Master degree in Computer Engineering

# Performance Evaluation of Computer Systems and Networks

UNIVERSITÀ DI PISA

## Opportunistic Cellular Network

February 2019

Federico Cappellini, Alberto Lunghi, Lorenzo Susini

# 1 Modeling

The aim of this project is to simulate and analyze a simple cellular network. In particular, a study of its throughput and response time has been developed in this project.

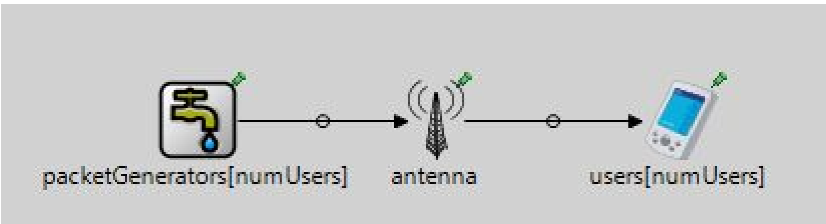
An antenna is transmitting packets to *n* users. Each user has its own FIFO queue on the antenna. On each timeslot, users report to the antenna a Channel Quality Indicator (CQI), a number that determines how many bytes the antenna can pack into a Resource Block (RB). RBs cannot be shared among user but a single RB can pack bytes of one or more different packets of the same user. The number of RBs is limited to 25. The antenna serves its users by descending CQI (*opportunistic policy*). When a user is considered for a service, its queue is emptied if the number of unallocated RBs is large enough.

The overall system can be modeled using the following modules:

- **N Packet Generators**, which will generate packets. Each Packet Generator models the traffic of one user. The destination user is imposed by the Packet Generator's ID. The size in bytes of the packets is randomly chosen from a uniform distribution with 75 bytes as maximum value. In this way, the largest packet dimension is such that it fits a frame at the minimum CQI. Packets are generated every t seconds, where t is an exponential RV with a chosen and configurable mean.

- **1 Antenna:** which will receive the packets from the N Packet Generators, enqueuing them in the right users' queues, and will forward packets to the N Users in the network according to the *opportunistic policy*. CQIs values are randomly chosen from two different distribution, depending on the scenario to be analyzed. The CQIs distribution can be Uniform of Binomial. CQI is a number in the range [ 1, 15 ], so both distribution must return a value in that range.

- **N Users:** which will receive the packet forwarded by the antenna. Users are exploited to compute the response time in the simulation.




A quick look at the network model thanks to the Omnet++ GUI

# 2 Simulation

## 2.1 Introduction and Tuning Factors

The simulation was repeated in two distinct scenarios:

- **Uniform Scenario**: exponential inter-arrivals of the packets, uniform service demands, uniform CQI
- **Binomial Scenario**: exponential inter-arrivals of the packets, uniform service demands, binomial CQI (chosen so that the mean CQI of different users will be sensibly different)

Other parameter had been chosen according to realistic use cases:

- the length of the timeslot set to 1ms (similarly to LTE)
- the packets arrival rate to 0.005s, 0.010s, 0.020s (similarly to common VoIP calls), 0.030s
- for the number of users, the network was simulated using very different numbers to better understand its characteristics: 10, 50, 100, 200, 300, 500, 1000, 1500.

Note that:

- the packets has a minimum dimension of 1 byte and a maximum dimension so that the packet fits an entire frame with the minimum CQI ( $25RB \cdot 3B/RB = 75 \ bytes$ )
- the number of repetitions for each simulation had been set to 35
- the maximum capacity for this network is, in both scenarios, equal to *2325 bytes / time-slot*. The value of time slot is *ts = 1ms*, thus ***C = 2.325 MB/s*** or ***18.6 Mb/s*** .

