

# Cargo handling processes during containerized shipments: mathematical modeling, optimality criteria

Igor Rusinov<sup>1</sup>, Irina Gavrilova<sup>1\*</sup>, and Aleksandr Bersenev<sup>1</sup>

<sup>1</sup>Admiral Makarov State University of Maritime and Inland Shipping, Dvinskaya st., 5/7, Saint-Petersburg, 198035, Russia

**Abstract.** Containerized shipment is the most convenient and cost-saving method for transportation of goods. In this article the processes of containerized cargo handling based on a queuing model is formalized. Different cargo handling models are described, and the process modeling based on local mutual assistance is considered. Analytical models for process performance indicators in steady-state conditions are defined. Optimization issues in designing and operating transshipment terminals are identified and the ways of resolving these issues are offered. Analytical models of process performance indicators, helping to optimize cargo handling processes in special conditions, are defined.

## 1 Problem setting (Introduction)

A cargo container is designed to ensure that cargo remains safe when shipped in any combination of air, water and land modes of transport. Thus, cargo handling processes speed up, costs are minimized at transport hubs and monitoring is simpler to carry out.

The main transport hub for containerized shipments is a specialized container terminal. Goods delivered by rail, sea and river vessels are handled at the specialized terminal.

Cargo handling channels (sea and river berths and piers) are an integral part of any cargo terminal. When designing and operating the terminals it is necessary to define the performance indicators for cargo handling processes specifying an average waiting time for vessels and an average time vessels stay in a terminal as well as an average number of vessels staying in queue or in a terminal, which is proportional to the above average values.

These criteria characterize the quality of services as waiting in queue can result in substantial losses for shipping companies. As a result, despite an increase in cargo turnover, an extra increase in the channel (berth) cargo load factor becomes commercially unprofitable. Therefore, losses grow with an increase in the number of ships queuing to berth.

---

\* Corresponding author: [gavrilova.irina.a@gmail.com](mailto:gavrilova.irina.a@gmail.com)

## 2 Mathematical modeling of cargo handling processes

Today deterministic cargo handling models have come into common use. When applying these deterministic models, it is assumed that arriving of marine and river vessels at a terminal is a regular flow of events, where the vessels arrive one after another on a strict schedule at regular intervals. However, the deterministic models do not take into account how cargo terminals actually function. In reality arrivals of vessels at a terminal is an irregular flow of events. The time sea and river vessels and ships reach a terminal, deviates in many cases from the estimated one for a great number of reasons, not provided for by the schedule. These may include weather conditions, route updates and a number of other reasons. All these deviations are, as a rule, of random nature. Moreover, it should be considered that the cargo handling time, which itself depends on a number of random factors, is also random. In this connection, the cargo handling process in a cargo terminal takes place irregularly, which in some cases can result in ships queuing or, in other cases, in channels (berths) staying idle. Therefore, probability models need to be employed to describe cargo handling processes at cargo terminals. The processes taking place during cargo handling operations move from one state to another at random time moments. In addition, the number of ships staying in queue and the number of busy channels is not constant. One process state changes into another, when a new ship approaches a terminal or one of the channels becomes available. The system contains a  $n+1$  variety of states:  $E_0, E_1, E_2, \dots, E_n$ , where  $n$  is the number of ships in the system, in other words, both ships staying in queue and those being handled are taken into account. As it is known, the total flow of events when sea and river vessels and ships arrive at a terminal can be viewed as a sum of ships that are owned by different companies and deliver different goods [1]. At the same time, it is known that overlapping a great number of similar steady flows with practically any consequence results in a flow, which is arbitrary close to the Poisson (simplest) flow. Let us assume that the time required to handle a ship is distributed exponentially. Taking into consideration these assumptions on the Poisson flow of ship arrivals and exponential distribution of cargo handling time, it is possible to apply a queuing model to describe the processes occurring in terminals. Using the queuing theory allows us to describe the cargo handling process in a terminal by means of linear differential equations and to represent the probabilistic indices for process quality in an analytical form.

