A MODEL OF THE DIMENSIONING OF THE NUMBER OF SERVICE PLACES AT PARKING LOT ENTRANCES BY USING THE QUEUING THEORY

Ljudevit Krpan, Robert Maršanić, Marin Milković

Original scientific paper

The issue of the optimal management of the traffic flow and the ensuring of an adequate number of parking spaces is present in the majority of highly developed urban areas of the European Union, including the Republic of Croatia. The formation of an adequate raster of the transport network, conditioned by the historical and existing buildings and areas is an extremely challenging task. Similarly, ensuring an adequate availability of free parking capacities in times of rush hours in the very centres of urban areas is also very demanding. However, what often occurs, even when adequate parking capacities are ensured, is a halt on the input and output terminals of parking lots and garage facilities that are derived by a disproportion of their capacities and traffic needs at a specific time. Precisely for that reason, this paper presents a potential model for determining of an optimal number of input service places (terminals) by using the queuing theory.

Keywords: queue; queuing system; parking; traffic modelling; traffic planning

Model dimenzioniranja broja uslužnih mjesta na ulazima u parkirališta primjenom teorije redova čekanja

Izvorni znanstveni članak

Problematika optimalnog vođenja prometnog toka te osiguranja odgovarajućeg broja parkirnih mjesta prisutan je u većini visokorazvijenih urbanih sredina Europske unije pa i Republike Hrvatske. Formiranje odgovarajućeg rastera prometne mreže, uvjetovane povijesnim i postojećim građevinama i zonama izuzetno je izazovan zadatak. Isto tako, zahtjevno je i osiguranje odgovarajuće dostupnosti do slobodnih parkirnih kapaciteta u vremenima vršnih opterećenja u samim središtima urbanih područja. No, vrlo su česti slučajevi u kojima, čak i u trenutku kada se osiguraju odgovarajući parkirni kapaciteti, dolazi do zastoja na ulaznim i izlaznim terminalima parkirališta i garažnih objekata koja su derivirana nesukladnostima njihovih kapaciteta i prometnih potreba u određenom vremenskom trenutku. Upravo iz tog razloga, u radu se predstavlja potencijalni model određivanja optimalnog broja ulaznih uslužnih mjesta (terminala) primjenom teorije redova čekanja.

Ključne riječi: parking; prometno modeliranje; prometno planiranje; red čekanja; sustav opsluživanja

1 Introduction

As a rule, traffic problems of urban areas are conditioned by transport demands and limited traffic areas that are conditioned by geomorphological features and previously constructed zones, i.e. objects. Precisely in order to ensure the maximum utilization of available traffic capacities it is necessary, among other things, to optimize the flows of traffic management. One of the key elements is to ensure adequate reception capacities at the entrances to the existing closed parking lots and garage facilities.

The purpose of this paper is to present a theoretical model of the application of the queuing theory on the optimization of the number of entrance ramps at closed parking lots and garage facilities.

The aim of the paper is to propose, through theoretical research, a model of determining the number of input terminals at the entrances to closed parking lots and garage facilities.

Based on the established purpose and objective of the research, the following hypothesis is proposed: "By applying mathematical method of the queuing theory it is possible to conduct optimization within the planning and construction of entrance terminals to the parking lots for motor vehicles".

The paper is structured into seven chapters. After the Introduction, the basic characteristics of service system are introduced in the second part. The third part, titled Parking Lot as a Service System, determines the framework of possible application of service system to parking (garage) spaces with closed systems of payment. The fourth part demonstrates the rules that apply during arrivals of vehicles at the parking lot and during vehicles being serviced while waiting at the entrance to the parking area.

The fifth part is the Conclusion which concisely presents the results of the research and confirms the research hypothesis. The sixth part consists of footnotes while the seventh part includes the bibliography used in preparation of the paper.

When analysing the quality of the service of input terminals in the system, the costs of a potential introduction of a new input terminal will not be analysed. Likewise, the paper does not take into consideration the elements of the parking space congestion conditioned by insufficient parking capacities.

2 Overview of queuing systems

The queuing theory is one of the methods of operational research that studies the problems of queuing, whose content is the serving of randomly received units or requests for a service. The queuing theory then uses mathematical models by means of which the following are determined: interdependence between the arriving units, their wait for a service, the service, and at the end the exit of the unit from the system.

The emergence of the queue is the result of the non-compliance of the capacity of service places with the requirements of service users. The capacity of a service place is expressed through a number of units (users) that can be served in a unit of time. If the average number of units requesting a service is greater than the capacity of a service place, the existing capacities cannot meet the requests, so the units are forced to wait on the service in a

queue. If the average number of units is lesser than the capacity of service places, the units will be served immediately, but if we have a case of over-dimensioned capacities, what will occur is the "waiting" of service places that will not be fully utilized.

