

An Algorithm for Customer Order Fulfillment in a Make-to-Stock Manufacturing System

D. Lečić-Cvetković, N. Atanasov, S. Babarogić

Danica Lečić-Cvetković, Nikola Atanasov, Sladan Babarogić

University of Belgrade, Faculty of Organizational Sciences

Jove Ilića 154, 11000 Belgrade, Serbia

E-mail: {danica,nikola.atanasov,sladjan}@fon.rs

Abstract: In the competitive environment, many manufacturers are increasingly focusing on designing the systems that help them to manage variable demand and supply situations. Dynamic allocation of demands is very important in case of customer order allocations. Order promising and allocation can be based on the simple sequence that enables a manufacturing company to receive orders unless there are some other priority orders. Manufacturing company can also manage allocations of supply to key customers and channels, thereby ensuring that they can meet contractual agreements and service levels in the priority that yields better profit. This paper will focus on a Make-to-Stock order fulfillment system facing random demand with random orders from different classes of customers. Available-to-promise (ATP) calculating from master production schedule (MPS) exhibits availability of finished goods that can be used to support customer order allocation. This order allocation system is adapted in MTS (make-to-stock) production model and all orders are treated according to maximization of customer service policy. It allows incoming purchase orders as well as existing inventory on hand to be selected and allocated to customer sale orders and back orders. The system then automatically allocates the available stock to the selected sales orders. We developed an integrated system for allocation of inventory in anticipation of customer service of high priority customers and for order promising in real-time. Our research exhibits three distinct features:

- (1) We explicitly classified customers in groups based on target customer service level;
- (2) We defined higher level of customer selection directly defined according to company strategy to develop small and medium customers;
- (3) We considered backorders that manufacturing company has to fulfill in order to maximize overall customer service for certain customers.

Keywords: manufacturing system, order allocation, customer service.

1 Introduction

Production companies are facing the dilemma on whether to increase or decrease the capacities of production lines and/or the capacities of production plants. Everyday, production managers are facing the dilemma on whether to buy a new machine, to increase the number of shifts within the production process, or to employ new workers, in spite of the fact that such decisions also affect the middle-term and long-term production planning, and should not be made hastily. Sales forecasting and identification of an increase in demand represent the starting point for production capacity planning, and also for making a decision on capacity expansion. Until the available capacities are increased, the production company has a goal of satisfying the market

demand by means of available production capacities. The allocation of available finished products to customer orders requires an efficient system of allocation aimed at improving the overall operating efficiency. Operating efficiency is directly related to quantities of goods produced and the profit from sales. In addition to profit-oriented decision on selection of orders to be fulfilled, one needs to consider the customer service, since the overall operating of a production company depends on customers. Customers, who account for a large share in the sales results, require special attention, and the fulfillment of every order. There is also a group of customers who constantly increase their orders, and thus expect adequate and better service. Small but numerous customers, thus shape the overall sales of a production company. Some of them represent the future potential for the increase in sales and also for the increase in incomes of the production company. The abovementioned facts highlight the importance of making a decision while selecting the orders to be fulfilled, and the necessity of having an algorithm that will efficiently and effectively perform the allocation of limited quantities of available finished products.

This paper is structured as follows: the second chapter presents relevant approaches of other authors in solving the observed problem. The third chapter defines the problem of allocation of limited resources in production companies. The problem of allocation refers to the distribution of limited amount of production to customer orders, with the goal of maximizing the percentage of fulfilled orders. Aiming to define efficient and effective allocation, the fourth chapter describes the procedure of the development of algorithms for fulfilling the orders in the make-to-stock production system. In their conclusion, authors cited the main advantages of the proposed algorithm for allocation of limited production, as well as the possible directions of development and upgrading the algorithm aimed at improving the performances of the production company.

