Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

BUFFER UTILIZATION BASED SCHEDULING OF MAINTENANCE ACTIVITIES BY A SHIFTING PRIORITY APPROACH – A SIMULATION STUDY

Maheshwaran Gopalakrishnan
Anders Skoogh

Dept. of Product and Production Development
Chalmers University of Technology
SE 421 96 Gothenburg, SWEDEN

Christoph Laroque

University of Applied Sciences Zwickau
Scheffelstraße 39
08056 Zwickau, GERMANY

ABSTRACT

Machine breakdowns and improper maintenance management cause production systems to function inefficiently. Particularly, breakdowns cause rippling effects on other machines in terms of starved and blocked states. Effective planning of maintenance can lead to improved production system efficiency. This paper aims at improving system throughput through prioritization of maintenance work orders by continuously monitoring buffer levels. This paper proposes and tests a new approach to determine the machine priorities for dynamic scheduling of maintenance work orders by identifying buffer utilization. The approach is exemplified in an industrial use-case. The results have shown to increase throughput in comparison to a first-come-first-served approach for executing maintenance work orders. This new approach relies on simple data collection and analysis, which makes it a viable option for industries to implement with minimal effort. The results can suggest that systems view for maintenance prioritization can be a powerful decision support tool for maintenance planning.

1 INTRODUCTION

Discrete production systems are not utilized to their full capacities due to their complex nature. A large problem that exists in industries today is that the Overall Equipment Effectiveness (OEE) of machines are 15 to 25% below the targeted level (Parida et al. 2014). Apart from machine failures, which cause direct downtimes, failure of one machine can also cause starvation in the machines downstream and blockage in the machines upstream in a production system. Hence, there are buffers in discrete production systems particularly to mitigate these losses due to machine blockage and starvation (also termed as system losses). Despite buffer allocation, machines are still not utilized to their full capacities and system losses occur. These losses directly affects the productivity, increases energy consumption, and incurs excessive costs to the industry. For example, an industrial study has shown that about 30% of the total energy consumption has been due to repair, idle, and stand-by states of machine (Skoogh, Johansson, and Hansson 2011). Therefore, maintenance of machines is not only important for solving direct downtime (machine level problem) but also important in order to indirectly solve starved and blocked states of other machines (system level problem). Consequently, the field of maintenance management is one of the major field of study within production systems research.

Currently, planning of maintenance has been carried out highly on a machine level. For example, the current planning tools such as Failure Mode and Effects Analysis (FMEA) and Root Cause Analysis (RCA) do a good job in increasing the availability of individual machines. However, increased availability of an individual machine might not necessarily increase the efficiency of an entire system, particularly in discrete manufacturing. The production efficiency of the system could be compromised due to another machine which is critical and that machine could make the other machines starve or blocked. Maintenance costs are usually high, for example one study has shown that about 15 to 70% of total production costs are spent on maintenance (Muthu et al. 2000). This implies that maintenance needs to be channelized to the critical machines of the system. More importantly, decisions in maintenance planning needs to become fact-based. Data-driven decision making has emerged as an important quality for performance improvement (Li, Chang, and Ni 2009). Research within this area has indicated that existing Computerized Maintenance Management System (CMMS) are inadequate for the dynamic requirements of maintenance operations and new Decision Support Systems (DSS) are needed (Ni and Jin 2012). Another research points out important issues in existing maintenance techniques, scheduling, and information systems and new shift in maintenance paradigm i.e. an integrated maintenance approach is highlighted (Garg and Deshmukh 2006). Therefore, a systems view is needed to improve the production efficiency and effectively plan maintenance (Ylipää et al. 2016).

A systems view will enable the identification of critical machine with respect to the performance of the system. Particularly in flow-oriented type discrete production systems, the performance in focus is the production rate, i.e. throughput. Therefore, the machine that impedes the throughput of the system in the strongest manner is the bottleneck (throughput critical) machine (Kuo, Lim, and Meerkov 1996). Additionally, bottlenecks are not static and move between machines from time to time (Roser, Nakano, and Tanaka 2002). According to Theory of Constraints (TOC), the bottleneck machines needs to be constantly tracked and continuously improved (Goldratt 1990). From a maintenance perspective, prioritizing the maintenance work orders for bottleneck machines have shown to increase the throughput of the entire production system (Langer et al. 2010; Li et al. 2009).

Despite the opportunities to increase throughput through maintenance prioritization, it is hard for the companies to identify the bottleneck machines with certainty. The existing bottleneck detection methods can be classified into analytical and simulation-based methods (Li et al. 2009). Currently, companies prioritize maintenance work orders non-factually; and another study shows maintenance technician creating maintenance work orders sets the priority depending on situations (Gopalakrishnan et al. 2015). This clearly shows a gap between bottleneck identification and maintenance prioritization within manufacturing companies. In order to bridge this gap, a systems view for scheduling maintenance is needed. A systems view can enable identification of the throughput critical machine. One way of identifying throughput criticality is through queue length (i.e. number of parts in buffer) (Lawrence and Buss 1994).

