Project Report

Control Tower

Performance Evaluation
of Computer Systems and Networks

Federico Fregosi, Tommaso Billi, Antonio Nunzio Pio Di Noia

Academic Year 2021/2022

Contents

1 Introduction

1.1 Overview

Consider a system that mimics an airport. The airport consists of one runway, a parking area and a control tower. The runway is used for landing/take-off and can be occupied by only one plane at a time, whereas the parking area has infinite capacity. The control tower manages the air traffic of the airport; each airplane follows the following steps to enter and leave the airport:

1. it queues for landing;

2. it lands on the runway;

3. it moves in the parking lot and stay here for a certain time;

4. it queues for take-off;

5. it takes-off and leave the airport.

Landing and take-off operations, as well as parking, require a specific time for each airplane. In particular, for each plane we have:

- t_l = time necessary to perform the landing operation;

- t_p = time that the airplane will spend in the parking area;

- t_o = time necessary to perform the take-off operation.

When the runway is occupied, planes that are required to land or take off are placed respectively in the landing queue or in the take-off queue. The queues are FIFO, and their capacity is assumed to be unlimited. When the runway is unoccupied instead, the control tower serves one airplane from a queue, prioritizing take-off over landing; planes in the landing queue are served only when the take-off queue is empty. Whenever a plane takes off, it leaves the system.

1.2 Objectives

In this analysis, the following objectives will be investigated:

- relationship between the service times and the waiting time in both landing and take-off queues: in the system we assume that airplanes have infinite fuel, while in a real environment airplanes have instead a finite quantity of fuel. Studies in the waiting time can be used to estimate the quantity of fuel needed in order to stay in the air until clearance for landing is received;

- optimal value for the size of the parking area: we want a good estimation of the space required to accommodate parked planes, without wasting too much of it.

1.3 Performance indexes

In order to fulfill the above-stated objectives, the following performance indexes will be taken into consideration:

- waiting time for the landing queue

- waiting time for the take-off queue

- number of airplanes in the parking area

To evaluate the state of the system and distinguish the transition state from the steady state, we also used as an index the throughput of the runway.

1.4 Scenarios

The following scenarios will be taken into consideration:

- constant interarrival times, constant service times (only for code verification since in this scenario queuing is of little interest);

- exponential distribution of interarrival times, exponential distribution of service times.

2 Model

2.1 Description

In fig. 1 you can see our model of the system, where:

- λ_V is the average arrival rate for airplanes that need for landing;

- the path is unique: the only possible outgoing arc from the Runway is the one with the same colour of the incoming arc;

- the average service rate of the Runway is μ_l for the landing operation (blue lines) and μ_o for the take-off operation;

- both queues are FIFO;

- the Parking Area is a delay center with infinite space, where μ_p average service rate.
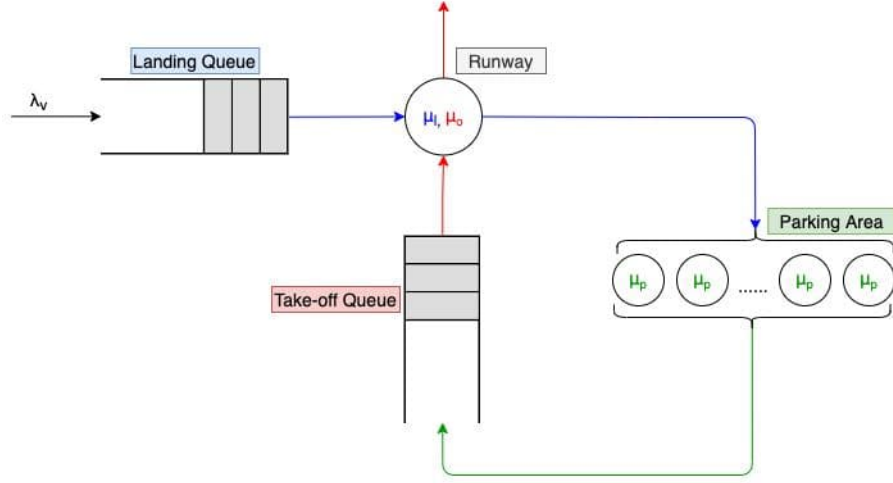
Figure 1: Schematic representation of the model of the system

2.1.1 Assumptions

The following assumptions were made when modeling the system:

- infinite queuing space: there is no upper bound in the number of planes in a queue;

- infinite parking space: the parking area is composed of infinite servers;

- no overtaking: planes cannot change their position in queue;

- all the service times of a planes are independent RVs.