2 Related work

Providing the ordered quantities of products represents the key function in planning operations within the entire supply chain. The system of providing stocks is recognized as a key challenge in production companies. The basic goal of the process of providing stocks ATP/CTP (ATP – available-to-promise, and CTP – capable-to-promise) is providing a reliable response to customer demands, taking into consideration the wide range of information and limitations that exist in the distribution channel network. Key measures of performances within production companies, according to [6] are recognized in measuring the level of customer service, and customer satisfaction. Traditional approaches to order provision, adjusted to MTS production system, were described in [1] while considering the available stocks of finished products for order fulfillment in line with the principle First Come – First Served (FCFS), without awarding priority to either customers or orders.

The priority of received orders also represents a significant factor in providing stocks for order fulfillment. Optimization of orders is carried out based on the maximization of profit for the entire operating, through the system of production planning represented in [9]. Priority is given to orders in the following way: (1) orders are given the priority if they are already in the forecast of sales, i.e. (2) orders are given high-priority if the delivery of such orders should bring higher profit to operating of a production company. The general application of ATP/CTP decision system was implemented with an aim to improve the profit and system performances, in line with [6]. By considering the Advanced Planning Systems (APS), it is possible to identify various approaches in defining the priority of customers, based on which it will be possible to fulfill the incoming orders. The essential idea of the approach described in [9] is the segmentation of customers with the goal of increasing the overall income of the production company through order acceptance, and delivery of orders with the greatest profit. ATP model allows for such system, since it provides quantities in advance in accordance with certain customer segments, by

satisfying only the orders of priority customers.

ATP system is based on mixed-integer programming, with the goal of optimizing the utilization of limited production capacities, in order to provide timely information regarding the fulfillment of a customer's order. In addition, the abovementioned problem is according to [2] classified amongst the dynamic models of order management with limited capacities based on the profitability analyses. If a production company has limited production capacities, it is clear that the company will decide to reject some of the incoming orders, which will directly affect the profitability of operating. The decision on rejecting orders is based on the comparison of orders, since the company rejects those who yield lower profit.

ATP system that is based on providing quotes for important customers in a defined timeframe, in order to provide timely information on the delivery date, is described in [9]. ATP defined in a way that orders yielding higher profit for the producer must be fulfilled with priority, in spite of the fact that there are less profitable orders still 'on hold'. Meyr focused ATP on MTS system through the assumption that the supply of finished products is fixed with available stocks and the ongoing production, which will be available in a short period of time. The model of allocation of limited quantities of finished products through customer segmentation, developed by Kilger and Meyr, is based on profit assessment that is generated by fulfillment of orders in accordance with their priority. New orders could be fulfilled by the allocation of quantities granted to the group of customers they belong to, i.e. if there is no 'free stock' in that group from the quantities granted to lower-priority groups. Such method prevents orders of low-priority customers to be fulfilled prior to orders of high-priority customers, who help in generating higher profit. The ATP allocation system is based on defining the priority classes with the goal of maximizing the overall profit of the production company. On the grounds of various researches, according to [9], it was determined that the FCFS system for customer orders bring the best results with limited production capacities, if case the production is realized based on the forecast accuracy. If there are classes defined in accordance with the priority of customers, the quantity awarded to certain class within a cycle, if not allocated, stays booked for the same class in the following cycle. The same author in [10] presents that the provisional allocation of stocks for certain customer classes within a defined period of time could give a significant contribution if the customer demand within the class could be anticipated with guaranteed accuracy.

3 Order fulfillment problem definition

Production companies operating on markets with irregular demand often face the problem of insufficient stock of finished products. Until it makes a decision on the expansion of production capacities, a production company has the goal of fulfilling as much orders as possible, only with available production capacities and quantities of finished products. The allocation of finished products to customer orders is realized through the process of allocation within the MTS (make-to-stock) production system. The problem of allocation of available products to incoming orders is listed among problems of the allocation of limited resources. In case the incoming customer orders within one cycle do not exceed the available stocks of finished products, the allocation is complete and all orders are fulfilled. In the opposite case, when the sum of overall orders is greater than the quantity of available stock, there is a need to define the way of product allocation, i.e. rules must be introduced to help the allocation of products to incoming customer orders. With the goal of maximizing the effect of allocation of available products, the need for the development of allocation algorithm has been perceived, with which to enable the fulfillment of orders in line with their priority. The fulfillment of orders of high-priority customers contributes to the maintenance of preferred service for those customers who rank high because of their share in incomes and the profit of the production company. The order fulfillment algorithm refers to

products of FMCG (Fast Moving Consumer Goods) industry, and with minor modifications, it is possible to apply it to the problem of allocation of products in other industries too.

