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

EXPERIMENTAL STUDY ON VARIATIONS OF WIPLOAD CONTROL IN SEMICONDUCTOR WAFER FABRICAITON ENVIRONMENT

Appa Iyer Sivakumar

Singapore-MIT Alliance (SMA)
Manufacturing Systems and Technology (MST) Programme
N3.2-01-36, 65 Nanyang Drive, 637460 Singapore
msiva@ntu.edu.sg

Chao Qi

Department of Control Science and Engineering
Huazhong University of Science and Technology
1037 Luoyu Avenue, Wuhan, Hubei 430074 China
qichao@mail.hust.edu.cn

Andy Darwin Kasan Hidayat

Singapore-MIT Alliance (SMA)
Manufacturing Systems and Technology (MST) Programme
N3.2-01-36, 65 Nanyang Drive, 637460 Singapore
andydarwin@alum.mit.edu

ABSTRACT

WIPLOAD Control has been proposed and evaluated as an effective job release control methodology. This paper presents a further simulation experimental study on two variations of WIPLOAD Control in semiconductor wafer fab environment. The first variation mechanism is to control separate WIPLOAD for each part type at a specified level under the situation when the type of job going to be released next can be controlled. The second variation is a mechanism with additional release control points along the production line. The performance of the WIPLOAD variations are compared with that of the original WIPLOAD Control and other existing release methods including CONWIP and Constant release in semiconductor manufacturing environment. The experimental results indicate that by breaking down the control of overall fab WIPLOAD into controlling WIPLOAD level for each part type, cycle time performance can be improved especially with respect to the standard deviation of cycle time.

1 INTRODUCTION

Job release control is an essential part of scheduling system, determining the type, amount and time point of release of new jobs into a manufacturing system. Since 1970s, an ever growing attention has been devoted into job release control especially with the emergence of complex manufacturing systems such as semiconductor wafer fabrication (fab). The significant impact of job release control on cycle time performance of wafer fab has been demonstrated by Wein (1988) and Glassey and Resende (1988).

Job release control methodologies can be generally classified into two categories: open-loop and closed-loop methodology. Open-loop release methodologies make release decisions regardless of any current system information; release is usually scheduled based on exogenous information such as prediction and demand; and the release time is not modified according to what is happening in the production line before serious performance degradation has been caused. In contrast, closed-loop release methodologies take into account the dynamic shop floor information according to specific objectives.

Queueing theory suggests that reducing the variability in the input process will improve the performance of an open-loop methodology. Therefore Constant release (CONST), which releases new jobs into manufacturing system with a constant rate, is the best open-loop release method. Similar release control method, e.g. a certain number of new lots are released at the beginning of every shift/day/week, is often observed in real life wafer fabs.

The majority of the existing closed-loop release methodologies adopt the idea of "WIP cap" (Hopp and Spearman 2000) to control the start of new jobs by limiting the workload. CONWIP (Spearman et al. 1990) and Workload Regulating (WR) (Wein 1988) are two representative closed loop job release control methodologies that are widely discussed in semiconductor manufacturing environment. The principle of CONWIP is to maintain a constant WIP level by starting new jobs whenever the WIP level has fallen below a specific level. WR is based on the Theory of Constraints (TOC), with the principle of limiting the workload of bottleneck workstation at a prescribed level. However the implementation of WR is constrained by the fact that bottleneck is constantly shifting in real life wafer fabs due to the change of product mix and other dynamic factors, while CONWIP does not take into consideration the distribution of jobs along the production line when release decisions are made.

In order to overcome the limitations of these existing approaches, Qi (2006) defined a new measure: system WIPLOAD, to measure the overall workload on the shop floor, as the sum of the remaining processing times of all the jobs on the shop floor. WIPLOAD Control (WIPLCtrl) (Qi 2006) is a closed-loop job release methodology proposed based on the concept of WIPLOAD to control system WIPLOAD at a reference level. The performance of WIPLCtrl was evaluated in different manufacturing system configurations and appears to outperform CONST, CONWIP, WR and some other existing release methodologies in terms of cycle time as well as due date performance (Qi et al. 2007). Its interactions with other production control factors as well as system factors were also investigated in semiconductor wafer fab environment (Qi et al. 2008).