- messages exchanged between the control tower and the planes have no delay;

- no additional runway for taxiing: landing and take-off operations include taxiing procedures.

2.1.2 Stochastic model

Defining an accurate queuing theory model for the system proved to be difficult, as it is made of two distinct queues that share a service center (with different service rates), and whose interarrival rates are non-independent.

For this reason, to study the waiting times in the queues we designed a simplified model, that can be obtained from the real model in figure 1 by applying these two modifications:

- removing the parking area from the picture, which acts as a buffer between the queues;

- consider a mean interarrival rate at the take-off queue equal to λ_v.

With these two considerations, and by taking into account the priority that the take-off queue has over the landing queue, we can predict the performance indexes for the two queues, as long as the utilization of the runway ρ is sufficiently low.

We denote $\rho_i = \dfrac{\lambda_v}{\mu_i} = \dfrac{t_i}{t_v}$ the runway utilization per queue. Since the take-off queue is not serving planes only in case the landing queue is serving one already, its response time can be computed as the following:

$$E[R_T] = \frac{1}{\mu_o - \lambda_v} + \rho_l \cdot \frac{1}{\mu_l}, \tag{1}$$

where the first term is the mean response time in an M/M/1 system, and the second is the probability that the runway is occupied by a landing plane, multiplied by the average landing time. The mean waiting time can be computed as $E[W_T] = E[R_T] - t_o$.

By applying Little's Law, we can compute $E[N_T] = E[R_T] \cdot \lambda_v = \dfrac{E[R_T]}{t_v}$. By observing that the landing queue is serving planes only when the take-off queue is empty, and the runway is free, its mean response time can be estimated as the following:

$$E[R_L] = \frac{1}{\mu_l - \lambda_v} + E[N_T] \cdot E[R_T] \cdot \sum_{n=0}^{+\infty} \rho_t^n = \frac{1}{\mu_l - \lambda_v} + E[N_T] \cdot E[R_T] \cdot \frac{1}{1 - \rho_t}, \tag{2}$$

where the second term is accounting for the fact that an arriving plane will have to wait that all planes in the take-off queue have been served, before it can land. During this time, more planes could queue up for take-off (this expected number can be computed by multiplying $E[N_T]$ by ρ_t), which would still have priority over landing planes. Again, during this wait, even more planes could earn priority over landing ones, and so on. Since we're interested in the behavior of the system in stable conditions $\rho_t < 1$, which means that the infinite sum always converges.

Finally, the waiting time can be trivially computed as $E[W_L] = E[R_L] - t_l$.

2.2   Factors

The factors that can affect the modeled system are:

- interarrival time of airplanes: $t_v = \frac{1}{\lambda_v}$

- landing service time: $t_l = \frac{1}{\mu_l}$

- parking service time: $t_p = \frac{1}{\mu_p}$
- take-off service time: $t_o = \frac{1}{\mu_o}$

3 Implementation

3.1 Code Overview

The model is composed by three nodes:

- **Airspace**: the node in which planes are created and initialized before reaching the control tower asking for a landing operation

- **Control tower**: the main node of the system. It is responsible for the manage of both landing and take-off queues and operations

- **Parking area**: node in which planes spend the t_p time

Note that the runway isn't a node itself, but it is included into the control tower.

Planes are modeled as messages. They contain information regarding the service times for the three operations that it will need to perform, that are generated whenever a new Airplane message is spawned by the Airspace module.

Although the probability of multiple events happening at the same time is almost negligible in stochastic scenarios, such eventuality must be taken into account. To make sure that the priority between the two queues is maintained even in these case, a different priority was assigned to each type of message that is sent among the modules of the system (from highest to lowest):

1) airplane message leaving the parking area and moving to the take-off queue;

2) message from the control tower giving the OK and starting the take-off operation;

3) airplane message spawned from the airspace and moving to the control tower;

4) message from the control tower giving the OK and starting the landing operation;