The purpose of this paper is to enable increased throughput using effective maintenance planning. The paper particularly aims increasing throughput by continuously monitoring the buffer levels to identify machine priorities for maintenance scheduling. This paper proposes and tests a new approach for setting machine priorities dynamically by identifying the utilization of buffers preceding the machines. The scope of the paper is delimited towards flow-oriented production lines and reactive maintenance work orders. A Discrete Even Simulation (DES) experimentation of an automotive industrial use-case is used to evaluate the new dynamic maintenance scheduling approach. Additionally, the results are also compared with the results from static and dynamic maintenance prioritization and discussed.

The remainder of the paper is structured as follows. Section 2 briefly reviews literature on buffer based bottlenecks. Section 3 describes the methodology of the paper, which includes the use-case description, experimental plan, and simulation implementation. Section 4 presents the results of the buffer utilization based maintenance prioritization. Section 5 discusses the results along with presenting the contributions and future research. Lastly, section 6 presents the concluding remarks of the paper.

2 BUFFER UTILIZATION AND BOTTLENECK MACHINES

Buffer allocation is a vast field of study focusing on different production system performance improvement measures. For example Roser, Nakano, and Tanaka (2003a) describes a buffer allocation prediction model of performance measures for different buffer sizes. Buffers can help a bottleneck machine in two ways:

firstly, buffer located subsequent to a bottleneck can provide free space for part flow i.e. reduce blockage and secondly, buffer located preceding to a bottleneck machine can provide additional parts for part flow in the system i.e. reduce starving (Roser Nakano, and Tanaka 2003a). Along with supporting a bottleneck machine, the parts in the buffer can also be used to identify the bottleneck machine. One of the early methods on bottleneck detection describes that the buffer having the highest queue length (number of parts in a buffer) indicates that the machine subsequent is a bottleneck machine (Lawrence and Buss 1994). The waiting time of parts in the buffer can also be used to detect bottlenecks through identifying the longest waiting time for parts in a buffer (Roser, Nakano, and Tanaka 2003b). However, these methods comes with delimitations in the capacity of buffer to be infinite and is applicable only for serial production lines. Even though this paper focuses on flow-oriented production systems the buffer capacities are high but not infinite (see Section 3.1). More recently, a "turning point" method for identifying bottlenecks based on the blockage to starvation trend changes to starvation and blockage in buffers was proposed (Li, Chang, and Ni 2009). However in this paper, an extension of the queue length based bottleneck detection is used in this paper. In order to eliminate the problem of infinite capacity problem, the utilization of buffer is used to identify the throughput critical machine. The continuous monitoring of buffer utilization and the implementation of maintenance work order prioritization is described in detail in section 3.2.

3 METHODOLOGY

A DES methodology suggested by Banks et al. (1996) is followed in this paper to evaluate the maintenance scheduling approaches. This paper presents an empirical study, in which an industrial use-case with real data was used for experimentation.

3.1 Use-Case Description

In this paper a specific production line within an automotive industrial company producing engine components for cars is used for evaluating the maintenance scheduling approaches. The production line, which consists of 11 highly automated machines are connected in series. The layout of the production line is shown in Figure 1 where the parts flow from machine M1 to M11. All the machines execute different manufacturing operations with different cycle times. Particularly, machine M10 is a special operation which can process 100 parts at the same time (i.e. high capacity machine). All machines are decoupled using buffers B1 to B9 with capacities ranging from 70 to 170, except machines M5 and M6 (Figure 1). Buffers B3 and B4 has a capacity of 100; buffers B2 and B9 has a capacity of 170; and the rest has a capacity of 70 each. The first machine M1 receives parts from an unlimited supply (M1 never starves) and the last machine M11 delivers the finished part to an unlimited demand (M11 never blocked).

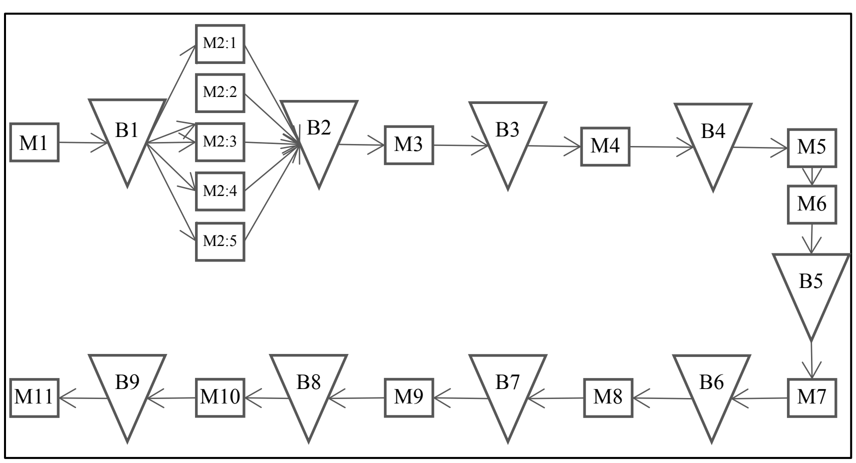

Figure 1: Layout of the automotive industrial production line (use-case).