In this paper, two variation mechanisms of WIPLCtrl are investigated through simulation experimental study of a semiconductor wafer fab. The first variation is to prescribe a reference WIPLOAD level for each part type with the assumption that the type of the job going to be released next can be controlled, instead of controlling an overall system WIPLOAD level as the original WIPLCtrl. The second variation is not to only control the job release point before the whole production line; instead, an additional control point is added along the production line to re-release the jobs queueing there with the principle of WIPLCtrl.

In section 2, the original WIPLCtrl is briefly introduced. It is followed by the introduction on the two WIPLCtrl variations in sections 3 and 4, respectively. In section 5, simulation experimental study is presented to evaluate the performance of the proposed variation release methods. Conclusions are included in section 6.

2 WIPLOAD CONTROL

Assume there are M types of parts processed in a wafer fab, for the original WIPLOAD Control (WIPLCtrl), we defined System WIPLOAD as the sum of theoretical remaining processing times of all the jobs on the floor, i.e. $L(t) = \sum_{m=1}^{M} \sum_{j=1}^{Jm} W_{m(j)}(t) R_{m(j)}(t)$, where $W_{m(j)}(t)$ is the number of lots of part type m at its j^{th} operation step at time t, and $R_{m(j)}(t)$ is the remaining processing time of the lot of part type m at its j^{th} operation step at time t. The involvement of the remaining processing times takes into account more system information when the shop load is measured. The WIP located at the front-end of the production line is thus considered to cause higher load for the shop floor in comparison to the WIP at the back-end. In this sense, WIPLOAD improves conventional shop load measure such as CONWIP that treats all WIP with the same assigned weight. The current WIPLOAD, *L(t)*, is updated in real time when a lot exits from a machine, and fed back to compare with a reference WIPLOAD level, L^*, the value of which is determined by the expected throughput level of the system. The lot with the highest priority will be released if system WIPLOAD will not exceed L^* with the additional load (i.e. the overall processing time of the lot) caused by the lot.

3 WIPLCTRL VARIATION 1: CONTROL WIPLOAD FOR EACH PART TYPE

The first variation of WIPLCtrl is derived with the assumption that the type of the lot going to be released next can be controlled when the release decisions are made. With this assumption, the fluctuation of system WIPLOAD can be further controlled by setting a reference WIPLOAD level for each part type, with L_m^* denoting the reference WIPLOAD level for part type m. The system WIPLOAD level of part type m at time t, $L_m(t)$, is calculated by $L_m(t) = \sum_{j=1}^{Jm} W_{m(j)}(t) R_{m(j)}(t)$ in real time. Then the principle of WIPLCtrl is applied for each part type to make the release decisions. As a result, the overall system WIPLOAD is also controlled under a hedging level corresponding to curtain expected system throughput level.

The first WIPLCtrl variation takes into consideration the additional information about the distribution of each type of lots along the production line, which is helpful as different part type has very different processing routes involving different set of equipment in a typical wafer fab. In this sense, when the total WIPLOAD is lower than our specified limit, this variation mechanism is capable of releasing a correct type of lot into the system.

4 WIPLCTRL VARIATION 2: ADDITIONAL RELEASE CONTROL POINT IN LINE

The manufacturing process in a typical semiconductor wafer fab is very complex consisting of hundreds of operation steps. It is very commonly observed that a released lot stays in the fab as WIP for more than a month before it leaves the fab as a finished product. To enhance the control of lots along the production line, the second variation mechanism of WIPLCtrl is proposed. An additional release control point is added in the production line, where the queueing lots will be re-released based on the principle of variation mechanism 1 for a certain number of downstream machines. Figure 1 illustrates the framework of variation 2. The second release control has its own reference sub-WIPLOAD level for each part type, which is denoted as $L2^*_m$. The sub-WIPLOAD levels are updated for the sub set of stations.

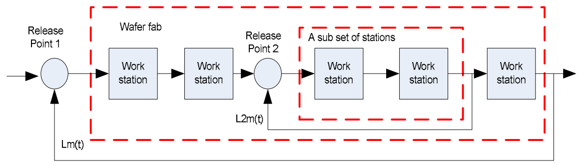

Fig. 1: WIPLCtrl Variation with Additional Control Point

5 SIMULATION EXPERIMENTAL STUDY

A Simulation experimental study is conducted using SIMUL8 to investigate the performance of two proposed WIPLCtrl variation mechanisms in comparison with the original WIPLCtrl.