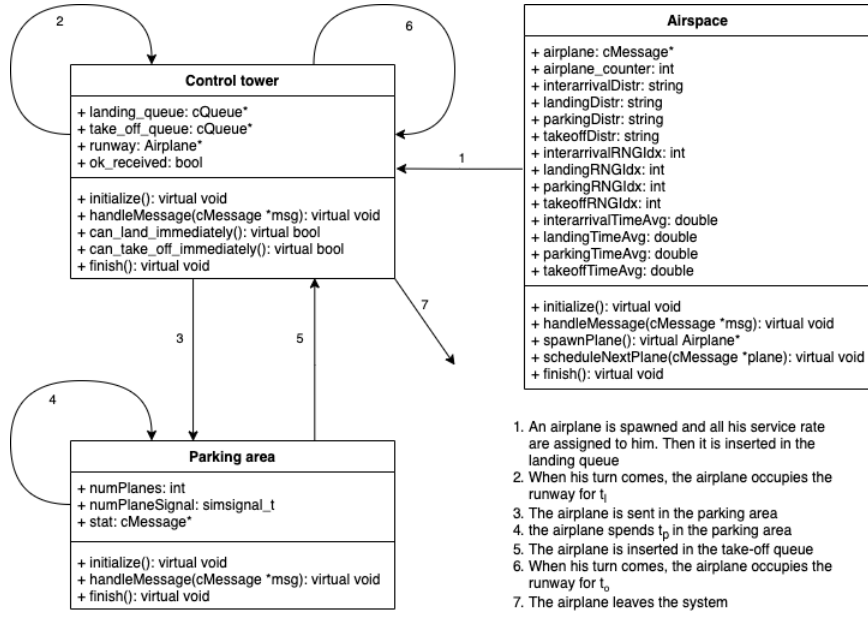5) self-messages for checking the queues after a take-off or landing operation.

Figure 2: Life-cycle of each airplane

3.2 Validation

3.2.1 Continuity Test

This test is about checking that if $input_i$ produces $output_i$ then slightly changing the input does not change the output wildly. As input it was considered the interarrival time of the airplanes, as output their response times. Before actually carrying out the continuity test the value of interarrival time for which the system reaches a steady state was found, by fixing all the service times (t_l, t_p, t_o) to constant reasonable, i.e. "realistic" values:

- t_l: 900 seconds = 15 minutes

- t_p: 3600 seconds = 1 hour

- t_o: 900 seconds = 15 minutes

For the system to be stable, the interarrival time must be $t_v > t_l + t_o$, so, taking some safety margin, the lowest value used to carry out this test was 2700 seconds (45 minutes). The interarrival time range upon which the test was carried out was the following:

$$iat = \{2700 \text{ s}, 2760 \text{ s}, 2820 \text{ s}, 2880 \text{ s}, 2940 \text{ s}, 3000 \text{ s}, 10000 \text{ s}\}$$

6

where the value 10,000 was inserted only to show the monotonic behavior of the output, as it should be: as the interarrival time increases there will be less planes competing at the same time for the single runway, so the response time decreases.
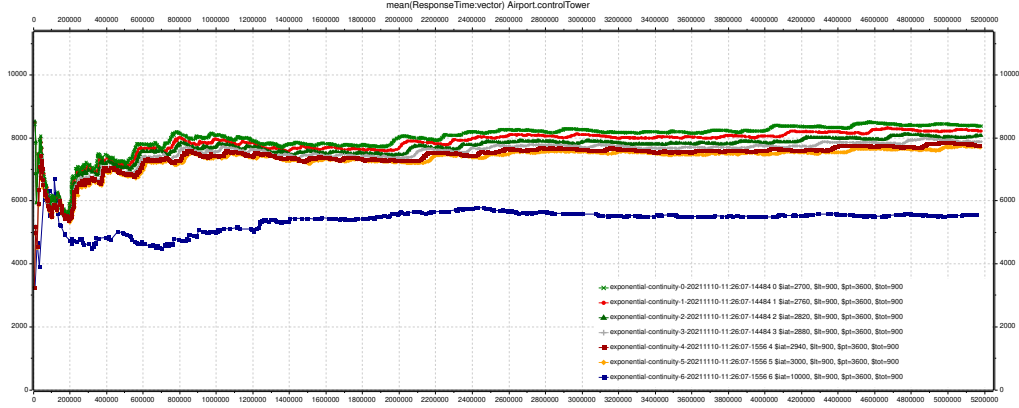
The result is shown in Figure 3.

Figure 3: Continuity test.

3.3 Preliminary calibration

In order to define the set of parameters to be used in the following simulations, we studied the mean response time as a function of the mean interarrival time of planes, for different combinations of service times.