The problem of queuing is the process of serving that can be the subject of the theory of general systems, where the name "queuing system" comes from. To solve the problem of queuing means to determine the optimal number of service places for which the time of waiting in line or the expenses (losses) caused by waiting will be minimal. It follows that by solving the problem of queuing, it will not be possible to completely eliminate the waiting, but the losses caused by waiting will be brought to a minimum. So, queueing theory consists of three parts: input process, queuing rules and service agencies [1, 2].
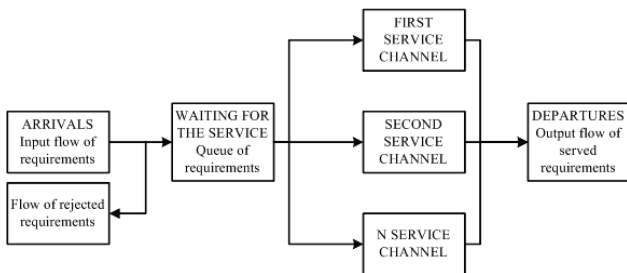

Figure 1 The process of waiting for more service places (channels) [2], adapted by the authors

The service system is then defined by the following elements [3]: 1) the distribution of the arrival time of units (request for a service), 2) the distribution of the service time, 3) the number of service places (serving channels), 4) the capacity of the service system, 5) the order of serving the units, i.e. the discipline of the queue, and 6) the number of service degrees.

The distribution of the arrival time of units is defined by the time between two successive arrivals of units to the service system. Given the size of the interval between the two arrivals, they can be: 1) with equal intervals of time, 2) with unequal, but pre-defined (determined) time intervals, and 3) with irregular intervals of time that are not known in advance, but are random and with a known probability distribution.

According to the source of the arrival of units, the problem of queuing is **open or closed**. In open systems of serving, the intensity of arrivals does not depend on the state of the system, and the source of the service user (the place where units come from) is located outside the system. In closed systems of serving, the intensity of the arrivals of units depends on the state of the system, and the sources where units come from are not outside the system, but are the internal elements of the system.

In order to analyse the functioning of the given problem of queuing, it is necessary to determine the number of units that come into the system, i.e. that are looking for a service or treatment per unit of time. Because of random arrivals, the number of units is not equal in each time unit, and therefore, what is collected is the average number of units (users) in a unit of time, which is also called the intensity of the flow of arrivals and is denoted by the Greek letter λ.

The arrivals of units in the system can be deterministic, and in such cases it is sufficient to know how much is λ (the average number of units received by the system in a unit of time) or t_{arr} (the time interval between two consecutive arrivals). However, in practice it is usually the case that the arrivals of units are random (stochastic), and then in addition to λ and t_{arr}, it is necessary to know their probability functions. Realistically, an unlimited number of different distributions of arrivals is possible. The distributions of arrivals are usually classified according to the type of arrivals to [4]: 1) properly distributed arrivals, 2) completely random arrivals (exponential distribution of arrivals), 3) arrivals distributed according to the Erlang distribution of the k order, and 4) generally independently distributed arrivals.

Properly distributed arrivals (mark D) relate to the distribution of arrivals at the time when units (users) arrive at regular intervals. The intensity of arrivals can be calculated with the help of the relation

$$\lambda = \frac{1}{t_{arr}},\qquad(1)$$

where t_{arr} is the time interval between two successive arrivals.

Completely random arrivals (mark M) relate to units that randomly arrive to the system with an average time interval between two successive arrivals $\bar{t}_{arr}$, or an average number of units λ. Since this is the case of random arrivals, it is necessary to determine the function of probability.

Determining the type of distribution according to which the arrivals of entities in the system and the service time behave, is carried out by statistical tests, i.e. by testing the hypothesis on the compliance of the assumed theoretical probability distribution with the empirical probability distribution. Consent, i.e. the coincidence of empirical and theoretical distributions, is examined so that a response on whether empirical frequencies adapt to some theoretical distribution is acquired. If the answer is positive, it means that in the future it is not necessary to carry out experiments (recording, monitoring) for each case, but that it is sufficient to take that theoretical distribution according to which the assigned empirical set (sample) behaves, and all conclusions about the observed phenomenon can be made on the basis of the theoretical distribution whose features are predefined and known in statistics.

In the process of examining the compliance of the empirical distribution with theoretical distribution, a hypothesis is set:

a) H_0 (null hypothesis) – the data of the empirical distribution behave according to the laws of the selected theoretical distribution,

b) H_1 (alternative hypothesis) – the data of the empirical distribution do not behave according to the laws of the selected theoretical distribution.

The decision relates to the acceptance of H_0 (meaning the rejection of H_1) or the acceptance of H_1 (meaning the rejection of H_0).