However, using the existing queuing models to determine the probabilistic characteristics of cargo handling processes does not seem appropriate as the specified models inadequately describe these processes in real operating conditions.

So, the classical queuing theory involves research into a multichannel system, where the number of servers  $S$  is equal to the number of channels. Each channel is served by one server, independently of other channels (a QS without mutual assistance between the servers). Otherwise, channels can also be serviced by all vacant facilities or by some of these vacant facilities (a QS with full or partial mutual assistance). The probabilities for the system to move from  $E_n$  state into  $E_{n-1}$  state, in other words, the probability of servicing one order depends on the number of service channels in operation. The resulting intensity of service in  $n$ -th state is determined according to the principle of linear superposition, i.e. it is equal to the total intensity of all servers and is divisible by the estimated intensity of one facility  $\mu_0$ . Thus, the resulting service intensity in this situation cannot exceed  $S \mu_0$ , i.e.,  $\mu_p \leq S \mu_0$  and the service intensity by one facility  $\mu_0$  does not change, depending on the QS state. Moreover, the service process is considered unpredictable and uncontrollable, in other words, the QS administrator does not know the number of orders that will come into the system in the nearest time and cannot change the intensity of service facilities depending on the QS state.

In real operating conditions of cargo terminals, the cargo handling processes do not comply with the given assumptions. Therefore, it is necessary to consider a centralized system of cargo handling, which is controlled by the terminal administrator. The

administrator determines the queuing and service disciplines, and monitors the distribution of human and technical resources between individual channels. If required, when the ship queue grows significantly, the administrator can allocate additional resources thus increasing the cargo handling intensity by individual channels. However, in most cases due to the limited area of transshipment operation, the output intensity of the system is less than the total intensity of separate technical means. Thus, in real conditions the output intensity of cargo handling, as a rule, can be divisible by an average intensity of cargo handling  $\mu_0$  and, in some cases, can exceed  $S\mu_0$ .

The given properties of the centralized cargo handling system require that the probability models to be developed take into account possible changes in the intensity of separate service facilities depending on the QS state. However, the classical queuing theory operates with constant values of intensity shown by individual servers. Let us consider a container terminal including  $S$  of channels (berths) with input intensity  $\lambda$  of the simplest flow of ships (orders). The estimated handling intensity of each berth is equal to  $\mu_0$ . However, the output intensity of the process can change depending on its state. The output intensity of cargo handling in condition  $E_n=r_n \mu_0$ ,

Where  $r_n$  is the factor of handling intensity, which can be both an integer and a rational number. As a rule, when all berths are busy, in other words, when  $n \geq S$  it, is assumed that  $r_n = \text{const}$  (usually  $r_n = r_{\text{max}}$ ). Then  $n+1$  differential equations describing the cargo handling process can be presented as follows:

$$\begin{aligned} P'_0(t) &= -\lambda P_0(t) + r_1 \mu_0 P_1(t) \\ &\dots \dots \dots \\ P'_n(t) &= -\lambda P_{n-1}(t) + (\lambda + r_1 \mu_0) P_n(t) + r_{n+1} \mu_0 P_{n+1}(t); \quad n = 1.2.3.\dots \end{aligned} \quad (1)$$

Let us assume that the cargo handling process is Markovian, random and ergodic.

