

An Order Release Control Mechanism based on Self-Adaptive Neural Fuzzy Inference System and Theory of Constraints

Chuangdong Zhan, Jiangming Kan*

School of Technology, Beijing Forestry University, Beijing 100083, China

*Corresponding author, e-mail: kanjm@bjfu.edu.cn

Abstract

Order release is the key premise for the semiconductor wafer fabrication system to perform well, which is also one of the paramount significant components in the scheduling strategies. Most order release strategies merely have focused on the workload but failed in considering the remarkable influence on cycle time of common orders that is brought by unexpected rush ones. In this paper an on-line mechanism based on Theory of Constraints for lot release using self-Adaptive Neural Fuzzy Inference System models was presented which is able to adjust the release rhythm dynamically according to dynamics of fabs. In our approach, an ANFIS model was established to predict the ratio between hot and common lots in wafer fab to perform adjustments on the order release schedule in advance. Simulated experiments based on the HP24 model were carefully performed and experimental results proved a better performance of common lots than original TOC on a large scale, especially when it comes to the situation of disturbance.

Keywords: ANFIS, order release, rush order, theory of constraints, cycle time

Copyright © 2013 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Semiconductor wafer fabrication system (SWFS) is considered to be one of the most sophisticated manufacturing systems of high-tech, which is known for multiple re-entrants, long production cycle time, complicated manufacturing steps, high cost of investment, and huge uncertainties [1-4]. Therefore, an efficient and effective strategy of order release is paramount significant for the reason that an efficient strategy will lead to reducing cost and living up to customers' expectations; otherwise it even can curtail the market share [5, 6]. Especially, rush orders caused by variation of real market demands subject manufacturing enterprises to a huge range of pressure and risks of profits reducing for the reason that rush orders are able to disrupt the stability of fabrication system, increase inventory level, rise cost and decrease throughput which can result in failure of meeting customers' demands [7].

Early researches of scheduling strategies for the semiconductor wafer manufacturing system date back to the late 80s of the 20th century. Uzsoy et al. ventured that researches of scheduling strategies are mainly consisting order releasing and dispatching [8]. As for order release, constant level of Work-In-Process (CONWIP) [9] regulation is a simple approach based on the amount of lots in processing, which has positive effect on cycle time. Extended from WIP control, Workload Control (WLC) release strategies such as workload regulation (WR) [10] and constant load (CONLOAD) [11] have attracted much attention which decide lot release according to the fab workload.

Though it is of great significance to cope with rush orders in job shops in reality, this topic hasn't received quite enough attention as it should have in semiconductor research field, which holds true for order release control. A precious contribution made by Wu and Chen was a model to calculate the production cost of a rush order in an assemble-to-order context [12]. Wang and Chen advocated the application of ANFIS into forecasting rush orders for regulating the capacity reservation mechanism in advance [13]. However, few embedded rush orders into their release control policies, which is right what our work did.

In this paper, we presented an on-line control strategy for order release concerning hot lots that is consisted of an self-adaptive neuro-fuzzy inference system (ANFIS) [14] and Theory of Constraints (TOC) [15-17]. An ANFIS model was built for predicting the ratio of hot lots of

next five release episodes through Matlab Fuzzy Toolbox. TOC was utilized to identify and elevate the bottlenecks of a simulated semiconductor wafer fabrication system-the Hewlett-Packard-24(HP24) model from Hewlett-Packard Technology Research Center Silicon Fab [10]. The ODBC interface connected to the SQL Server 2005 was utilized to implement data exchange between Matlab 2008a and eM-Plant 8.1 for the purpose of updating the ANFIS parameters on-line.

This paper is organized as follows. Related methods used in this paper are introduced in Section 2. The detailed structure of the combined strategy for release control will be presented in Section 3 and Section 4 will cover the simulated experiments performed to check and verify the scheduling strategy with corresponding results and analysis. In the end, conclusions will be given in Section 5.

