International Journal of Production Research, 2018
Vol. 56, No. 22, 6956–6970, https://doi.org/10.1080/00207543.2017.1421789

Differentiated service policy in smart warehouse automation

Zijian He, Vaneet Aggarwal* and Shimon Y. Nof

School of Industrial Engineering, Purdue University, West Lafayette, IN, USA

(*Received 24 July 2017; accepted 18 December 2017*)

Smart warehouse automation has emerged as an effective, competitive solution for suppliers and distributors. With the increasing demand for physical storage and distribution services, suppliers and service providers are challenged to respond not only effectively, but with minimal latency. Differentiated service levels for different classes of customer orders have not yet, however, been developed for physical storage and retrieval. In this paper, in the context of smart warehouse automation services, a novel policy, called Differentiated Probabilistic Queuing (DPQ) is developed for servicing customers' orders by Automated Guided Vehicles (AGV). Applying the DPQ policy, the average overall latency of each customer order, the mean overall processing time of this customer's orders in the smart warehouse automation system, is characterised under Poisson customer order arrival patterns. The weighted average latency of all customer orders is optimised over the choice of (1) storage assignment and (2) DPQ policy. Due to the existence of two types of variables, Alternating Minimisation method is applied to solve this joint optimisation problem. Compared with a combination of the classical turn-over rate storage assignment method and FCFS policy, the new approach yields 19.64% lower (better) objective function value with statistical significance. Numerical analysis results also indicate, as expected, that when the smart warehouse system resources become more limited, and the price difference among different classes of customer orders increases, the improvement becomes even more significant.

Keywords: joint optimisation; order picking planning; storage assignment; supply network; weighted queuing

1. Introduction

Smart warehouse automation has emerged as an effective, competitive solution for suppliers and distributors (Trebilcock 2003; Zhong et al. 2015; Ries, Grosse, and Fichtinger 2017; Zhou et al. 2017). It enables cost-effective end-to-end warehouse services with better control and execution, more efficient storage and retrieval (De Koster, Van der Poort, and Wolters 1999; Lee and Srisawat 2006; Bozer and Kile 2008; Yu and De Koster 2012), reduced errors and minimal delay (Wang, McIntosh, and Brain 2010; Li, Moghaddam, and Nof 2015; Culler and Long 2016).

With better effectiveness and control enabled by smart warehouse automation, better customer-centric services are possible, and can serve as strategic advantages. One of the greatest challenges for suppliers and service providers is the need to respond with minimal latency. A useful way to distinguish suppliers is by offering a higher level of customer-oriented services: It creates a practice of diversified pricing strategies for delivery. For those customers who are more sensitive to price, they will be offered a lower price with relatively long delivery times, and vice versa. Warehouse, as an important element of a supply network, has major impact on the delivery time. A critical connection between the processing time of an order in warehouse, and delivery time is that the faster an order can be retrieved, (1) the sooner it is available for shipping and (2) it lowers the probability for any potential delay in the delivery network (De Koster, Le-Duc, and Roodbergen 2007). As part of smart warehouse automation, most online retailers have changed their warehouse pick-and-pack operations from human-to-goods to goods-to-human; they use robotic, Automated Guided Vehicles (AGVs) in an effort to speed up operations. With this type of automation, the first approach for providing price-based differentiated services for the incoming orders from different classes of customer orders is developed in this article. We assume that there are differentiated classes of customer orders (price based), and wish to optimise the service latency, which is a combination of queue waiting time and the AGV retrieval time.

We develop here a novel strategy, Differentiated Probabilistic Queuing (DPQ), for scheduling orders to AGVs. In this strategy, each order is probabilistically assigned (with certain probabilities which can be optimised) to be served by an AGV. As a special case, some AGVs serve only certain classes of customer orders, but having this probabilistic approach enables more flexibility based on the order frequencies of different classes of customer orders to optimise the

*Corresponding author. vaneet@purdue.edu

overall system objectives. Therefore, in general, orders from customers who are willing to pay more will have shorter expected waiting time than orders from those who are willing to pay less. Thus, service quality can be controlled fairly by optimising the probability terms.

Under the DPQ policy, the average overall latency of each customer order is defined as follows. It is the mean overall processing time of the given customer's orders in the warehouse automation system, under Poisson order arrival process assumption. This measure is derived under certain simplifying assumptions explained below. The overall processing time is a function of both the storage assignment and the probability terms of the DPQ policy. The key idea is that the orders with a higher class can be placed closer, as well as obtain priority from the AGVs. We consider system objective function, which is the weighted processing time for each customer order, where the weight is based on the pricing strategy and order frequencies. A joint optimisation of the system objective is formulated over both the storage assignment and the probability parameters of the DPQ policy. In order to solve this joint optimisation, Alternating Minimisation (AM) algorithm is used.

DPQ policy probabilistically assigns the orders to different AGVs, where the probabilities can be optimised, thus giving more flexibility as compared to the standard priority queueing (which is a special case by having the weight of the lower priority class infinitely smaller than that of the higher priority class). As an example, if one customer pays \$100 and other pays \$99, the weighted scheme will be able to provide weighted resources rather than a strict priority where lower paying user does not have to wait significantly larger time, and thus is able to achieve weighted fairness. Such improved flexibilities by weighted schemes are also depicted in Xiang, Aggarwal, et al. (2017) and Xiang, Lan, et al. (2017), where it is shown that extremely large waiting times for lower service classes can be alleviated using optimised weighted scheduling approaches. This leads the proposed approach to achieve better weighted latency as compared to priority queueing. Further, we note that the closed form results obtained for the policy make the approach more amenable for performing optimisation.

The main contributions of this research are as follows. First, a novel order picking planning policy of warehouse automation system, DPQ policy, is developed for differentiating service levels for different classes of customer orders. Using the DPQ policy in warehouse automation system, the average overall processing time of the customers' orders is derived. Second, a joint optimisation of the weighted latency of all customer orders is formulated over the parameters of the storage assignment and the probability terms of DPQ policy. This optimisation problem is solved using an alternating minimisation algorithm, where each sub-problem is solved using Simulated Annealing algorithm.

Compared with a combination of the classical turn-over rate storage assignment method, and First Come First Serve (FCFS) policy, the new approach yields 19.64% lower (better) objective function value (on average), with statistical significance. Numerical analysis results demonstrate that the joint optimisation over the two sets of variables yields significant improvements, over choosing one as a baseline strategy and choosing the other by our optimisation method.

The remainder of the article includes: Contributions and limitations of previous research, summarised in Section 2. The system model and assumptions of the warehouse automation system are described in Section 3, with the new DPQ policy and formulation of the joint optimisation problem. Next, the AM algorithm is presented for solving the joint optimisation problem. The performance of the developed algorithm is evaluated numerically in Section 4. Section 5 includes the conclusions.

2. Related work

This article considers the joint optimisation of both the storage assignment and the order picking planning policy. Various types of storage assignment method in warehouse automation system have been studied in previous research, e.g. based on bill of material information (Xiao and Zheng 2010); based on data mining (e.g. Pang and Chan 2017). Three typical approaches (Table 1) are: Full-turnover storage (e.g. Yu and De Koster 2009), class-based storage (e.g. Petersen and Aase 2004), and product affinity-based heuristic (e.g. Li, Moghaddam, and Nof 2015). None, however, has addressed physical storage considering classification of ordered products for differentiated service levels in physical product retrieval, which is the focus and original contribution of this article.

For order picking planning policy, previous researchers have developed routing policies for picker-to-part systems (e.g. Hwang, Oh, and Lee 2004); K-means Batching (e.g. Hsieh and Huang 2011); order picking based on storage assignment and travel distance estimation (e.g. Tsai, Liou, and Huang 2008; Battini et al. 2015; Cheng, Liao, and Hua 2017; Epp, Wiedemann, and Furmans 2016; Buonamico, Muller, and Camargo 2017); human-factors based algorithms (e.g. Grosse, Glock, and Neumann 2017), and Self-Organisation Map Batching (e.g. Ma and Zhao 2014) to improve the system performance in total travel distance and average picking vehicle utility. Inspired by the Collaborative Control Theory (Nof 2007), Zhang, Zhong, and Nof (2015) developed a fuzzy collaborative intelligence-based algorithm (following Azadeh et al. 2015) to optimise the assignment plans. There is no prior work in warehouse automation systems

Table 1. Storage Automation Policies: Previous Research (sample), and the New Approach.

