Industrial Engineering and Applications
L.-C. Tang (Ed.)

doi:10.3233/ATDE230026

Production System Configuration Design for an Unmanned Manufacturing Factory

Yuting LUO [a], Maoning CHEN [a], Long ZENG [b] and Canrong ZHANG [a,c,1]

[a] *Research Center for Modern Logistics, Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China*

[b] *Division of Advanced Manufacturing, Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China*

[c] *Department of Industrial Engineering, Tsinghua University, Beijing 100084, China*

Abstract. In Unmanned Manufacturing Factory (UMF), we consider a workstation consisting of a machine and several workbenches, each assigned an AGV to perform the parts transporting service. Every workpiece requires multiple processes at the same workbench, and the time for the AGV to transport a part is greater than the time for the machine to assemble it. This operating model is complex, as several workpieces corresponding to several workbenches are in service at the same time. We extract the production system into a Semi-Open Queuing Networks (SOQNs) model. We use four parameters to describe the states of the system and construct the transition rate matrix. We find that the matrix has a particular structure that enables us to solve it with Matrix Geometric Method (MGM). Following a First-Come-First-Served (FCFS) policy, the performance of this production system with unlimited queuing space is evaluated in terms of service intensity, queue length, sojourn time, and throughput. The numerical experiments demonstrate a significant reduction in the sojourn time of the workpiece in the system when the number of workbenches increases to a certain value. Our work can provide important suggestions for designing UMF.

Keywords. A Workstation with a Machine and several Workbenches, Semi-Open Queueing Network, Matrix Geometric Method

1. Introduction

In 2015, the Chinese government proposed the “Made in China 2025” program to promote upgrading the country’s manufacturing sector. Unmanned Manufacturing Factories (UMF) based on information technologies are leading the change in manufacturing especially when facing the challenge of population aging [1] and supply chain variation. Automation is a major feature of UMF, where human work is replaced by intelligent devices. At the same time, unmanned manufacturing represents the most important direction in today’s efforts to increase productivity and flexibility in medium-volume production because of individual and variable demands [2].

[1] corresponding author; crzhang@sz.tsinghua.edu.cn

In UMF, workstations and material handling systems (MHS) can be flexibly configured to meet changing demands. A workstation consisting of a machine and several workbenches can process different products. MHS can do the following operations: loading/unloading, transportation, inspection, and storage/retrieval [3]. In this case, we only consider MHS with transportation function, which is dominated by automatic guided vehicles (AGVs) and conveyors in the workplace. A workpiece requires multiple processes at the same workstation, and the time for the AGV to transport materials may be greater than the time for the machine to assemble it. Existing studies have analyzed cases where only one workbench is configured for the machine, and several AGVs serve one workbench simultaneously. In this paper, we study the case where several workbenches are configured for the machine. Each workbench is served by only one AGV, meaning there will be several workpieces in service at the workstation. This new scenario will be more complex, as shown in Figure1, since more workpieces will be in service simultaneously.

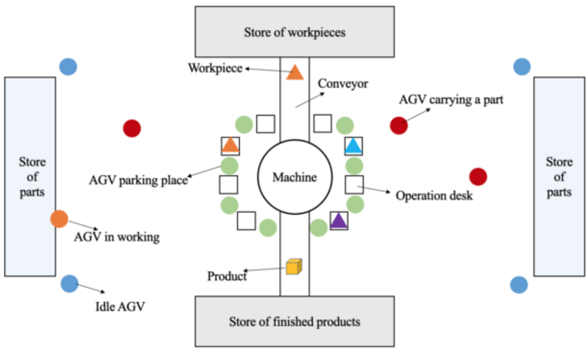

Figure 1. Production system in UMF

Queueing networks are a common method to study the facility layout and performance evaluation of manufacturing systems [4–6]. To handle these problems and derive a better model, we introduce Semi-Open Queueing Networks (SOQNs) to model UMF production processes. SOQNs have been commonly used to evaluate system performance in many fields, including production systems [7], storage systems [8], hospitals [9], restaurants [10], container terminal [11] and interterminal transportation [12]. The current methods for solving SOQN (both single- and multi-class network) can be roughly classified into four categories: (1) matrix geometric method (MGM), (2) aggregation, (3) decomposition, and (4) performance bounds [13]. We solve this problem using MGM because of their special structure in the transition rate matrix [14].