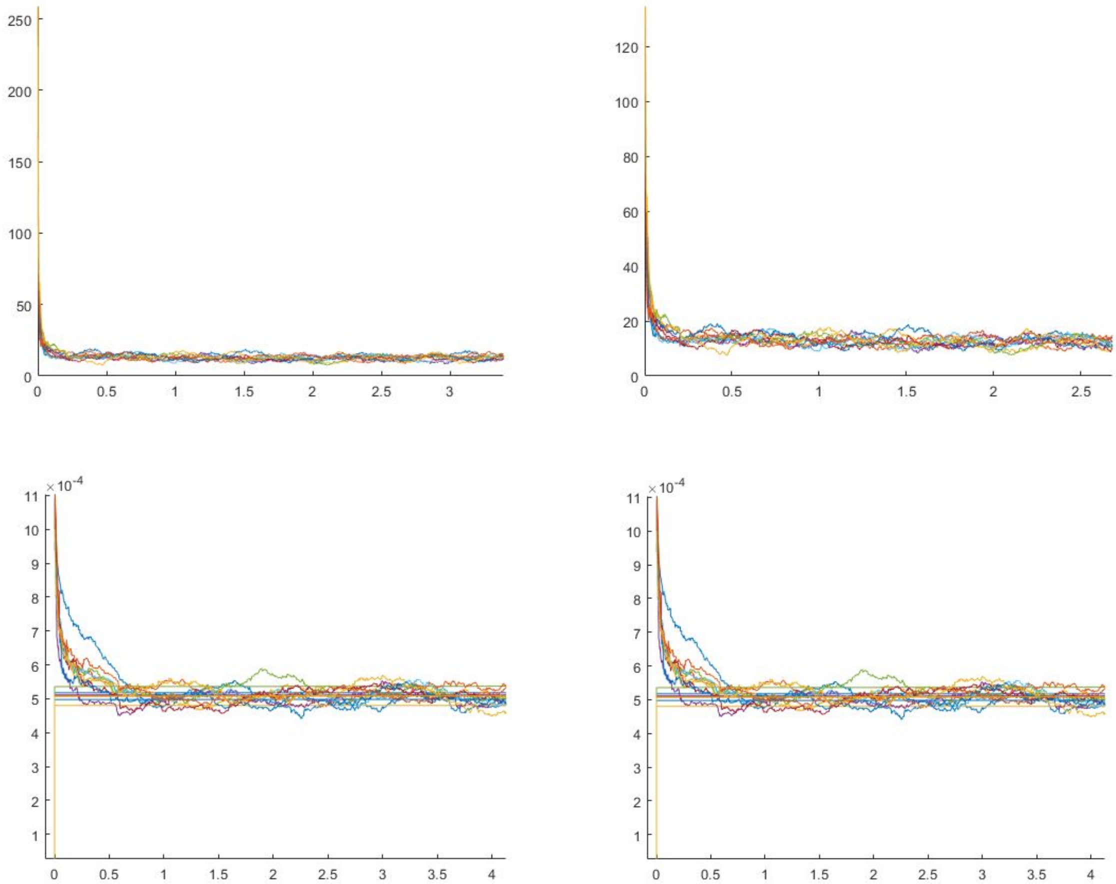
## 2.2 Warm-up Period Estimation

In order to collect meaningful results, a warm-up period must be considered for the simulation. In this way, the simulation starts recording data from a certain point in time where we are confident that the system has already reached a steady state.

Taking into account the previous considerations on the use cases of this transmission system, for both scenario the graphs (throughput) were plotted for the worst case that was characterized by the following parameters:

- packet arrival rate ( $\lambda$ ): 0.030 packets per time slot
- time slot: 0.001 seconds
- number of users: 10

For each repetition a sliding moving average (with a window of 200 elements) was applied to the data so that it was possible to identify the time it took to stabilize around a certain value. Finally, the warm-up period was chosen as **the worst value** of that time (the greatest one).