Storage automation policies		Classify orders for differentiated services?	Jointly optimise storage assignment and order picking planning?	Physical or digital storage system?
Storage Assignment	Full-turnover storage (Yu et al. 2009)	No	No	Physical
	Class-based Storage (Petersen and Aase 2004)	No	No	Physical
	Product Affinity-based Storage (Li, Moghaddam, and Nof 2015)	No	No	Physical
	Assignment considering Bill of Material (Xiao and Zheng 2010)	No	No	Physical
	Data Mining Based Location Assignment (Pang and Chan 2017)	No	No	Physical
Order Picking Planning	Routing Policies for Picker-to-Part System (Hwang, Oh, and Lee 2004)	No	No	Physical
	K-means Batching (Hsieh and Huang 2011)	No	No	Physical
	Self-Organisation Map Batching (Ma and Zhao 2014)	No	No	Physical
	Fuzzy Collaborative Planning (Zhang, Zhong, and Nof 2015)	No	No	Physical
	Order Picking by Storage Assignment and Travel Distance (Battini et al. 2015)	No	Yes	Physical
Networking and Cloud Storage	Start-time Fair Queuing (Goyal, Vin, and Chen 1997)	No	No	Digital
	Efficient Fair Queuing (Stiliadis and Varma 1998)	No	No	Digital
	Joint Optimisation of Placement of Encoded Chunks and Scheduling Policy (Xiang et al. 2016)	No	Yes	Digital
	Weighted Queuing for Differentiated Services in Erasure-coded Storage System (Xiang, Aggarwal, et al. 2017; Xiang, Lan, et al. 2017)	Yes	Yes	Digital
Approach of This Article	Joint Optimisation of Storage Assignment and Differentiated Queuing Policy in Warehouse Automation System	Yes	Yes	Physical

(Table 1) that deals with physical storage and provides differentiated service levels to different classes of orders in order picking planning, as developed in the research described in this article.

The probabilistic queuing policy and the differentiated objective function in this research are motivated by the policies in networking and cloud storage research, which focuses on digital storage. For instance, a start-time fair queuing algorithm for packet switching network (Goyal, Vin, and Chen 1997), an efficient fair queuing algorithm for packet switching network (Stiliadis and Varma 1998), three-level approaches for differentiated services in measuring Web quality of service (Lee and Park 2009), joint optimisation of encoded chunks placement while optimising scheduling policy of erasure-coded storage system with arbitrary service time distribution, and consisting of multiple heterogeneous files (Xiang et al. 2014, 2016). Probabilistic scheduling is used by the latter, and the placement and access of contents are optimised for minimising access latency in distributed storage systems. This particular idea is extended in this article for physical warehouse systems, where we use a probabilistic strategy for order pickup. Differentiated services have been considered in many areas, including wireless networks (Chen and Mohapatra 1999; Veres et al. 2001; Le, Hossain, and Alfa 2006), cloud computing (Grit and Chase 2008; Li 2009; Rao et al. 2013), distributed scheduling (Jin, Chase, and Kaur 2004; Shue, Freedman, and Shaikh 2012; Aggarwal et al. 2017; Xiang, Lan, et al. 2017), machine scheduling (Lenstra, Kan, and Brucker 1977; Weng, Lu, and Ren 2001; Murray, Chao, and Khuller 2016), supply chain (Morash and Clinton 1998; Hilletofth 2009), smart grids (Deshpande, Kim, and Thottan 2011; Bitar and Low 2012; Negrete-Pincetic and Meyn 2012). The differentiated services have been studied even dating back to 1950s (Smith 1956).

An analytic upper bound of erasure-coded digital storage system using weighted queuing to provide differentiated services (Xiang, Aggarwal, et al. 2017, Xiang, Lan, et al. 2017) is also useful for our new approach: Probabilistic queuing approach is developed for order picking planning, while jointly optimising the storage assignment and order picking planning, differentiating service levels for different classes of customer orders in the warehouse automation system.

3. Formulation of joint optimisation of storage assignment and differentiated probabilistic queuing policy

3.1 *System model and assumptions of a warehouse automation system*

The assumptions about the warehouse automation system considered in this article are as follows.

(1) There are differentiated classes of customer orders (price based). Incoming orders of a product from each customer arrive following a Poisson process. Poisson distribution is the common assumption for random request arrival patterns, and has been used widely for incoming orders of online retailers, see for example Xu (2005); Schneider and Klabjan (2013). This assumption can hold in general for shorter time durations (e.g. few hours), as seen in Hill, Seifbarghy, and Smith (2007). Thus, the results in this paper can be run repeatedly for shorter durations where the arrival rates can be predicted.
(2) Inventory is always available to fill in all the available rows of each shelf. Further, the replenishing time to refill an empty space is not considered. This is assumed because the key focus of this work is on the order retrievals when the inventory is not a bottleneck.
(3) We assume that each incoming order requests for a single product. This can be extended to the case when an order requests for multiple products, by sub-dividing the order into multiple sub-orders, where each sub-order is for a single product.
(4) An order will wait in queue until the assigned AGV is able to serve it. This is assumed because the key bottleneck considered in the paper is the number of AGVs. Thus, if the AGVs are serving other orders, the incoming orders have to wait in the queues.
(5) One AGV can process only one order at a time. This is a simplifying assumption for the analysis, and considering further extensions are left for future work.
(6) Even though real warehouse automation system has a 3D layout, we assume that the warehouse automation system has flat layout. This assumption simplifies the analysis, while can be readily extended. This assumption is also made in many other works in the area, such as the research works by Fuentes Saenz (2011), Liu (2012) and Li, Moghaddam, and Nof (2015).
(7) Loading and unloading times are negligible since we consider AGVs availability as the main bottleneck such as the research assumptions by Fuentes Saenz (2011) and Liu (2012). Further, this time can be directly incorporated in the proposed formulation by adding these times in the service time of the order.

Note that our assumptions, while simplified for practicality, are realistic based on several on-going research projects on warehouse logistics with companies, see for example Wurman, D'Andrea, and Mountz (2008), Hu and Chang (2010), Vliet et al. (2014), Fox13 (2015), Rowley and Battles (2015), and Vliet et al. (2016). Based on the above assumptions, more detailed parameters of this system are defined as follows:

(a) The system has in total D classes of customer orders with different priorities, based on the price the ordering customer has paid. The set of classes of customer orders is denoted by $\boldsymbol{C}$, with $|\boldsymbol{C}| = D$.
(b) Each row of each shelf can only store one type of product, and one type of product can only be stored in a given location. Suppose the warehouse has a total of M shelves and each shelf has in total of J rows, this warehouse has in total $M \cdot J = I$ types of products. The set of shelves is denoted by $\boldsymbol{H}$. The set of rows is denoted by $\boldsymbol{R}$. The set of products is denoted by $\boldsymbol{P}$, where $|\boldsymbol{H}| = M$, $|\boldsymbol{R}| = J$ and $|\boldsymbol{P}| = I$. The different set notations, their cardinalities, and the set elements are also represented in Table 2.
(c) The system has N AGVs, the speed of all AGVs are the same, which is denoted by s. The set of AGVs is denoted by $\boldsymbol{A}$.
(d) The distance between the loading zone and the closest rows of all the shelves is fixed, L. The length of the loading zone is assumed to be included in L. Distance between two rows is b.

Average processing time is the expected time from the arrival of an order, to the time when the requested product of the order is successfully fetched back to the loading zone by the AGV. Let $\mathbf{O} = \{\mathbf{o_1}, \ldots, \mathbf{o_{DI}}\}$ be the set of orders, where the order is $o_{p+(d-1)I}$ corresponds to order of customer class $d = \{1 \ldots, D\}$ requesting product $p = \{1 \ldots, I\}$. Let L_{o_f} represent the average processing time of a type of order $o_f \in \mathbf{O}$. This is the sum of expected waiting time and expected retrieval time, and is given as

$$L_{o_f} = t_{waiting,o_f} + t_{retrieval,o_f}, \tag{1}$$

where $t_{waiting,o_f}$ is the expected waiting time in the queue, and $t_{retrieval,o_f}$ is the average time for retrieving the order using an AGV.