This paper is organized as follows. Section 2 describes the scenario and problem, and the SOQN model is developed in section 3. Numerical experiments are implemented in section 4, and section 5 discusses the conclusion.

2. Problem Description

In the UMF, the workstation consists of one service machine and several workbenches, as shown in Figure.2. The number of workbenches is flexibly configured according to processing requirements, and each workpiece has one AGV to transport parts. The task is to assemble a limited number of parts for the workpiece. The workpieces are transported by conveyer, and parts are transported by AGV.

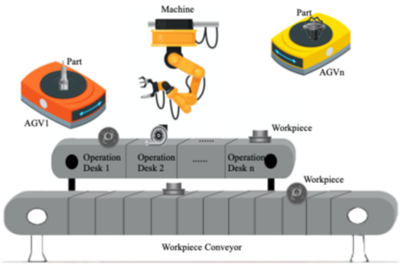

Figure 2. The workstation

Due to the separation of the storage area and the workspace, the time to transport a part is greater than the time of service. In order to improve the productivity of the machine, equip the machine with several workbenches, where some workpieces are waiting for parts, and the machine can serve the workpieces whose parts are available.

2.1. Process Flow

For the workstation, the process will not end until all the workpieces leave the system, and the steps are as follows:

1. The system assigns workpieces to different workstations.
2. The workpiece enters the external queue and waits.
3. When the machine has an idle workbench/AGV, the workpiece will be placed on the idle workbench, and a relationship will be established with the workbench and AGV.
4. The related AGV will only transport the part for the workpiece.
5. AGV carrying a part waits to be served in the part queue when the machine is serving other workpieces.
6. When the machine gets the part from the AGV, it is judged whether it is the last one to be assembled. If it is, when finishing the service, the AGV unbinds the workpiece immediately. Otherwise, return to step 4, and the AGV will continue transporting the part.
7. Repeat the above steps until all the workpieces are served.

2.2. Key Nodes

Binding: workpiece arrives and enters the external queue. When the workbench is free, bind the workpiece at the top of the external queue to the free workbench and AGV. This AGV transports parts for this workpiece only since the ratio of AGV to the workbench is 1:1.

Service: the AGV carries a part into the part queue and waits. The machine takes the first part of the part queue and assembles it on the related workpiece. If the machine is idle, the part queue is empty.

Unbinding: when all parts are assembled, the workpiece leaves the workstation immediately. The workbench and AGV are released and can serve the next workpiece.

3. Semi-Open Queuing Network Model

3.1. Assumptions

Our SQON model is based on the following assumptions:

1. The workpieces are available, and assumed to be i.i.d. The arrival interval follows an exponential distribution.
2. The size of external queue is assumed to be infinite.
3. The workpiece is bound to one AGV only, and the AGV will transport all required parts for it.
4. The time for an AGV to transport a part is assumed to be i.i.d and follows an exponential distribution.
5. Only one event happens at a time, and the system will be stable eventually.
6. All queues follow the FCFS policy.
7. Each workpiece is assumed to require the same number of parts.
8. When assembling different parts, the setup time of the machine is ignored.
9. The capacity of a workbench is set to 1. This means that a workbench can only load one workpiece at a time.
10. The capacity of AGV is set to 1. This means that an AGV can transport only one part at a time.
11. The time to transport a part is greater than the time of a service.
12. The AGV is fully charged and will never break down, and the machine will not break down.
13. The parts are never out of stock.
14. The workpieces sent in or out through the conveyor are pickup and placed on the workbench by the machine. The processing time of this operation can be ignored.