Top line: throughput, bottom line: response times (uniform scenario on the left, binomial on the right)

Remember that:

- x-axis: time
- y-axis: mean throughput for the throughput graph and mean response time for the response time graph

The warm-up period is:

- 2 seconds for the uniform scenario
- 2 seconds for the binomial scenario

## 2.3 Confidence Intervals

Assuming the IIDness of the average values of the throughput across the various simulations and considering a number of repetitions equal to 35, a general throughput mean value and its corresponding 95% confidence interval were calculated. The repetition number was large enough ( $n \geq 30$ ) to make reasonable to use the following formula to compute the confidence interval:

$$\left[\overline{X} - \frac{S}{\sqrt{n}} * z_{\frac{\alpha}{2}} \; , \; \overline{X} + \frac{S}{\sqrt{n}} * z_{\frac{\alpha}{2}}\right]$$

Where $z_{\alpha/2}$ is the $(1 - \alpha/2)$ percentile of the Standard Normal and S is the sample standard deviation.

As for response times, it is not possible to extract a mean value and its corresponding confidence interval in every test case. Indeed, in some test cases, especially the binomial ones, the network exhibits a saturation state and thus it cannot reach a steady state; both the mean value and the variance tend to increase with the simulation time (or the number of measurements). In the extreme case both the mean value and the variance are infinite because some users never obtain their packets.

| Measurements count | Variance |
|---|---|
| 30000 | 0.0333 |
| 60000 | 0.0614 |
| 90000 | 0.0993 |
| 120000 | 0.1453 |
| ... | ... |
| eventually | ∞ |

For this reason, the response times are not analyzed in case of saturation. In all other cases, both a mean value and a confidence interval are calculated in the same way of the throughput.

In the graphs in this document all the confidence intervals for each measurement are so small that it was almost impossible to draw them, so we decided to cut them for graphical reason and also because they are not significant in order to make a comparison

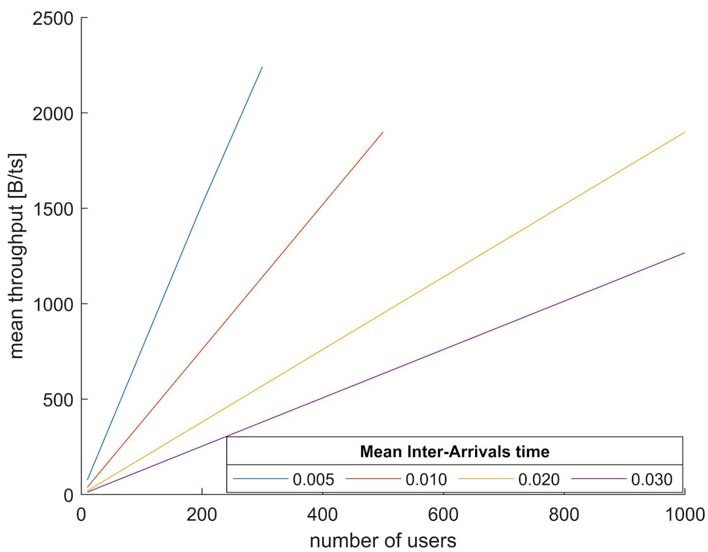
# 3. Scenario: Uniform CQIs

For this scenario we choose these values for the simulation parameters:

- Exponential interarrivals times for packets
- Uniform service demands (the largest packet dimension is such that it fits a frame at the minimum CQI)
- Uniform values for CQIs

## 3.1 Throughput

We have analyzed the mean throughput using different values of *mean inter-arrival time* and number of users. For each value, a 95% Confidence Interval was computed for the mean throughput.  The overall result is shown here:

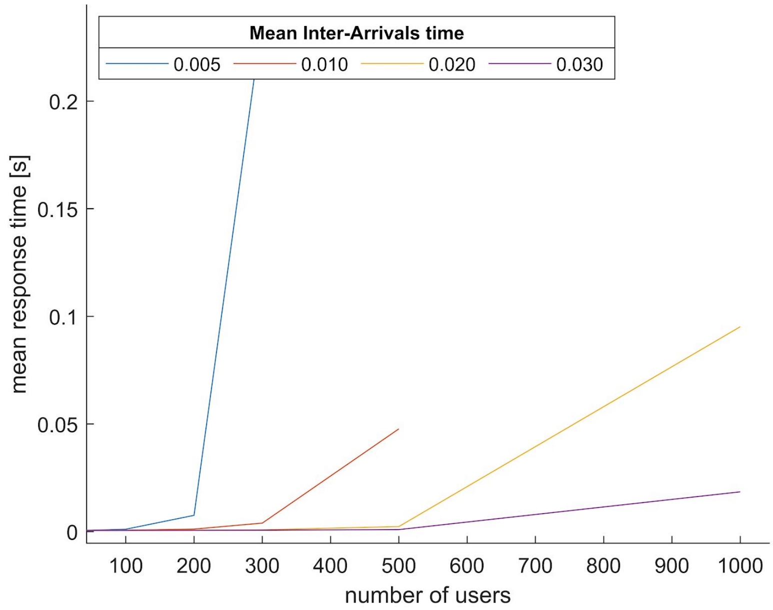



Mean Throughput for the Uniform scenario

As shown by the above plot, the mean throughput rises when the number of users in the network increases. Obviously, the throughput depends on the value of the arrival rate. Indeed, for a given number of users, the more the arrival rate increases, the more the throughput increases.

## 3.2 Response Time

We have analyzed the mean throughput using different values of *mean inter-arrival time* and number of users. The overall result is shown here:

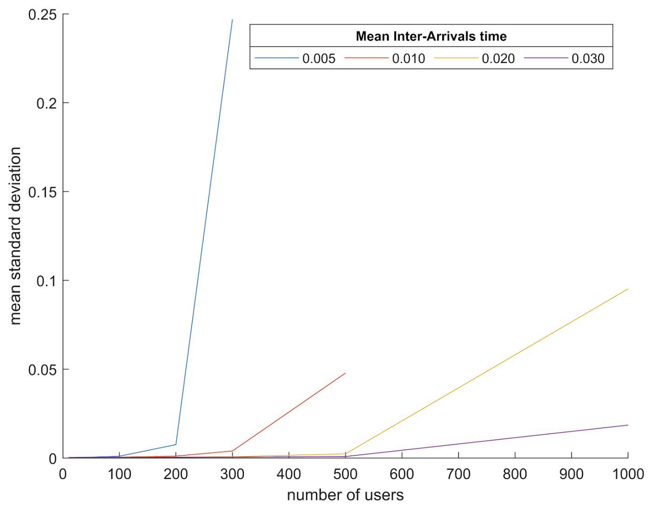



Mean Response Time for the Uniform scenario

As shown by the graph, for small number of user all the configurations are acceptable in term of response time. In fact for a number of users equal to 200 and mean interarrival time equal to 0.005 even the blue quick curve shows a result of about 0.25 s.

For a great number of users, some configurations are unacceptable in term of delay because the network exhibits a very high response time on average because it is in a saturation state. So we decided to cut the results for those configurations.

In order to have an empirical information about the network jitter, the mean response time variances (and their corresponding 95% confidence interval) were computed for each configuration of number of users and mean inter-arrival time.




Mean response time standard deviation as function of users number and mean inter-arrival rate

The variance is practically not relevant for a number of users under 200 ( $\lambda = 0.005s$ ). Over 200, it starts to be more relevant: with 300 users the mean response time can not be considered fully stable for every modern network application. With a lower inter-arrival rate ( $\lambda = 0.020s$ ) it starts becoming relevant around 1000 users and with an even more relaxed inter-arrival rate, the network has practically no jitter.