4 Algorithm

The problem of allocation of limited quantities of finished products to customer orders could be solved by creating an algorithm that systematizes the allocation process. The proposed allocation algorithm tends to maximize the customers service that is expressed by the number of fulfilled orders and the percentage of order fulfillment, all this through the processing of orders in accordance with previously defined groups and partitions of customers. The criterion of dividing customers into groups is based on the consideration of revenue per customer within the company, profit per customer, development potential, service rate per customer, strategic partnership with the customers, etc. With the goal of providing strategic potential for the development of a customer network, customers are classified into partitions that are provided with guaranteed quantity of products for certain groups of customers, which would otherwise be marginalized. The proposed algorithm has polynomial complexity.

Table 1: Variable descriptions

Variables	Description
PT_l	Total available production in l -iteration ($l = 1, \dots, r$)
OT_l	Total order in l -iteration ($l = 1, \dots, r$)
P_k	Partition ($k = 1, \dots, p$)
G_j	Group ($j = 1, \dots, m$)
C_i	Customer ($i = 1, \dots, n$)

Algorithm is applicable only when $OT_l > PT_l$. In the beginning of the year, or at the beginning of iteration:

1. Form the list of customers with parameters ($i = 1, \dots, n$)

The list comprises basic data on customers and their business indicators for the previous year, for the last three years, and last five years.

2. Define the number of groups ($j = 1, \dots, m$)

Every identified group has its priority. Smaller ordinal number of a group brings higher priority. The recommendation is to form three groups, in spite of the fact that the algorithm functions even if only two groups are present.

3. Clustering customers into groups

At the beginning of the year, a company disposes of the list of customers, and it revises business results from the previous period for every respective customer, which is actually the base for classification into groups (Figure 1a). Every customer within the same group is equal with other members of the group when it comes to products' allocation. Clustering (classification) into groups could be done based on results of applying methods such as ABC method of prioritizing.

$$C_i \in G_j, \quad \text{for } i \in \{1, \dots, n\} \text{ and } j \in \{1, \dots, m\}$$

Information on which customer belongs to what group are kept in the matrix customer group (MCG), as shown on Figure 1b.

If a new customer emerges at the very beginning of new iteration, he is been added to the system and attached to some of the existing groups.

$$C_{n+1} \in G_j, \quad \text{for } j \in \{1, \dots, m\}$$

4. Define the number of partitions with protective percentage quotes ($k = 1, \dots, p$)

Partitions are being introduced with an aim to protect small customers with an opportunity to growth in the future. Granting a quote to the partition with groups containing small customers will provide a certain quantity of products for small customers who have great potential for development and also for the increase in quantities to be ordered in the future. The protective quote, i.e. the percentage portion of the overall production in l -iteration is defined for every partition, and it will be available to the observed partition.

$$0 \leq KP_k \leq 1, \text{ for } k \in \{1, \dots, p\}, \quad \sum_{k=1}^p KP_k = 1$$

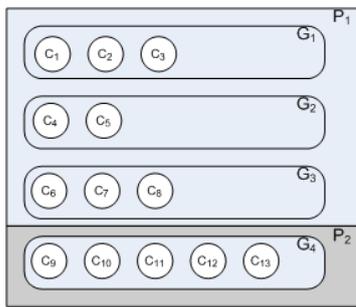
In addition, the number of partitions is smaller or equal to the number of groups. The concept of partition will lose sense (purpose) if only one partition has been defined.

5. Clustering groups into partitions

Every defined partition must have at least one group, even though such case counts as if every group has been awarded the protective quote (Figure 1a).

$$G_j \in P_k, \text{ for } j \in \{1, \dots, m\} \text{ and } k \in \{1, \dots, p\}$$

Information on which group belongs to what partition are kept in the matrix group partition (MGP), as shown on Figure 1c.