There are three maintenance technicians in this production line who are responsible for executing the reactive maintenance work orders on a first-come-first-served basis. The technicians were divided based on their knowledge of the reactive maintenance in order to execute the work orders. Technician T1 is responsible for machines M1 to M2 (all five M2 machines); technician T2 is responsible for machine M3 to M6; and technician T3 is responsible for machines M7 to M11. The production line is simulated using a discrete event simulation software AutoMod for a time period of one work week containing three shifts a day and 7 working days (i.e. 168 hours). A warmup period of 8 working hours has been removed before the 168 hours in order to achieve steady-state. Additionally, all the simulation models were run for 50 replications and the simulation results were extracted at 95% confidence interval.

3.2 Experimental Plan

The goal of the experimental plan was to setup simulation models of the use-case to study and to gather results for the maintenance scheduling method based on buffer utilization. The models given below describe the execution of maintenance work orders. Additional models are used to compare the results of the buffer utilization based prioritization with the static and dynamic maintenance prioritization.

3.2.1 Reference Model

This model will be used as the reference model for the experiments carried out. In this model the reactive maintenance work orders are executed in a first-come-first-served approach.

3.2.2 Buffer Utilization Model

In this model the maintenance work orders are prioritized in a dynamic manner. The simulation model will continuously monitor the buffer levels and calculate its utilization (Utilization = number of parts divided by buffer capacity). After comparing the utilization of the buffers in the production line, the highest priority is given to the machine subsequent to the buffer having the highest buffer utilization and low priorities are given to the machines subsequent to the buffer with low utilization. Eventually, all machines will have a priority number based on the utilization of the buffer preceding the machines. The priority setting using buffer utilization is visualized in Figure 2. This figure is an example excerpt at a time T1 during the run of a production line consisting of three machines and two buffers in-between them. As seen from the Figure 2, utilization of buffer B1 is 60%, whereas the utilization of buffer B2 is 75%. Despite having different buffer capacities, due to the high utilization of buffer B2 machine M3 claims priority 1.

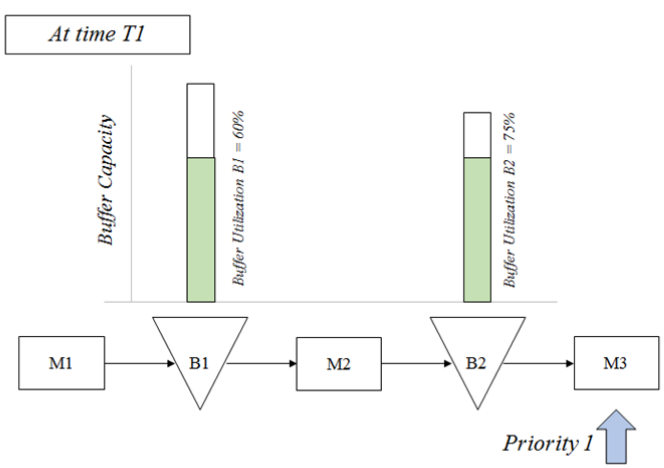

Figure 2: Buffer utilization based priority shifting.

Moreover, priorities continuously (i.e. shifting priorities) shift as the buffer utilization changes within the runtime of the model. The priorities are independent of the work ordered generated from the machines. therefore upon attaining priorities, if two work orders arises then the maintenance technician chooses the work order of the machine which has high priority rather than the time of occurrence of failure. Additionally, if a high priority machine creates a maintenance work order during the work order execution of a low priority machine then the maintenance technician stops executing the work order at the low priority machine and completes the work order of the high priority machine. This implies that the high priority machines' work orders are always executed first (i.e. reduce/eliminate maintenance waiting time). This routine is carried out throughout the end of the simulation run.

3.2.3 Additional Models

There are two additional models used to compare the buffer utilization approach with static and dynamic prioritization approach. These approaches were previously published by the authors. (1) Static priority model, which prioritizes maintenance work orders for machines based on static bottleneck approach (Gopalakrishnan, Skoogh, and Laroque 2013). The static bottleneck approach used was active period percentage method suggested by Roser, Nakano, and Tanaka (2001). (2) Dynamic priority model, which prioritizes maintenance work orders for machines through a shifting bottleneck approach (Gopalakrishnan, Skoogh, and Laroque 2014). The shifting bottleneck approach used was the momentary bottleneck detection method suggested by Roser, Nakano, and Tanaka (2002). The execution of maintenance work orders after receiving priorities are same as buffer utilization model (see Section 3.3.2.).