# 4. Scenario: Binomial CQIs

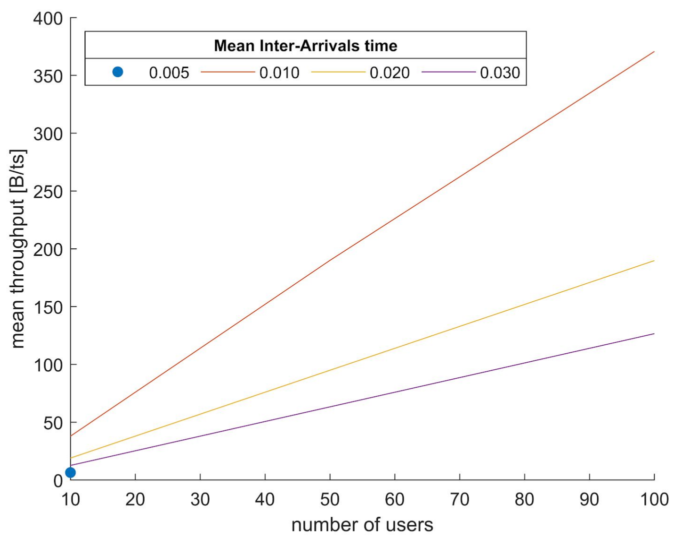
For this scenario, these values were chosen as simulation parameters:

- Exponential interarrivals times per for packets
- Uniform service demands (the largest packet dimension is such that it fits a frame at the minimum CQI)
- Binomial values for CQIs

## 4.1 Throughput

We have analyzed the mean throughput using different values of *mean inter-arrival time* and number of users.

The overall result is shown here:



Mean throughput for the Binomial Scenario

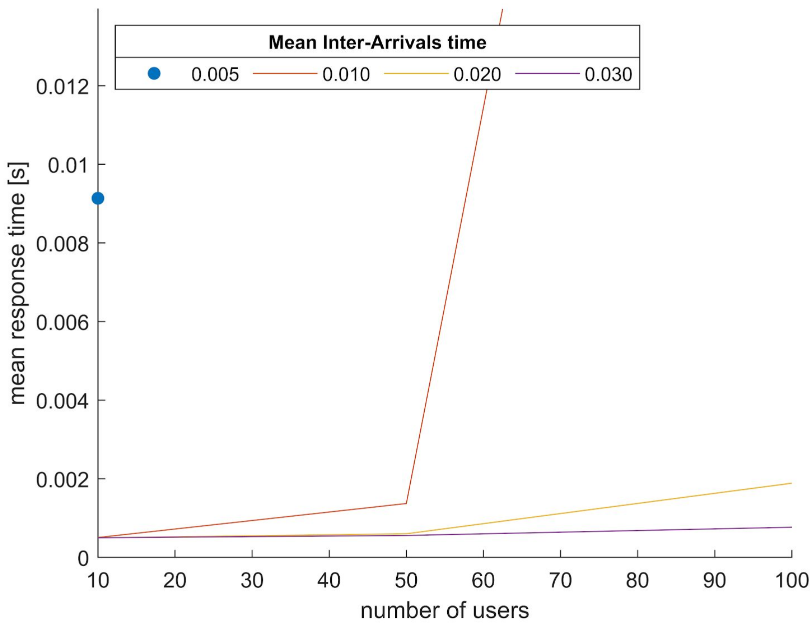
As shown by the above plot, even in this scenario, the mean throughput rises when the number of users in the network increases.

Obviously the throughput depends on the value of the arrival rate. Indeed for a specified number of users, the more the arrival rate increases, the more the throughput increases.

The curve with the interarrival time of 0.005 is cutted from the graph because even for low number of users the antenna experiments a situation of saturation.