5.1 Model Description

The simulation model is a simplified wafer fab model with 55 workstations, among which 30 workstations are highly utilized while the remaining 25 are modelled as a "delay" workstation. Out of the 30 highly utilized stations, 19 of them are represented each by one single machine. Nine "Diffusion" stations consist of five parallel machines each. Those five machines only can take in a certain product types with a minimum batch size of one and a maximum of six. Five part types are simulated with different processing routes representing five typical products in silicon foundry. Delay work station is a simplification of low utilized facility that we do not want to model in detail. To capture more accurate behaviour of the factory, the stations of STEPPERA and STRIPPER are modelled in the machine level because both of them are the most utilized facilities. The system performance measures of interest in this study are the average cycle time and the standard deviation associated with it as cycle time is the number one performance metric in semiconductor industry (Fowler and Robinson 1995).

5.2 Performance of Original WIPLCtrl

The performance of the original WIPLCtrl is first evaluated using the simulation model in comparison with CONST and CONWIP. The average values of 11 independent replications are used as the results of the experiments. Simulation length of each test run is set to 31,536,000 seconds (1 year time). The output of the initial warm-up period of 15,768,000 seconds is discarded to remove the transient state when the results are analyzed. The length of warm up period is determined according to Welch's procedure (Welch 1960). The dispatching rule used is FIFO.

The results showed in Figures 2 and 3 agrees with what had been claimed in our previous related studies (Qi 2006; Qi et al. 2007; Qi et al. 2008) that WIPLCtl is able to outperform CONST and CONWIP simultaneously in terms of average cycle time and standard deviation of cycle time for a certain throughput level.

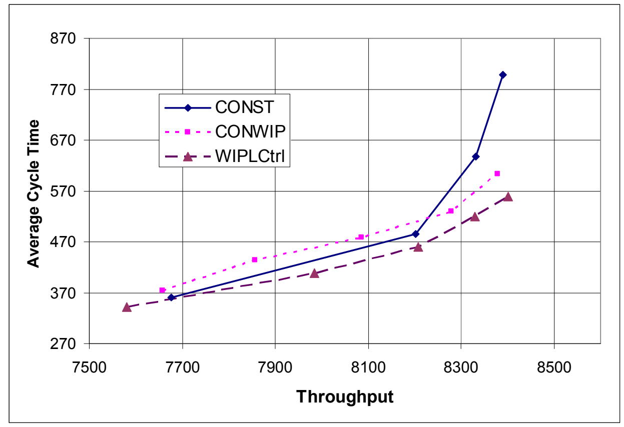

Fig. 2: WIPLCtrl Performance on Average Cycle Time

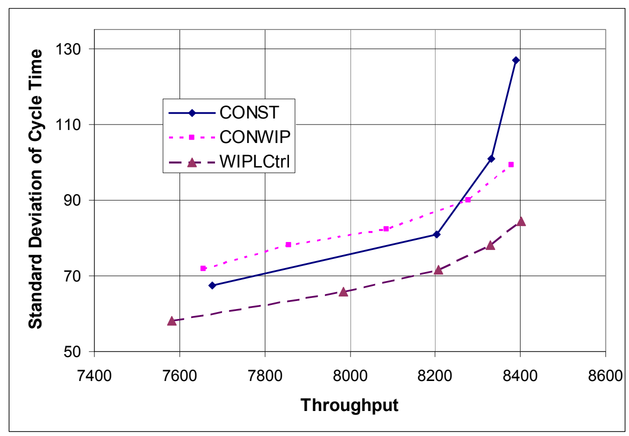

Fig. 3: WIPLCtrl Performance on Standard Deviation of Cycle Time

5.3 Performance of WIPLCtrl Variation 1

The performance of the first WIPLCtrl variation is simulated with 11 replicas of different levels of WIPLOAD in comparison with the original WIPLCtrl. The results are shown in the Figures 4 and 5. It is observed that variation mechanism 1 is able to outperform original WIPLCtrl. WIPL-n-Ctrl can reduce average cycle time and a significant reduction in standard deviation for the same production output as compared to WIPLCtrl.

The improvement from WIPLCtrl variation 1 verifies the effect of controlling type of the job going to be released next, which is able to efficiently reduce the variability of WIPLOAD flow.

Although the performance of original WIPLCtrl and variation 1 is close on average cycle time, the results of student t-test verify that with 99% confidence interval, there is a statistically significant mean different between the original WIPLCtrl and variation 1.