3.3 Simulation Implementation

The reference simulation model was built first and the other models were subsequently built over the reference model with changes to the maintenance work order execution. The use-case of the paper consists of three maintenance technicians executing the maintenance work orders of various machines. Therefore, the priorities were set for the machines according to the maintenance technician's responsible machines. For example, maintenance technician T2 is responsible for machines M3, M4, M5, and M6. Hence, technician T2 will have four priorities for those machines. Similarly, technician T3 will have five priorities for the machines M7, M8, M9, M10, and M11 as that technician is responsible for those machines. It must be noted that the technician T1 has not been modeled for executing maintenance work orders according to priorities. Technician T1 is responsible for machines M1 and M2, where M1 is the first machine (non-bottleneck) and M2 is a group of parallel machines. Therefore, the maintenance work orders of machines M1 and M2 have been executed only on a first-come-first-served approach. Particularly for the buffer utilization model, the priorities are set for machines based on the buffer utilization preceding the machines. All machines have a buffer preceding them except Machine M6 (Figure 1). Therefore, the machine M6 has been modelled to have the same priority of the preceding machine (M5).

4 RESULTS

The throughput analysis for the different maintenance scheduling approaches were carried out and are presented below. Firstly, the results from buffer utilization based maintenance prioritization is presented. Additionally, a thorough analysis of the buffer utilization, machine states, the shifting of priorities, and throughput are presented. Lastly, the throughput results of the buffer utilization based prioritization are compared with the static and dynamic based prioritization.

4.1 Buffer Utilization Based Prioritization

In order to analyze the buffer utilization based prioritization of scheduling maintenance, it is important to see the current state of the machines in the production line. Table 1 below shows results from the simulation of the reference model. The mean buffer utilization and the machine state information such as mean utilization and mean downtime are shown (machines M1 and M2 are not presented as they were not considered for setting priorities). From Table 1, the mean utilization shows the mean amount of time each machine has been processing parts; the mean downtime shows the combined mean breakdown time and the mean waiting time for repairs; and the remaining percentage of the time the machines are in a state blocked and starved for parts. On comparing the buffer utilization and machine utilization together, it can be observed that the machine with high utilization tends to have high utilization on the buffer preceding it. This phenomenon is observed in the technician T3's machines, where machine M7 has the highest utilization (60.4%) and the buffer preceding that machine (B5) has the highest buffer utilization (58.3%). However with technician T2's machines, all the four machines have machine utilization values close to each other. In addition, there is no buffer preceding the machine M6. Therefore, the buffers preceding these machines also have buffer utilization values close to each other. Hence, this phenomenon suggests that the buffer utilization is an indicator of the machine utilization in the production system. With respect to the system performance, the reference model produces a mean throughput of 24685 parts during the run time of 168 hours i.e. about 147 parts per hour (see Table 2) from executing maintenance work orders in a first-come-first-served approach.

Table 1: Buffer and machine states results from the reference model.

Buffers	**Utilization %**	**Machine**	**Utilization %**	**Downtime %**
B2	90.0	M3	47.7	16.7
B3	91.0	M4	50.2	14.7
B4	88.9	M5	49.4	21.4
-	-	M6	50.6	19.4
B5	58.3	M7	60.4	11.5
B6	44.6	M8	52.6	19.0
B7	29.2	M9	44.1	18.3
B8	14.3	M10	24.5	0.3
B9	37.6	M11	32.7	14.6

On continuously monitoring and shifting the priorities, the buffer utilization model has produced a mean throughput of 25554 parts during the run time of 168 hours i.e. about 152 parts per hour. Hence, an increased throughput than the reference model is achieved. The throughput comparisons between the two models along with the confidence interval are shown in Table 2. From this table, it can be observed that the throughput increase of the buffer utilization model is statistically significant with the confidence intervals at 95% are not overlapping to that of the reference model. Additionally, the confidence intervals of buffer utilization model is narrower than the reference model, i.e. the production system is more stable. Additionally, the improvement achieved is about 3.5%.

Table 2: Throughput comparison.

	System Throughput	**Confidence Interval**	**Percentage Improvement**
Reference Model	24685	+/- 340.04	-
Buffer Utilization Model	25554	+/- 283.72	3.5 %

On further analysis of throughput critical machine shifting, the amount of time each machine claimed priority 1 is analyzed. The Figure 3 shows the percentage amount of time the machines in the production line has claimed priority 1 for maintenance work order execution, i.e. the amount of time each machine has been the throughput critical machine according to buffer utilization approach. Please note that Figure 3 shows two charts, as the machines of the two maintenance technicians are given priorities individually.