2. Methods

2.1. ANFIS

This section covers a typical adaptive network with a fuzzy inference system[14]. Sugeno model is one of the most widely used fuzzy inference models. For a first-order Sugeno fuzzy inference system with two inputs, a common set rule set with two fuzzy if-then rules is the following:

Rule 1: if x is A_1 and y is B_1 , then $f_1 = p_1x + q_1y + r_1$

Rule 2: if x is A_2 and y is B_2 , then $f_2 = p_2x + q_2y + r_2$

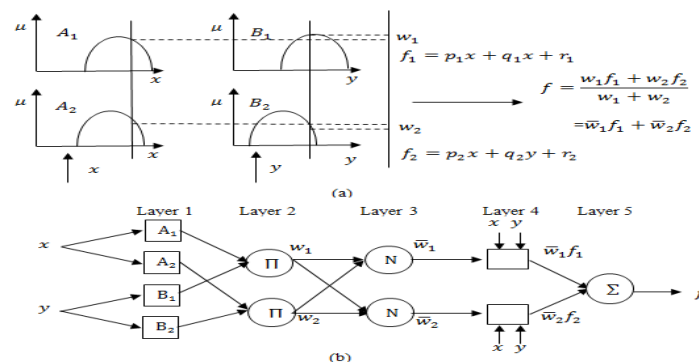


Figure 1. (a) A Two-input First-order Sugeno Fuzzy Inference System with Two Rules, (b) An Equivalent ANFSI Architecture

An ANFIS network of five layers is demonstrated in Figure 1 with the equivalent Sugeno fuzzy inference system above. Learning of ANFIS applied in this paper consists of structure-learning in the first place and then parameters-learning. Structure-learning includes space classifying of fuzzy input and rule-extracting. According to the statistical distribution fuzzy c-means (FCM) clustering by extracting a set of rules that models the data behavior was utilized to classify the training sample space. If the space is clustered into n_i classes, then there will be corresponding n_i fuzzy rules. Therefore, initial input parameters of membership functions for each class are determined by the clustered center coordinates and its radius length [18].

ANFIS parameters are divided into two parts: antecedent parameters and consequent parameters, so ANFIS Parameters-learning is surely a procedure made up of identification and adjustment of the two kinds of parameters. As the ANFIS model was established through Matlab Fuzzy Toolbox, a hybrid learning method was applied to configure ANFIS parameters: Least Squares (LS) method for the learning of consequent parameters and Gradient Descent method for antecedent parameters.

2.2. Procedure of Building ANFIS Prediction Model

Initially, simulate HP24 Fab with eM-Plant for an adequate long time so that it is enough for the fab stabilizes and works well. At the same time, properties of the operating fab will be

collected for training and learning parameters of the ANFIS prediction model and evaluation as follows:

- $z_{1,i}$: mean order release interval in i week;
- $z_{2,i}$: mean amount of lots before the constraint device in i week;
- $z_{3,i}$: mean WIP level in i week;
- $z_{4,i}$: total input amount of common lots in i week;
- $z_{5,i}$: total input amount of urgent lots in i week;

Input parameters are 7 in all: $\ln \sum_{j=i-4}^i z_{1,j}$, $z_{2,i-4}$, $z_{2,i-3}$, $z_{2,i-2}$, $z_{2,i-1}$, $z_{2,i}$, $z_{3,i-4}$; Output parameter is $\sum_{j=i}^{i+4} z_{5,j} / \sum_{j=i}^{i+4} z_{4,j}$.

Then, through training of input and output data a Sugeno Inference System which is able to simulate the given data behavior will be established with the least number of fuzzy rules required. In our work, Matlab Fuzzy Toolbox was used to finish that job, the detailed steps are below.

- Load the input and output data which has been normalized
- Classify the data set through fuzzy c-means clustering (structure-learning)
- Select the hybrid learning method and set training cycle and limits of error
- Train the ANFIS model with the data set (parameters-learning)
- Test the ANFIS model

The experiments performed in our work divided into two parts: building ANFIS Prediction Model and testing the on-line release control mechanism. Both building and testing were carried out coherently under two conditions: without and with disturbance of ratio between common and urgent lots. Dynamic Bottleneck Dispatching (DBD) [19] was used for dispatching during simulation.