3.2. Working Flow of UMF

Through the cooperation of the machine, conveyor, workbenches and AGVs, the workstation assembles several parts on the workpieces. The workpiece arrives and enters the external queue to wait for service until there is an idle workbench and an AGV. The workpiece enters the workpiece buffer waiting for service, and the AGV bound to it enters the part buffer after getting the part. Moreover, the service order of workpieces follows the FCFS policy of parts. If the workpiece still has parts to be assembled, it returns to the buffer. Otherwise, it leaves the workstation and releases the workbench and AGV.

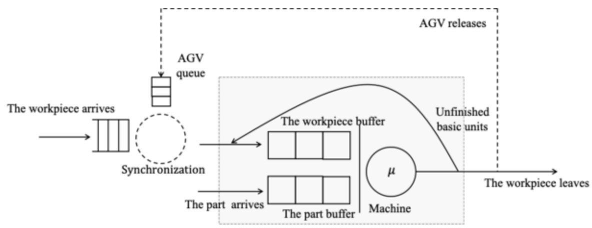

Figure 3. Semi-open queueing network model

Throughout the operation, the flow of workpieces is continuous. There is a machine, and the number of workbenches and AGVs will remain the same since they are determined. Because internal resources are limited and external orders are unlimited, orders need to wait for available resources to be served. We construct SOQN as shown in Figure.3, which clearly and intuitively shows the workflow of the workstation.

3.3. States of the Workstation

We assume that a workstation has one machine and x workbenches, so there are at most x AGVs working in the system simultaneously. Meanwhile, the machine assembles y parts for each workpiece. Both the arrival of workpieces and parts follow Poisson Processes. The workpieces arrive at a rate λ_m. When there is only one AGV transporting parts, the parts arrive at a rate λ_l. The machine has an exponential service time of $1/\mu$. The number of orders is sufficiently large, so the external queue will never explode. The state of the workstation is denoted by a 4-parameter vector (om, l, m, sta), where $om \geqslant 0$ is the number of workpieces in the external queue, and $0 \leqslant l \leqslant x+1$ is the number of parts in the workstation, and $0 \leqslant m \leqslant x$ is the number of workpieces in the workstation, and sta shows the number of parts left to be assembled per workpiece. Specifically, sta is denoted by a sub-vector $(s_1, s_2, \ldots, s_i, \ldots, s_m)$, where s_i means the number of parts remaining to be assembled of the ith workpiece. We analyze the changes of states from the following three situations:

- Workpiece arrival. When $m < x$, the arriving workpiece directly enters the workpiece buffer, and m becomes $m+1$, and the sta changes accordingly. When $m = x$, the arriving workpiece enters the external queue, and om becomes $om+1$, other parameters remain unchanged.
- Part arrival. The arriving part directly enters the part buffer, and the number of parts l becomes $l+1$. Other parameters remain unchanged.
- Machine serving. The machine is busy only when $l > 0$. When the part to be assembled is the last of the workpiece ($\exists s_i = 1$), the workpiece will leave after completion. At this time, if $om = 0$, this workbench will be idle, and the number of workpieces in the system, m, will shift to $m-1$, and $sta(s_1, ..., s_m)$ will be $sta(s_1, ..., s_{m-1})$. If $om > 0$, the workpiece of the external queue enters the workpiece buffer immediately, and om will become $om-1$, and $sta(s_1, ..., s_m)$ will be $sta(s_1, ..., s_m^{'})$. When the part to be assembled is not the last of the workpiece, the $sta(s_1, \ldots, s_i, \ldots, s_m)$ turns into $sta(s_1, \ldots, s_i - 1, \ldots, s_m)$, and other parameters remain unchanged.

3.4. Generator Matrix

The transition rate caused by the arrival of workpieces is λ_m, and the transition rate caused by the arrival of parts or services varies according to the current state.