	G ₁	G ₂	G ₃	G ₄
C ₁	1	0	0	0
C ₂	1	0	0	0
C ₃	1	0	0	0
C ₄	0	1	0	0
C ₅	0	1	0	0
C ₆	0	0	1	0
C ₇	0	0	1	0
C ₈	0	0	1	0
C ₉	0	0	0	1
C ₁₀	0	0	0	1
C ₁₁	0	0	0	1
C ₁₂	0	0	0	1
C ₁₃	0	0	0	1

	P ₁	P ₂
G ₁	1	0
G ₂	1	0
G ₃	1	0
G ₄	0	1

Figure 1a: Customers, groups, partitions

Figure 1b: Matrix customer group

Figure 1c: Matrix group partition

If we mark the customer i demand with OC_i , then we could say that the total order for products in l -iteration is:

$$OT_l = \sum_{i=1}^n OC_i \quad \text{or} \quad OT_l = \sum_{j=1}^m \sum_{i=1}^n OC_i \cdot MCG_{ij}, \quad \text{for } l \in \{1, \dots, r\}$$

Variable AP_k is then introduced with the goal to keep the amount of products allocated to k -partition in l -iteration:

$$AP_k = PT_l \cdot KP_k, \quad PT_l = \sum_{k=1}^p AP_k, \quad k \in \{1, \dots, p\} \text{ and } l \in \{1, \dots, r\}$$

where PT_l is the total production in l -iteration.