2.3. Our Approach

As shown in Figure 2, the HP24 fab transmits input data consisting of production properties that are exactly needed for the trained ANFIS model firstly. Then, the model predicts the ratio of next five weeks and sends it back to the running fab. At the same time, the ANFIS model trains itself so as to keep up with the updating data behaviors of fab. In the end, the fab adjusts its release interval according to the forecasted ratio. The adjustment value is computed as following:

$$\Delta Interval_i = C \times (Ratio_{i+1} - Ratio_0) \times BB_i \quad (1)$$

$\Delta Interval_i$: The incremental value of the adjusted release interval;

C : The incremental coefficient of the adjusted release interval according to the fab;

$Ratio_{i+1}$: The predicted ratio between urgent and common lots of next five week;

$Ratio_0$: The set ratio between urgent and common lots;

BB_i : The work time before the first constraint device.

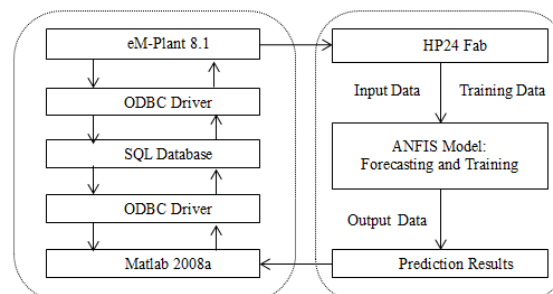


Figure 2. Data Flow Structure of ANFIS Model

3. Experimental Design

Experiments were carried out under two conditions: without and with disturbance of the ratio between urgent and common lots. Under the circumstance of none disturbance, the ratio

between two types of lots was set 20%. On the other hand, the ratio under disturbance floated from 15% to 25% according to Gaussian distribution with the mean value of 20%. The following three approaches of order release control are implemented and will be compared with each other under two conditions coherently:

- TOC: TOC regulation is based on TOC concerning the working ability of Capacity Constraint Resources (CCRs).
- CONWIP: CONWIP regulation is simply based on the WIP level in the fab, which is a representative of Workload Control.
- On-line ANFIS based on TOC: Adjust the release interval on-line that is based on TOC according to the forecasted value provided by the ANFIS model.

As the HP24 fab only produces one particular type of lot, each of the three approaches above has been tested considering: total amount of lots in and out, mean throughput per day, mean cycle time of common and urgent lots, and mean WIP level. The key results we concentrated on were mean throughput per day and mean cycle time of two types of lots. Mean throughput per day illustrates the performance of lots after release and allows us to evaluate the performance of the fab. Mean cycle time, which incorporates the pool delay, describes the whole performance of lots across the fabrication system and captures working percentages of lots. According to the Little's Law, WIP level is strongly proportionally connected with cycle time. Thus, WIP level allows us to estimate the performance of fab and mean cycle time of lots in another way. Each experiment is consisting of 10 simulation years of running: results are collected per week unit; the warm-up period is set 15 week units to avoid start-up effects.

4. Results

4.1. Building ANFIS Prediction Model

Results of training and testing of the ANFIS prediction model for the HP24 fab are shown as Figure 3 and Figure 4. Figure 3 depicts the built fuzzy inference system (FIS) architecture; Figure 4 demonstrates training results of the ANFIS model. As shown in Figure 3, the data set was clustered into 10 classes responding to 10 fuzzy rules in ANFIS and the total amount of fuzzy rules was 70 (10 *ins* × 7). Therefore, initial values of input parameters of membership functions for each class (initial values of the antecedent parameters) are determined by the clustered center coordinates and its radius length. Depicted in Figure 4, after 1000 episodes of training, training errors reached the expected limits and test output values of the ANFIS model kept in pace with fab test values. Through computing prediction results of the test values, total mean error was 0.0348016 within the range of permissible error and output of the ANFIS model is stable.

