Proceedings of the 2008 Winter Simulation Conference
S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds.

A NEW METHOD FOR BOTTLENECK DETECTION

Sankar Sengupta
Industrial and Systems Engineering Department
654 Science and Engineering Bldg.
Oakland University
Rochester, MI 48306, USA

Kanchan Das
Technology Systems Department
245 Slay Hall
East Carolina University
Greenville, NC 27858, USA

Robert P. VanTil
Industrial and Systems Engineering Department
653 Science and Engineering Bldg.
Oakland University
Rochester, MI 48306, USA

ABSTRACT

This paper presents a new method to identify and rank the bottlenecks in a manufacturing system. The proposed method is based on performance related data that are easy to capture, offers low computational burden and less prone to be in error due to its simplicity. The proposed method analyzes inter-departure time from different machines to identify and rank the bottlenecks. In a follow-up paper the authors plan to present a method to allocate loss of production to different machines based on analysis of inter-departure time. This paper also proposes a set of rules that may be used to improve data integrity. The proposed method may be used to analyze both steady state as well as non-steady state data and can be extended easily to analysis of a job shop.

1 INTRODUCTION

Detecting bottlenecks in a production line is not a trivial task. The study of the problem comes under the heading of Optimum Production Technology, lately labeled as Theory of Constraints. The problem of detecting the bottleneck is further aggravated by the fact that the manufacturing lines are inherently variable. The cause of this variation might be due to random events (machine failure) or long term changes in the system (seasonal variation of demand, new product launch, machine load changes). A production system may have a dominant bottleneck which will appear as the bottleneck most often during analysis or it may have momentary bottlenecks that keep shifting in the system with time. Shifting of bottleneck happens due to random event such as machine failure. Literature on bottleneck analysis refers to both momentary and dominant (average) bottleneck. There are various methods available for detecting momentary and average bottlenecks. A partial list includes the following:

- Shifting bottleneck detection based on the duration of machine being active without interruption,
- Bottleneck detection based on utilization of machines, the machine with the highest utilization is considered bottleneck,
- The machine with the longest average up-stream queue length is considered bottleneck,
- Analytical approach based on estimation of the blockage and starvation probabilities of a machine,
- A graph theoretic approach to detect bottleneck in order to optimize the scheduling in a shifting bottleneck procedure.

2 LITERATURE REVIEW

This section presents a summary of some recent research findings on identification and ranking of bottlenecks in a manufacturing system.

The method proposed by the Toyota Research Laboratory (Roser, Nakano, and Tanaka 2001, 2002) specifically applies to a discrete event system. At any given time, it is made sure that all the machines are assigned to any one of the predefined discrete states. These states are then grouped into active and inactive categories. The duration for which machines were active without interruption by an inactive state is measured. Consecutive active states are considered as one active state. The machine with the longest average active period is considered to be the bottleneck as this machine is likely to be interrupted by other machines and is most likely to dictate the overall system throughput. Benefits of this method can be stated as follows: (1) No knowledge of the structure of the system or order of processing is needed, and (2) It can calculate the level of confidence of the solution recommended. Five distinct states are recognized for each machine. These are: Working, Waiting, Blocked, Tool Change, and Under Repair. For analysis, Waiting and Blocked were considered inactive. In general, with conventional bottleneck detection methods, it is hard to identify the distinct bottleneck, but with the largest duration of active period method, a clear distinction can be established. Using this method, both momentary and average bottlenecks can be located.

A rigorous mathematical approach is proposed by Meerkov and his colleagues (Chiang, Kuo, and Meerkov 1998, 2002; Kuo, Lim, and Meerkov 1996). The machine with the largest sensitivity of the system's performance index to this machine's production rate in isolation is defined as the bottleneck machine. Production rate is considered as the performance index of the system. Both empirical and analytical approaches are presented for estimating blockage and starvation probabilities of different machines. A graph theoretic approach to detect the bottleneck is presented by Adams, Balas, and Zawack (1988). Uzsoy and Wang (2000) compared the theory of constraints with shifting bottleneck procedure.