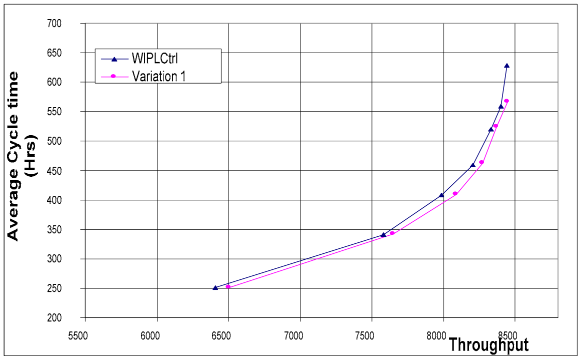

Fig. 4: Performance of WIPLCtrl and Variation 1 on Average Cycle Time

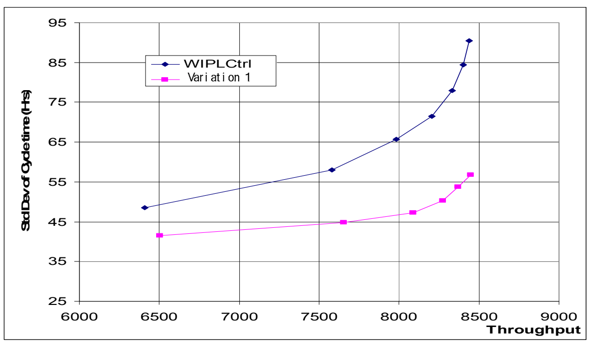

Fig. 5: Performance of WIPLCtrl and Variation 1 on Standard Deviation of Cycle Time

5.4 Performance of WIPLCtrl Variation 2

For WIPLCtrl variation 2, we add an additional release control point on step 100 of the process. All part types should go through the second release control when they reach step 100 before they can go to the downstream processes. Both the first (i.e. release control at the beginning of the production line) and the second (i.e. the additional release control point) release control the WIPLOAD level for each part type at a specified level. The second release control has its own reference sub-WIPLOAD level (i.e. $L2_m^*$) for each part type *m*. The calculation of the second stage WIPLOAD for part type *m* at time *t* is done starting from step 101 to the end as $L2_m(t) = \sum_{j=101}^{Jm} W_{m(j)}(t) R_{m(j)}(t)$. If $L2_m(t)$ is lower then $L2_m^*$, a job will be allowed to go on to the step 101.

When $L2_m^*$ is set to be a very large number, WIPLCtrl variation 2 will perform exactly the same as variation 2. In our experiment, when we decrease $L2_m^*$, the performance of the factory decreases as well in terms of mean and standard deviation of cycle time. The additional release control has the potential to re-arrange the sequence of the job flows along the production line; however it creates additional delay in the system when it does not allow the jobs to flow. The lower the $L2_m^*$, the higher the additional delay that was created.

It is understood that in order to choose which type of job to release, a queue of jobs should exist in from of the controller. Therefore, it is proposed to use the existing queue that is always built up such as in front of STRIPPER station due to its high utilization. In this sense, it will behave as a job sequencer in front of STRIPPER station. It will choose which part type to be passed on by choosing the lowest $L2_m(t)$ whenever STRIPPER is available.

The simulation results are shown in Figures 6 and 7 comparing the performance of WIPLCtrl variation 1 and variation 2 on mean and standard deviation of cycle time. WIPLCtrl variation 2 becomes effective under the relatively high production throughput levels. However, the decrease on mean cycle time is achieved with the cost of increase on standard deviation of cycle time. The ineffectiveness under low throughput is because there are probably no lots to be re-sequenced or re-released when there is not much queue built up.

The deterioration on the standard deviation of cycle time can be understood by considering that when the second release controller re-sequences the lots, it might stop the flow of particular items that have relatively high $L2_m(t)$ level and keep letting other types of lots keep flowing. As a result some lots with longer cycle time might take further longer time to be finished because they are given lower priority by the re-release controller. On the other hand those lots usually with short cycle times are moving even faster on the line because they are given higher priority.