## 4.2 Response Time and Fairness

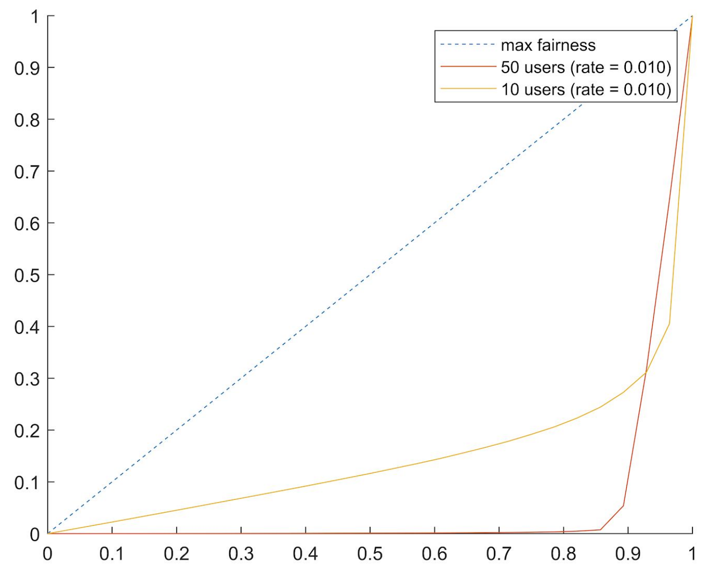
We have analyzed the mean response time using different values of *mean inter-arrival time* and number of users. The overall result is shown here:




Mean Response Time for the Binomial scenario

A quick look at the plot bring us to the conclusion that each mean response time is acceptable only for a limited number of users. In fact, as the number of users increases, the mean response time starts to be very big if compared with the timeslot, causing instability in the system.

In order to retrieve information about the fairness for the users in the network, a Lorentz Curve plot was realized based on per-user mean response time.




Lorenz-curve in order to measure the response time fairness between users in Binomial Scenario

As the plot shows, the level of user-fairness is very low, in particular, as the number of users rises the fairness level decreases because: the more users the longer the queues of the users with a low CQI and so the higher their mean response time.

So, some users are experiencing very short response times, others are experiencing much longer response times in according to the opportunistic policy and the particular choice of CQIs.

In order to have an empirical information about the network jitter, the mean response time variances (and their corresponding 95% confidence interval) were computed for each configuration of number of users and mean inter-arrival time.

Unlike the uniform scenario, in this case the jitter starts to be relevant with a much smaller number of users. With a “fast” mean inter-arrival rate of 0.010s, already with 50 users the network experiences a high jitter. With more than 100 users, the network begins experience a saturation state even before an higher jitter.

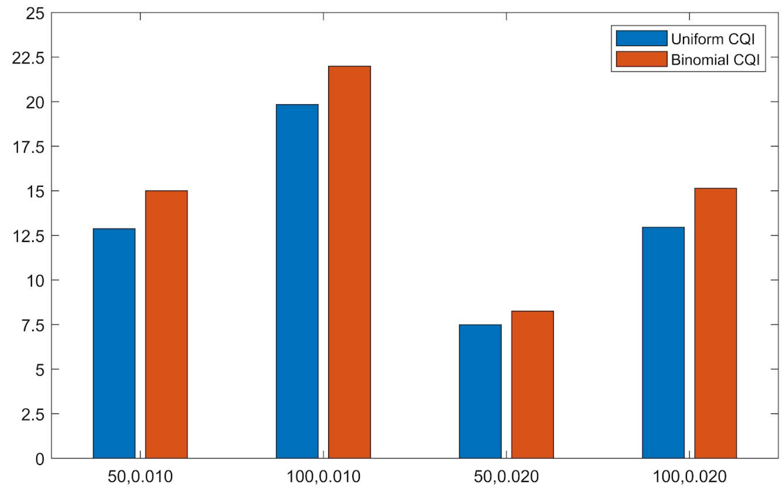
# 5. Resource Blocks Utilization

Another important statistical data about the system is the resource blocks utilization.