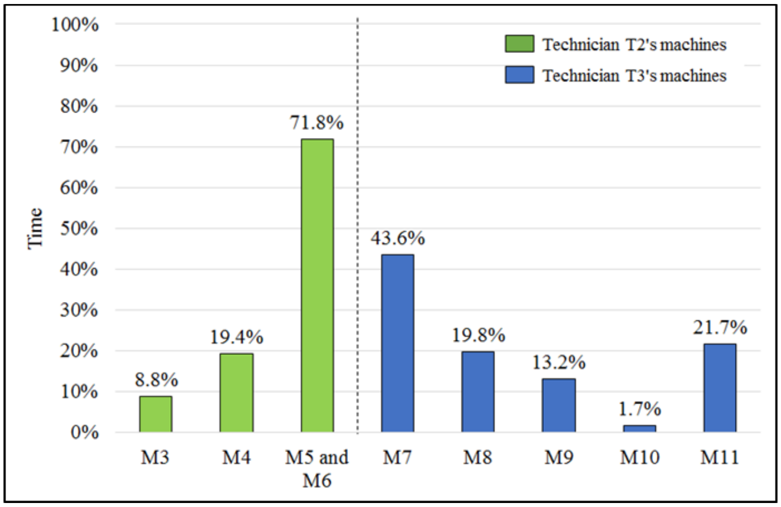

Figure 3: Percentage of time the machines claiming priority 1.

From Figure 3 (technician T3's machines), it can be observed that machine M7 has claimed priority 1 the longest period of time (about 43.6%). From Table 1, machine M7 also has the highest machine utilization. Therefore, it can be argued that machine M7 is the long-term bottleneck machine. Additionally, from Figure 3 (technician T2's machines), the machines M5 and M6 (both machines had same priority) claimed joint priority 1 the highest (combined 71.8%) but these machines do not work as high as machine M7 as indicated by low utilization than M7 (Table 1). Hence, it can be argued that machine M3 and M4 are not primary bottlenecks but has been critical to throughput of the system. Contrarily, machines M3 and M10 have claimed priority 1 the least amount of time. Machine M3 is in the beginning of the production line and machine M10 is a high capacity machine (see Section 3.1), which are the reasons for having reduced utilization on the preceding buffers. However, all machines (Figure 3) have claimed priority 1 at some point in time. This implies that all machines have been throughput critical at certain times.

4.2 Throughput Comparison with Static and Dynamic Bottleneck Based Priorities

The additional models were used to compare the mean throughput of the buffer utilization based prioritization with the mean throughput of static and dynamic bottleneck based prioritizations. The results of the mean throughput comparison are shown in Figure 4. The buffer utilization model has produced 3.5% throughput increment, whereas the static prioritization model based on utilization has produced 5% throughput increment and dynamic prioritization model based on momentary bottleneck has produced 2.2% throughput increment. In Figure 4, the different models are presented based on the static and dynamic priorities. The models on the left (green color bars) are static priority models, whereas the models on the right (grey color bars) are shifting priority models. Therefore, the models on the left (reference and static) have no shifting of priorities, whereas the model on the right has high level of priorities being shifted. The buffer utilization approach for prioritization presented in this paper is not as dynamic of the momentary bottleneck based prioritization and produced 3.5% throughput increment.

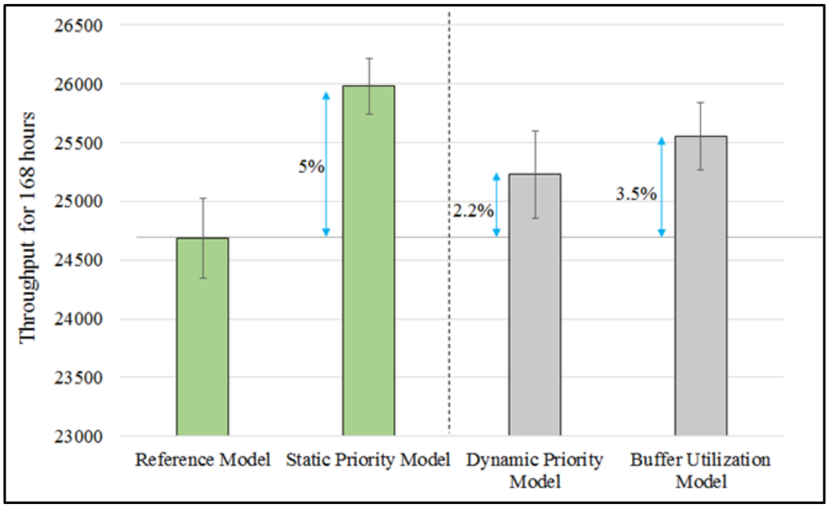

Figure 4: Comparison of different prioritization schedules.