The operation mode of cargo terminal, in which probabilities  $P_n$  for the system to be in  $n$  state do not depend on the time, is called a steady-state condition. To determine probabilities of individual states in steady-state conditions, it is necessary to set the values of derived states to zero, in other words, the left parts of the system of equations (1) as well as to transfer one of the term to the left hand side in each equation. As a result, we obtain:

$$\begin{aligned} r_1 \mu_0 P_1 &= \lambda P_0 \\ r_2 \mu_0 P_2 &= (\lambda + r_1 \mu_0) P_1 - \lambda P_0 \\ r_3 \mu_0 P_3 &= (\lambda + r_2 \mu_0) P_2 - \lambda P_1 \\ &\dots \dots \dots \\ r_n \mu_0 P_n &= (\lambda + r_{n-1} \mu_0) P_{n-1} - \lambda P_{n-1} \end{aligned} \quad (2)$$

### 3 Identification of analytical models to measure process performance

Let us introduce a parameter  $\psi = \frac{\lambda}{\mu_0}$ , the given flow density of ship arrivals. Then, having solved the system of equations (2) and considering that all states represent the full group of events, we obtain an equation for so called rated condition:

$$P_0 \left[ \sum_{n=0}^{S-1} \frac{\psi^S}{\prod_{n=1}^S r_n} + \frac{\psi^S}{\prod_{n=1}^{S-1} r_n} \sum_{d=0}^{\infty} \left( \frac{\psi}{r_{\max}} \right)^d \right] = 1 \tag{3}$$

where  $d=n-S$  is the number of ships waiting in queue.  
 Let us determine an average number of ships in a queue.

$$\bar{d} = \sum_{n=S+1}^{\infty} (n-S)P_n = P_0 \frac{\psi^S}{\prod_{i=1}^S r_i} \frac{r_{\max}}{\left(1 - \frac{\psi}{r_{\max}}\right)^2} = P_0 \frac{\psi^{S+1}}{\prod_{i=1}^{S-1} r_i (r_{\max} - \psi)^2} \tag{4}$$

Therefore, an average number of ships staying in the terminal is:

$$\bar{n} = \sum_{n=0}^{\infty} nP_n = \sum_{n=0}^S nP_n + \sum_{n=S+1}^{\infty} nP_n = \bar{d} + S - \sum_{n=0}^S (S-n) \frac{\psi^n}{\prod_{n=1}^S r_n} P_0 \tag{5}$$

An average time of ships waiting in a queue and an average time of a ship staying in the terminal is determined by the Little's formulas.

An average time of a ship waiting to be unloaded or loaded due to busy channels is:

$$\bar{T}_{waiting} = \frac{\bar{d}}{\lambda} = \frac{P_0}{\lambda} \frac{\psi^{S+1}}{\prod_{i=1}^{S-1} r_i (r_{\max} - \psi)^2} \tag{6}$$

An average total time of a ship staying in the terminal is:

$$\bar{T}_{\Sigma} = \frac{\bar{n}}{\lambda} = \frac{\bar{d}}{\lambda} + \frac{S}{\lambda} + \frac{1}{\lambda} \left[ S - \sum_{n=0}^S (S-n)P_n \right] = \bar{T}_{oz} + \frac{1}{\lambda} \left[ S - \sum_{n=0}^S (S-n)P_n \right] \tag{7}$$

Consequently, an average time required to handle one ship is:

$$\bar{T}_{andling} = \frac{1}{\lambda} [S - \sum_{n=0}^S (S-n)P_n] \tag{8}$$

The cargo handling processes in specialized conditions are optimized based on the given analytical expressions of probability measures.

## 4 Solving optimization problems when designing and operating transshipment terminals

When operating and designing transshipment terminals, a number of optimization challenges need to be considered, such as the following:

- determination of an optimal number of berths and a number of ships based on the planned shipment traffic of coastal cargo;
- planning of optimal load for a transshipment terminal, in particular, determination of an optimal output intensity flow of coastal vessels arrivals at the port;
- validation of a development program for capacity building of transshipment terminal;
- monitoring terminal operating effectiveness [4].

Formalization of the optimization issues can be challenging, given that coastal cargo handling processes depend on multiple criteria and in most cases there is no information required to obtain the quantitative solutions to these problems.