The results that emerged from our study, as we expected, showed that if $t_v < t_l + t_o$, $E[R]$ rapidly increases, as a consequence of the landing queue growing indefinitely. If instead $t_v > t_l + t_o$, the mean response time tends to quickly stabilize around a constant value. Another result that emerged from our tests is that the average time spent in the parking area acts as an additive constant for $E[R]$. This is exactly what we would expect, as no queuing occurs in the parking area, and the mean service time for a delay center is t_p.

Having fixed the t_l and t_o service times at 900s, to find a lower bound for the mean interarrival time we computed 95% confidence intervals for the mean response time over 20 independent repetitions, and found that $t_v \geq 2400$s yields a relative margin of error below 2%, hence we picked $t_v = 2400$s as the mean interarrival time for the worst case in the following studies.

To calibrate the system, the following factor ranges were defined:

- $t_l \in [300\text{s}, 900\text{s}]$

7

- $t_p \in [3600s, 36000s]$

- $t_o \in [300s, 900s]$

- $t_v \in [2400s, 7200s]$

3.3.1 Calibration of warm-up period and simulation duration

In order to define the required warm-up period, we plotted the sliding window average of the response time as a function of the simulation time for 20 different runs. The runs were configured in order to represent a worst-case scenario: t_l=900s, t_p=36000s, t_o=900s, $\frac{1}{\lambda_v}$=2400s, where the value chosen for lambda is explained in section 3.3. We carried out the warm-up analysis both on the response time and on the throughput, anyway the the value produced through the study on the response time was larger and so it was chosen as worst case, and it is the one shown in Figure 4.
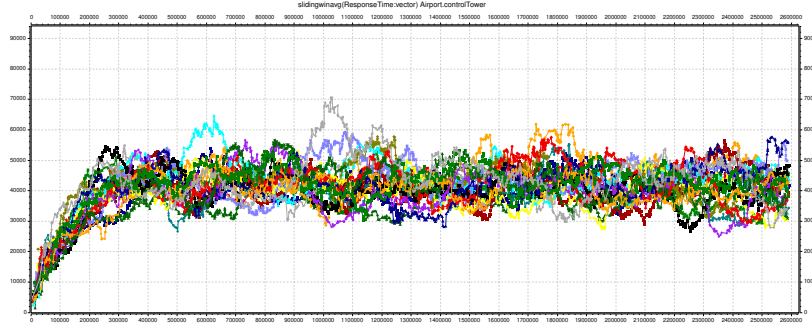

Figure 4: Study on warm-up.

As it can be noted, the plot starts converging at 200.000s, and so, taking a safety margin, a value of 432.000s (= 5 days) was chosen to be used in all the subsequent experiments.

As for the simulation time, a value of 100 days was chosen. This value has to be chosen as a trade-off between the required number of data, and of the time and memory consumption. Since the simulator reported in this work is not too much memory/time demanding to be run, a large value for the simulation time was preferred. Also, a study on the standard deviation has been carried out on the response time and other performance indexes, where of course a smaller and more stable standard deviation for large values of simulation time has been observed in each of them.

3.4   2kr analysis

Before starting with the real experiments, and in order to reduce their numbers, it was decided to perform a $2^k r$ (k=4, r=5) factorial analysis

for each of the metrics on which the experiments will be carried on to understand which are the factors that are more relevant for that metric. The hypothesis on the residuals were visually checked using QQ-plots and scatter plots, and 90% confidence intervals were computed for each of the effects. Moreover, the unexplained variation was computed for each of the performance indexes, and in all the cases it was found to be negligible, always $< 2\%$.

In these analyses we considered the following factors, and their ranges, taken from the exponential scenario:

- A: Interarrival time [2400s, 7200s]

- B: Landing time [300s, 900s]

- C: Parking time [3600s, 36000s]

- D: Take-off time [300s, 900s]

For the **number of airplanes in the parking area**, the most relevant results are the following:

- *Parking time* has a positive impact: 4.46 ± 0.11, accounting for 60.22% of the variation. An increase in the parking time of course would cause that while elapsing the parking time, more planes will enter the system and so the parking area.

- *Interarrival time* has a negative impact: -2.81 ± 0.11, accounting for 23.98% of the variation. An increase in the interarrival time would decrease the number of planes in the parking area since planes enter the system slowly, thus more planes have enough time to leave the parking area before new planes enter the area.