3 BOTTLENECK IDENTIFICATION AND RANKING METHOD

The proposed method analyzes inter-departure time data from each machine and looks for signature in the data to identify the bottleneck machine. The proposed method recognizes four valid states for each machine. These states are: (1) cycle, (2) blocked-up, (3) blocked-down and (4) fail states. A machine carries out the value-added task during each cycle. The cycle time decides the time spent in this state during each visit. A machine either delivers a completed job to the next machine or enters the blocked-down state. A machine enters a blocked-down state either due to a full down-stream buffer or a failed or a blocked-down down-stream machine. Once a machine clears a blocked-down state, it either enters a blocked-up (idle) state or a cycle state with a new part. An unreliable machine fails from time to time. A failed machine is repaired and put back into service. Both time between failure and time to repair are modeled by appropriate statistical distributions. These distributions are developed by following the formal method of input modeling of historical data. It is assumed that a machine is restored to its original state of health following each repair. This analysis does not consider age dependent failure of a machine.

A machine can be only in one of the four states as defined earlier at any point in time. The periods of preventive maintenance or other planned outages of each machine are ignored. In case of completely reliable machines connected as a serial flow line, the machine with the largest cycle time (slowest) is the bottleneck machine. However, in a typical manufacturing system machines are unreliable and each machine goes through a failure followed by a repair cycle from time to time. Based on authors' interactions with on-the-field practitioners, a popular current practice is to identify the machine, that has the largest combined % of residence time in cycle and fail states, as the bottleneck machine. We must recognize that a fast machine may turn out to be the bottleneck machine due to very poor reliability. The bottleneck machine sets the pace for other machines in the system and each machine attains the same average throughput rate in steady state.

The proposed method is based on the assumption that the bottleneck machine is least affected by other machines in the system. A review of the four states of a machine defined earlier indicates that blocked-up and blocked-down states represent the impact of a system on an individual machine. A fast machine being fed by a slow machine will often be in a blocked-up state. Same reasoning applies to blocked-down state if a fast machine feeds a slow machine down stream. The proposed method defines the machine with minimum combined % of residence time in blocked-up and blocked-down state as the bottleneck machine. In our opinion this definition of a bottleneck machine offers a better understanding of "bottleneck" compared to the rule based on combined residence time in cycle and fail state. The readers may notice that the proposed new rule may be derived easily from the rule described earlier.

The proposed method uses inter-departure time data from individual machines and analyzes it to identify the bottleneck machine. This paper considers deterministic cycle time only.

The following steps summarize the proposed method.

1. Collect inter-departure time data from each machine for a specified time period,
2. Identify the failure cycles only for each machine,
3. Filter the data collected in step 1. by eliminating the failure cycles,
4. Estimate the combined time spent by each machine in blocked-up and blocked-down states.

In step 2, time stamps of begin and end of every failure cycle of each machine are to be collected. The authors feel that a machine transits from a well defined cycle state to one of the many well defined failed states during failure and returns to the cycle state following repair. It is also expected that there will be fewer failure cycles compared to the other states due to improved machine reliability.

Based on the new rule described earlier, the machine with minimum combined total time spent in blocked-up and blocked-down states is the number one bottleneck machine. The same measure may be used to rank the bottlenecks. The following section shows the results from a simulation model of a flow line followed by the results from a job shop.

3.1 Supporting Numerical Examples

A four machine three buffer flow line has been simulated using Arena 10.0 software in this study. The system parameters are shown in the following table.