In practice, the compliance of empirical distributions with theoretical distributions is performed by applying non-parametric tests: Pearson's χ^2 - test, Kolmogorov - Smirnov test and other tests that are based on the comparison of empirical and theoretical frequencies, i.e. frequencies that are expected when the observed empirical distribution behaves according to a particular theoretical distribution.

One of the most often used tests is the χ^2 - test (chi-square test). The value χ^2 is calculated according to the formula

$$\chi^2 = \sum_{i=1}^{n} \frac{f_i - f_{t_i}}{f_{t_i}}, \chi^2 \geq 0, \qquad (2)$$

where:
f_i − is the empirical frequency, i.e. the real number of days with n vehicles,
f_{t_i} − is the expected (theoretical) frequency, i.e. the expected number of days when n vehicles will be at the parking lot,
$i = 1,..., n$-number of groups, i.e. pairs of values f_i and f_{t_i}.

The size of χ^2 is distributed according to the laws of Γ – distribution with k degrees of freedom. The number of freedom degrees is calculated from the relation

$$k = n - r - 1, \qquad (3)$$

for the Poisson distribution, $k = n - 1$ for normal distribution $k = n - 3$.

The value χ^2 calculated for the empirical distribution is compared to the value χ_0^2 from statistical tables which contain maximum deviations between empirical and theoretical frequencies, depending on the number of freedom degrees and the significance level. **Three cases are possible:**

$$\chi^2 = \chi_0^2, \ \chi^2 > \chi_0^2 \ \text{and} \ \chi^2 < \chi_0^2. \qquad (4)$$

If χ^2 is in the interval $[0, \chi_0^2)$, **the null hypothesis is accepted**, and if χ^2 is outside that interval, i.e. in the interval $[\chi_0^2, \infty)$, **the null hypothesis is rejected and the alternative hypothesis is accepted.**

The chi-square test is based on the assumption that $N \rightarrow \infty$, which means that in practice N has to be large enough (according to some authors, the order of magnitude of at least 100 frequencies). If the calculated value of theoretical frequency is $f_{t_i} < 5$, then data is re-grouped by adding two or more theoretical frequencies until f_{t_i} exceeds the value of 5. The value χ^2 cannot be lower than zero (because the sum of the square is in the numerator), theoretically it can be equal to zero, but that rarely happens in practice because that would mean that the empirical frequencies are completely equal to theoretical frequencies.

In the case where the intervals between successive arrivals and/or the duration of the service are irregular, but also of an unknown probability distribution, the queuing theory cannot be applied.

The distribution of service time is defined by the duration of the service, i.e. the length of time it takes for one service place to serve a unit. The duration or the time of carrying out a service can be: a) constant (always the same amount of time), b) variable, but known in advance (determined), and c) random when the service time is not known, but it is possible to determine the probability distribution. The service time is expressed through the number of time units required for serving one unit, i.e. for carrying out a certain service. With the help of that indicator, the throughput of the service channel is expressed. Service time can be analysed similarly to the arrivals of units, but there is a substantial difference between those two processes. The arrivals of units are distributed within the entire observed time interval, while the service is defined only when units are in the system; if there is no unit, the channel is unused.

If $\bar{f}_{ser}$ is the average time of the duration of a service, the average number of users served in a unit of time is

$$\mu = \frac{1}{\bar{f}_{ser}}. \qquad (5)$$

The service time can be: a) constant (mark D), b) distributed according to the exponential distribution (mark M), c) distributed according to the Erlang distribution of the k order (mark E_k) whose is the density of the shape, and d) distributed according to the general distribution (mark G).

The arrival of vehicles over a particular period of time does not depend on the number of vehicles that previously arrived to the parking lot. Therefore, it is said that the arrivals of vehicles are flows without consequences. This feature of the flow makes sense only if the vehicles are arriving from different directions, which is most often the case in parking lots.

A service place has a capacity that is expressed through a number of services that a service place can do in a unit of time. Due to the different duration of a particular service, here it is also necessary to take the average number of services per unit of time. That number implies the average number of units (users) that can be served in a unit of time and is called the intensity of serving per channel with a mark of the Greek letter μ.

The intensity of the flow of arrivals (λ) and **the intensity of serving** (μ) are basic parameters that determine a particular problem of queuing; λ indicates the capacity of the source from where units are arriving to the system, and μ indicates the capacity of service places. Both values are calculated as arithmetic means with the condition that the data about the arrivals of units and the number of services in the observed time unit are known.