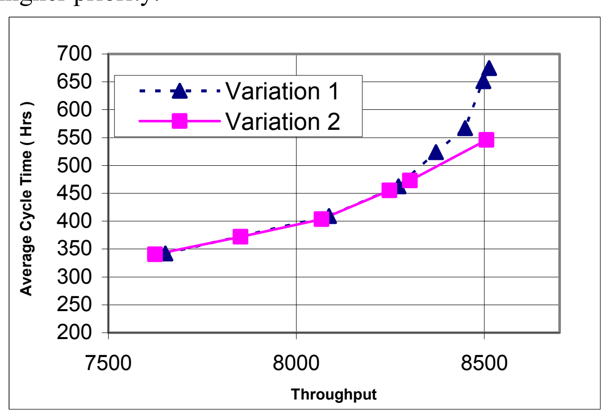

Fig. 6: Performance of WIPLCtrl Variation 1 and Variation 2 on Average Cycle Time

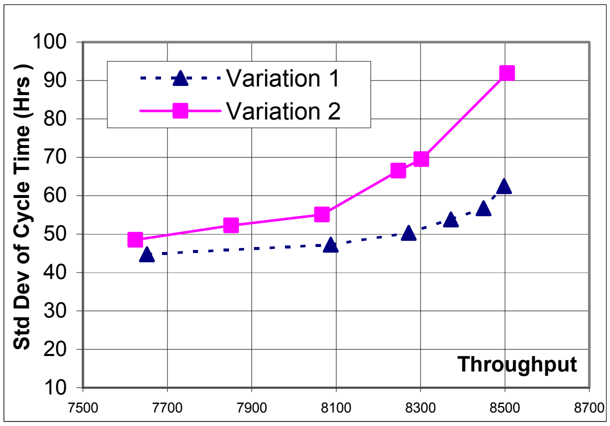

Fig. 7: Performance of WIPLCtrl Variation 1 and Variation 2 on Standard Deviation of Cycle Time

5.5 Variation 1 with Dispatching Rules

In this study we also investigate the interaction of WIPLCtrl variation 1 with dispatching rule. Two commonly used dispatching rules are considered: First-In-First-Out (FIFO) and Earliest Starting Time (EST) incorporated to variation 1. The performance of the original WIPLCtrl working with FIFO is considered as the comparison benchmark. It is observed from Figures 8 and 9 that WIPLCtrl variation 1 with EST results in higher average cycle time compared to variation 1 with FIFO. However Variation1-EST could maintain extremely low standard deviation of cycle time because EST reduces the possibility of old lots waiting at the bottom of queue and as a result avoids those lots from getting further delay.

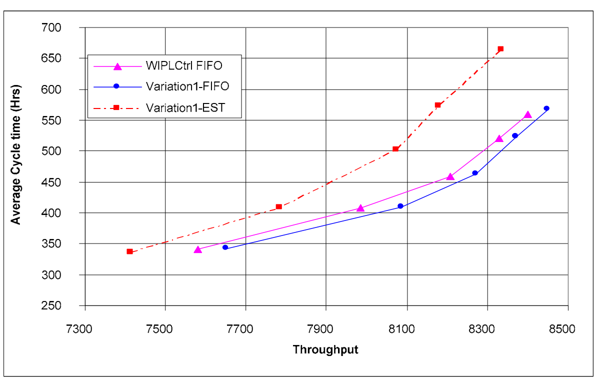

Fig. 8: WIPLCtrl Variation 1 with Dispatching Rules on Average Cycle Time

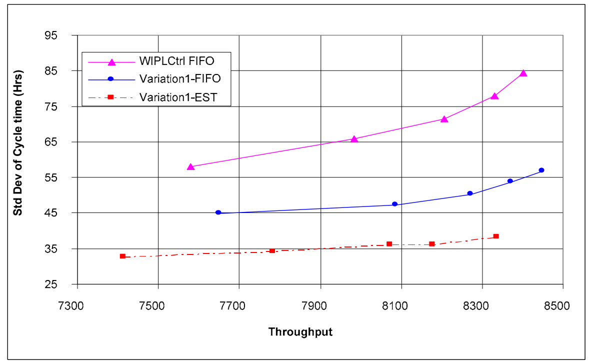

Fig. 9: WIPLCtrl Variation 1 with Dispatching Rules on Standard Deviation of Cycle Time

6 CONCLUSIONS

This paper further verifies that WIPLOAD Control is able to outperform other release control methodology such as COWIP and CONST. In this study, system WIPLOAD is broken down into WIPLOAD level for each part type, which creates extra information that can be used to determine what type of job should be released. By using this extra information the fab cycle time performance can be improved, especially with respect to the standard deviation of cycle time.