Table 1: System Parameters

Machine #	1	2	3	4
TP rate	75 JPH	72 JPH	69 JPH	66 JPH

TBF	Expo(120)	Expo(90)	Expo(25)	Expo(60)
TTR	Expo(3)	Expo(3)	Expo(3)	Expo(3)
Buffer capacity	#1 – 8 units	#2 – 8 units	#3 - 8 units	

Machine throughput rates (TP) follow a pattern typically seen in an automotive assembly line where an over speed is added to each machine as we move upstream. The machines have identical repair behavior. The time between failures (TBF) as well as repair time (TTR) for each machine is expected to follow Exponential distribution. The mean of each distribution is set at a value based on experience with real world data. Time is measured in minutes. Any one of the four machines can be made the number one bottleneck by lowering its mean time between failures thereby letting it fail more often. In this study mean of the time between failures distribution of a machine is changed to force it to be the bottleneck.

3.1.1 Baseline Model

Three different configurations of the base line model have been simulated in this study. The model parameters are shown in the next table. Steps 1 through 4 described earlier have been followed to analyze simulation results. The proposed method identifies a bottleneck machine based on the combined time spent in blocked-up and blocked-down states. The results are compared with bottleneck ranks based on the current popular practice (used in the automotive industry) which focuses on time spent in cycle and fail states. The results show agreement between the two methods.

Table 2: Typical Sample Data

Idt	0.8	0.8	0.8	0.8	0.8	1.17	0.8
Fct	1.17	6.32	4.02	6.57	1.89	3.21	1.96

(Idt- Inter departure time, Fct – Idt with a failure cycle included)

The first row shows a few inter-departure times from machine 1 with a cycle time of 0.8 minute. The second row shows the first eight inter-departure times from machine 1 with a failure cycle included in each data element. The sixth data element in the first row includes a failure cycle.

Table 3: Trial Configurations.

Config.#1	TBF M1-Expo(120)	TBF M2-Expo(90)	TBF M3-Expo(60)	TBF M4-Expo(25)
Config.#2	TBF M1-Expo(120)	TBF M2-Expo(90)	TBF M3-Expo(25)	TBF M3-Expo(60)
Config.#3	TBF M1-Expo(120)	TBF M1-Expo(25)	TBF M1-Expo(90)	TBF M1-Expo(60)

(TBF- time between failures, Expo (120)-Exponential distribution with a mean =120 minutes).

The rest of the system parameters are set at the base values as shown in Table 1. A simulation model based on Arena software is used for this study. Each simulation is run for 600 minutes.

Column 3 of each table shows data generated based on the proposed method. Column 5 of each table shows data used directly from Arena output. The advantages of the proposed method are described in later section.

Table 4: Results from Configuration #1

M-1	B-up+B-down (minutes)*	155.2	(C+F)[1]	72.57%	4
M-2	B-up+B-down	150.2	C+F	79.34%	2
M-3	B-up+B-down	151.5	C+F	75.72%	3
M-4	B-up+B-down	102.2	C+F	96.1%	1

Table 5: Results from Configuration #2

M-1	B-up+B-down	152.2	C+F	77.42%	4
M-2	B-up+B-down	133.2	C+F	78.2%	3
M-3	B-up+B-down	86.6	C+F	95.08%	1
M-4	B-up+B-down	102.8	C+F	87.41%	2

Table 6: Results from Configuration #3

M-1	B-up+B-down	163	C+F	76.26%	4
M-2	B-up+B-down	111.5	C+F	88.99%	2
M-3	B-up+B-down	128.7	C+F	78.48%	3
M-4	B-up+B-down	108.6	C+F	93.84%	1

*B-up and B-down imply Blocked-up and Blocked-down respectively). (C+F)[1] represents combined cycle and failure states. The last column represents bottleneck rank.

3.1.2 Consistency of Inter-departure Time

One of the objectives of this research is to identify a descriptive statistic based on inter-departure time that can be used directly to identify and rank the bottlenecks. The authors feel that consistency of inter-departure time may be correlated to the severity of bottleneck. Results of the following simulation study provide support to this thought.

The four machine three buffer system was run with the following time between failures parameters:

Table 7: Time Between Failures Parameters

M1- expo(120)	M2-expo(100)	M3-expo(80)	M4-expo(60)

The results of a 600 minutes and a 1200 minutes simulation run are shown in the next table.