The quotient of the intensity of the flow of arrivals and the intensity of serving is an indicator of the load of the service centre and is marked with the Greek letter ρ.

$$\rho = \frac{\lambda}{\mu}. \qquad (6)$$

Given the number of service places, it is important to distinguish between ρ and ρ_S; ρ is the size that refers to one service place (channel), and ρ_S refers to the system with more channels. When $\rho < 1$, the user will be served sooner or later, depending on the situation, but for $\rho \geq 1$ accumulation will increase over time and the system will not be able to function normally. In the long term, there would be a congestion of the system. Therefore, in order for the system to be stable, ρ must not be greater than or equal to 1, which means that λ should be smaller than μ. If that is not the case, the number of service places (channels) should be increased in order to satisfy the condition of the stability of the system

$$0 < \rho_S < 1, \text{ where } \rho_S = \frac{\rho}{S} \qquad (7)$$

However, this condition applies to systems with queues of unlimited length, in contrast to the queuing order with a limited queue length where the congestion of the system cannot happen since queuing is limited to m places, so the $(m + 1)$ unit receives a "cancellation" and has to leave the queue.

The number of channels is the number of service places in the serving system on which units can be simultaneously served. According to the number of channels, there are single-channelled and multi-channelled problems of queuing. The number of service places is either known in advance, or the number of necessary service places is determined on the basis of the elements that determine a certain waiting process. Single-channelled systems have one service place, while multi-channelled places have a couple of them. With multi-channelled systems, the queue can be: a) joint, meaning that for every individual application, a channel on which the unit will be served is not defined, and up until the moment of serving, that can be any of the existing channels, and b) a separate queue for each service place when a request, having entered in a certain queue, and selected according to a particular criterium, chooses a particular service place.

The capacity of the service system is the maximum number of units that are waiting in line to be served and of the units that are being served. When the service system is fully occupied, the unit that arrived cannot enter the system, i.e. it will enter the queue. Those systems that are of high capacity are generally accepted as systems with an infinite capacity. According to the possibility of the occurrence of queuing, two types of queuing systems are: 1) systems with units waiting in the queue and 2) systems with the cancellation of units. [5]

Systems with the waiting process consist of a "waiting room" and the service channel. If all channels are busy, the unit lines up and waits for the service until one channel is free. Restrictions on systems with unit queuing can be a particular number of places in the queue (limited length of the queue), or a maximum allowed time spent in the queue and similar.

The system with the cancellation of units is a system in which units leave the service system if all channels are occupied. In practice, those are systems that do not have the devices for the formation of queues (e.g. telephone exchange).

Queuing systems, according to the possible length of the queue (the number of places in a queue), are divided into "pure" and "mixed" queuing systems [3]. The system is pure when the number of places in the queue and the waiting time are not limited. In such systems, the unit does not get cancelled; it will sooner or later be served. The system is mixed when units can wait in the queue, but can also be cancelled because of the limited number of places in the queue or because of the limited waiting time that service users have at their disposal. Units that cannot wait to be served in the queue or service channels are called "impatient" units.

The discipline of the queue is the way in which units (service users) from the queue access the service channel. There are several ways:

1. FIFO (first in - first out) is the way which takes into considerationthe order of arrivals: who comes first will be served first.

2. LIFO (last in - first out) is the way which gives priority to the unit that entered the queue last: last in - first served.

3. PRIOR is a mark of the service priority that favours particular units for the service.

4. SIRO refers to a random selection which gives each unit the same service probability regardless of the time of its arrival to the queue.

5. GD is a mark for any other waiting discipline.

Taking into consideration the listed elements that describe an individual process of queuing, there is a need to mark the category and type of the queuing problems. To this end, **Kendall's notation** has been accepted:

$$v / w / x / y / z \qquad (8)$$

according to which the given letters indicate:
v − the distribution of the arrival rate of units in the system,
w − the distribution of the service time of units,
x − the number of service channels,
y − the capacity of the service system,
z − the discipline of the queue.

For v and w, standardized labels for distribution are used:[1]
D − the deterministic time between the arrivals and the service,
M − the exponential distribution,
E_k − the Erlang distribution of the queue k ($k = 1,2,\dots$),
G − any distribution (including M, E_k and D).

If y and z are not specified, it is implied that $y = \infty$, and $z = FIFO$.

---

[1] Instead of the symbol E that would be expected for the exponential distribution, the symbol M is used to avoid confusion because of a similar symbol E_k that indicates the Erlang distribution of the order k; that symbol comes from Markov's dependence of the exponential distribution.

In practice, a large number of categories and types of queuing problems appears because the six elements that are used to describe the service process can occur in a large number of variations. According to the certainty of the process, i.e. the way in which the units arrive to the system and the duration of the service, **deterministic and stochastic queuing problems are distinguished**. In a deterministic system, the arrival time of units and the service time, i.e. the processing time, are known and determined in advance. Therefore, it is possible to solve such problems with very simple computation processes. In practice, there will be no congestion of the system because organizational measures can deploy the arrival and departure of units. As opposed to deterministic queuing problems, with stochastic systems, the arrivals of units and the service time are random variables. For this reason, stochastic problems are the object of the study of the queuing theory.