This number is computed as the mean number of RBs used at each timeslot and expresses how much the system is used: a low value indicates that the system is working, but it could globally perform better (in terms of throughput) with an higher workload, whereas a high value indicates that the system is near to the saturation state.

This parameter can be also associated to the results obtained for response times. In fact, an high value of mean number of used RBs means that the throughput is high too, but also the response time starts increasing because the system becomes fully used. For instance, the reader can notice that in the case of 100 users and a mean inter-arrival time of 0.010s, the system saturates and the mean number of used RBs is close to the maximum.

It is also interesting to notice that the mean number of used RBs increases as the number of users in the system increases. When the same number of users and the same rate for packets is considered, the binomial scenario is always using a number of RBs that is greater than the uniform scenario. This means that, generally, the binomial scenario experiences worse configurations of CQIs. This consideration is in accordance with the fact that the binomial scenario saturates with a much lower number of users and the uniform one can achieve much higher throughputs.




Resource block utilization for the uniform and the binomial scenario

# 6. Another Scenario: VoIP Calls

An interesting scenario for this cellular network is to model a traffic of packets made up of bursts, like VoIP calls. VoIP calls normally consists of a stream of packets with a constant rate, typically 0.020s, but the stream is not homogeneous: normally a VoIP user does not speak the all the call long. The application do not forward any packet if there is a silence. Assuming calls made of questions and answers, it is possible to say that a user speak for a given arc of time (creating a data stream) and he remains quiet for about the same amount of time (interrupting the data stream), and then speak again just after (resuming the data stream).

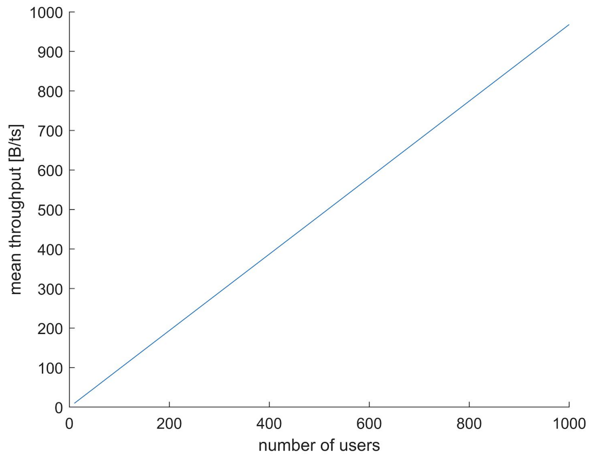
A good model for this scenario can be derived from the previously described one with just two modifications:

- inter-arrival times do not follow an exponential distribution but the arrivals are separated by a constant amount of time (0.020s),
- there is an "on-off modulation" on the packet generators: every node switches its state from on (speaking) to off (not speaking) and viceversa following an exponential distribution with mean *0.5s.* This means that a timer is set on the packet generators that is able to switch the state on/off, modeling a typical voice call.

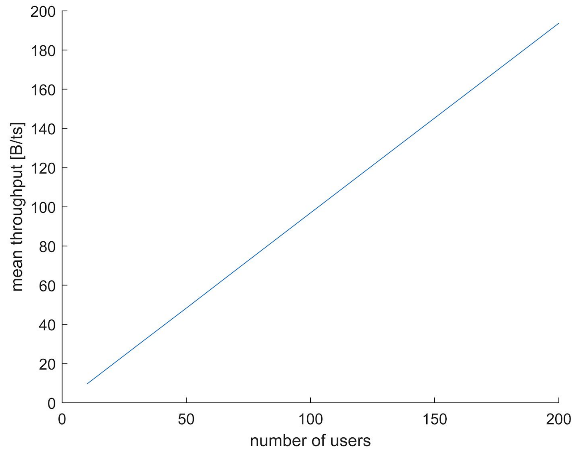
On average $E[t_{ON}] = E[t_{OFF}]$ for each user. This means that $P\{stream\ is\ on\}$ can be considered a bernoullian PMF with a mean of $1/2$ and thus, after a long time, $P\{number\ of\ active\ streams\}$ can be considered a binomial PMF (sum of independent bernoullian RVs) with a mean value of $np = n/2$.

