A sample path-based method for the digital twin prediction update synchronization problem of unreliable production lines

Yi-Xing Gou
Department of Industrial Engineering, School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, P. R. China, pacpos.yxgou@gmail.com
Andrea Matta
Department of Mechanical Engineering, Politecnico di Milano, Milano 20156, Italy, andrea.matta@polimi.it
Barış Tan
Faculty of Engineering & Faculty of Business, Özyeğin University, Çekmeköy Campus Nişantepe District, Orman Street, Çekmeköy, Istanbul 34794, Türkiye, baris.tan@ozyegin.edu.tr

The digital twin prediction update synchronization problem determines whether or not to update the performance prediction from the digital twin at each observation period depending on the observed state of the physical system. Existing approaches provide solutions for the prediction update problem, but they can be applied only to simple systems. In this study, we propose a sample path-based method to solve the prediction update problem for unreliable production lines composed of multiple machines and finite buffers. The method estimates the performance measure for each synchronization decision by partially observing the state of the system. An optimal state-dependent synchronization policy is determined based on the observed state to balance the prediction bias and the synchronization cost. The results show that partially observing the state of the bottleneck machine, rather than fully observing the state of the system, is efficient for solving the problem without requiring a very long sample path.

Key words: Digital twin synchronization; Prediction update; Sample path; Unreliable production lines

1. Introduction

Digital twins are considered important parts of smart manufacturing to monitor and control physical systems more accurately and optimize manufacturing processes (Tao et al. 2019). The development of IoT technologies makes data collection and communication more efficient, which enables digital twins to capture the behavior of physical systems in increasing detail. A digital twin of a production system involves a lot of variables that reflect the dynamics of the physical system, which increases the difficulty of aligning the digital twin with its physical system. Synchronizing all variables of the digital twin with the data from the physical system may take much time and resources (Modoni et al. 2019, Sargent 2013). Not synchronizing the digital twin with the actual state of the physical system results in significant prediction errors that may cause production loss and management challenges (Zipper and Diedrich 2019, Zipper 2021). Consequently, it is necessary to determine when and how to synchronize the digital twin with its physical system in the most efficient way.

The optimal digital twin synchronization problem for production systems is defined as a stochastic control problem, and the prediction update synchronization problem of an unreliable machine is solved analytically in Tan and Matta (2024). The prediction update problem aims to decide whether or not to update the prediction from the digital twin depending on the observed state of the physical system. A sample path-based method is introduced in Tan and Matta (2024) that can be used to solve the problem. The method requires estimating the system performance for each synchronization decision by observing the state of the system from the sample path. A state-dependent synchronization policy is determined based on the observed state of the system. In the case of a single unreliable machine, the state of the system is fully observable and is either up or down. For an unreliable production line, the state of the system is composed of states of all machines and buffers. If we fully observe the state of the system, the available data in a given sample path that can be used to estimate the performance prediction is limited. This will result in an inaccurate prediction when the number of system states is large. In this study, partial observations of the production lines are used to solve the problem. The main contribution of this work is a sample path-based method that estimates the performance prediction and determines the optimal state-dependent policy based on the partially observed state of the system.

2. Prediction update problem of unreliable production lines

A production line consisting of M unreliable machines and $M-1$ finite buffers is considered in this study. Machine $m_i, i=1,2,\dots,M$ follows the geometric reliability model. Buffer $b_i, i=1,2,\dots,M-1$ has a finite capacity denoted by B_i. The state of the system at time t is a set of the states of all machines and buffers, denoted by $y_t=(\alpha_t^1,\dots,\alpha_t^M,\beta_t^1,\dots,\beta_t^{M-1})$. α_t^i denotes the state of machine $m_i, i=1,2,\dots,M$ at time t, which is 1 if machine m_i is up and 0 if it is down. β_t^i denotes the buffer level of buffer b_i at time t, $\beta_t^i\in\{0,1,\dots,B_i\}$. If not all machines and buffers are observable or not all observations are used to determine the synchronization policy, y_t includes only the states of the machines and the buffers observed or used.

A digital twin, which uses a discrete event simulation model of the physical system, is adopted to predict the throughput of the production line based on the available history. The throughput is evaluated each Δ cycles at times $\Delta,2\Delta,\dots,N\Delta$ until the end of the planning observation period. The throughput within a given time interval $[t,t+\tau)$ is the expected number of parts produced by the production line during this interval denoted by $\mathbb{E}[TH_{t,t+\tau}]$.

Synchronizing the digital twin with the current observations allows obtaining a better prediction of the throughput. This comes at a cost that includes retrieving the data from the production line, executing a simulation experiment, and changing the production resources based on the updated prediction. The synchronization decision at time $n\Delta$ is $\mathbf{u}_n=\mathrm{H}_n$, in which H_n is equal to 1 if the digital twin is synchronized at time $n\Delta$ and 0 otherwise. The synchronization cost of decision $\mathbf{u}_n=\mathrm{H}_n$ at time $n\Delta$ is calculated with $C_{\mathrm{DT}}(\mathbf{u}_n)=c_{\mathrm{H}}\mathrm{H}_n$, where c_{H} is the cost for each synchronization. We consider a state-dependent policy that the synchronization decision is taken depending on both the evaluation period n and the observed state of the system y_n. Not synchronizing the digital twin results in inaccurate prediction, which incurs a cost of prediction bias. The bias cost at time $n\Delta$ denoted with $C_{\mathrm{B}}(n,\mathbf{S}_n,\mathbf{u}_n)$ depends on the difference between the best estimation $R_n^*(y_n)$ with the most recent observations y_n and the estimation from the digital twin based on synchronization decision at time $n\Delta$, $R_n^{\mathrm{DT}}(\mathbf{S}_n,\mathrm{H}_n)$, i.e., $C_{\mathrm{B}}(n,\mathbf{S}_n,\mathbf{u}_n)=c_{\mathrm{B}}(R_n^{\mathrm{DT}}(\mathbf{S}_n,\mathrm{H}_n)-R_n^*(y_n))^2$. The prediction update synchronization problem is to find a policy that determines when to update the throughput prediction from the digital twin to minimize the expected total cost of prediction bias and of synchronizations over N observation periods:

$$\min_{\{\mathrm{H}_n\}} \quad \sum_{n=1}^{N}(c_{\mathrm{H}}\mathrm{H}_n+c_{\mathrm{B}}\mathbb{E}[(R_n^{\mathrm{DT}}(\mathbf{S}_n,\mathrm{H}_n)-R_n^*(y_n))^2]). \tag{1}$$