Table 2. Notations for different sets used in this paper.

Name	Notation	Cardinality	Set element notation
Class of customer orders	C	D	d
Set of shelves	H	M	h
Set of rows	R	J	r
Set of products	P	$JM = I$	p
Set of AGVs	A	N	a
Set of orders	O	DI	o_f
Set of weights	W	DI	w_{o_f}

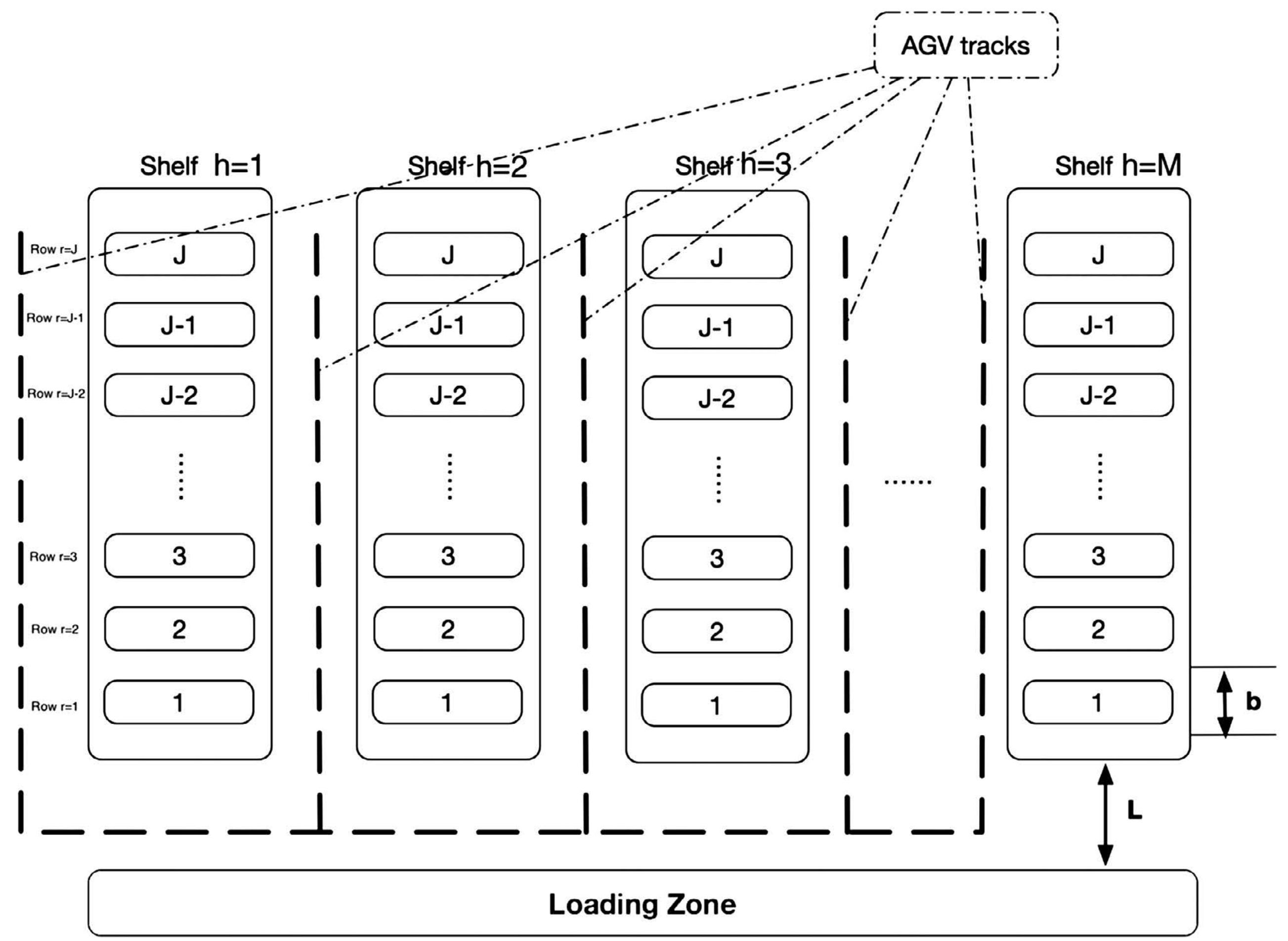

Figure 1. This figure demonstrates the distance between the loading zone and the location of products.

(e) For simplicity, it is assumed that the travelling distance of each AGV is only the vertical distance between specific row and the loading zone. The distance between row $r \in \{1, \cdots, J\}$ and the loading zone is $(r-1) \cdot b + L$, as shown in Figure 1.

3.2 *Retrieval latency*

Based on the above assumptions and information, the mathematical model for retrieval time of each order, which depends on the storage assignment, is constructed in this section.

For each order $o_f, f \in \{1, \cdots, DI\}$, its arrival rate is denoted by λ_{o_f}, which is assumed to be given or predicted based on historical data. Let $B_{p,r,h}$ be an indicator variable for the location of product $p \in \mathbf{P}$:

$$B_{p,r,h} = \begin{cases} 1, & \text{if product } p \text{ is stored at row } r \text{ of shelf } h \\ 0, & \text{if product } p \text{ is not stored at row } r \text{ of shelf } h \end{cases} \tag{2}$$

Based on the assumptions, we have the following constraints on $B_{p,r,h}$. A type of product can only be stored in one location (a row of a shelf):

$$\sum_{r=1}^{J} \sum_{h=1}^{M} B_{p,r,h} = 1 \tag{3}$$

For a row of a shelf, it can only store one type of product:

$$\sum_{p=1}^{I} B_{p,r,h} = 1 \tag{4}$$

Once a product has been assigned to a storage location, its service time is fixed. Based on the above assumptions, the expected retrieval time for product $p \in \mathbf{P}$ is:

$$\eta_p = \sum_{r=1}^{J} \sum_{h=1}^{M} B_{p,r,h} \cdot \frac{(r-1) \cdot b + L}{s} \tag{5}$$

Since the set of orders is the multiplication of the set of product and the set of classes of customer orders, the expected retrieval time for an order o_f is defined as follows:

$$\frac{1}{\mu_{o_f}} = \eta_p, \text{where } f = p + (d-1)I, d \in \{1, 2, \cdots, D\}, p \in \{1, \cdots, I\} \tag{6}$$

Thus, the expected service time $\frac{1}{\mu_{o_f}}$ for those orders of the same product type are the same and deterministic, because they are only related to the placement of products. As discussed above, the expected retrieval time is only related to the storage assignment.

$$t_{retrieval,o_f} = \frac{1}{\mu_{o_f}} \tag{7}$$

3.3 *Differentiated probabilistic queuing policy*

Based on the above assumptions, the warehouse automation system has in total $|\mathbf{O}| = D \cdot I$ types of orders. The order set is the Cartesian product of the set of classes of customer orders and the set of products, thus containing the class of customer order information. An ideal order picking planning policy must consider the queue state, seeing what all orders are not yet processed. Thus, a Markov Decision Process with multiple states can be used. However, there are a lot of states and this approach does not give expressions that can be optimised for the decision process and decisions on the storage assignment. In order to come up with an approach to jointly consider an optimisation of storage assignment and order picking policies, this paper proposes a feasible stochastic approach.