Table 8: Simulation Results

(cycle+ fail)	M1-83.76%	M2-86.72%	M3-93.46%	M4-96.64%
Bottleneck rank	4	3	2	1
(cycle+ fail)	M1-91.27%	M2-93.27%	M3-93.97%	M4-97.20%

Bottleneck ranking is assigned based on the rule that the most severe bottleneck will have the highest combined % residence in cycle and fail states. The inter-departure time from each machine is analyzed and a measure of consistency is estimated for each machine. The results are shown in the next table.

Table 9: Consistency Measure of Inter Dep. Time

Average absolute deviation of inter-departure time(minutes)	M1- 0.29	M2- 0.21	M3– 0.17	M4– 0.16
Bottleneck Rank	4	3	2	1
Average absolute deviation of inter-departure time(minutes)	M1- 0..22	M2- 0.18	M3 – 0.13	M4 – 0.09

Average absolute deviation is a measure of variability in a data set and smaller numerical value indicates smaller variability or higher consistency. This example presents a very specific configuration where the machine reliability degrades progressively as we move down stream. Combined with progressively slower cycle times as we move down stream makes this configuration unique. However authors are encouraged by this result and research is in progress to establish this finding on a stronger foundation. Other measures of consistency are to be investigated. The authors plan to present additional results during the conference since this is an ongoing research.

3.1.3 Job Shop Model

In this study, a job shop with five machines processing four parts has been simulated using Arena 10.0 software. There is random arrival of parts to the system that follows a specified mix. Each part has unique process plan. The cycle time on each machine is assumed to be deterministic and machines are unreliable. Each machine is served by a buffer of finite capacity of 6 units. The following tables show relevant system parameters.

Table 10: System Parameters

	Part 1	Part 2	Part 3	Part 4
Proc. Plan	2,4,2,1,2	1,3,2,4,2	5,2,3,2	2,3,2,4,5
Part Mix	25%	35%	15%	25%

The process plan for a part shows the visitation sequence of the part to different machines.

Table 11: Machine Failure and Repair Parameters

	M1	M2	M3	M4	M5
TBF	Expo(45)	Expo(30)	Expo(30)	Expo(75)	Expo(90)
TTR	Expo(3)	Expo(3)	Expo(3)	Expo(3)	Expo(3)

Time between failures and time to repair for each machine has been modeled by exponential distribution. Machines have different reliability as shown by different mean time between failures. However, each machine has identical repair pattern.

The job shop has been simulated for 1800 minutes. The result is shown in the next table.

Table 12: Simulation Results

Machine	M1	M2	M3	M4	M5
Method 1	15.23	51.98	13.12	16.96	5.8
Rank	3	1	4	2	5
Method 2	1559.9	938.8	1570.1	1540.8	1695.4
Rank	3	1	4	2	5

Result from Method 1 shows the combined % of residence time in cycle and fail state for each machine. Result from Method 2 shows the combined residence time in blocked-up and blocked-down states for each machine. The results show bottleneck rank of each machine generated by the two methods described earlier. The results show a complete agreement between the methods. Additional results will be presented at the conference.

Research is in progress on how to best utilize the "inter-departure" time signal for performance analysis and improvement in system performance.

4 DATA INTEGRITY ISSUE

One of the motivations behind the proposed method is to offer a process that is less affected by data error. Based on authors' experiences, analysis of performance related data of a manufacturing system is challenged by data error. The analyst has to take special care to filter the data to improve data accuracy. Data errors are primarily due to: (1) Incorrect definition of machine states, (2) Errors in sensing data, and (3) Software error. A set of rules is proposed here that may help to identify the data elements in error. These rules are based on the valid states of a machine defined earlier in this paper.

Rule #1 - The sum of percentage residence time in different states must be equal to 100%.

Rule #2 - A fail state must be preceded as well as succeeded by a cycle state.

Rule #3 - A blocked-up state must be followed by a cycle state.

Rule #4 - A blocked-up state must be preceded by either a cycle or a blocked-down state.