In order to analyze the high capacity buffer effects the downtime of machine M7 was analyzed further to identify its rippling effects on other machines upstream and downstream. Machine M7 was chosen as this is the primary bottleneck of the system and had claimed priority 1 the longest (Figure 3). Machine M7 generated a mean of 60 maintenance work orders during the simulated time of 168 hours. An analysis was carried out to check the number of times the machines located upstream were blocked and machines located downstream were starved due to the rippling effects of downtime of machine M7. Therefore, the upstream machines were checked for changing its state from "working" to "blocked" and the state change from "working" to "starved" was checked for downstream machines during the downtime of Machine M7. Table 3 shows the mean frequency of starve and blocked states of machines due to the failure of M7.

Table 3: Rippling effect analysis of downtime of Machine M7.

Upstream machines of M7	**Frequency of blockage due to M7's downtime**	**Downstream machines of M7**	**Frequency of starvation due to M7's downtime**
M6	24.5	M8	19.7
M5	24.4	M9	14.1
M4	20.6	M10	10.3
M3	17.5	M11	17.9

From the results presented in Table 3, it can be observed that downtime of M7 causes blockage for all the machines upstream and starvation for all the machines downstream. Particularly, the upstream machines M5 and M6 have the highest impact and the downstream machine M8 has the largest impact. It also shows that the blockage effect reduces on the upstream machines gradually from M6 to M3. Similarly, the starvation reduces gradually from M8 to M10, but increases for machine M11. This rippling effects shows that the downtime of bottleneck machine considerably affects all the other machine in the system, but do not affect as much as it would affect in a single piece flow.

5 DISCUSSION

The paper presents and tests a new approach for scheduling maintenance work orders through prioritizing machines based on continuous monitoring of buffer levels. This approach was exemplified using an automotive industrial use-case. The proposed approach was shown to increase the throughput by 3.5% in comparison to a first-come-first-served approach of executing maintenance work orders. The throughput increment was achieved through prioritizing the maintenance work orders for the machines, which has the highest utilization of buffers preceding them. Information from the buffers can indicate the constraint of the production system. Particularly, bottleneck machines can be identified by analyzing the number of parts in buffer (queue length), waiting time of parts in buffer, and "turning point" in buffers (Lawrence and Buss 1994; Roser, Nakano, and Tanaka 2003b; Li, Chang, and Ni 2009). The approach proposed in this paper to identify throughput critical machine is an extension of the queue length method of bottleneck detection. In the queue length method the number of parts in the buffer is monitored, whereas in this approach utilization of buffers (number of parts per capacity) is monitored.

The achieved increase in throughput is an indicator of the potential for improving the production system performance without additional efforts. Efforts such as increase in personnel, resources, or reduction in repair time are not needed. However, communication channels for sharing criticalities and work order settings are needed. Additionally, the increment is achieved through smartly working with an limited downtime of machines (10 to 15% of the total time – see Table 1). Previously, maintenance prioritization based on bottleneck machines have also showed to increase throughput (Langer et al. 2010; Li et al. 2009). However, one of the major problems is to have high quality and availability of machine data to identify bottlenecks. On the other hand, data from the buffers are fairly easy to collect and analyze, which is the only data source needed for this approach. This approach is also a dynamic way to prioritize maintenance, where the machines has been given priority as it affects the system performance the most (Figure 3). Modern maintenance management strategies calls for such type of data driven (fact-based) and dynamic planning for maintenance (Li, Chang, and Ni 2009; Ni and Jin 2012).

A critical aspect of this approach arrived when it was compared to the static and dynamic approaches for prioritizing maintenance. Firstly, the authors tried a static prioritization approach based on utilization and found 5% increase in throughput (Gopalakrishnan, Skoogh, and Laroque 2013). However, dynamic prioritization tends to produce higher throughput than that of static prioritization (Langer et al. 2010). Therefore, a dynamic approach based on momentary bottlenecks was tried, where only 2.2% throughput increment was achieved (Gopalakrishnan, Skoogh, and Laroque 2014). The main reason for this difference is the dynamics of the production line studied in this paper. Due to the relatively high capacity of buffers, continuously prioritizing maintenance for the momentary bottlenecks does not ensure increased throughput. However, it is only logical to identify the right level of priority shifting in order to achieve the highest possible throughput improvement as bottlenecks move between machines. Therefore in this paper, the buffer utilization approach was proposed, which has produced higher throughput increment than that of dynamic approach, but lower throughput increment than that of static bottleneck approach (Figure 4). Therefore, it can be argued that the proposed approach is still not the best, but certainly an effective approach as it produces substantial throughput increase. In this approach the buffers protect the long-term bottlenecks from blocking and starving while a momentary bottleneck is prioritized (see Figure 3, long-term bottleneck M7 is prioritized only 43.6%). This is when it loses parts in the long-term bottleneck machine in comparison to the static priority approach, which continuously prioritizes the long-term bottleneck machine. Despite this, the authors believe that the continuous monitoring of production system (machines and buffers states) will enable identifying shifting of critical machines but the right approach is needed to be found. The approach of this paper has obviously captured the critical machines better than the momentary bottleneck approach, hence the better throughput increment (Figure 4). Therefore, the results of this approach suggests that the authors are in the right direction towards identifying the best approach for scheduling maintenance as a better dynamic approach than momentary bottleneck has been identified.