Let us consider the process performance indicators, which need to be considered to identify the optimal solutions. The most well-known and traditional performance indicator for terminals is the berth productivity representing the amount of cargo that can be handled per unit time, corresponding to a unit length of the berthing line. However, given that the intensity of coastal cargo handling is determined by material and technical resources, personnel qualification, warehouses' location and access to roads, the maximum berth productivity can be achieved only when one ship comes after another on a non-stop basis, in other words, when the berth load factor is close to one. It is possible only when there is constant queue of ships, i.e. when increasing waiting time of ships (idle time), thus greatly deteriorating the quality of services rendered to shipping companies, which may result in most cases in fines, in compliance with contract terms and conditions. If the transshipment terminal and ships belong to the same company, for example, as in case of Norilsk Nickel MC PJSC, then such an idle time results in direct losses for the company. Therefore, the berth productivity is the most important indicator. However, it is not the only indicator to measure a transshipment terminal performance. In addition, calculating this indicator requires a huge amount of initial data, which are not always known. Therefore, in most cases the berth load indicator can be used instead, which provides a relatively correct indirect evaluation of the berth productivity, when the intensity of coastal ship handling is given.

Indicators, such as an average waiting time, an average time of ships staying in the terminal, as well as an average number of ships being in queue or in the terminal, which are proportional to the first two, are particularly important. These indicators measure the quality of services, since ships waiting in a queue may lead, as described above, to significant losses. The amount of such losses depends on the operating costs and losses in ship payload delivery rate and is evaluated in terms of penalties or the amount of direct loss, as in the case of Norilsk Nickel MC PJSC.

As a result, despite an increase in cargo turnover, an excessive increase in channel (berth) cargo load deems to be commercially unprofitable. Moreover, losses grow with an increase in the number of ships waiting to berth.

The third group of indicators are economic indicators, including capital, current costs and costs incurred, as well as the indicators reflecting the size of current financial revenues. The latter include net present value, real income, and, finally, terminal's net income. The analysis of these economic indicators shows that the net income is the indicator, which reflects the transshipment terminal's commercial activity the most comprehensively, because the maximum profit is possible only when the terminal functions in the most optimal way. If the terminal's handling planning is set (the number of ships and the inflow rate of these ships are known), in other words, the expected net present values are known, then the minimum costs incurred will correspond to the maximum net income. This can be used, for example, to determine the optimal number of berths. Considering all these groups of indicators can be challenging. This problem is addressed most effectively, when the terminal's profit or the costs incurred can be expressed through the berth load and an average number of ships waiting in queue. However, if there is no sufficient amount of initial data or the obtained optimal solutions to the optimization problem under consideration are inadequate, the heuristic methods can be applied helping to cover all aspects of the tasks in a more flexible way.

The optimal solutions obtained as a result of calculations can be taken as a basis for any of the above-mentioned tasks of optimizing the cargo handling processes [4].

The decision-making theory focuses on formalizing the tasks of defining optimal solutions based on several performance indicators (multi-criteria optimization tasks).

In such a case, the final number of solutions is assumed to be known. Then, the task of multi-criteria optimization is reduced to a comparative evaluation of such solutions based on an optimality criterion (a preference function), which is usually understood as a tool helping to ensure that various solutions are compared and an optimal solution is selected. For example, the optimal number of ships or berths is determined. This rule should consider all main requirements imposed on the quality of cargo handling processes. The preference function is based on preferability conditions for compared solutions.

When using mathematical methods to make decisions, it is necessary to select a specific mathematical model of preference function, which is a multi-criteria optimization. In multi-criteria optimization, a comparative, heuristic evaluation of options by preferences is conducted not immediately but using the set of quality indicators that represents numerical functions  $K_1, K_2, \dots, K_l$  from optimized parameters  $x_1, x_2, \dots, x_n$ . Such functions are usually called the performance indicators or criteria. Their specificities in relation to the containerized cargo handling process are described above.