- *The interplay of interarrival time and parking time* has a negative impact: -2.27 ± 0.11, accounting for 15.63% of the variation.

For the **waiting time in the landing queue** instead the most relevant results are the following:

- *Interarrival time* has a positive impact: -627.65 ± 70.04, accounting for 28.13% of the variation. A decrease in the interarrival time would of course bring more planes in the system in less time, so more planes will queue up in the landing queue.

- *Take-off time* has a positive impact: 510.60 ± 70.04, accounting for 18.62% of the variation. This is due to the fact that we have just one runway, and to the policy of the airport of giving priority to planes that need to take-off, so if the take-off time increases, the runway will be occupied for a longer period by taking-off planes.

- *Landing time* has a positive impact: 481.4 ± 70.04, accounting for 16.55% of the variation. Of course, if the landing time of planes increases, planes will experience a longer queuing time in the whole system, and so also in the landing queue.

Finally, for the **waiting time in the take-off queue** the most relevant results are the following:

- *Interarrival time* has a positive impact: -180.99 ± 9.98, accounting for 36.68% of the variation. This is due to the same reason as in the previous metric.

- *Take-off time* has a positive impact: 148.45 ± 9.98, accounting for 24.67% of the variation. Again, more time needed to take-off means that planes will have to wait for a longer period before starting their take-off.

- *Landing time* has a positive impact: 132.69 ± 9.98, accounting for 19.71 % of the variation. This because there's the probability that a plane arrives at the airport finding the take off queue empty, so it starts landing; while landing a parked plane has to take-off but the runway is occupied by the landing plane, so the lower the landing time, the lower the waiting time in the take-off queue.

Given these results, it was possible to restrict the number factors to be tuned during the experiments to the most relative ones, that are reported in this section.

4 Experiments

4.1 Waiting time in the queues

We wanted to study how the waiting times in both queues were affected by the service times of the landing and take-off operations. To do this, we chose t_v =3000s, and in turn we fixed one service time to 600s, while varying the other in the range [300s, 900s]. We then compared the data we collected from the simulations with those predicted using the formulas for the simplified model explained in section 2.1.2.

The results can be seen in figure 5. As it can be noted, waiting times in both queues are more sensible to variations of t_o rather than t_l, confirming the results obtained from our $2^k r$ analysis. As it can be seen in figure 5a, when we fixed t_l, the values for $E[W]$ in both queues computed using the formulas almost perfectly approximate the average wait collected in our simulations. In figure 5b, where we used a constant value of t_o while changing t_l, the same formulas tend to underestimate $E[W]$.
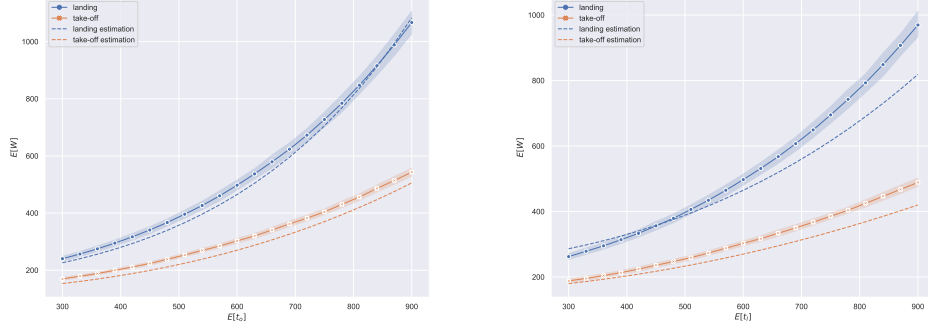
(a) $t_l = 600, t_o = 300..900$ step 30   (b) $t_l = 300..900$ step $30, t_o = 600$.

Figure 5: Waiting time in landing and take-off queues with different t_o (5a) and t_l (5b) values ($t_v = 3000, r = 30$).