For this reasons this scenario is more incline to achieve a higher number of users to serve.

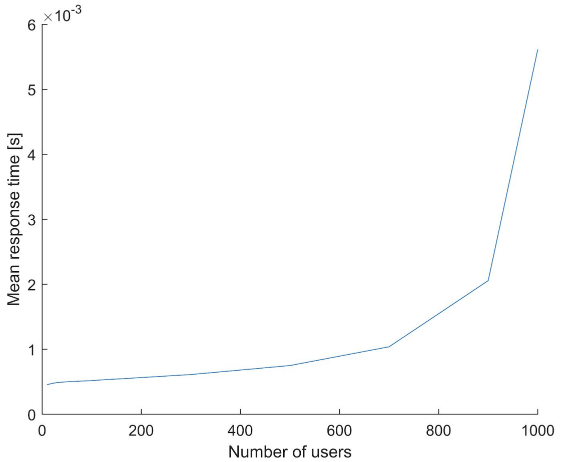
This model was simulated and studied in the same way of the other two: considering the CQIs uniform of binomial and with 35 repetition (to be able to compute the 95% confidence interval in the same way).



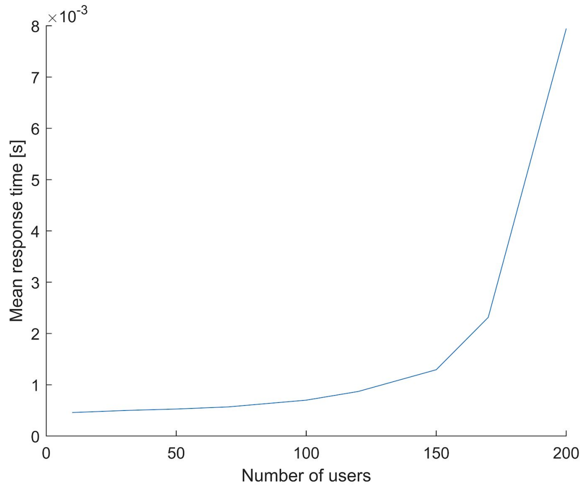

Uniform mean throughput for the VoIP scenario




Binomial mean throughput for the VoIP scenario




Uniform mean response time for the VoIP scenario




Binomial mean response time for the VoIP scenario

Note that also in this case, the study was done only out of network saturation conditions.

In order to provide an analytic expression for the throughput and the response time in the VoIP scenario, a 4 model fittings were performed. The throughput is in a linear relation with the number of users while the response time has an exponential relation, in particular these four equation were fond:

- Binomial throughput: $y = 0.9691x - 0.0885$
- Uniform throughput: $y = 0.9678x - 0.0090$
- Binomial response time: $y = (8.83 * e^{0.0329x} + 928) * 10^{-6}$
- Uniform response time: $y = (3.33 * e^{0.0072x} + 294) * 10^{-6}$

In all cases the regression is very close to 1 and both the error plot (the plot that sees the error as function to its order number) and the homoskedasticity plot confirm the correctness of the model.

# 7. Conclusion

As the plots show, in the uniform scenario the network performs better than in the binomial one in terms of response times and throughput. However, the two scenarios give us information in two precise use cases. In reality, an antenna for cellular communications can experience many scenarios even in a short period. So, considering the summarized results from both scenarios, it is possible to say that this kind of network can be used for applications that do not require high throughput and particularly stable response times. However, applications with an inter-arrival time such as VoIP (or lower), if not in combination with a large number of users, can be used successfully. For instance, this cellular network could be used for IoT: a centralized antenna for sending commands to a network of fixed sensors, if the channels can be approximated with the binomial CQIs scenario.