The process performance indicators are called homogeneous, if they have a common scale. However, in reality the process performance indicators, as a rule, are presented on the ratio basis and have different scales [2]. Thus, to solve the multi-criteria optimization tasks the heterogeneous scales are converted into homogeneous using the normalization principle (rated). Today more or less different normalization methods of performance indicators are used. Let one-sided constraints be imposed on the value of comparative performance indicators. Then, depending on the selected performance indicator; its extreme value can be minimum and or maximum. First of all, let us consider the normalization issues of minimized process performance indicators (a total time of staying in the terminal, waiting time) and then by analogy - the maximized indicators (berth load).

Let us define the normalization rules for minimized indicators. Let us assume that the following ratio  $Z^i = 1$ , when  $K^\rho = K^{\rho \max}$  and  $Z^i = 0$ , when  $K^\rho = K^{\rho \min}$  are true for the normalized values of performance indicators. Let us consider the most common way for proportional normalization of performance indicators. In this case the normalized values of minimized indicators will be determined by the following expression:

$$Z_p = \frac{K_{\rho \max} - K_\rho}{K_{\rho \max} - K_{\rho \min}} \quad (9)$$

The normalized value of maximized indicators respectively:

$$Z_p = \frac{K_\rho - K_{\rho \min}}{K_{\rho \max} - K_{\rho \min}} \quad (10)$$

Thus, in a number of cases, to select optimal solutions, their normalized values  $Z_1, Z_2, \dots, Z_l$  considered instead of indicator values  $K_1, K_2, \dots, K_l$ . The normalized values are considered homogeneous if they have a common scale, in other words, the same closed multitudes of normalized indicators are determined for each indicator  $Z_p (p = 1, 2, \dots, m)$ .

Let us consider an infinite multitude of options  $R$ , which are understood to mean a direct product of multitudes  $Z_1, Z_2, \dots, Z_l$ , i.e.,  $R = Z_1 * Z_2 * \dots * Z_l$ , where the given product is a multitude of various vectors is  $[Z_1, Z_2, \dots, Z_l]^T$ . The preference function  $Y$  for initial situation, which has not been formalized, can be constructed using a variety of methods, so that these functions would fully cover preferences, when selecting the optimal solutions.

## 5 Optimality criteria of cargo handling processes

Let us assume that the following one-sided constraints are imposed on the cargo handling performance indicators:

$$\begin{aligned} K_{\rho}(x_1, x_2, \dots, x_n) &\geq K_{\rho \min} & \rho = 1, 2, \dots, l_1 \\ K_{\rho}(x_1, x_2, \dots, x_n) &\leq K_{\rho \max} & \rho = l_1 + 1, l_1 + 2, l_2 + \dots + l \end{aligned} \quad (11)$$

Moreover, the two-sided constraints are imposed on the values of optimized process parameters:

$$X_{i \min} \leq X_i \leq X_{i \max} \quad (12)$$

Constraints (1.5) and (1.6) determine a multitude (a region) of permissible parameter values. An optimal solution is selected from all permissible solutions. It is assumed that it has the best vector value  $\bar{Y}$  of normalized performance indicators in terms of accepted optimality criterion.

In this setting, the solution  $S'$  is the best for the multitude  $S$  if inequality is true:

$$Z_{\rho}(S') \geq Z_{\rho}(S''); S'' \in S; \rho = 1, 2, \dots, l \quad (13)$$

If there are similar solutions, then they are equivalent. If the inequality (13) is strictly fulfilled for at least one indicator, then the best solution is only one. If the best solution is technically feasible and meets constraints (11) and (12), then it belongs to a permissible multitude of solutions  $L_p$ . Let us split a permissible multitude of system solutions into two subsets. The inefficient (worst) solutions would be in the first subset. Solution  $X(S')$  is called inefficient, if multitude  $L_p$  has at least one solution  $X(S'')$  with the best vector value of normalized performance indicators in terms of unconditional optimality criterion. Otherwise, let us call a solution ineffective (not worst) and classify it as the second subset of effective options  $L_e$ , providing that all its solutions cannot be improved from the standpoint of Pareto optimality principle.