The other WIPLCtrl variation by adding an additional release control point on the production line also applies to reduce the average cycle time when the fab capacity is highly utilized. It is suggested to locate the additional release control point in front of bottleneck machine in order to be able to regulate the work flows. The experimental results also show that adding a sub WIPLOAD limit to close the flow of the jobs when the factory load is high at the middle of production line will deteriorate the factory performance.

The interaction with dispatching rules also affects the cycle time performance of the first WIPLCtrl variation. For the two commonly used dispatching rules (i.e. FIFO and EST) considered in this study, variaion1-FIFO results in reduced average cycle time, but with the cost of very high standard deviation of the cycle time. On the other hand, variaion1-EST results in higher average cycle time but it can maintain very low standard deviation of the cycle time.

Our ongoing work focuses on the identification of the reference WIPLOAD/sub-WIPLOAD level and the location of effective release control points on the production line.

ACKNOWLEDGMENTS

This research work is supported by Singapore-MIT Alliance (SMA).

REFERENCES

Fowler, J. and J. Robinson. 1995. Measurement and Improvement of Manufacturing Capacity (MIMAC) Designed Experiment Report. Technology Transfer 95062860A-TR. SEMATECH.

Glassey, C. R., and M. G. C. Resende. 1988. Closed-Loop Job Release Control for VLSI Circuit Manufacturing. *IEEE Transactions on Semiconductor Manufacturing*. 1: 36—46.

Hopp, W. J., and M. L. Spearman. 2000. *Factory Physics*. Irwin/McGraw-Hill.

Qi, C. 2006. *Closed-Loop Job Release Based on WIPLOAD Control in Semiconductor Wafer Fabrication*. PhD Dissertation. Nanyang Technological University, Singapore.

Qi, C., Sivakumar, A. I., and Gershwin, S.B. 2007. An Efficient New Job Release Control Methodology. *International Journal of Production Research*, iFirst, 1-29.

Qi, C., Sivakumar, A.I., and Gershwin, S.B. Aug 2008. Impact of Production Control and System Factors in Semiconductor Wafer Fabrication. to appear in *IEEE Transactions on Semiconductor Manufacturing*.

Spearman, M. L., D. L. Woodruff, and W. J. Hopp. 1990. CONWIP: A Pull Alternative to Kanban. *International Journal of Production Research*. 28: 879—894.

Wein, L. W. 1988. Scheduling Semiconductor Wafer Fabrication. *IEEE Transactions on Semiconductor Manufacturing* 1:115—130.

Welch, P. D. 1983. *The Computer Performance Modeling Handbook*. New York: Academic Press.

AUTHOR BIOGRAPHIES

Dr. APPA IYER SIVAKUMAR is an Associate Professor in the School of Mechanical and Aerospace Engineering (MAE) at the Nanyang Technological University, Singapore and a Faculty Fellow of Singapore - Massachusetts Institute of Technology (MIT) Alliance (SMA-MST programme) He was at Gintic Institute of Manufacturing Technology, Singapore prior to this appointment. His research interests are in the area of OR, Optimization, Advanced Manufacturing Systems engineering, Discrete Event Simulation, Scheduling, Logistics, Supply chain design, and Research Methodology. He is the Chairman of the NTU Undergraduate Research Programme (URECA) and the Chairman of the MAE Engineering Innovation and Design (EID) programme. He has many years of industrial and research experience in the UK prior to the post in Singapore in 1993. He received a Bachelors of Engineering in Manufacturing Systems Engineering and a PhD in Manufacturing Systems Engineering from University of Bradford, UK. He has made many contributions at international conferences and journals. He has trained and supervised a number of PhD students and many Master's level students. He has been the technical chair and co-edited the proceedings of the 3rd and 4th International Conference on Computer Integrated Manufacturing (ICCIM '95 and ICCIM'97), Singapore.

Dr. CHAO QI is a faculty member of the Institute of Systems Engineering, Department of Control Science and Engineering, Huazhong University of Science and Technology, China. She was a postdoctoral research fellow of Singapore-Massachusetts Institute of Technology (MIT) Alliance (SMA) before this position. She received the Ph.D. degree in systems and engineering management from Nanyang Technological University, Singapore, in 2006. Her research interests include planning and scheduling in manufacturing systems and logistics systems.

Andy Darwin Kasan Hidayat was a Master student in Singapore-MIT Alliance (SMA). He is interested in the research on supply chain management and scheduling in manufacturing systems.