Of all the possible mass service systems, the ones that usually occur in practice are the queuing systems. Those are queuing problems in which units wait for a service if their number exceeds the number of channels, or they wait for service places (channels) if the number of units is smaller than the number of channels. From the types of queuing service systems, the following are of special significance in traffic [3, 6]:

1. single-channelled system with an unlimited number of places in a queue ($S = 1$, $m = \infty$),
2. multi-channelled system with an unlimited number of places in a queue ($S > 1$, $m = \infty$),
3. single-channelled system with a limited number of places in a queue ($S = 1$, m is the final number),
4. multi-channelled system with an unlimited number of places in a queue ($S > 1$, m is the final number).

3   Parking lot as a service system

It is said that many phenomena that involve unpredictability and volatility are stochastic because the parameters that characterize them are of random nature. The stochastic process is a process that takes place during the course of an experiment or observation in such a way that it cannot be accurately predicted how that process will unfold. Since the arrivals of cars can be seen as random (stochastic) variables, it is possible to approximate them with the corresponding theoretical distributions. Flows of the arrivals of vehicles and the service time are simple random flows, for which theoretical distributions according to which the empirical data is assigned can be determined or approximated.

The key parking characteristics are occupancy, turnover, duration of stay per user, acceptability of parking charges, as well as potential identification of the modal split across alternative means of transport. [7]. That is why parking systems can be analysed as mass service systems. For that reason, it is possible to, by applying analytical methods, determine the characteristics of the input terminals and examine their permeable power for a different number of parking lot entrances.

Traffic entities (hereinafter referred to as vehicles) access the service place at random points in time. For the majority of cases, those points in time are mutually independent. The concept of the arrival of a vehicle can be identified with the event that is carried out at the entrance to the service system. Therefore, the input flow of events can be discussed.

Homogenous and inhomogeneous flows of events can be distinguished, depending on whether the units belong to "one" type or different types of vehicles. The presence of larger vehicles such as trucks, buses, vans, etc. in a traffic flow on the road reduces the capacity of available resources. For the purpose of an easier application, inhomogeneous flows are converted into homogeneous flows in a way that larger vehicles are converted into passenger car equivalents (the so-called PCE's).

Events in a homogeneous flow differ only according to the moments of appearance, which is why such a flow can be graphically displayed as a set of points t_1, t_2,...,t_n on a number line t, where points t_1, t_2,..., t_n represent the moments of appearance.

Figure 2 Homogenous flow of arrivals to the service system [3]

What is interesting in the queuing theory are the simple random flows that, other than random arrivals (in randomly selected moments), have the following characteristics: **stationariness, flow without consequences, and ordinariness** [1, 8, 9].

Stationariness is a feature that shows accidental fluctuations around the mean value. This feature can be accepted for parking systems, so the intensity of the flow of the arrivals of vehicles does not depend on time, but is constant and represents the average number of vehicles arrived in the unit of time (λ).

A flow without consequences (arrival of vehicles to the parking lot) means that the arrival of one unit (user) does not depend on other units, i.e. that the arrival of units to the system does not depend on the arrival or non-arrival of other, "new" units. At the parking lots, the flows of events are usually irregular (random) because the requirements for a service do not occur according to a predetermined order.

The flow is ordinary if in a particular arbitrarily small time interval Δt only one unit or no unit can arrive. According to the time of appearance, it can be accepted that the flows of events in parking lots are ordinary, which means it is highly unlikely that two vehicles will appear at the same time and with the same request, but it is accepted that vehicles enter the parking lot one after another. The waiting processes with the given properties are called simple Markov chain systems. [6]

4   Regularities of the arrival flow and the service time of vehicles at parking lots

If the number of vehicles and the length of service time are random variables, it is not possible to pre-determine their values, but it is possible to determine their probability distributions. To be able to calculate the probabilities of the realization of a random variable that represents the number of the arrivals of vehicles and the number of served vehicles, it is necessary to pay special attention to the methodology of determining the regularities of the arrival flow and the service time of vehicles at the parking lot.

To determine the regularities of the arrivals of vehicles, first it is necessary to gather the input data, then to conduct a statistical analysis, and finally to examine the compliance of the empirical distribution with the theoretical distribution. The phase of gathering the data is followed by the grouping of the data, which secures the empirical distributions of vehicles, the number of entry ramps, and all other data related to the production of a parking service. For those distributions, basic statistical parameters are calculated: the arithmetic mean, standard deviation, range of variation, measures of asymmetry and roundness. In the case of samples, it is necessary to apply the sampling method, especially the interval evaluations of particular parameters and the testing of the set up hypotheses. The relationship between certain variables is examined by the correlation method, linear or nonlinear.

Based on the parameters calculated for the assigned empirical distribution, the procedure of determining the distribution type is continued by calculating the theoretical frequencies for the selected theoretical distribution, and finally, the compliance of the empirical with the theoretical distribution is examined. Theoretical frequencies are acquired based on the formula:

$$f_{t_i} = N P_n, \qquad (9)$$

where:
N − the total number of units;
P_n − the probability for the current value of the random variable.