Another thing we wanted to study was the relation between the waiting times in the queues and the utilization of the runway. In order to perform such study we observed that $\rho = \rho_l + \rho_t = \lambda_v t_l + \lambda_v t_o$. With $t_l = t_o = 900s$, we varied t_v to obtain a target ρ value between 0.1 and 0.9, with $t_v = \frac{t_l + t_o}{\rho}$.
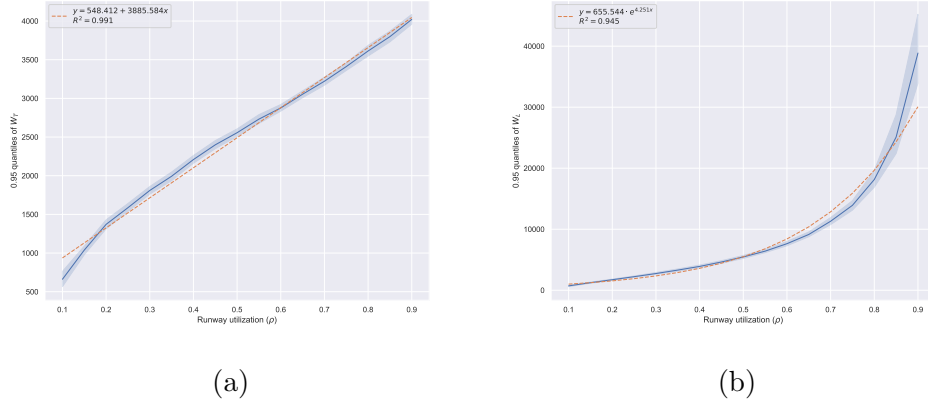
(a)   (b)

Figure 6: Regression of waiting times as a function of ρ ($t_l = 900, t_o = 900, r = 30$).

Figure 6a shows the waiting times in the take-off queue as ρ varies. The 95% percentiles of waiting times that we obtained in our simulations tend to grow almost linearly with the runway utilization, although for lower values of ρ the waiting time rapidly goes to 0, as the probability of an airplane not queuing up increases.

We performed a similar study for the landing queue as well. Figure 6b

11

shows the trend for the 95% percentiles of waiting times in the landing queue. In this case, we tried to compute an exponential regression for the obtained results, by performing a linear regression on the natural logarithm of the waiting times. The results we obtained show that an exponential can approximate this trend quite well, as long as ρ is not too close to 1. In particular, values of W_L grow faster than the fitted function for values of ρ above ~ 0.8. This result is critical, as it gives us an upper bound for the time that planes will need to wait in the air before landing (thus the amount of excess fuel required) in 95% of the cases, at least when the runway utilization does not exceed ~ 0.8.

As a proof of this, we studied the advantage of taking-off planes over landing planes as a function of the runway utilization, computed as $1 - \frac{E[W_T]}{E[W_L]}$. The plot in figure 7 shows that the advantage grows linearly with the runway utilization; this means that as ρ approaches 1, a plane would theoretically spend infinitely more time waiting to land than to take-off. Thus, we can conclude that the fairness between the two queues decreases with ρ.
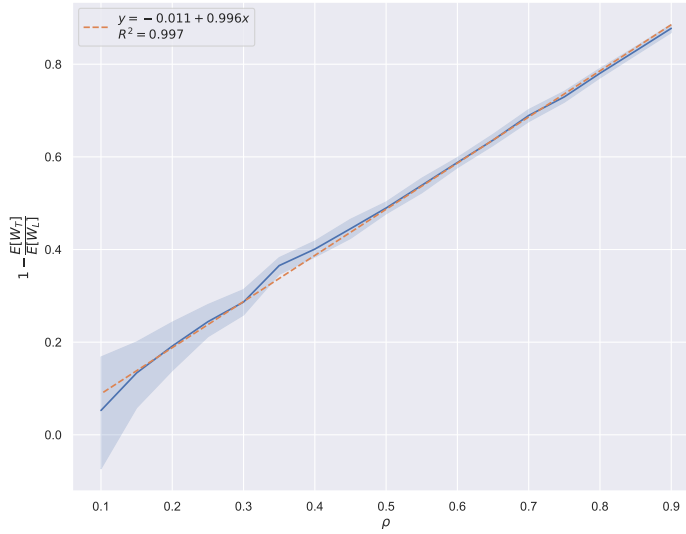

Figure 7

4.2 Parking occupancy

4.2.1 Deterministic study

For a deterministic workload, it's possible to find the exact size of the parking lot, such that we are sure to never have space problems, depending

12

on the parameter of the scenario. In a stable state the rate at which airplanes enter the system is the same at which airplanes enter and exit the parking lot (by definition of stable state). Therefore, the parking alternates an entering and a leaving airplane. As such, if we compute the number of airplanes that enter the parking lot before the first one leaves we will have the maximum size reachable by the parking lot.