- The transition rate caused by the arrival of part is denoted by $k\lambda_l$, which depends on the number of AGVs that transport parts to workstations. Parameter k is the number of workbenches in use minus the number of AGVs waiting in the part buffer.

- The transition rate caused by the service is denoted by $P_s\mu$, which depends on the proportion of different quantities of parts in sta with the notation P_s. Because of several workpieces in the workstation, the next state will be multiple (at most m types, according to the number of workpieces in the workstation). We calculate the possibility of each sta in the next state. For example, $sta(3,2,2)$ means three workpieces are in the system. One of the workpieces has three parts to be assembled, while the other two have two remaining parts. Therefore, if the next state is assembling parts, there are two kinds of sta in the next state where $sta(2,2,2)$ is one-third of the probability and $sta(3,2,1)$ is two-thirds of the probability.

By adjusting the order of each state, it is found that the generator matrix has a special block structure. Since the generator matrix is infinite, three block matrices are constantly repeated after the initial stage. The infinite generator matrix:

$$Q = \begin{pmatrix} B_{00} & B_{01} & 0 & 0 & \cdots \\ B_{01} & A_1 & A_2 & 0 & \cdots \\ 0 & A_0 & A_1 & A_2 & \cdots \\ 0 & 0 & A_0 & A_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

in which the B_{00} is square, the dimension of B_{01} and B_{10} are related to the dimension of B_{00} and A_1; and submatrices A_0, A_1, A_2 are square with the same dimension and need not have the same size as B_{00}. Furthermore, the submatrix on the diagonal in Q represents that the number of workpieces does not change, the submatrix on the right side of the diagonal represents an increase in the number of workpieces, and the submatrix on the left side of the diagonal represents a decrease in the number of workpieces. The submatrices B_{00}, B_{01}, B_{10} describe boundary conditions when the number of workpieces in the workstation is less than the number of workbenches or even equal to 0, $0 \leqslant m < x$. The submatrices A_0, A_1, A_2 reflect that the transition rates are the same from level to level when workbenches are not idle.

3.5. Matrix Geometric Method (MGM)

This queueing network is a quasi-birth-death (QBD) process because the generator matrix Q has the same block tridiagonal structure. We get a numerical solution to this problem by using MGM proposed by Neuts [15].

First, we need to ensure that this queueing network is ultimately stable, which means that the probability of increasing the number of workpieces is strictly less than that of reducing the number of workpieces. A similar condition exists for a QBD process to be ergodic [9]. Assuming π_A denotes the stationary distribution of the infinitesimal generator $A = A_0 + A_1 + A_2$. For a QBD process to be ergodic, the condition $\pi_A A_2 e < \pi_A A_0 e$. must hold:

As mentioned above, A_2 represents an increase in the number of workpieces, while A_0 represents a decrease in the number of workpieces.

We obtain stationary distribution from $\pi Q = 0$, and π has the same dimension of Q: $\pi = (\pi_0, \pi_1, \pi_2, ..., \pi_i, ...)$, where π_i is a sub-vector, and π can be calculated via $\pi e = 1$.

4. Numerical Simulation and Result

4.1. Notations

The notations are shown in table 1.

Table 1. NOTATIONS AND PARAMETER DEFINITIONS

Notations	Definitions
$P_m(tm)$	Probability of tm workpieces in the system
$P_l(l)$	Probability of l parts in the system
L_m, L_l	Average number of workpieces/parts in the system
$\widetilde{L_m}$	Average number of workpieces in the external queue
$\widetilde{L_l}$	Average number of parts in part buffer
W_m, W_l	Average time that workpieces/parts spend in the system
$\widetilde{W_m}$	Average time that workpieces spend in external queue
$\widetilde{W_l}$	Average time that parts spend in part buffer
N	Maximum number of parts in the system, $N = x + 1$
a	Number of AGV transporting parts, $a = N - l$
$\lambda_l(a)$	Arrival rate of parts when there are a AGVs transporting parts
ρ	Probability that machine is busy
TH	Throughput of the system

4.2. Analytic Measures

- Service intensity of the machine

Service intensity of the machine means the probability that the machine is busy, reflecting the machine utilization. The machine will be idle when there are no parts in the system. Machine busy probability ρ is denoted by $\rho = 1 - P_l(0)$.