Table 2: Variable descriptions

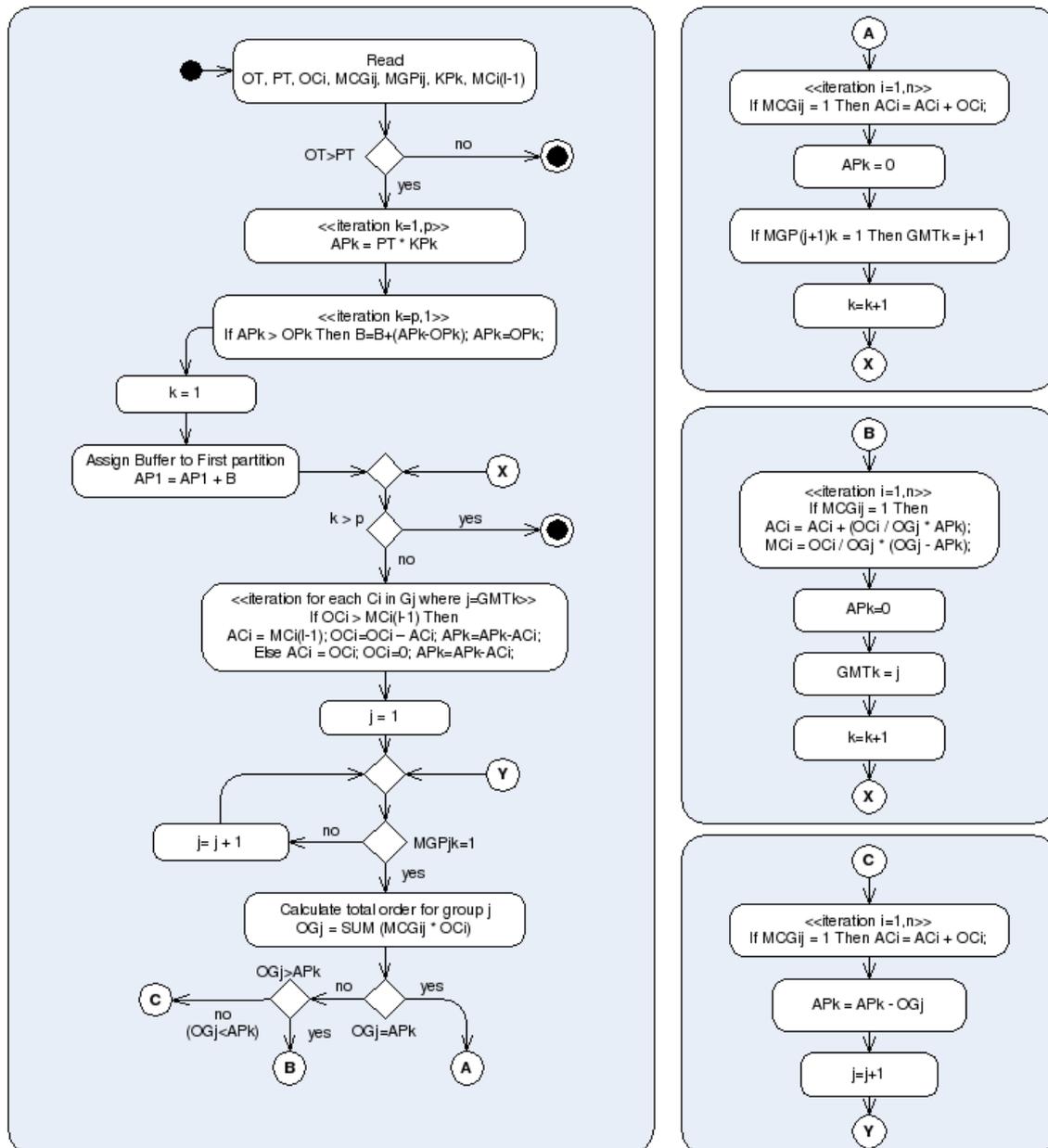


Figure 2: Algorithm represented as UML 2.0 Activity Diagram

Variables	Description
OC_i	Order of customer i in l -iteration
AC_i	Allocated quantity of customer i in l -iteration
$NC_j = \sum_{i=1}^n MCG_{ij}$	Number of customers in group j
$OG_j = \sum_{i=1}^n MCG_{ij} \cdot OC_i$	Total order of group j
$NG_k = \sum_{j=1}^m MGP_{jk}$	Number of groups in partition k
$OP_k = \sum_{j=1}^m \sum_{i=1}^n OC_i \cdot MCG_{ij} \cdot MGP_{jk}$	Total order of all customers in partition k
B_l	Buffer in l -iteration
MC_i	Memory - remembered quantity of unfulfilled part of the order of customer i in iteration l
GMT_k	Order number of group in partition k where token is set

The diagram on Figure 2 shows that the allocation algorithm in every iteration starts from checking whether $AP_k > OP_k$. Checking goes from the last partition to the first one, and if this condition has been met, then all allocated surplus is transferred to a joint buffer, whereas the allocated amount of finished products transferred to k -partition now equals the overall demand in given partition (first set $B_l = B_l + (AP_k - OP_k)$, and then assign $AP_k = OP_k$). This situation might happen in partitions that gather customers with small orders in case that higher quote than necessary is given in the beginning. This is a sign that one has to reaccess the analysis of partition of 'small' customers, and award smaller and more reasonable protective quote (*fine tuning*). Within the algorithm, all amounts transferred to buffer are added to the allocated amount of finished products of the first partition ($AP_1 = AP_1 + B_l$).

One of the most important concepts of the proposed algorithm is group memory token (GMT). It is introduced in order to designate a group of customers whose order has not been fulfilled during an observed iteration. There could be only one group memory token (pointer) in each partition. In the end of one allocation iteration, token stays in a certain group with unfulfilled demand, which gives a priority to that group in the following iteration, so that the customers would get the unfulfilled order in the following iteration as a priority. In order to remember the unfulfilled order of group of customers with GMT, the variable MC (Memory of customer) is introduced. We need to stress that MC obtains the value for every customer from the groups

that kept token in the observed iteration, and it is used only during the allocation process in the following iteration.

Formula for calculating the unfulfilled part of an order is:

$$MC_i = \frac{OC_i}{OG_j} \cdot (OG_j - AP_k) \cdot MCG_{ij}$$

where MC_i is remembered quantity of unfulfilled part of an order of i -customer in l -iteration that is being transferred to the following iteration.