Another critical aspect is the delimitation of the approach, where the machines were given priorities only based on the buffers located preceding the machine. One of the main purpose of buffer allocation is the allocate space (mitigate starving) and allocate parts (mitigate blocking) due to the failure of other machine (Roser, Nakano, and Tanaka 2003a). In this paper as well, the failure of machine M7 (long-term bottleneck) not only caused blockage on upstream machines, but also caused considerable starvation for machines downstream (Table 3). Since this approach focuses only the preceding buffers, the problem of starvation was not included. Upon considering information from both the buffers (preceding and subsequent) around machines, the throughput critical machine can be identified with even greater certainty. Prioritizing those machines for maintenance work order execution might enable achieving throughput even greater than that of static priority.

Previous study has proposed effective decision support tools, such as short-term bottleneck analysis and maintenance prioritization are needed to fulfill the demands of dynamic maintenance requirements (Ni and Jin 2012). Additionally, maintenance planning has been shown to be on an equipment level and systems view is needed to increase the scope of existing maintenance (Ylipää et al. 2016). The approach presented in this paper aligns itself with the directions of the recent researches in this area. Particularly, a systems view is applied while prioritizing maintenance while analyzing the utilization of buffers. This approach can be an effective decision support tool for maintenance planning in the industry.

5.1 Contributions and Future Research

The main scientific contribution of this paper include, a new approach for scheduling maintenance work orders based on continuous monitoring of buffer levels. This approach is a new addition to the field of maintenance prioritization studies, where previously bottlenecks and system values were the focus for prioritization (Li et al. 2009; Ni and Jin 2012). In addition to this, the applicability of this approach is another contribution. The requirements it sets on the data collection and data analysis for identifying and prioritizing critical machines is considerably less than that of static and dynamic priority approaches used in this paper. Therefore, this approach is a viable option for the industries to implement data-driven (fact-based) and dynamic decision making without additional efforts and gain significant throughput increase. Direct industrial contributions include potential to improve system efficiency, particularly throughput. Developing this approach into a decision support tool can be a future work towards industrial implementation.

The future research opportunities for the authors arising from this research are aplenty. The main future work of this research will be on the generalizability of the results, since a single use-case was used in this paper. Therefore, a multiple use-case with different production set-ups are needed to validate the results of this research. Particularly from the approach presented where the buffers preceding the machines were monitored, an interesting future work is to include monitoring of buffer subsequent to the machines to enable better identification of throughput critical machine. Finally, the preventive maintenance planning can also be included in this priority based approach to enable higher potential for throughput increase.

6 CONCLUSION

The aim of this paper is to propose and test a new dynamic maintenance scheduling approach in order to increase system throughput with particular intentions to make priority decisions through continuously monitoring buffer levels. From the simulation study of an automotive industrial use-case, a significant throughput increment of about 3.5% was achieved in comparison with a first-come-first-served based maintenance scheduling for a specific flow-oriented production line. The increment was achieved through a new approach for maintenance scheduling, in which the throughput critical machines are identified and prioritized based on highest utilization of buffers. The requirements on the data collection and data analysis are minimal in comparison to both static and dynamic bottleneck priority approaches. The results show that data-driven (fact-based) and dynamic decision making can be implemented for planning maintenance through easily available data and without additional efforts. In addition, the results show that a system view for maintenance enables effective maintenance planning and more efficient production system.

ACKNOWLEDGEMENTS

This work has been performed within Sustainable Production Initiative and the Production Area of Advance at Chalmers. The support is appreciated greatly.

REFERENCES

Banks, J., J. S. Carson, and B. L. Nelson. 1996. *Discrete-Event System Simulation*. 2nd ed. Prentice-Hall, Upper Saddle River, New Jersey.

Garg, A, and S. G. Deshmukh. 2006. "Maintenance Management: Literature Review and Directions." *Journal of Quality in Maintenance Engineering* 12(3): 205-238.

Goldratt, E. M. 1990. *What is This Thing Called Theory of Constraints and How Should it be Implemented?* New York: North River Press.