- Average number of parts

The parts queue can be considered a capacity-constrained queueing system with a single service station, where the number of parts in the system is l and the maximum value of l is $N(N = x+1)$.If the number of part buffers equals x, this means all the AGVs are waiting in the part buffer carrying a part, and no parts will arrive at this time.

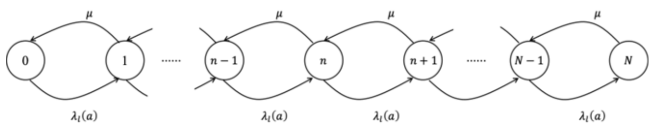

Figure 4. Transition relationship when the number of parts changing

The arrival of parts is influenced by the current state of the system, especially related to the number of AGVs $a(a = N - 1)$. The parts arrive at a rate $\lambda_l(a)$ following a Poisson Process. When the system is stable, the transition relationship between the states with different number of parts in the system is shown in Figure.4.

The average number of parts in the system: $L_l = \sum_{l=0}^{N} lP_l(l)$
The average number of parts in part buffer: $\widetilde{L_l} = \sum_{l=1}^{N}(l-1)P_l(l) = L_l - \rho$

- Average time of parts spend

From Little's law, the average length of the queue is denoted by: $L_l = \lambda_e W_l$ In the above equations, λ_e is the average arrival rate of parts, and since the service intensity ρ is defined as: $\rho = \frac{\lambda_e}{\mu} = 1 - p_l(0)$
So the average arrival rate is given by: $\lambda_e = \mu \times (1 - P_l(0))$
The average time that parts spend in the system is denoted by: $W_l = \frac{L_l}{\lambda_e} = \frac{L_l}{\mu(1-p_l(0))}$
The average time that parts spend in part buffer is: $\widetilde{W_l} = \frac{\widetilde{L_l}}{\lambda_e} = \frac{\widetilde{L_l}}{\mu(1-P_l(0))}$

- Average number of workpieces

The average number of workpieces in the system is: $L_m = \sum_{m=0}^{\infty} mP_m(m)$, while The average number of workpieces in external queue can be denoted by$\widetilde{L_m} = \sum_{m=x+1}^{\infty}(m - x)P_m(m)$

- Average time of parts spend

The average time that basic parts spend in the system is denoted by the following equation: $W_m = \frac{L_m}{\lambda_m}$ whiled in external queue this time is denoted by: $\widetilde{W_m} = \frac{\widetilde{L_m}}{\lambda_m}$

- Throughput of the system

In a steady state, throughput is the average number of workpieces completed service per unit of time, and it also reflects the departure rate of workpieces. The last action before all the workpieces leave the system is that the machine assembles the last part of the workpiece. Since multiple parts are to be assembled for each workpiece, the time interval between workpieces leaving is not uniform. From the timeline as shown in Figure.5, d_i refers to the departure of the ith workpiece, and the small time interval before d_i in the figure represents the service time of the last part of the ith workpiece.

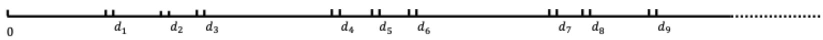

Figure 5. Time line

In addition, we identify the state that the workpiece is about to complete its service, which means that there are parts in the system and at least one workpiece has only one part left to assemble. Since the external queue is infinite, when the system gets a stationary distribution, we assume that the probability of the nth (n is large enough) workpiece leaving the system is P_d. The average time for the machine to assemble a part is $1/\mu$ according to the assumptions in section 3. When the nth workpiece leaves, the total machine running time is $T = n * \frac{1/\mu}{P_d}$,the average time interval t_d for the departure of the workpieces is $t_d = \frac{T}{n} = \frac{1/\mu}{P_d}$ and the throughput is $TH = \frac{1}{t_d} = \mu P_d$.