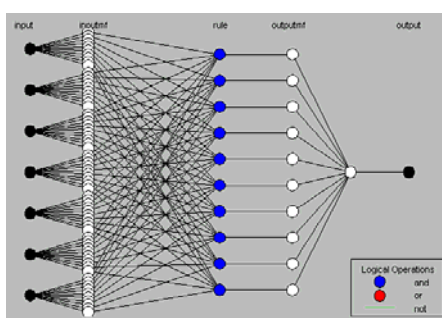


Figure 3. Built FIS Architecture

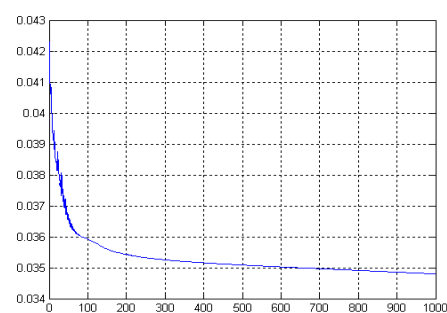


Figure 4. Training Error of the ANFIS Model

4.2. On-line Order Release Control

The on-line feed control was operating with three other strategies: CONWIP, Fixed Interval and DBR. Each type of strategy ran twenty-four hours a day for 3655 days in simulation units with properties of the wafer fab being collected and written into eM-Plant tables. Simulation results are shown in Figure 5 to Figure 11 revealing that our proposed on-line release control is of possibilities to perform better than original TOC mechanism, especially contending with

common lots under certain perturbation which may be applied in a real wafer fab for better achievements in further future somehow as well.