A sample path-based method is proposed to determine the optimal state-dependent policy for the model described in Equation (1). The method estimates the throughput based on the observed state of the last synchronization at time $k\Delta$, $r_{n,k}^{\mathrm{DT}}(y_k)=\mathbb{E}[TH_{n\Delta,(n+1)\Delta}|y_k]$. A sample path with L periods consists of the states of all machines and buffers during $L\Delta$ cycle times. The state of the system at time t in the sample path is denoted with o_t. If the state of the system at time $l\Delta$ in the sample path is equal to the state of the digital twin at the last synchronization at time $k\Delta$, i.e., $o_{l\Delta} = y_k$, the throughput during the next period after $n-k$ periods at time $l\Delta$ in the sample path can be used to estimate the throughput at time $n\Delta$ based on state y_k. Then $r_{n,k}^{\mathrm{DT}}(y_k)$ can be obtained by using the average of the observations from the sample path. Accordingly, the estimate from the digital twin based on the synchronization decision H_n at time $n\Delta$, $R_n^{\mathrm{DT}}(\mathbf{S}_n, \mathrm{H}_n)$ is evaluated as

$$R_n^{\mathrm{DT}}(\mathbf{S}_n, \mathrm{H}_n) = \begin{cases} r_{n,n}^{\mathrm{DT}}(y_n), & \mathrm{H}_n = 1 \\ r_{n,k}^{\mathrm{DT}}(y_k), & \mathrm{H}_n = 0 \end{cases}. \tag{2}$$

The best estimate of the throughput at time $n\Delta$ is obtained by taking synchronization action with $\mathrm{H}_n = 1$, i.e., $R_n^*(y_n) = r_{n,n}^{\mathrm{DT}}(y_n)$. Then, an optimal state-dependent policy is determined to solve the problem formulated in Equation (1).

3. Numerical results and conclusions

To validate the performance of the sample path-based method, a simulation-based method is also proposed in this work. The method differs from the sample path-based method in that $R_n^*(y_n)$ is evaluated using simulation based on the current full observations of the system at time $n\Delta$. The performance of both methods is compared in two test cases. Some results are shown in Figure 1.

The numerical results show that observing the bottleneck machine is more critical to determining an optimal state-dependent policy than observing other machines in unreliable production lines. As the length of the sample path increases, the average cost and the average number of synchronizations obtained using the sample path-based method almost stay stable. Longer sample paths are not required for the cases analyzed. The simulation-based method yields policies with lower average costs than the sample path-based method, but it takes much more time as the sample path becomes longer. For more complex systems, the sample path-based method is more efficient to apply in solving the prediction update synchronization problem. This will be studied in the future.

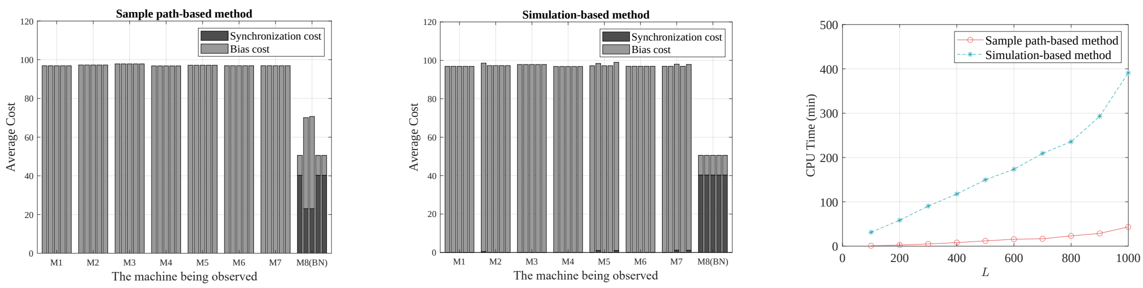

Figure 1. Effect of the observed machine on the average cost obtained using both methods over five experiments ($c_{\mathrm{H}} = 20, L = 1000$), and CPU time conducting one experiment using both methods for different values of L

References

Modoni, G. E., E. G. Caldarola, M. Sacco, W. Terkaj. 2019. Synchronizing physical and digital factory: benefits and technical challenges. *Procedia CIRP* **79** 472–477.

Sargent, R. G. 2013. Verification and validation of simulation models. *Journal of Simulation* **7**(1) 12–24.

Tan, B., A. Matta. 2024. The digital twin synchronization problem: Framework, formulations, and analysis. *IISE Transactions* **56**(6) 652–665.

Tao, F., Q. Qi, L. Wang, A.Y.C. Nee. 2019. Digital twins and cyber-physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering* **5**(4) 653–661.

Zipper, H. 2021. Real-time-capable synchronization of digital twins. *IFAC-PapersOnLine* **54**(4) 147–152.

Zipper, H., C. Diedrich. 2019. Synchronization of industrial plant and digital twin. *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 1678–1681.