In case GMT in several iterations stays in the same group of the first partition of important customers, then the necessity of expanding capacities becomes clear.

5 Conclusion

The goal of this paper was to create an algorithm for solving the problem of allocation of limited stocks to incoming orders. Order fulfillment is carried out by letting the allocation algorithm grant stocks in line with the priority of customers, with the goal of providing high level of customer service. The algorithm proposed in this paper provides following: (1) classification in groups will provide the order of allocation with primarily focus to satisfy customers that are important for the company, in accordance with previously defined criteria, (2) application of partitions will help certain groups of lower priority within protected partitions to be involved in allocation so that the low-ranked customers would be at least partially satisfied, which keeps all the customers in the system and sends a useful signal to company's management saying that there is still unsatisfied demand, and they need to enlarge the production capacities, (3) introduction of GMT allows all customers within a group to be delivered the backorders from the previous cycle, with the extended delivery lead time, with which they improve the overall customer service.

The main differences between the allocation algorithm presented in this paper, and the models described in the related work are: (1) this allocation algorithm directly affects the long-term operating results by accomplishing the customer service as the primary goal, unlike other models that focus only on short-term profit generation, and (2) there is a tendency of bounding customers with different priority for a long-term period, i.e. the tendency of keeping the customers within the system. In addition, there are also differences in basic goals of models: the primary goal of advanced algorithm is the maximization of customer service, unlike other systems that are primarily orientated towards profit maximization.

Further research will be focused on testing and improving algorithm based on real data obtained in other companies and also from different industries. In addition, there are possibilities for further improvement of the allocation algorithm by creating a report system that would support management in making timely decisions on the change in operating strategy.

Bibliography

- [1] Cederborg O., Rudberg M.: Customer Segmentation and Capable-to-promise in a Capacity Constrained Manufacturing Environment, 16th International Annual EurOMA Conference, Göteborg, Sweden, June 14-17, 2009.
- [2] Chan F.T., Chung S.H.: A modified multi-criterion genetic algorithm for order fulfillment in manufacturing network, Proceedings of the 9th Asia Pacific Industrial Engineering & Management System Conference, APIEMS, Indonesia, 2008.

- [3] Chen J.H., Lin J., Wu Y.S.: Order Promising Rolling Planning with ATP/CTP Reallocation Mechanism, IEMS, APIEMS, Vol. 7, No.1, 2008.
- [4] Kaschel H., Bernal L.M.S.: Importance of Flexibility in Manufacturing Systems, International Journal of Computers, Communications & Control, Vol. I, No. 2, 2006.
- [5] Lawrence G.: Introducing APS: Getting Production in Lock Step with Customer Demand, Automotive Manufacturing&Production, Vol. 110, Issue 5, 1998.
- [6] Lin J., Chen J.H.: Enhance order promising with ATP allocation planning considering material and capacity constraints, Journal of the Chinese Institute of Industrial Engineers, Vol.22, No.4, 2005.
- [7] Lupse V., Dzitac I., Dzitac S., Manolescu A., Manolescu M.-J., CRM Kernel-based Integrated Information System for a SME: An Object-oriented Design, *International Journal of Computers Communications and Control*, ISSN 1841-9836, Suppl. S, 3(S): 375-380, 2008.
- [8] Makatsoris H.C., Chang Y.S., Richards H.D.: Design of a distributed order promising system and environment for a globally dispersed supply chain, Int. J. Computer Integrated Manufacturing, Vol. 17, No. 8, 2004.
- [9] Meyr H.: Customer segmentation, allocation planning and order promising in make-to-stock production, OR Spectrum, Vol 31, No 1, 2009.
- [10] Rudberg W., Wikner J.: Mass customization in terms of the customer order decoupling point, Production Planning & Control, Vol. 15, No. 4, 2004.
- [11] Wikner J., Rudberg M.: Introducing a customer order decoupling zone in logistics decision-making, International Journal of Logistics: Research and Applications, Vol 8, No. 3, 2005.