To provide differentiated service levels in order picking planning policy, we describe the DPQ policy as follows:

(1) Each AGV has its own queue, the system has N queues in total.
(2) All orders are served under First Come First Serve policy in each queue.
(3) Orders from $o_f, f \in \{1, \cdots, DI\}$ are assigned to the queue of AGV $a \in \mathbf{A}$ with probability:

$$P_{o_f,a} \geq 0 \tag{8}$$

(4) For any order o_f,

$$\sum_{a \in \mathbf{A}} P_{o_f,a} = 1 \tag{9}$$

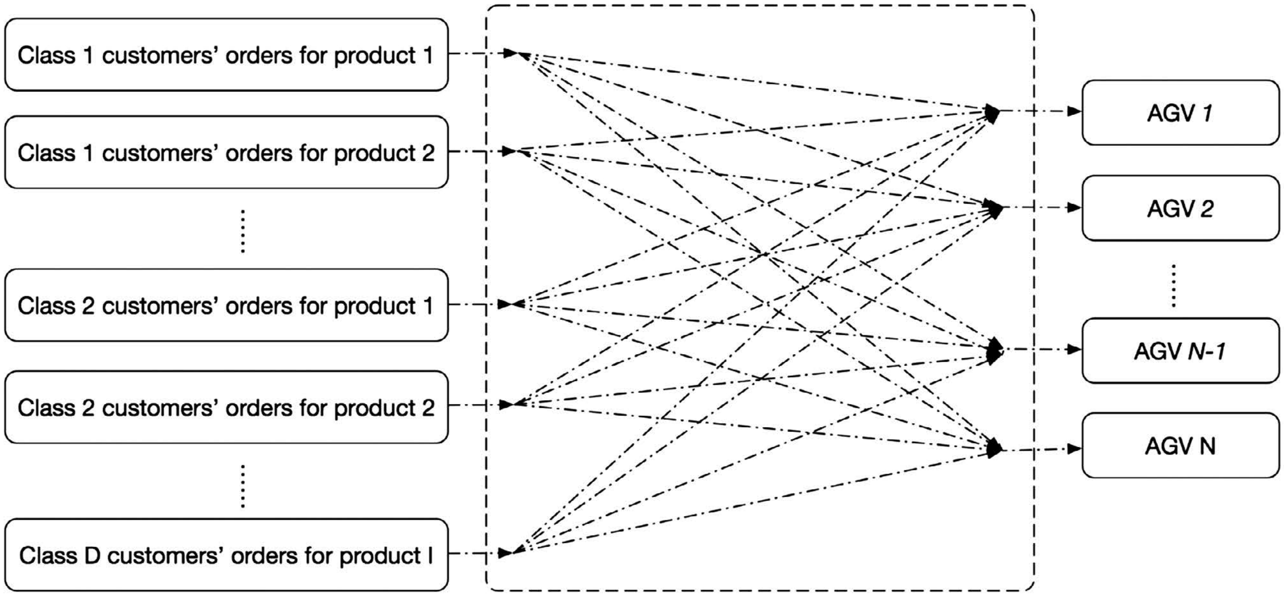

Figure 2. Processing of orders under DPQ policy. Different orders are probabilistically assigned to AGVs.

The orders' flow using DPQ policy is shown in Figure 2. Suppose the system has two classes of customer orders, two types of products, and two AGVs (denoted by a_1 and a_2). Then, there are four possible types of incoming orders to this system: (1) orders for product 1 from class 1 customer orders, (2) orders for product 2 from class 1 customer orders, (3) orders for product 1 from class 2 customer orders and (4) orders for product 2 from class 2 customer orders; denoted by o_1, o_2, o_3, o_4, respectively ($\mathbf{O} = \{\mathbf{o_1}, \mathbf{o_2}, \mathbf{o_3}, \mathbf{o_4}\}$). When each type of those orders arrives at the system, they will be assigned to the queue of the two AGVs with certain probabilities. For example, in this situation, o_1 will be assigned to the queue of a_1 and a_2 with probability P_{o_1,a_1} and P_{o_1,a_2}, respectively. Note that $P_{o_1,a_1} + P_{o_1,a_2} = 1$, which ensures that the order can be served by one and only one of those AGV. Within each queue, the orders are processed under First Come First Serve policy.

3.4 *Waiting latency formulation based on DPQ policy*

Based on the DPQ policy, for each AGV, it is a M/G/1 system with incoming orders from all queues (with different probabilities) and with service rates related to the average travel distance of the requested products. The overall arrival rate of the orders at AGV $a \in \mathbf{A}$ is:

$$\Lambda_a = \sum_f \lambda_{o_f} P_{o_f,a} \tag{10}$$

For each order $o_f \in \mathbf{O}$, the expected service time $\frac{1}{\mu_{o_f}}$ is deterministic, based on the storage assignment result. The expected service time at AGV $a \in \mathbf{A}$ is the conditional expectation of the service time $\frac{1}{\mu_{o_f}}$ of each order $o_f \in \mathbf{O}$ with probability $\frac{P_{o_f,a} \cdot \lambda_{o_f}}{\Lambda_a}$. Thus, the expected service time of AGV $a \in \mathbf{A}$ is:

$$E[S_a] = \sum_f \frac{P_{o_f,a} \cdot \lambda_{o_f}}{\Lambda_a} \cdot \frac{1}{\mu_{o_f}} \tag{11}$$

According to the queuing stability condition, the total arrival rate of AGV $a \in \mathbf{A}$ should be less than its expected service rate:

$$\Lambda_a < 1/E[S_a], \forall a \in \mathbf{A} \tag{12}$$

Since the expected service time is deterministic, which is only related to the location of a product, the second moment of expected service time is:

$$E[S_a^2] = \sum_f \frac{P_{o_f,a} \cdot \lambda_{o_f}}{\Lambda_a} \cdot \frac{1}{\mu_{o_f}^2} \tag{13}$$

For a given AGV $a \in \mathbf{A}$, since it is a M/G/1 system, the expected waiting time in the queue is (Ross 2014):

$$T_a = \frac{\Lambda_a E[S_a^2]}{2(1 - \Lambda_a E[S_a])} \tag{14}$$

Using conditional expectation, the expected waiting time for any order $o \in \mathbf{O}$ can be formulated as follows:

$$t_{waiting,o_f} = \sum_{a \in \mathbf{A}} P_{o_f,a} T_a \tag{15}$$

By now, we have the closed loop function of the expected waiting time of each type of order, under the DPQ policy. Combining Equations (1), (7) and (15), we have the following equation for the average processing time of an order:

$$L_{o_f} = \sum_{a \in 2A} P_{o_f,a} T_a + \frac{1}{\mu_{o_f}} \tag{16}$$

3.5 *Joint optimisation formulation of storage assignment and queuing policy*

To differentiate service level for different classes of customer orders, a price constant c_d is introduced to represent the price that class $d \in \mathbf{C}$ customer have paid for their orders. The differences among c_d represent the differences of the paid prices among different classes of customer orders.

We define a set of weights $\mathbf{W} = \{\mathbf{w_{o_1}}, \mathbf{w_{o_2}}, \cdots, \mathbf{w_{o_{DI}}}\}$ for each type of order, where

$$w_{o_f} = c_d \cdot \lambda_{o_f}, \text{where} d = \frac{o_f}{I} \tag{17}$$

w_{o_f} combines a price constant and the arrival rate of each type of order, which is actually the weighted arrival rate of each type of order. It measures the importance of a specific type of order by two dimensions, both the price paid by the respective customer, and the arrival rate of this type of order. Based on these weights, we wish to minimise the weighted processing time of the system, and thus the objective function of the system is defined as:

$$\sum_f w_{o_f} L_{o_f} \tag{18}$$

The objective function measures the overall service-level based response by measuring the weighted average processing time of all incoming orders. Based on the weight functions, the more a customer has paid for his/her order, the less time s/he should wait. Further, high arrival rate orders would have large impact, thus motivating a weight proportional to the given order's arrival rates.

As mentioned above, $L_{o_f} = t_{waiting,o_f} + t_{retrieval,o_f}$, where the retrieval time depends only on storage assignment, but the waiting time is related to the queuing policy and its formula contains retrieval time. Thus, this optimisation problem is a joint optimisation problem. Combining the equations in the previous sections, the joint optimisation of the storage assignment *and* DPQ policy is formulated as follows:

$$\min_{B_{p,r,h}, P_{o_f,a}} \sum_f w_{o_f} L_{o_f} \tag{19}$$

Subject to Equations (3–6), (8–14), (16), (17)

$$\text{var}\, P_{o_f,a}, \forall f \in \{1, \cdots, DI\}, a \in \mathbf{A}$$

$$\text{var}\, B_{p,r,h} = 0 \text{ or } 1, \forall p \in \mathbf{P}, \mathbf{r} \in \mathbf{R}, \mathbf{h} \in \mathbf{H}$$

3.6 *Optimisation algorithm*

Jointly optimising the storage assignment and probabilities in the DPQ policy has integer constraints for the storage assignment, thus making the problem hard. To solve this problem, Alternating Minimisation (AM) algorithm is applied. The sub-linear convergence properties of the AM algorithm for convex problems with two sets of variables were proven by Beck (2015).

Table 3. Alternating minimisation algorithm for joint optimisation.