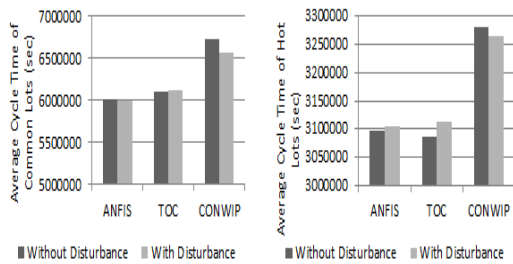


Figure 5. Mean Cycle Time of Common and Hot Lots

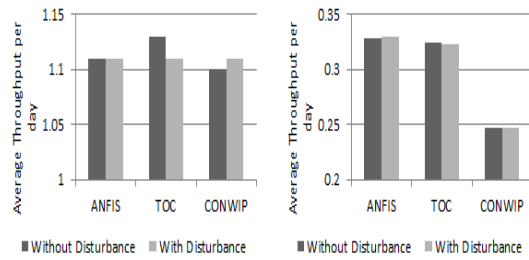


Figure 6. Average Throughput and Lots Working Ratio

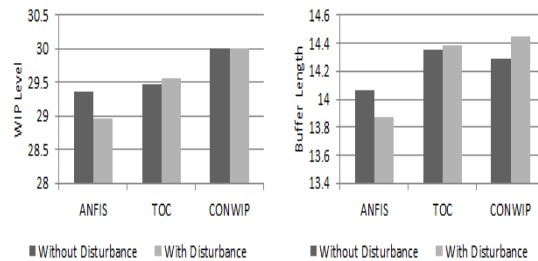


Figure 7. WIP Level and Buffer Lots Amount before CCR

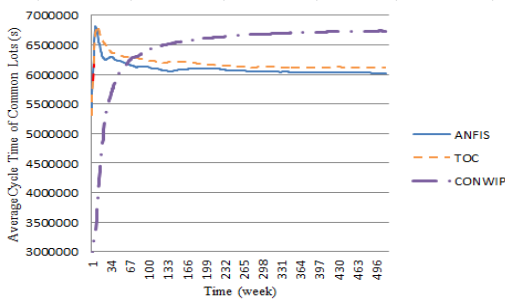


Figure 8. Trend of Average Cycle Time of Common Lots in Fab without Disturbance

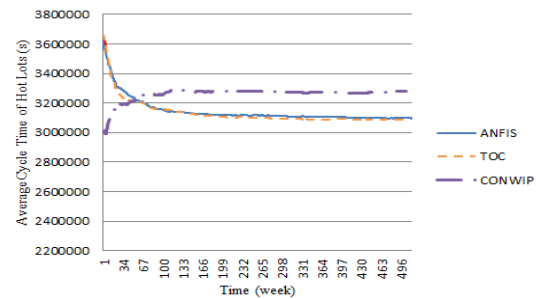


Figure 9. Trend of Average Cycle Time of Hot Lots in Fab without Disturbance

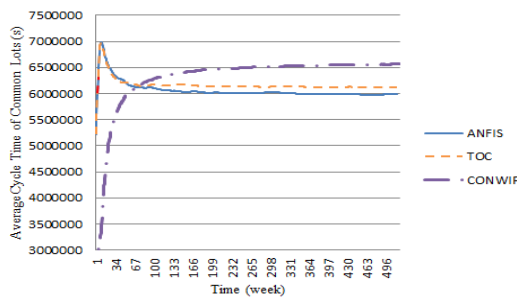


Figure 10. Trend of Average Cycle Time of Common Lots in Fab with Disturbance

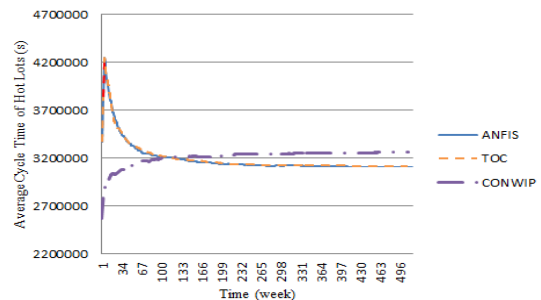


Figure 11. Trend of Average Cycle Time of Hot Lots in Fab with Disturbance

As shown in Figure 5 and Figure 7, average cycle time of common lots under ANFIS regime is much lesser than that of TOC and CONWIP under both conditions, which also holds true for buffer lots before CCR. Average cycle time of common lots of ANFIS is 93139 seconds (1.078 days) lesser than that of TOC under the condition without Gaussian disturbance and 138641 seconds (1.605 days) lesser under the condition with Gaussian disturbance. Though average cycle time of hot lots under ANFIS regime is slightly higher than that of TOC under the condition without Gaussian disturbance, when it comes to situation with disturbance, its performance is better than that of TOC.

Figure 6 shows that the three approaches of order release control approximately share the same throughput and output ratio under both two different conditions meaning that production capacity under each approach is almost equal; that is to say that our approach is able to produce an equal number of lots in the same time which could be proven in Figure 7. However, Figure 5 and Figure 6 show that average cycle time of both common and hot lots and average throughput under ANFIS are more stable than those under either TOC or CONWIP. Especially, average cycle time of common lots in the TOC fab rises up when Gaussian disturbance is added in, while that in the ANFIS fab almost does not fluctuate at all. That means performances of common lots in the ANFIS approach are much less prone to variations of the fabrication system caused by rush orders to some degree. Figure 8 to Figure 11 capture that, as simulation goes, average cycle time of common lots in the ANFIS fab decreases more than that in the TOC fab and stabilizes, meaning that our approach behaves better than the original TOC in the aspect of average cycle time of common lots under two different conditions in the long run.

Compared to the original order release control based on TOC with concerning almost the same WIP level and average throughput per day, our approach reduces 1.52% of average cycle time of common lots under the condition of none disturbance and 2.26% under condition of disturbance with Gaussian distribution. Besides, the variation of average cycle time of hot lots under two different conditions under ANFIS mechanism is much smaller than that under original TOC mechanism: 8221 seconds for the former and 26859 seconds for the latter, revealing that our approach holds better performances of stability against disturbance with Gaussian distribution.

5. Conclusion

This paper presented an on-line order release control strategy based on ANFIS for handling rush orders in SWFS. An ANFIS model was built for predicting the ratio of hot lots of next five episodes through Matlab Fuzzy Toolbox with its parameters being updated on-line and real-time. TOC was utilized to identify and elevate the bottlenecks of the HP24 fab, and decide order release interval as well. We compared the fab performance under the regime of ANFIS, TOC or CONWIP in two conditions of ratio between hot and common lots: without and with Gaussian disturbance. ANFIS outperforms the other strategies with respect to achieving the desired level of average throughput. The reason lying in the good performance of our approach is that: when the predicted ratio between hot and common lots is lower than 20% more common lots will be released; otherwise, lesser common lots will be released; as a result, through adjusting the number of common lots to be released in advance, average cycle time of common lots won't rise and stabilize under variations brought by hot ones with maintaining the expected WIP level in all. Simulation results show that this on-line strategy is of possibilities to perform better than contemporary ones which neglect the variation of rush orders, especially contending with common lots under acceptable disturbance of hot lots.