4.3. Numerical Experiment

The algorithm is programmed in MATLAB and runs on a computer with dual-core Intel Core i5, a 2.7GHz processor, and 8GB RAM.

When assembling three parts for each workpiece $(y = 3)$, keeping other parameters unchanged and only changing the number of workbenches, the arrival rate of workpieces, and the arrival rate of parts, respectively, the results are shown in the following three tables.

From table 2, we find the number of workpieces in the external queue and the average time that workpieces spend in the system decrease as the number of workbenches increases, especially when the number of workbenches increases from 3 to 4. It is shown from table 3 that the service intensity of the machine increases with the arrival rate of the workpieces, but the number of workpieces in the external queue increases, especially when the workpiece arrival rate increases from 1/8 to 1/7.5. When the arrival rate of parts increases, as shown in table 4, the number of workpieces in the external queue decreases. From the experimental data in the three tables, the throughput is equal to the workpieces arrival rate.

Table 2. Results With Varied Number Of Workbench

$y = 3, \lambda_m = 1/11, \lambda_l = 1/7, \mu = 1/3$										
x	ρ	L_l	$\widetilde{L_l}$	W_l	$\widetilde{W_l}$	L_m	$\widetilde{L_m}$	W_m	$\widetilde{W_m}$	TH
3	0.8182	1.8382	1.0201	6.7402	3.7402	38.8256	35.8824	427.0821	394.7062	0.0909
4	0.8182	2.2428	1.4226	8.2236	5.2236	6.6203	3.2699	72.8238	35.9689	0.0909
5	0.8182	2.6337	1.8156	9.6571	6.6571	5.5901	1.8468	61.4908	20.3153	0.0909
6	0.8182	2.9949	2.1767	10.9812	7.9812	5.4485	1.3433	59.9331	14.7759	0.0909

Table 3. Results With Varied Arrival Rate Of Workpieces

$y = 3, x = 4, \lambda_l = 1/7, \mu = 1/2$										
λ_m	ρ	L_l	$\widetilde{L_l}$	W_l	$\widetilde{W_l}$	L_m	$\widetilde{L_m}$	W_m	$\widetilde{W_m}$	TH
1/9	0.6667	1.4656	0.799	4.3969	2.3969	4.8124	1.6664	43.3116	14.9975	0.1111
1/8.5	0.7059	1.5924	0.8866	4.5119	2.5119	6.1275	2.7599	52.084	23.4588	0.1176
1/8	0.75	1.7395	0.9895	4.6387	2.6387	9.3325	5.7111	74.6602	45.6887	0.125
1/7.5	0.8	1.9115	1.1115	4.7788	2.7788	35.6036	31.6891	267.0271	237.6682	0.1333

Table 4. Results With Varied Arrival Rate Of Parts

$y = 3, x = 3, \lambda_m = 1/11, \mu = 1/3$										
λ_l	ρ	L_l	$\widetilde{L_l}$	W_l	$\widetilde{W_l}$	L_m	$\widetilde{L_m}$	W_m	$\widetilde{W_m}$	TH
1/7	0.8182	1.8382	1.0201	6.7402	3.7402	38.8256	35.8824	427.0821	394.7062	0.0909
1/6.5	0.8182	1.8756	1.0574	6.877	3.877	15.4991	12.6508	170.4906	139.1586	0.0909
1/6	0.8182	1.9157	1.0975	7.0243	4.0243	10.2186	7.462	112.4041	82.0815	0.0909
1/5.5	0.8182	1.9591	1.1409	7.1833	4.1833	7.882	5.2137	86.7023	57.3511	0.0909

5. Conclusion

As demand becomes more individualized, the production changes from large batches of fewer items to small batches of multiple items. In addition, to save land resources, factories tend to use multifunctional machines, which means that multiple products can be processed on the same machine by adjusting parameters. We combine multifunctional machines with MHS to build workstations in UMF. We extract the problem into a SOQN model and solve it by MGM. The performance of the workstation is analyzed by indicators such as machine service intensity, the average number of workpieces, the average number of parts, and throughput. It provides managerial insights for the design of UMF.