The number of airplanes that managed to enter before the first left can be computed as $S = \left\lceil \frac{t_p}{t_v} \right\rceil$. Calculated that, the parking lot occupancy will oscillate between $S - 1$ and S. Thus, S is the maximum number of airplanes that the parking lot will have to contain at the same time.

For our system, the maximum parking lot dimension reached in our worst case scenario ($t_p = 36000$ and $t_v = 2400$) is $S = 15$.

4.2.2 Exponential study

In this case, we cannot reliably measure the maximum occupancy of the parking lot, so we settled at measuring a quantile of it and at how the number of airplanes in the parking area varies in relation with system's factor.

From the $2^k r$ analysis, we saw that both the landing time t_l and the take-off time t_o don't affect the number of planes in the parking lot, as we also saw in the deterministic study above: even if t_l changes, the number of airplanes that manage to get in the parking lot before an airplane comes out of it is determined only by the ratio $\frac{t_p}{t_v}$, while t_l may only delay the flux of airplanes of a random amount, which has almost no influence.

So we studied the variation of the number of planes in the parking lot with various combination of the parking time t_p and the interarrival time t_v. Results are shown in Figure 8.
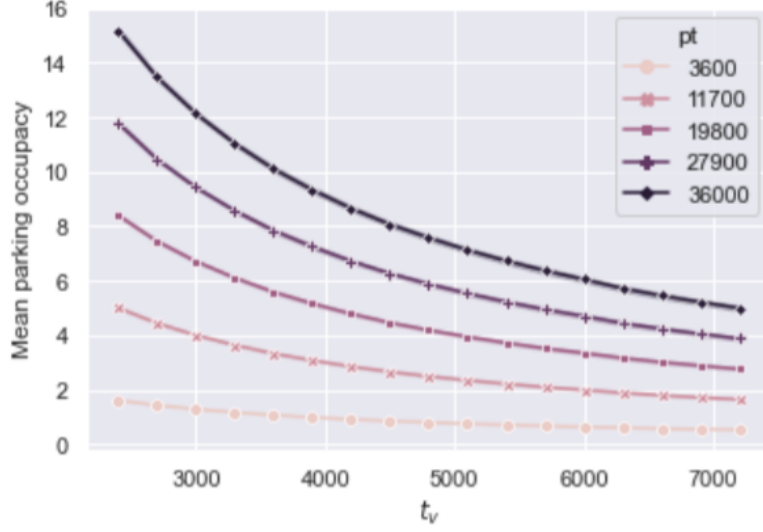
Figure 8: Mean number of airplanes in the parking lot with different $t_v = 2400..7200$ step $300, t_p = 3600..36000$ step 8100 values $(t_l, t_o = 600, r = 30)$.

We tried to fit those results in a well-known distribution and we have found that these points fit, in a very precise way $(\chi^2 \sim 10^{-4})$, a power law function e.g. $f(x) = ax^b$ with a and b values showed in Table 1. This is true also for lower values of t_v, even the unstable ones, since $\rho < 1$. When $\rho \geq 1$ instead, the mean stabilize to a constant $\sim \frac{t_p}{t_o+t_l}$. In Figure 9, the results of the fitting for $t_p = 36000$ is shown.
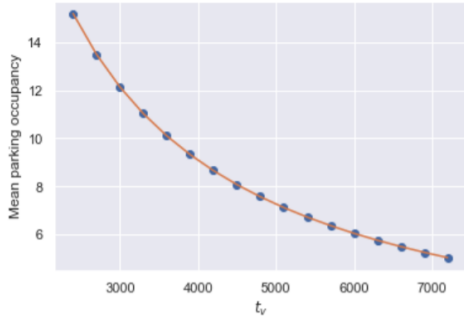

t_p	a	b
3600	3967	-1.0025
11700	12566	-1.0056
19800	21174	-1.0062
27900	29762	-1.0064
36000	38401	-1.0067

Figure 9: Fitting using a power law function of the mean number of airplanes in the parking lot with $t_p = 36000$.

Table 1: Values of a and b derived from the fitting for the power law function $f(x) = ax^b$ for various parking times.