Algorithm 1 Alternating Minimisation Algorithm for Joint Optimisation

input Joint optimisation problem $f\left(\{P_{o_f,a}\},\{B_{p,r,h}\}\right)$, a set of weights $W=\{w_1,w_2,\ldots,w_{D\cdot I}\}$ for orders, number of maximum iteration $k_{\max}$, threshold value δ

Generate a feasible solution for $f\left(\{P_{o_f,a}\},\{B_{p,r,h}\}\right)$, denoted by $\{P_{o_f,a}\}_0, \{B_{p,r,h}\}_0$

$\{P_{o_f,a}\} \leftarrow \{P_{o_f,a}\}_0$

$\{B_{p,r,h}\} \leftarrow \{B_{p,r,h}\}_0$

$C_{\text{old}}, C_{\text{new}} \leftarrow f\left(\{P_{o_f,a}\},\{B_{p,r,h}\}\right)$

$i \leftarrow 1$

while $C_{\text{old}} - C_{\text{new}} > \delta$ or $i \le k_{\max}$:

  $C_{\text{old}} \leftarrow C_{\text{new}}$

  call *Algorithm 2* to solve $\min f\left(\{P_{o_f,a}\}\right)$ with fixed $\{B_{p,r,h}\}$

  $\{P_{o_f,a}\} \leftarrow \arg\min f\left(\{P_{o_f,a}\}\right)$

  call *Algorithm 2* to solve $\min f\left(\{B_{p,r,h}\}\right)$ with fixed $\{P_{o_f,a}\}$

  $\{B_{p,r,h}\} \leftarrow \arg\min f\left(\{B_{p,r,h}\}\right)$

  $C_{\text{new}} \leftarrow f\left(\{P_{o_f,a}\},\{B_{p,r,h}\}\right)$

  $i \leftarrow i+1$

end while

output $\{P_{o_f,a}\},\{B_{p,r,h}\}$

The basic idea of AM algorithm (Table 3) is to first fix one set of variables, and find the optimised solution of the other set of variables. Then, fix the set variables that have just been optimised and find the optimised solution of the first set of variables. This iteration repeats until the maximum times of iteration $k_{\max}$ is reached, or the difference between new solution and old solution is smaller than a predefined threshold value δ. Since each of the sub-problem is non-convex, Simulated Annealing (SA) algorithm (Metropolis et al. 1953; Kirkpatrick, Gelatt, and Vecchi 1983) is applied to find the solution of each set of variables during each iteration of the AM method (Table 3).

A feasible solution is found with both sets of variables generated randomly at the beginning (Table 4). The initial value of the objective function is calculated and stored as old cost. Then, the storage assignment result $B_{p,r,h}$ is fixed and passed to the SA algorithm with other fixed parameters. When the SA algorithm optimises the objective function, the variables in the function are $P_{o_f,a}$. After this optimisation, a relatively good solution of $P_{o_f,a}$ under the fixed $B_{p,r,h}$ is passed back to the main function. Then, this set of $P_{o_f,a}$ is fixed and passed to the SA algorithm with other fixed parameters. When the SA algorithm optimises the objective function, the variables in the function are $B_{p,r,h}$. After this optimisation, a relatively good solution of $B_{p,r,h}$ under the fixed $P_{o_f,a}$ is passed back to the main function. By then, one iteration of alternating minimisation is finished, and a new solution with better objective function value is obtained. If the difference between new cost and old cost is smaller than a predefined threshold value δ, the alternating minimisation will stop. Otherwise, it keeps iterating until the maximum times of iteration $k_{\max}$ is reached, or the difference between new solution and old solution is smaller than δ.

We note that the optimisation over storage assignment may not be possible for smaller time scales, and can be performed at longer times when the distributions of the order arrivals have changed significantly (in a few months). However, the order picking policy can be optimised more often.

4. Numerical results

The system performance of the developed joint optimisation of the storage assignment and the DPQ policy is compared with three baseline systems. The system performance is measured by the objective function, which is the overall weighted processing time of all incoming orders. As comparisons, we use three baseline strategies for the combinations of storage assignment and order picking.

(1) Baseline1 system: In this system, we use the Turn-over Rate storage assignment method for storage assignment and the FCFS policy for order picking.
(2) Baseline2 system: In this system, we use the Turn-over Rate storage assignment method for storage assignment and the DPQ policy for order picking.
(3) Baseline3 system: In this system, we use the optimised storage assignment, while use the FCFS policy for order picking.

Table 4. Simulated annealing algorithm for optimising single set of variables.

Algorithm 2 Simulated Annealing Algorithm for Optimising Single Set of Variables

input Optimisation problem $f(\{X\})$, a set of weights $W = \{w_1, w_2, \ldots, w_{D \cdot I}\}$ for orders, a set of fixed parameters $\{Y\}$, initial temperature T_o, number of maximum iteration $k_{\max}$, threshold value δ, cooling rate α

Generate a feasible solution for $f(\{X\})$ denoted by $\{X\}_0$

$\{X\} \leftarrow \{X\}_0$

$T \leftarrow T_o$

$C_{\text{old}}, C_{\text{new}} \leftarrow f(\{X\})$

while $T > T_{\min}$

  $i \leftarrow 1$

  while $C_{\text{old}} - C_{\text{new}} > \delta$ or $i \le k_{\max}$:

    $C_{\text{old}} \leftarrow C_{\text{new}}$

    Pick a random neighbouring solution $\{X\}'$

     $C_{\text{new}} \leftarrow f(\{X\}')$

     if $C_{\text{new}} < C_{\text{old}}$ **then**

       $\{X\} \leftarrow \{X\}'$

     else

     if $e^{\frac{C_{\text{old}} - C_{\text{new}}}{T}} > random(0, 1)$ **then**

       $\{X\} \leftarrow \{X\}'$

     else

       $C_{\text{new}} \leftarrow C_{\text{old}}$

     end if

   end if

     $i \leftarrow i + 1$

   end while

$T \leftarrow T \cdot \alpha$

end while

output $\{X\}$

Numerical tests are performed under Matlab environment, with the baseline settings as follows. The number of classes are chosen to be 3, where higher the class number, higher is the priority of the class. The order frequencies λ_o are generated randomly by the triangular distribution with the range of (0, 6) and peak of 3. The price function is chosen as $c_d = d^2$. The number of rows are chosen to be 20, number of shelves are chosen to be 10 and the number of AGVs are chosen to be 15. The experiments are repeated 50 times for better averaging.

4.1 *Impact of number of AGVs*

The impact of changing the number of AGVs can be observed as follows. Under the baseline settings, where the number of AGVs is 15, joint optimisation of the storage assignment and DPQ policy has 19.64, 4.79 and 17.16% improvement in system performance as compared with the baseline1, baseline2, baseline3 systems, respectively. The results are shown in Table 5 and Figure 3. It can be seen that as the number of AGVs decreases, the improvement percentage increases. This is because with less AGVs, more orders wait in the queues. Thus, the relative impact of efficient orders scheduling can increase. Therefore, the joint optimisation of the storage assignment and DPQ policy will have relatively larger improvement when there are fewer AGVs.

Even though more AGVs in the system result in a relatively lower improvement in system performance, when the number of AGVs increases to 21, joint optimisation of storage assignment and the DPQ policy still has 13.24, 3.66 and 10.56% relative improvement in system performance, compared with the baseline1, baseline2 and baseline3 system, respectively (Table 5). The pairwise t-test results in Table 5 indicate that in all the tested scenarios, the improvement in system performance brought about by applying joint optimisation of the storage assignment and DPQ policy is significant, compared with the baseline1 system. In summary, the heavier traffic caused by reducing the number of AGVs leads to relatively higher percentage system improvement. Furthermore, joint optimisation of storage assignment and DPQ policy results in greater relative gains than taking any combination of one of the mixed baseline strategies, baseline 2 and baseline 3.