Once the arrival of vehicles and the service rate act according to a certain theoretical distribution, it is possible to apply the queuing theory to calculate the indicators of the functioning of an assigned system, depending on the distribution of the arrivals of vehicles and the distribution of the length of the service time.

A Poisson stream of arrivals corresponds to arrivals at random. In a Poisson stream successive customers arrive after intervals which independently are exponential distributed. The Poisson stream is important as it is a convenient mathematical model of many real life queuing systems and is described by a single parameter-the average arrival rate. [2] Based on the above-mentioned facts, it is accepted that the arrivals of vehicles to the parking lot represent an easy flow distributed according to the **Poisson distribution**, and that the service time of each vehicle is a random variable distributed according to the **exponential distribution**.

The calculation of theoretical frequencies depends on the chosen theoretical distribution. It is stated in statistical literature that the tabulation of probability for $n > 20$ in the Poisson distribution is impractical, considering the fact that the Poisson distribution for $n = \infty$ becomes less and less asymmetric and tends to a normal curve. Therefore, in such cases $P(\lambda)$ **can be approximated by a normal distribution**.

The characteristics of the intensity of the flow of the arrivals of vehicles and the length of the service time at the entrance to the parking lot are [8, 9, 10]:
− arrivals of vehicles take place one after another in randomly determined moments in an observed time interval and they represent an input flow,

− it is not possible to accurately predict the arrival time of each vehicle,
− the input flow of the arrival of vehicles to the parking lot is inhomogeneous because it is different with regard to the vehicle type, and is calculated into the number of personal vehicles,
− flows of the arrivals of vehicles to the parking lot are random (irregular) because the demand for a service does not occur on the basis of a predetermined queue,
− it is not possible to accurately predict the length of the parking time,
− according to the service time, it can be accepted that at the entrance to the parking lot the flows of events are ordinary, which means that it is highly unlikely that the two vehicles will appear at the same time, i.e. only one vehicle at the time can approach the ramp,
− the arrival of vehicles that are arriving to the parking lot at a given time is independent of the number of vehicles that have previously arrived to the parking lot, which is why it can be considered that the arrivals of vehicles to the parking lot entrance are flows without consequences.

The consequence of these facts is the unbalanced use of the parking lot capacities, because if the number of vehicles that arrived to the parking lot is greater than the number of vehicles that the existing parking lot capacities can serve in a given time, vehicles then appear in the queue; and in the opposite case, the vehicles are not waiting, but parking lot capacities are not fully utilized. The aim must be a more balanced relationship between the parking capacities and the number of vehicles.

On the basis of defining the input terminal of the parking lot as a mass service system, the indicators of the functioning of the observed input terminal are calculated. For the selected system of input terminals to the parking lot, the **intensity of the flow of arrivals** λ represents the average number of vehicles that arrive to the parking lot in the observed time unit (hour, day, ...).

Likewise, for the chosen system of input terminals to the parking lot, the service intensity μ represents the average number of vehicles that can be served in an observed time unit (hour, day, year). The service intensity is the reciprocal value of the average time needed for a vehicle to enter the parking lot and of the length of the service time ($\mu = 1/t_{ser}$). The average number of vehicles that arrive to the parking lot can be used, i.e. the arithmetic mean that represents the intensity of the flow of arrivals λ (the average number of vehicles per day that arrive with the aim of entering the parking lot).

The service time (t_{ser}) is expressed through a number of time units required for the serving of one vehicle, i.e. for the entrance of vehicles to the parking lot. With the help of that data, the throughput of the service channel is expressed. The average time needed for the vehicle to enter the parking lot is calculated as the arithmetic mean of the total time (the sum of all individual waiting times of vehicles on entrance ramps) needed for the vehicle to enter the parking lot at all entrance ramps. The total service time consists of: the time needed for the driver to stop in front of the input terminal, presses the button or scans the barcode card, the time needed for the entrance

ramp to lift, and the time the driver needs to proceed and enter the parking lot after the ramp has been lifted.

The relation between the intensity of the flow of arrivals and the intensity of the flow of serving the vehicles is the degree of the load of the entry ramp $\rho = \lambda/\mu$. The coefficient of the utilization of more input terminals is calculated as $\rho_S = \lambda/(S\mu)$, where S stands for the number of input terminals. On the basis of the afore-mentioned for input/output parking terminals, it can be concluded that if the input ramps are occupied, the vehicles wait in the queue until served. Considering the fact that the length of the queue is limited by the length of the space on which vehicles can halt while waiting to enter the parking area, the parking lot is defined as a multi-channelled service system with a limited length of the queue. In order to simplify the model, it is assumed that all ramps can serve all types of motor vehicles.

