Communication Network Design - AY 2021/2022
Homework *1: Traffic Sources*

2052409

1 Development

1.1 Introduction

In this document I'm going to describe and model the "received traffic" parameters (**Interarrival Times** $\{A_i\}$, **Batch Size** $\{B_i\}$ and **Workload** $\{W_i\}$) of a *Zoom* call of approximately 40 seconds performed between a laptop (destination) and a smartphone (source) connected to the same WiFi network.

The traffic data have been collected using wireshark.

The results have been obtained with Python and in particular with a Jupyter Notebook which is available in the folder of the homework. If you don't have it installed in your computer you can run it in your browser at https://jupyter.org/try-jupyter/lab/.

In order to estimate the parameters of interest, I extracted from the csv file generated by wireshark only 2 quantities: "Arrival Time" and "Length" of the **received** packages.

1.2 Interarrival Times

Given the vector of the Arrival Times $T = [t_0, t_1, t_2, \ldots, t_{n-1}, t_n]$, I derived the Interarrival Times vector as $A = [t_1 - t_0, t_2 - t_1, \ldots, t_n - t_{n-1}]$.

Its histogram distribution is showed on the left of **Figure 1**. The histogram clearly resembles an exponential, so I extracted the mean of the vector μ and obtained the exponential parameter as $\lambda = 1/\mu$. The model (which CFD $F(x)$ is shown in blue in the right figure) is therefore:

$$p_x(x) = \lambda e^{-\lambda x}, \quad \lambda = \frac{1}{\text{Mean}} \approx \frac{1}{0.0086\ldots} \approx 116.272$$

$$F(x) = \mathbb{P}(X \leq x) = 1 - e^{-\lambda x}$$

1.3 Batch Size

With the sensitivity of wireshark ($1 \cdot 10^{-6}$ [s]) the Batch Size is constantly equal to 1 ($\{B_i\} = 1\ \forall i$).

This can be explained noticing that, differently to the example seen in class, this is a 1-to-1 communication occurring in the same Wifi network, while in class there were more than 20 people connected leading to a much greater overall traffic.

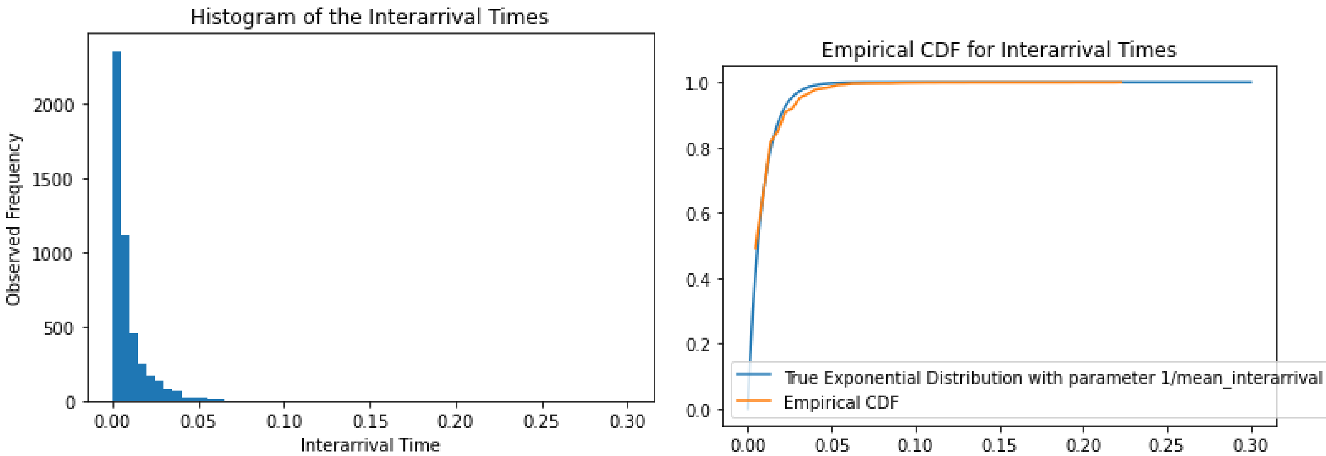
1.4 Workload

Since $\{B_i\} = 1\ \forall i$, then $\{W_i\} = \text{Length[i-th pck]}\ \forall i$.

The histogram of the packet sizes (see **Figure 2** for the **normalized** distributions) resemble 3 separate Gaussian distributions corresponding to 3 categories of packets:

- Small Size Packets.

Figure 1: Histogram Distribution of Interarrival Times (Left). Comparison between CFD and the proposed exponential model (Right). The model fits nicely except when the Cumulative Distribution approaches 1, where a small deviation occurs.



- Middle Size Packets.
- Large Size Packets.

I therefore isolated each one of these categories and extracted the following means and standard deviations:

$$\mu_1 \approx 140.11 \qquad \sigma_1 \approx 35.84$$
$$\mu_2 \approx 317.12 \qquad \sigma_2 \approx 26.06$$
$$\mu_3 \approx 1199.26 \qquad \sigma_3 \approx 68.93$$

I finally plotted 3 separate Gaussian models with these parameters which are shown in the same **Figure 2**.

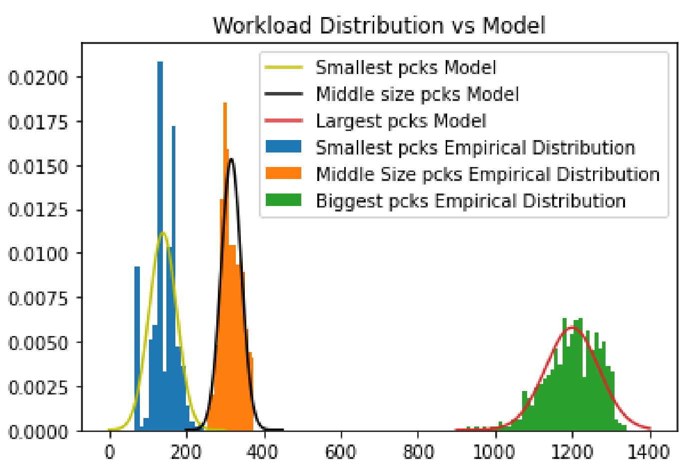


Figure 2: Empirical Normalized Histograms vs My 3 Proposed Gaussian Models

Regarding the last Gaussian (associated with the largest and hence most important packets) we can see that the model works nicely except for the right tail of the curve, since the empirical distribution is affected by some Negative Skewness, however the model seems still reasonable.