4.2 *Impact of price function*

In the second series of experiments, the price function of the system is changed, under the baseline settings. As shown in Table 6 and Figure 4, when $c_d = d$ (d^{th} class requests hasve d times more weight as compared to the first class), the

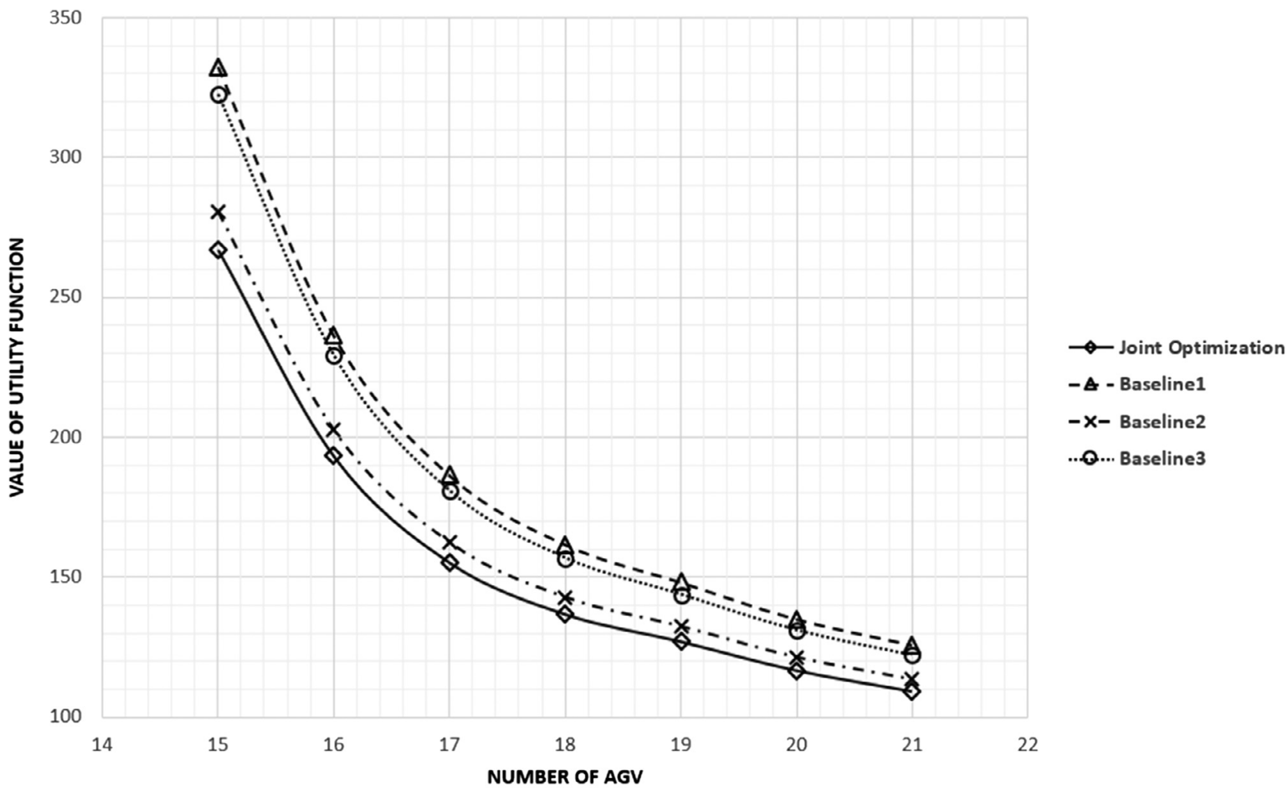

Figure 3. System performance with the proposed joint optimisation approach for different number of AGVs.

Table 5. Numerical results for joint optimisation and baseline systems as the number of AGVs changes under baseline settings.

Number of AGVs	Improvement in system performance (compared with baseline1) (%)	*P*-value of pairwise t-test results (95% confidence level, compared with baseline1)	Improvement in system performance (compared with baseline2) (%)	Improvement in system performance (compared with baseline3) (%)
15	19.64	2.82×10^{-47}*	4.76	17.16
16	18.20	2.13×10^{-55}*	4.58	15.68
17	16.76	2.43×10^{-62}*	4.4	14.18
18	15.45	3.04×10^{-68}*	4.21	12.84
19	14.34	1.74×10^{-67}*	4.03	11.7
20	13.65	2.29×10^{-68}*	3.85	10.98
21	13.24	5.87×10^{-68}*	3.66	10.56

Note: * indicates significant difference

Table 6. Numerical results for joint optimisation and baseline system as price function changes under baseline settings.

Price function	Improvement in system performance (compared with baseline1) (%)	P-value of pairwise t-test (95% confidence level, compared with baseline1)	Improvement in system performance (compared with baseline2) (%)	Improvement in system performance (compared with baseline3) (%)
$c_d = e^d$	20.42	5.46×10^{-23}*	4.76	17.39
$c_d = d$	15.99	6.95×10^{-44}*	3.85	14.28
$c_d = d^2$	19.64	2.82×10^{-47}*	4.76	17.16
$c_d = d^3$	26.86	1.78×10^{-48}*	6.54	23.81

Note: * indicates significant difference.

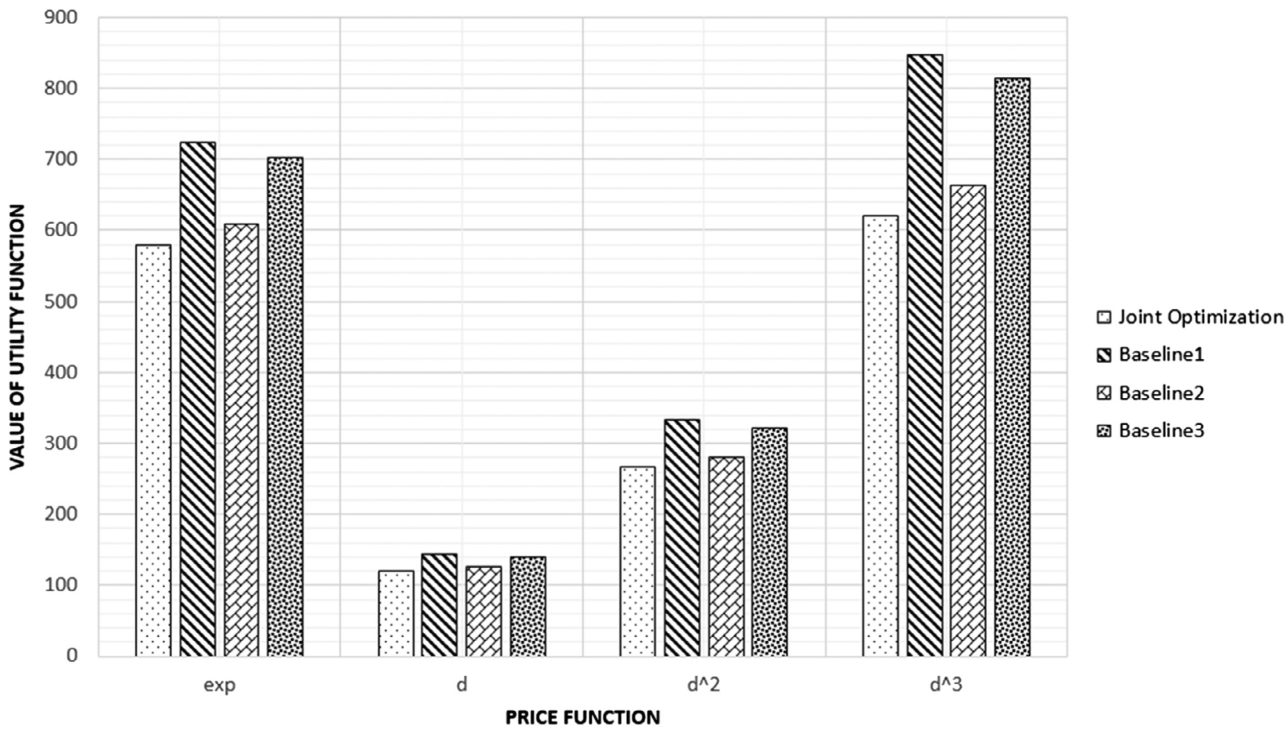

Figure 4. System performance with the proposed joint optimisation approach for different price functions.

relative improvement in system performance is 16.00, 3.85 and 14.68%, respectively, compared with those three baseline systems. This relative improvement increases to 19.64, 4.76 and 17.16% when $c_d = d^2$ and is 26.86, 6.54 and 23.81% when $c_d = d^3$. The results of pairwise t-test (Table 6) indicate the statistical significance of these improvements, compared with baseline1 system. Thus, the higher the difference in weights among the different classes, the higher is the advantage of using the differentiated service levels approach.

