Material", *Annals of Oper. Res.*, vol. 93, pp. 217-243, 2000.

[14] J. Li, "Performance Analysis of Production Systems with Rework Loops", *IIE Trans.*, vol. 36, pp. 755-765, 2004.

[15] J. Li, "Throughput Analysis in Automotive Paint Shops: A Case Study", *IEEE Trans. on Autom. Science and Eng.*, vol. 1, pp. 90-98, 2004.

[16] J. Li, "Overlapping Decomposition: A System-Theoretic Method for Modeling and Analysis of Complex Production Systems," *IEEE Trans. on Autom. Sci. and Eng.*, vol. 2, pp. 40-53, 2005.

[17] J. Li and N. Huang, "Modeling and Analysis of a Multiple Product Manufacturing System with Split and Merge," *Int. J. of Prod. Res.*, vol. 43, pp. 4049-4066, 2005.

[18] M. Colledani, F. Gandola, A. Matta and T. Tolio, "Performance Evaluation of Linear and Non-linear Multip-Product Multi-Stage Lines with Unreliable Machines and Finite Homogeneous Buffers," *IIE Trans.*, vol. 40, pp. 612-626, 2008.

[19] J.G. Shanthikumar, S. Ding and M.T. Zhang, "Queueing Theory for Semiconductor Manufacturing Systems: A Survey and Open Problems," *IEEE Trans. on Autom. Sci. and Eng.*, vol. 4, pp. 513-522, 2007.

[20] C.-T. Kuo, J.T. Lim and S.M. Meerkov, "Bottlenecks in Serial Production Lines," *Math. Prob. in Eng.*, vol. 2, pp. 233-276, 1996.

[21] S.-Y. Chiang, C.-T. Kuo and S.M. Meerkov, "c-Bottlenecks in Serial Production Lines: Identification and Application," *Math. Prob. in Eng.*, vol. 7, pp. 543-578, 2001.

[22] S. Ching, S.M. Meerkov and L. Zhang, "Assembly Systems with Non-exponential Machines: Throughput and Bottlenecks," *Nonlinear Anal.: Theory, Methods & Appl.*, vol. 69, pp. 911-917, 2008.

[23] S. Biller, J. Li, S.M. Meerkov and L. Zhang, "Bottlenecks in Bernoulli Serial Lines with Rework," to appear in *IEEE Trans. on Autom. Sci. and Eng.*, 2009.

[24] J. Li and S.M. Meerkov, *Production Systems Engineering*, Springer, 2009.

[25] J.W. Haige and K.N. Paige, *Learning Simul8: the Complete Guide*, Plain Vu Publisher, 2001.