For input/output parking terminals it can be concluded that [10, 11]:

- considering the fact that the arrival flow of vehicles is not an integral part of the system, the system is an open one,
- considering the number of input ramps, the system is a single-channelled or multi-channelled one,
- the arrivals of vehicles to the parking lot are distributed according to certain theoretical distributions (in this paper the Poisson distribution is the presumed one),
- the service time is also distributed according to certain theoretical distributions (in this paper the exponential distribution is the presumed one),
- the serving of vehicles is done according to the FIFO rule (first come, first served),
- considering the fact that the parking lot has a certain space where the vehicles waiting to enter it can wait, that means that the length of the queue is limited and that the vehicles located outside the waiting area for the entrance to the parking lot have to leave the queue.

Since the input terminals (ramps) at the parking lot are a mass service system, **with the application of analytical methods, the characteristics of input terminals can be determined, and their throughput for a different number of entrances can be examined**. The efficiency of the service system is determined with the help of the length of the stay of users in the system and the non-utilization of the service place. The value of these indicators is affected by the number of service places and the intensity of the serving of each of the service places, taking into account the assigned or the expected number of users. Due to the technical-technological limitations of input terminals, it is not possible to increase the intensity of the serving of service places, and only the number of entrance channels (terminals) can be considered as the main factor. For the system to be stable, ρ should not be larger than or equal to 1, which means that λ should be smaller than μ. If that is not the case, then the number of service places should be increased in order to satisfy the condition of stability, i.e. $0 < \rho_S < 1$. This condition is valid for the systems with an unlimited length of the queue, as opposed to the case where the length of the

queue is constrained, and in which there can be no system congestion since the queue is limited to m places and the $(m + 1)$ unit receives a "cancellation" and has to leave the queue. On the basis of defining the input terminal of the parking lot as a mass service system, the appropriate indicators of functioning such as the following are calculated [3, 10]:

1. The degree of the load of the service place/terminal (ρ).
2. The coefficient of the utilization of the parking lot (ρ/S).
3. The probability that there are no vehicles in the service system, i.e. that the capacity of the ramp is unused (P_0).
4. The probability that n vehicles are in the service system, i.e. that n vehicles are being served or are waiting in line to be served (P_n).
5. The probability that the vehicle entering the parking lot will not be served, i.e. that it will get cancelled; the probability of a cancellation (non-serving) (P_{can}). In theory P_{can} is not assigned, but it cannot be 0. In practice, in case of a large P_{can}, the vehicles will be re-oriented. If $\rho_S < 1$, P_{can} is smaller, while for $\rho_S \geq 1$, P_{can} is larger.
6. The number of units that will be served in a unit of time (Q_A).
7. The average number of vehicles in the queue (L_Q).
8. The average number of vehicles in the service system, i.e. the number of vehicles in the queue, plus the vehicles that are being served (L).
9. The average number of vehicles that are being served ($L_{ser} = L - L_Q$).
10. The average waiting time spent in the queue, i.e. the waiting time of the vehicle before being served (W_Q).
11. The average time spent in the service system, i.e. the waiting time of the vehicle spent in the queue and the service time (W).
12. The average service rate ($W_{ser} = W - W_Q$).
13. The average number of free (unoccupied) ramps ($S - \rho$).
14. The probability that all service places (ramps) are occupied, i.e. that the vehicles will wait $P(n \geq S)$.

Given the definition of the parking lot as a service system, the selected ones are: a single-channelled and multi-channelled system with a waiting line and the limited length of the queue. The indicators of the functioning of the selected queuing types are calculated according to the appropriate formulas depending on the type of the problem. For the queuing problem, a mathematical model in the form of a system of linear equations can be set, and with its help the probability P_n depending on λ, μ and the number of channels S is calculated, as is the case with the other indicators of the functioning of the service system for all values of stochastic variables (the number of input units and the number of served units).

5   Conclusion

The arrivals of vehicles and the length of their service time at the entrance to closed parking lots and garage facilities may be accepted as stochastic variables. The

empirical distribution of those variables can be approximated with corresponding theoretical distributions. Consequently, when dimensioning and optimizing the service system of vehicles in closed parking lots and garage facilities, it is possible to apply the analytical approach, i.e. the theory of queues.

This paper presents a theoretical model of the application of the theory of queues on the optimization of the number of service entrance ramps at closed parking lots and garage facilities. Additionally, what is suggested through theoretical research is a method of determining the number of input terminals at the entrances to garage-parking facilities that can be universally applied.