5. Conclusion

One of the significant advantages of smart warehouse automation is the ability to offer a variety of customer oriented services. Such services enable suppliers to distinguish themselves in an increasingly competitive market. Differentiated service is a useful and practical approach when a service provider has to respond effectively to many customers under time constraints, and some of the customers are willing to pay a higher price for the service, to obtain faster and/or better service quality. The research reported in this article is the first work that considers differentiated service levels for different classes of customer orders in a smart warehouse automation system. It provides an integrated vision on the delivery system of online retailers with multi-levels pricing strategy. A novel order picking planning policy of warehouse automation system, DPQ policy, is developed in this article for differentiating service levels for different classes of paying customers. This policy is more flexible compared to the standard priority queueing policy, since general weights for different customers can be used. Using DPQ policy, the average overall latency of each customer order is characterised. The weighted latency for different price-based classes of customer orders is optimised over storage assignment and the parameters of the DPQ policy. Numerical results indicate statistically significant performance improvements as compared to the traditional combination of Turn-over Rate storage assignment method, and the FCFS policy for AGV picking.

Future research will address relaxing some of the assumptions made (summarised in Table 4). Other interesting open questions to consider are the comparison of the DPQ policy in digital versus physical storage service cases, and different combinations of product pricing and service level pricing.

Nomenclature

AGV Automated Guided Vehicle
AM Alternating Minimisation
DPQ Differentiated Probabilistic Queuing
CCT Collaborative Control Theory
FCFS First Come First Serve
SA Simulated Annealing

Acknowledgements

The authors would like to thank Dr. Rodolfo Portillo, Senior Manager at Amazon, for helpful discussions on the automated storage and retrieval assumptions in this article.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported in part by Kimberly Clark Corp. through the Purdue PRISM Center for Production, Robotics, and Integration Software for Manufacturing & Management and the National Science Foundation under grant CNS-1618335.

References

Aggarwal, V., Y. R. Chen, T. Lan, and Y. Xiang. 2017. Sprout: A Functional Caching Approach to Minimize Service Latency in Erasure-coded Storage. *IEEE/ACM Transactions on Networking* 25 (6): 3683–3694.

Azadeh, A., M. Moghaddam, B. Doust, and F. Jalalvand. 2015. "Fuzzy and Stochastic Mathematical Programming for Optimisation of Cell Formation Problems in Random and Uncertain States." *International Journal of Operational Research* 22 (2): 129.

Battini, D., M. Calzavara, A. Persona, and F. Sgarbossa. 2015. "Order Picking System Design: The Storage Assignment and Travel Distance Estimation (SA&TDE) Joint Method." *International Journal of Production Research* 53 (4): 1077–1093.

Beck, A. 2015. "On the Convergence of Alternating Minimization for Convex Programming with Applications to Iteratively Reweighted Least Squares and Decomposition Schemes." *SIAM Journal on Optimization* 25 (1): 185–209.

Bitar, E., and S. Low. 2012. "Deadline Differentiated Pricing of Deferrable Electric Power Service." In *2012 IEEE 51st Annual Conference on Decision and Control (CDC)*, 4991–4997. Maui, HI: IEEE.

Bozer, Y. A., and J. W. Kile. 2008. "Order Batching in Walk-and-pick Order Picking Systems." *International Journal of Production Research* 46 (7): 1887–1909.

Buonamico, N., L. Muller, and M. Camargo. 2017. "A New Fuzzy Logic-based Metric to Measure Lean Warehousing Performance." *Supply Chain Forum: An International Journal* 18 (2): 96–111.

Chen, X., and P. Mohapatra. 1999. "Providing Differentiated Service from an Internet Server." In *Proceedings of Eight International Conference on Computer Communications and Networks, 1999*, 214–217. Boston, MA: IEEE.

Cheng, X., S. Liao, and Z. Hua. 2017. "A Policy of Picking Up Parcels for Express Courier Service in Dynamic Environments." *International Journal of Production Research* 55 (9): 2470–2488.

Culler, D., and J. Long. 2016. "A Prototype Smart Materials Warehouse Application Implemented Using Custom Mobile Robots and Open Source Vision Technology Developed Using EmguCV." *Procedia Manufacturing* 5: 1092–1106.

De Koster, M. B. M., E. S. Van der Poort, and M. Wolters. 1999. "Efficient Orderbatching Methods in Warehouses." *International Journal of Production Research* 37 (7): 1479–1504.

De Koster, R., T. Le-Duc, and K. J. Roodbergen. 2007. "Design and Control of Warehouse Order Picking: A Literature Review." *European Journal of Operational Research* 182 (2): 481–501.

Deshpande, J. G., E. Kim, and M. Thottan. 2011. "Differentiated Services QoS in Smart Grid Communication Networks." *Bell Labs Technical Journal* 16 (3): 61–81.

Epp, M., S. Wiedemann, and K. Furmans. 2016. "A Discrete-time Queueing Network Approach to Performance Evaluation of Autonomous Vehicle Storage and Retrieval Systems." *International Journal of Production Research* 55 (4): 960–978.

Fox13. 2015. "Behind the Scenes of an Amazon Warehouse." *Fox 13 News*, https://www.youtube.com/watch?v=Y-lBvI6u_hw

Fuentes Saenz, H. 2011. "Data Mining Framework for Batching Orders in Real-time Warehouse Operations." MS thesis, University of Iowa.

Goyal, P., H. M. Vin, and H. Chen. 1997. "Start-time Fair Queueing: A Scheduling Algorithm for Integrated Services Packet Switching Networks." *IEEE/ACM Transactions on Networking* 5 (5): 690–704.

Grit, L. E., and J. S. Chase. 2008. “Weighted Fair Sharing for Dynamic Virtual Clusters.” *ACM SIGMETRICS Performance Evaluation Review* 36 (1): 461–462.

Grosse, E., C. Glock, and W. Neumann. 2017. “Human Factors in Order Picking: A Content Analysis of the Literature.” *International Journal of Production Research* 55 (5): 1260–1276.

Hill, R. M., M. Seifbarghy, and D. K. Smith. 2007. “A Two-echelon Inventory Model With Lost Sales.” *European Journal of Operational Research* 181: 753–766.

Hilletofth, P. 2009. “How to Develop a Differentiated Supply Chain Strategy.” *Industrial Management & Data Systems* 109 (1): 16–33.

Hsieh, L. F., and Y. C. Huang. 2011. “New Batch Construction Heuristics to Optimise the Performance of Order Picking Systems.” *International Journal of Production Economics* 131 (2): 618–630.

Hu, K. Y., and T. S. Chang. 2010. “An Innovative Automated Storage and Retrieval System for B2C e-commerce Logistics.” *The International Journal of Advanced Manufacturing Technology* 48 (1–4): 297–305.

Hwang, H., Y. H. Oh, and Y. K. Lee. 2004. “An Evaluation of Routing Policies for Order-picking Operations in Low-level Picker-to-part System.” *International Journal of Production Research* 42 (18): 3873–3889.

Jin, W., J. S. Chase, and J. Kaur. 2004. “Interposed Proportional Sharing for a Storage Service Utility.” *ACM SIGMETRICS Performance Evaluation Review* 32 (1): 37–48.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi. 1983. “Optimization by Simulated Annealing.” *Science* 220 (4598): 671–680.

Le, L. B., E. Hossain, and A. S. Alfa. 2006. “Service Differentiation in Multirate Wireless Networks With Weighted Round-robin Scheduling and ARQ-based Error Control.” *IEEE Transactions on Communications* 54 (2): 208–215.

Lee, M. S., and C. H. Park. 2009. “The Three-level Approaches for Differentiated Service in Clustering Web Server.” *Active and Programmable Networks Lecture Notes in Computer Science* 230–235.

Lee, J., and T. Srisawat. 2006. “Effect of Manufacturing System Constructs on Pick-up and Drop-off Strategies of Multiple-load AGVs.” *International Journal of Production Research* 44 (4): 653–673.

Lenstra, J. K., A. R. Kan, and P. Brucker. 1977. “Complexity of Machine Scheduling Problems.” *Annals of discrete mathematics* 1: 343–362.