The problem of selecting a conventional optimality criterion usually comes down to mapping the vector of normalized performance indicators  $\bar{Z} = [Z_1, Z_2, \dots, Z_l]^T$  on a preference function (a target function)  $Y = F(Z_1, Z_2, \dots, Z_l)$ , which is a scalar function of optimized indicators. An optimal solution is selected by determining a permissible solution corresponding to the extreme value of preference function. The selection of a conventional quality criterion in every individual case is determined by specificities of the cargo handling process.

An optimal option of the process  $S_0$  should be an efficient option included in a multitude of efficient options  $L_e$ .

Today methods based on the principle of "flexible priority" are commonly accepted [3].

In general, this principle of "flexible priority" is reduced to additional normalization of indicator space, thus within reasonable limits the degree of preference of one indicator over another is taken into account. In this case, it is assumed that the influence of each  $\rho$  single indicator on the value of preference function depends not only on the normalized value of this indicator but on a weight factor  $M_{\rho}$ , i.e. the importance of the indicator. Then, the preference function can be presented as a function from the normalized values of single indicators  $Z_{\rho}$  and weight factors  $M_{\rho}$ . Now, there is a number of various types of preference

functions taking into account the principle of flexible priority, among which simple exponentials have gained widespread use:

$$Y = \left[ \sum_{\rho=1}^n M_{\rho} Z_{\rho}^V \right]^{1/V} \quad (14)$$

In this case, it is assumed that the weight factors are also normalized, so that:

$$\sum_{\rho=1}^m M_{\rho} = 1$$

Changing values  $V$  in equation (14) makes it possible to get different modifications of preference functions.

If  $V=1$ , then we will get an additive function:

$$Y = \sum_{\rho=1}^m M_{\rho} Z_{\rho} \quad (15)$$

This function based on an arithmetic mean estimate has gained most widespread use in solving the tasks of multi-criteria optimization. The latter is explained both by simple interpretation of this preference function and by the circumstance that its use may be substantiated by the additive utility theory [5]. In this case, as before mentioned, along with constraints on the range of change in parameters (12) constraints on the values of single simple exponentials (11) are considered as well. However, it may happen that an optimal solution will be at the boundary of a region to be determined by inequalities (11), in other words, an insufficient level of some indicators will be compensated by a high level of other indicators. This approach can turn out to be unacceptable in many practical cases [3].

A preference function based on an average geometric estimate is widely accepted for a family of equations (14):

$$Y = \prod_{\rho=1}^m Z_{\rho}^{M_{\rho}}, \quad (16)$$

which is obtained, when  $V \rightarrow 0$ .

From the equation (16), this preference function significantly decreases and then turns to zero with any indicator trying to reach zero.

## 6 Conclusion

One of the main scientific results of this article is the development of classical queuing theory considering the functional specificities of a cargo terminal, in other words, considering the possibility of changing the value of service intensity by individual servers depending on the QS state.

In this article it is also revealed that at present the integrated quality assessment methods recommend that a geometric average should be used for a high spread of normalized values of single indicator and an arithmetic average be used for a low spread. In a similar way it is possible to work out equations for the other types of preference functions that are characterized by different values  $V$ .

This article offers a criterion based on a non-linear non-additive preference function. Unlike the preference functions considered in this article, this function has high potential adequacy, which makes it possible to consider the preference function-single indicator non-linearity and mutual influence of various single indicators.

## References

1. B. V. Gnedenko, I. N. Kovalenko. *Introduction to the queuing theory*. (Moscow, Science, 1987)
2. L. T. Kuzin. *Basics of cybernetics, 2*, (Moscow, Energy, 1979)
3. I. A. Rusinov. *Handling and storage of refrigerated cargos at a specialized terminal*. (SPB, SPBII RAS, 2005)
4. I. A. Rusinov. PhD Thesis (Saint Petersburg, 2010)
5. I. A. Rusinov. Operation of maritime transport. **4 (58)**. 8-9 (2009)