From those results is easy to notice that value $a \sim t_p$ while $b \sim -1$ for every fit. Substituting these values in the power law function, we obtain
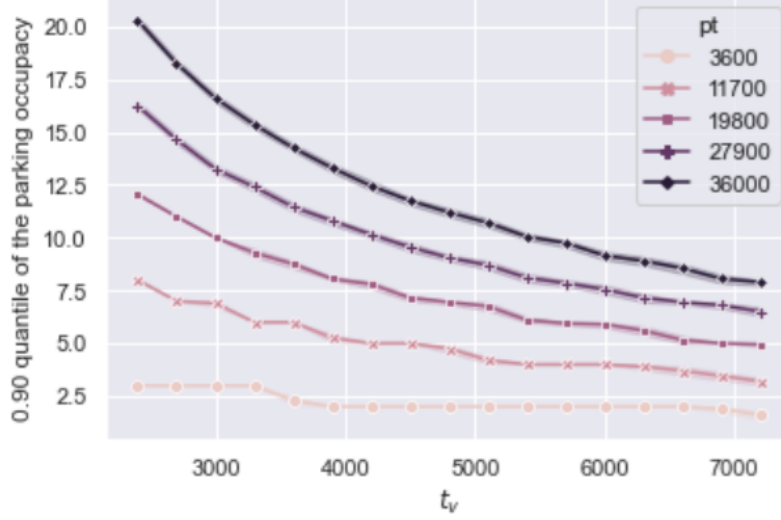
14

Figure 10: 0.90 quantile of the parking occupancy with different $t_v = 2400..7200$ step $300, t_p = 3600..36000$ step 8100 values $(t_l, t_o = 600, r = 30)$.

$numplanes = \frac{t_p}{t_v}$ that is the same function we found in the deterministic study. This is confirmed also from Queueing Theory: we know that the distribution of the size of the parking lot, which is a delay center, is a Poisson distribution with mean $\frac{\lambda_p}{\mu_p}$. So we can now safely say that his mean is equal to $\frac{t_p}{t_v}$.

For what concern the parking dimensioning, we decided to dimension it as the 0.9 quantile of the parking occupancy. This quantile allows us to have enough space for most of the time, without wasting to much space when the utilization is low. The pattern of the 0.9 quantile varying t_v and t_p is the same as the one of the mean but with higher values. So we tried to find a relation between the mean number of planes in the parking area and the 0.9 quantile of the maximum parking occupancy. Results for this analysis are shown in Figure 10 and Figure 11.

Results shows that there is a linear relation between the 2 quantities, fitted by the curve $f(x) = 1.29851883 \cdot x + 0.77097012$. Using the upper bound of the 99.9% confidence interval ($a = 1.37464547$ and $b = 1.22630968$) we can choose an a-priori dimension for the parking lot as $\left\lceil 1.38 \cdot \frac{t_p}{t_v} \right\rceil + 1$ such that the probability of not having enough space to contain the 0.9 quantile of the parking occupancy is close to 0.
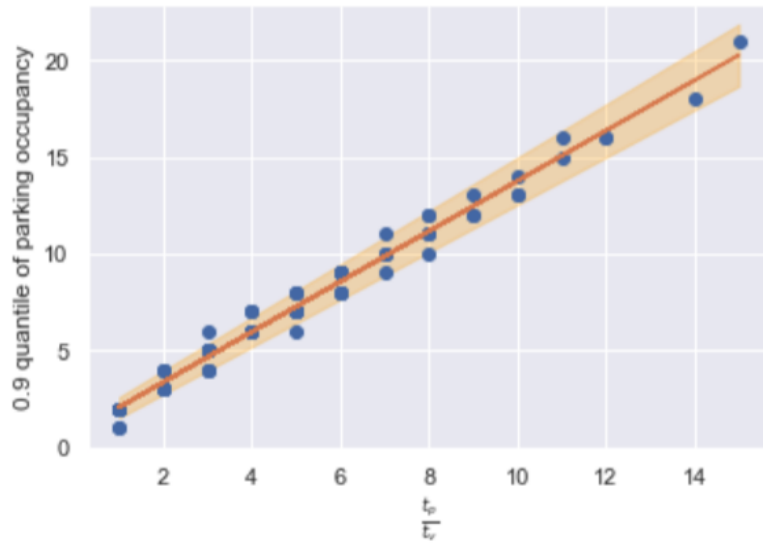
15

Figure 11: Linear regression of the relation between 0.90 quantile of the parking dimensioning and the mean number of airplanes in the parking area, with C.I.=99.9%