Acknowledgement

This work is supported by the National Key Research and Development Program of China under Grant No.2018AAA0101705, and the Shenzhen Municipal Science and Technology Innovation Committee under Grant No.WDZC20200821140447001.

References

[1] Chen R, Xu P, Song P, Wang M, He J. China has faster pace than Japan in population aging in next 25 years. BST. 2019;13(4):287-291. doi:10.5582/bst.2019.01213

[2] Zelený J, Praha V. Unmanned Technology and Control Strategies in Flexible Manufacturing Systems with Random Flow of Parts and Tools. CIRP Annals. 1985;34(1):391-395. doi:10.1016/S0007-8506(07)61796-X

[3] W. Muller, Integrated Materials Handling in Manufacturing. London: IFS Publications, 1985.

[4] Pourvaziri H, Pierreval H. Dynamic facility layout problem based on open queuing network theory. European Journal of Operational Research. 2017;259(2):538-553. doi:10.1016/j.ejor.2016.11.011

[5] Chen C, Tiong LK. Using queuing theory and simulated annealing to design the facility layout in an AGV-based modular manufacturing system. International Journal of Production Research. 2019;57(17):5538-5555. doi:10.1080/00207543.2018.1533654

[6] Zhang HY, Chen QX, Smith JM, Mao N, Liao Y, Xi SH. Queueing network models for intelligent manufacturing units with dual-resource constraints. Computers & Operations Research. 2021;129:105213. doi:10.1016/j.cor.2021.105213

[7] Heragu SS, Srinivasan M. Analysis of manufacturing systems via single-class, semi-open queuing networks. International Journal of Production Research. 2011;49(2):295-319. doi:10.1080/00207540903527378

[8] Duan G, Zhang C, Gonzalez P, Qi M. Performance evaluation for Robotic Mobile Fulfillment Systems with time-varying arrivals. Computers & Industrial Engineering. 2021;158:107365. doi:10.1016/j.cie.2021.107365

[9] Carmen R, Van Nieuwenhuyse I, Van Houdt B. Inpatient boarding in emergency departments: Impact on patient delays and system capacity. European Journal of Operational Research. 2018;271(3):953-967. doi:10.1016/j.ejor.2018.06.018

[10] Roy D, Bandyopadhyay A, Banerjee P. A nested semi-open queuing network model for analyzing dine-in restaurant performance. Computers & Operations Research. 2016;65:29-41. doi:10.1016/j.cor.2015.06.006

[11] Roy D, De Koster R, Bekker R. Modeling and Design of Container Terminal Operations. Operations Research. 2020;68(3):686-715. doi:10.1287/opre.2019.1920

[12] Mishra N, Roy D, van Ommeren JK. A Stochastic Model for Interterminal Container Transportation. Transportation Science. 2017;51(1):67-87. doi:10.1287/trsc.2016.0726

[13] Roy D. Semi-open queuing networks: a review of stochastic models, solution methods and new research areas. International Journal of Production Research. 2016;54(6):1735-1752. doi:10.1080/00207543.2015.1056316

[14] Stewart WJ. Probability, Markov Chains, Queues, and Simulation: The Mathematical Basis of Performance Modeling. Princeton University Press; 2009.

[15] M. F. Neuts, Matrix-Geometric Solutions in Stochastic Models. Bal- timore, MD: Approach. JohnsHopkins University Press, 1981.