References

- [1] JH Chen, LC Fu, MH Lin, AC Huang. Petri-net and GA-based approach to modeling, scheduling, and performance evaluation for wafer fabrication. *IEEE Transactions on Robotics and Automation*. 2001; 17: 619-636.
- [2] AR Moro, H Yu, G Kelleher. Hybrid heuristic search for the scheduling of flexible manufacturing systems using Petri nets. *IEEE Transactions on Robotics and Automation*. 2002; 18: 240-245.
- [3] TC Chiang, YS Shen, LC Fu. *Adaptive lot/equipment matching strategy and GA based approach for optimized dispatching and scheduling in a wafer probe center*. Proc. IEEE Int. Conf. Robot. Automat., New Orleans, LA. 2004; 3125-3130.

-
- [4] ZC Cao, F Qiao, QD Wu. Research Progress of Modeling Methods and Scheduling Strategies for Semiconductor Wafer Fabrication. *Acta Electronica Sinica*. 2006.
 - [5] HP Huang, TY Chen. *A new approach to on-line rescheduling for a semiconductor foundry fab – Systems, Man and Cybernetics*. SMC'06. IEEE International Conference.2006; 4727-4732.
 - [6] YD Kim, JU Kim, SK Lim, HB Jun. Due-date based scheduling and control policies in a multiproduct semiconductor wafer fabrication facility. *IEEE Transactionson Semiconductor Manufacturing*. 1998; 11: 155-164.
 - [7] B Ehteshami, RG Petrakian, PM Shabe. Trade-offs in cycle time management: hot lots. *IEEE Transactions Semiconductor Manufacturing*. 1992; 5: 101-106.
 - [8] R Uzsoy, CY Lee, LA Martin-Vega. A review of production planning and scheduling models in the semiconductor industry part II: shop-floor control. *IIE transactions*. 1994; 26: 44-55.
 - [9] ML Spearman, LW DAVID, J WALLACE. CONWIP: a pull alternative to kanban. *The International Journal of Production Research*. 1990; 28: 879-894.
 - [10] LM Wein. Scheduling semiconductor wafer fabrication. *IEEE Transactionson Semiconductor Manufacturing*. 1988; 1: 115-130.
 - [11] O Rose. *CONLOAD a new lot release rule for semiconductor wafer fabs*. Proceedings of the 31st conference on Winter simulation: Simulation-a bridge to the future. 1999; 1: 850-855.
 - [12] MC Wu, S Chen. A cost model for justifying the acceptance of rush orders. *International Journal of Production Research*. 1996; 34: 1963-1974.
 - [13] WP Wang, Z Chen. A neuro-fuzzy based forecasting approach for rush order control applications. *Expert Systems with Applications*. 2008; 35: 223-234.
 - [14] JSR Jang. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactionson Systems, Man and Cybernetics*. 1993; 23: 665-685.
 - [15] RE Fox. The constraint theory. NAA Conference. 1988.
 - [16] MG ELIYAHU. Computerized shop floor scheduling. *The International Journal Of Production Research*. 1988; 26: 443-455.
 - [17] EM Goldratt, RE Fox, G Grasman. *The race*. North River Press Croton. Hudson, NY. 1986; 179.
 - [18] AB Zhang. Hierarchical fuzzy system design based on subtractive clustering and Adaptive Neural-Fuzzy Inference System. *Control Theory and Applications*. 2004; 21: 415-518.
 - [19] H Zhao, Z Jiang, YF Lee, CP Ko, C Luke, LP Lim. *An approach of dynamic bottleneck machine dispatching for semiconductor wafer fab*. *Semiconductor Manufacturing*. ISSM 2007. International Symposium. 2007; 1-4.