Gopalakrishnan, M., J. Bokrantz, T. Ylipää, and A. Skoogh. 2015. "Planning of Maintenance Activities – A Current State Mapping in Industry." *Procedia CIRP* 30: 480-485.

Gopalakrishnan, M., Skoogh, A., and Laroque, C. 2013: "Simulation-Based Planning of Maintenance Activities in The Automotive Industry." In: *Proceedings of the 2013 Winter Simulation Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 2610-2621. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Gopalakrishnan, M., Skoogh, A., and Laroque, C. 2014. "Simulation-Based Planning Of Maintenance Activities by a Shifting Priority Method." In: *Proceedings of the 2013 Winter Simulation Conference*, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 2168–2179. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, 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(3): 233-276.

Langer, R., J. Li, S. Biller, Q. Chang, N. Huang, and G. Xiao. 2010. "Simulation Study of a Bottleneck-Based Dispatching Policy for a Maintenance Workforce." *International Journal of Production Research* 48(6): 1745-1763.

Lawrence, S. R., and A. H. Buss. 1994. "Shifting Production Bottlenecks: Causes, Cures, And Conundrums." *Production and Operations Management* 3(1): 21-37.

Li, L., Q. Chang, and J. Ni. 2009. "Data Driven Bottleneck Detection of Manufacturing Systems." *International Journal of Production Research* 47(18): 5019-5036.

Li, L., Q. Chang, J. Ni, and S. Biller. 2009. "Real Time Production Improvement through Bottleneck Control." *International Journal of Production Research* 47(21): 6145-6158.

Muthu, S., S. R. Devadasan, S. Ahmed, P. Suresh, and R. Baladhandayutham. 2000. "Benchmarking for Strategic Maintenance Quality Improvement." *Benchmarking: An International Journal* 7(4): 292-303.

Ni, J., and X. Jin. 2012. "Decision Support Systems for Effective Maintenance Operations." In: *CIRP Annals - Manufacturing Technology* 61(1): 411–414.

Parida, A., U. Kumar, D. Galar, and C. Stenström. 2014. "Performance Measurement and Management for Maintenance: A Literature Review." *Journal of Quality in Maintenance Engineering* 21(1): 2-33.

Roser, C., M. Nakano, and M. Tanaka. 2001. "A Practical Bottleneck Detection Method." In: *Proceedings of the 2002 Winter Simulation Conference*, edited by B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, 949–953. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Roser, C., M. Nakano, and M. Tanaka. 2002. "Shifting Bottleneck Detection." In: *Proceedings of the 2002 Winter Simulation Conference*, edited by E. Yücesan, C.-H. Chen, J.L. Snowdon, and J.M. Charnes, 1080-1086. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Roser, C., M. Nakano, and M. Tanaka. 2003a. "Buffer Allocation Model Based on a Single Simulation." In *Proceedings of the 2003 Winter Simulation Conference,* edited by S Chick, P. J. Sanchez, D Ferrin, and D. J. Morrice. 1238-1246. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Roser, C., M. Nakano, and M. Tanaka. 2003b. "Comparison of Bottleneck Detection Methods for AGV Systems." In *Proceedings of the 2003 Winter Simulation Conference,* edited by S Chick, P. J. Sanchez, D Ferrin, and D. J. Morrice, 1192-1198. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Skoogh, A., B. Johansson, and L. Hansson. 2011. "Data Requirements and Representation for Simulation of Energy Consumption in Production Systems." In: *Proceedings of the 44th CIRP Conference on Manufacturing Systems* 2011.

Ylipää, T., M. Gopalakrishnan, A. Skoogh, and J. Bokrantz. 2016. "Identification of Maintenance Improvement Potential using OEE Assessment." *Accepted for publication in the International Journal of Productivity and Performance Management* April 2016.

AUTHOR BIOGRAPHIES

MAHESHWARAN GOPALAKRISHNAN is a PhD student working at the Department of Product and Production Development, Chalmers University of Technology, Sweden. He has a Master's in Production Engineering and Management. His main research focus is maintenance planning and identifying criticality in Production Systems. His email address is mahgop@chalmers.se.

ANDERS SKOOGH. PhD, is an Associate Professor at the Department of Product and Production Development at Chalmers University of Technology. He is a research group leader for Production Service Systems & Maintenance. Anders is also the director of Chalmers' Master's program in Production Engineering and a board member of Sustainability Circle (www.sustainabilitycircle.se) with responsibilities for research collaboration. Before starting his research career, he accumulated industrial experience from being a logistics developer at Volvo Cars. His e-mail is anders.skoogh@chalmers.se.

CHRISTOPH LAROQUE studied business computing at the University of Paderborn, Germany. Since 2013 he is Professor for Business Computing at the University of Applied Sciences Zwickau, Germany. He is mainly interested in the application of simulation-based decision support techniques for operational production and project management. His email address is Christoph.laroque@fh-zwickau.de.