Li, L. 2009. “An Optimistic Differentiated Service Job Scheduling System for Cloud Computing Service Users and Providers.” In *Third International Conference on Multimedia and Ubiquitous Engineering, 2009*, 295–299. Qingdao: IEEE.

Li, J., M. Moghaddam, and S. Y. Nof. 2015. “Dynamic Storage Assignment with Product Affinity and ABC Classification – A Case Study.” *The International Journal of Advanced Manufacturing Technology* 84 (9–12): 2179–2194.

Liu, H. 2012. “Warehouse Redesign of Facility Layout for Wegmans Food Market.” MS Thesis, Pennsylvania State University.

Ma, T., and P. Zhao. 2014. “A Review of Algorithms for Order Batching Problem in Distribution Center." *Proceedings of the International Conference on Logistics, Engineering, Management and Computer Science (LEMCS 2014)*. Qingdao: Atlantis Press.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. “Equation of State Calculations by Fast Computing Machines.” *The Journal of Chemical Physics* 21 (6): 1087–1092.

Morash, E. A., and S. R. Clinton. 1998. “Supply Chain Integration: Customer Value through Collaborative Closeness Versus Operational Excellence.” *Journal of Marketing Theory and Practice* 6 (4): 104–120.

Murray, R., M. Chao, and S. Khuller. 2016. “Scheduling Distributed Clusters of Parallel Machines: Primal-dual and LP-based Approximation Algorithms.” In *24th Annual European Symposium on Algorithms (ESA 2016). LIPIcs-Leibniz International Proceedings in Informatics*, edited by P. Sankowski and C. Zaroliagis, vol. 57, 68:1–68:17. Dagstuhl: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. http://drops.dagstuhl.de/opus/volltexte/2016/6410.

Negrete-Pincetic, M., and S. Meyn. 2012. “Markets for Differentiated Electric Power Products in a Smart Grid Environment.” In *Power and Energy Society General Meeting, 2012 IEEE*, 1–7. San Diego, CA: IEEE.

Nof, S. Y. 2007. “Collaborative Control Theory for e-Work, e-Production, and e-Service.” *Annual Reviews in Control* 31 (2): 281–292.

Pang, K. W., and H. L. Chan. 2017. “Data Mining-based Algorithm for Storage Location Assignment in a Randomised Warehouse.” *International Journal of Production Research* 55 (14): 4035–4052.

Petersen, C. G., and G. Aase. 2004. “A Comparison of Picking, Storage, and Routing Policies in Manual Order Picking.” *International Journal of Production Economics* 92 (1): 11–19.

Rao, J., Y. Wei, J. Gong, and C. Z. Xu. 2013. “QoS Guarantees and Service Differentiation for Dynamic Cloud Applications.” *IEEE Transactions on Network and Service Management* 10 (1): 43–55.

Ries, J. M., E. H. Grosse, and J. Fichtinger. 2017. “Environmental Impact of Warehousing: A Scenario Analysis for the United States.” *International Journal of Production Research* 55 (21): 6485–6499.

Ross, S. 2014. *Introduction to Probability Models*. 11th ed. Oxford: Academic Press.

Rowley, P. D., and J. S. Battles. 2015. Continuous flow processing of packaged items at order fulfillment centers. U.S. Patent No. 9174758. Washington, DC: U.S. Patent and Trademark Office, Chicago.

Schneider, F., and D. Klabjan. 2013. “Inventory Control in Multi-channel Retail.” *European Journal of Operational Research* 227 (1): 101–111.

Shue, D., M. J. Freedman, & A. Shaikh. 2012. “Performance Isolation and Fairness for Multi-tenant Cloud Storage.” In *OSDI*, vol. 12, 349–362. Hollywood, CA.

Smith, W. E. 1956. “Various Optimizers for Single-stage Production.” *Naval Research Logistics Quarterly* 3 (1–2): 59–66.

Stiliadis, D., and A. Varma. 1998. “Efficient Fair Queueing Algorithms for Packet-switched Networks.” *IEEE/ACM Transactions on Networking* 6 (2): 175–185.

Trebilcock, B. 2003. “Warehouse Software Meets RFID.” *Modern Materials Handling* 58 (13): 32–34.

Tsai, C. Y., J. J. Liou, and T. M. Huang. 2008. “Using a Multiple-GA Method to Solve the Batch Picking Problem: Considering Travel Distance and Order Due Time.” *International Journal of Production Research* 46 (22): 6533–6555.

Veres, A., A. T. Campbell, M. Barry, and L. H. Sun. 2001. “Supporting Service Differentiation in Wireless Packet Networks Using Distributed Control.” *IEEE Journal on Selected Areas in Communications* 19 (10): 2081–2093.

Vliet, K. R., W. A. Kiser, M. A. Holland, & J. S. Battles (2014). “Modular Material Handling System for Order Fulfilment.” U.S. Patent No. 8892240. Washington, DC: U.S. Patent and Trademark Office, Chicago.

Vliet, K. R., M. A. Holland, P. T. Shigihara, and J. S. Battles. 2016. “Flexible Processing Module for Use in Order Fulfillment Centers.” U.S. Patent No. 9,409,664. Washington, DC: U.S. Patent and Trademark Office, Chicago.

Wang, Q., I. McIntosh, and M. Brain. 2010. “A New-generation Automated Warehousing Capability.” *International Journal of Computer Integrated Manufacturing* 23 (6): 565–573.

Weng, M. X., J. Lu, and H. Ren. 2001. “Unrelated Parallel Machine Scheduling With Setup Consideration and a Total Weighted Completion Time Objective.” *International Journal of Production Economics* 70 (3): 215–226.

Wurman, P. R., R. D’Andrea, and M. Mountz. 2008. “Coordinating Hundreds of Cooperative, Autonomous Vehicles in Warehouses.” *AI magazine* 29 (1): 9.

Xiang, Y., V. Aggarwal, Y. F. Chen, and T. Lan. 2017. “Differentiated Latency in Data Center Networks With Erasure Coded Files Through Traffic Engineering.” *IEEE Transactions on Cloud Computing*. doi:10.1109/TCC.2017.2648785.

Xiang, Y., T. Lan, V. Aggarwal, and Y. F. R. Chen. 2014. “Joint Latency and Cost Optimization for Erasure-coded Data Center Storage.” *ACM SIGMETRICS Performance Evaluation Review* 42 (2): 3–14.

Xiang, Y., T. Lan, V. Aggarwal, and Y. F. R. Chen. 2016. “Joint Latency and Cost Optimization for Erasure-coded Data Center Storage.” *IEEE/ACM Transactions on Networking* 24 (4): 2443–2457.

Xiang, Y., T. Lan, V. Aggarwal, and Y. F. Chen. 2017. “Optimizing Differentiated Latency in Multi-tenant, Erasure-coded Storage.” *IEEE Transactions on Network and Service Management* 14 (1): 204–216.

Xiao, J., and L. Zheng. 2010. “A Correlated Storage Location Assignment Problem in a Single-block-multi- Aisles Warehouse Considering BOM Information.” *International Journal of Production Research* 48 (5): 1321–1338.

Xu, P. J. 2005. “Order Fulfilment in Online Retailing: What Goes Where.” Doctoral diss., Massachusetts Institute of Technology.

Yu, M., and R. B. De Koster. 2009. “The Impact of Order Batching and Picking Area Zoning on Order Picking System Performance.” *European Journal of Operational Research* 198 (2): 480–490.

Yu, M., and R. B. De Koster. 2012. “Sequencing Heuristics for Storing and Retrieving Unit Loads in 3D Compact Automated Warehousing Systems.” *IIE Transactions* 44 (2): 69–87.

Zhang, L., H. Zhong, and S. Y. Nof. 2015. “Adaptive Fuzzy Collaborative Task Assignment for Heterogeneous Multirobot Systems.” *International Journal of Intelligent Systems* 30 (6): 731–762.

Zhong, R. Y., G. Q. Huang, S. Lan, Q. Y. Dai, X. Chen, and T. Zhang. 2015. “A Big Data Approach for Logistics Trajectory from RFID-enabled Production Data.” *International Journal of Production Economics* 165: 260–272.

Zhou, W., S. Piramuthu, C. Feng, and C. Chengbin. 2017. “RFID-enabled Flexible Warehousing.” *Decision Support Systems* 98: 99–112.