Conclusion obtained from this paper can be summarized as follow:
– arrivals of vehicles take place one after another in randomly determined moments in an observed time interval and they represent an input flow,
– it is not possible to accurately predict the arrival time of each vehicle,
– the input flow of the arrival of vehicles to the parking lot is inhomogeneous because it is different with regard to the vehicle type, and is calculated into the number of personal vehicles,
– flows of the arrivals of vehicles to the parking lot are random (irregular) because the demand for a service does not occur on the basis of a predetermined queue,
– it is not possible to accurately predict the length of the parking time,
– at the entrance to the parking lot the flows of events are ordinary, which means that it is highly unlikely that the two vehicles will appear at the same time, i.e. only one vehicle at the time can approach the ramp,
– the arrival of vehicles that are arriving to the parking lot at a given time is independent of the number of vehicles that have previously arrived to the parking lot, which is why it can be considered that the arrivals of vehicles to the parking lot entrance are flows without consequences,
– considering the fact that the arrival flow of vehicles is not an integral part of the system, the system is an open one,
– considering the number of input ramps, the system is a single-channelled or multi-channelled one,
– after changing queueing model from single line to multi line model it is found that waiting period is reduced tremendously. At the same time, it is also found that number of served cars also increased,
– the arrivals of vehicles to the parking lot are distributed according to certain theoretical distributions (Poisson distribution is the presumed one),
– the service time is also distributed according to certain theoretical distributions (Exponential distribution is the presumed one),
– the serving of vehicles is done according to the FIFO rule (first come, first served),
– the length of the queue is limited and the vehicles located outside the waiting area for the entrance to the parking lot have to leave the queue,
– for different combinations of input values, one can easily see adequate solutions of necessary numbers entrances for vehicles,
– estimation of car waiting time during the considered time period depends on the acceptable average waiting time/average service time ratio. This

parameter is obtained as the function of number of cars, number of entrances for vehicles, waiting and service times.
– analytical results should be obtained with M/M/S/m according to Kendall's notation,

Finally, the applied methodology is rather simple in the estimation of the existing state and planning requirements. Furthermore it provides better management and control as well as decision-making process related to parking management. With that the hypothesis of this research is approved.

6   References

[1] Yang, S.; Yang, X. The Application of the Queuing Theory in the Traffic Flow of Intersection. // International Journal of Mathematical, Computional, Physic, Electrical and Computer Engineering. 8, 6(2014), pp. 986-989.
[2] Jaroslav, M.; Camaj, J.; Nedeliakova, E. Application of the Queuing Theory in the Werehouse Optimization. // International Journal of Social, Behavioral, Educational, Business and Industrial Engineering. 9, 11(2015), pp. 3587-3591.
[3] Zenzerović, Z. Teorija redova čekanja, Stohastički procesi II. dio. Autorizirana predavanja, Pomorskifakultet u Rijeci, Rijeka, 2003.
[4] Žiljak, V. Simulacija računalom. Školska knjiga, Zagreb, 1982.
[5] Vukadinović, S. Elementi teorije masovnog opsluživanja. Naučna knjiga, Beograd, 1988.
[6] Bhat, U. N. An Introduction to Queuing Theory, Modeling and Analysis in Applications. Birkhäuser, London, 2015. DOI: 10.1007/978-0-8176-8421-1
[7] Basarić, V.; Mitrović, J.; Papić, Z. Passenger Car Usage for Commuting to Work as Function of Limited Stay at Car Parks. // Promet – Traffic & Transportation. 25, 4(2013), pp. 323-330. DOI: 10.7307/ptt.v25i4.322
[8] Maršanić, R.; Zenzerović, Z.; Mrnjavac, E. Planning model of optimal parking area capacity. // Promet – Traffic & Transportation. 22, 6(2010), pp. 449-457. DOI: 10.7307/ptt.v22i6.210
[9] Maršanić, R.; Zenzerović, Z.; Mrnjavac, E. Application of the queuing theory in the planning of optimal number of servers (ramps) in closed parking systems. // Ekonomska istraživanja. 24, 2(2011), pp. 26-43. DOI: 10.1080/1331677X.2011.11517453
[10] Krpan, Lj.; Maršanić, R. Optimizacija prometnih tokova na ulazu u parkiralište primjenom teorije redova čekanja. // Suvremeni promet. 26, 3-4(2006), pp. 227-230.
[11] Chen, Q.; Wang, Y. A Model for Dynamic Parking Choice of Drivers. // Promet – Traffic & Transportation. 27, 5(2015), pp. 365-377. DOI: 10.7307/ptt.v27i5.1594

Authors' addresses

izv. prof. dr. sc. Ljudevit Krpan,
Primorje-Gorski Kotar County,
Adamićeva 10, 51000 Rijeka, Croatia
E-mail: ljudevit.krpan@pgz.hr

dr. sc. Robert Maršanić
Rijeka promet d.d.,
Fiumara 13, 51000 Rijeka, Croatia

prof. dr. sc. Marin Milković
University North,
Trg dr. Žarka Dolinara 1, 48000 Koprivnica, Croatia