Rule #5 - A blocked-down state must be preceded by a cycle state.

These rules must be applied to the data collected by a data collection system in a manufacturing system. The data elements that are in error need to be analyzed further and corrective actions must be taken. The authors are working on development of an automated data correction procedure. New rules are to be added to the set described earlier if more states are defined for a machine.

5 ADVANTAGES

The major advantages of the proposed method are as follows:

1. The method uses inter-departure time data and failure cycle data only. The inter-departure time data are easy to collect and likely to be error free because of its simplicity. In addition, over a specified time period, the frequency of fail state is expected to be small compared to other states of a machine due to improved machine reliability. The authors strongly feel it is easy to collect failure cycle data since a machine transits from a well defined cycle state to one of the known fail states following a failure. The proposed method gains in replacing collection of cycle time data with inter-departure time data.

2. The study shows that under certain operating conditions a statistic (average absolute deviation) derived from inter-departure time data can be used directly to rank the bottlenecks. Research is in progress to extend this result to general operating conditions.

6 CONCLUSION

A method is presented in this paper which analyzes inter-departure time and failure cycle data to identify and rank bottlenecks in a manufacturing system. The method is based on a simple signal, offers less computational burden and is less sensitive to data error. The method is applicable to identification of both momentary as well as average bottleneck. The current method uses deterministic cycle time only. Research is in progress to extend the proposed method to variable cycle time.

Research is also in progress to develop other statistics derived from inter-departure time data that can be used directly to identify and rank the bottlenecks. We had limited success with "average absolute deviation" measure under restricted operating conditions. If successful, failure cycle data will not be needed for this analysis in future.

REFERENCES

Adams, J., E. Balas, and D. Zawack. 1988. The shifting Bottleneck Procedure for Job-Shop Scheduling. *Management Science* 34(3):391-401.

Chiang, S.-Y., C.-T. Kuo, and, S.M. Meerkov. 1998. Bottlenecks in Markovian Production Lines: A Systems Approach. *IEEE Transactions on Robotics and Automation* 14(2):352-359.

Chiang, S.-Y., C.-T. Kuo, and S.M. Meerkov. 2002. Bottlenecks in Serial production Lines: Identification and Application. *Mathematical Problems in Engineering* 2002.

Kelton, D. W., R.P. Sadowski, and D. A. Sadowski. 2007. *Simulation with Arena.* Fourth Edition, McGraw-Hill, Inc.

Kuo, C.-T., J.-T. Lim, and S.M. Meerkov. 1996. Bottlenecks in Serial Production Lines: A System-Theoretic Approach. *Mathematical Problems in Engineering* 2:233-276.

Law, A.M., and D.W. Kelton. 2000. *Simulation Modeling and Analysis*. McGraw-Hill, Inc.

Roser, C., M. Nakano, and M. Tanaka. 2002. Shifting Bottleneck Detection. In *Proceedings of the Winter Simulation Conference,* ed. E. Yucesan, C.-H. Chen, J.L. Snowdon, and J.M. Charnes, 1079-1086.

AUTHOR BIOGRAPHIES

SANKAR SENGUPTA Ph.D. is an associate professor in the Department of Industrial and Systems Engineering at Oakland University, Rochester, MI. Current research topics include modeling and analysis of manufacturing and service system, application of simulation in decision making, application of lean principles, quality control and robust design methods. His e-mail address is <sengupta@oakland.edu>.

KANCHAN DAS, Ph.D. is an Assistant Professor in the Technology Systems Department at East Carolina University, Greenville North Carolina. Current research topics include modeling and analysis of performance of a supply chain, reliability, design of manufacturing cell and simulation. His e-mail address is <dask@ecu.edu>.

ROBERT P. VANTIL, Ph.D. is a professor in the Department of Industrial and Systems Engineering at Oakland University, Rochester, MI. Current research topics include modeling and analysis of manufacturing system, development of simulation based predictive tools and application of lean principles to manufacturing and service systems. His e-mail address is <vantil@oakland.edu>.