

# Synthetic Fine-Grained Traffic Generation

Engineering Adaptation of the KTH Framework for Sparse Bouygues Datasets

Ahcene LOUBAR & Maryam BACHTTI

Laboratoire des Signaux et Systèmes (L2S) – CentraleSupélec

February 19, 2026

## Abstract

This technical report documents the implementation of a traffic augmentation pipeline designed to synthesize instantaneous (1-second) traffic loads from coarse (1-week) aggregate measurements. The methodology is based on the Interrupted Poisson Process (IPP) framework described in KTH [Section II-C].

**The Engineering Challenge:** The KTH reference model relies on second-moment statistics (variance) to calibrate traffic burstiness. However, the available Bouygues dataset provides only single weekly averages, rendering standard variance estimation impossible.

**The Solution:** We developed a *Hybrid Calibration Strategy*. We rigorously derive an analytical solution to preserve the first moment (mean) exactly, ensuring the synthetic data respects the weekly trends. For the second moment (burstiness), we replaced the KTH analytical solver with a grid-search optimization loop, tuning the IPP parameters  $(\tau, \zeta)$  to maximize realistic peak-to-mean ratios. The final output is a high-fidelity, event-driven traffic simulation suitable for LSTM and Chronos forecasting models.

## 1 PROBLEM FORMULATION

### 1.1 Available Dataset and Constraints

The Bouygues dataset provides a single traffic measurement per antenna sector every 5 minutes. Let

$$y_t \quad [\text{Mbps}]$$

denote the observed traffic rate at coarse slot  $t$ , where each slot has duration

$$T = 300 \text{ seconds.}$$

Unlike the KTH framework (Section II-C), the dataset does not provide repeated daily observations for each time index. Therefore:

- First-order statistics (mean) are observable.
- Second-order statistics (variance across days) are not observable.

The objective is to reconstruct a fine-grained traffic process

$$x_{t,k}, \quad k = 1, \dots, \frac{T}{dt}$$

with resolution  $dt = 5$  seconds (or 1 second), such that:

1. The slot-average equals the observed coarse value:

$$\frac{1}{T} \int_0^T x_t(s) ds = y_t.$$

2. The signal exhibits realistic bursty ON/OFF behavior.

## 2 INTERRUPTED POISSON PROCESS (IPP) MODEL

Following KTH Section II-C, traffic arrivals are modeled using an Interrupted Poisson Process (IPP).

### 2.1 ON/OFF Continuous-Time Markov Chain

The system alternates between two states:

$$\begin{aligned} \text{OFF} &\xrightarrow{\tau} \text{ON}, \\ \text{ON} &\xrightarrow{\zeta} \text{OFF}. \end{aligned}$$

The stationary probability of being in the ON state is:

$$P_{\text{ON}} = \frac{\tau}{\tau + \zeta}.$$

### 2.2 Arrival Process During ON

During ON periods:

- Packet arrivals follow a Poisson process with rate  $\lambda$ ,
- Packet sizes  $\psi_j$  are exponentially distributed.

Let  $U$  denote the number of arrivals during one slot of duration  $T$ .

From KTH Eq.(5):

$$E[U] = \lambda P_{\text{ON}} T.$$

The aggregate traffic volume in one slot is:

$$\Psi = \sum_{j=1}^U \psi_j.$$

From KTH Eq.(9):

$$E[\Psi] = E[U]E[\psi].$$

## 3 ADAPTATION TO SPARSE BOUYGUES DATA

### 3.1 Limitation of the Original KTH Moment Matching

In the original KTH model, parameters  $\lambda$  and  $E[\psi]$  are derived by matching both:

$$\begin{aligned} E[\Psi] &= E(O_i), \\ Var(\Psi) &= Var(O_i), \end{aligned}$$

leading to the closed-form expressions (KTH Eq.(11)-(12)).

However, in the Bouygues dataset:

- Only one observation per time index exists,
- $Var(O_i)$  cannot be reliably estimated.

Therefore, solving Eq.(11)-(12) is impossible.

### 3.2 First-Moment Exact Calibration

We enforce exact preservation of the coarse traffic rate.

Since  $y_t$  is a rate:

$$E[\Psi_t] = y_t T.$$

Using:

$$E[\Psi_t] = \lambda P_{\text{ON}} T \cdot E[\psi_t],$$

we obtain:

$$E[\psi_t] = \frac{y_t}{\lambda P_{\text{ON}}}. \quad (1)$$

This guarantees:

$$\frac{1}{T} E[\Psi_t] = y_t.$$

Thus, the slot average is exactly preserved.

The best-performing parameters were:

$$\tau = \frac{1}{15}, \quad \zeta = \frac{1}{60}, \quad \lambda = 2.0.$$

These correspond to:

- Average OFF duration: 15 seconds,
- Average ON duration: 60 seconds,
- Arrival intensity during ON: 2 arrivals/sec.

## 4 VALIDATION RESULTS

### 4.1 Mean Preservation Check

The fine-grained signal is aggregated back to coarse resolution:

$$\hat{y}_t = \frac{1}{T} \int_0^T x_t(s) ds.$$

The reconstruction error satisfies:

$$\text{MAPE} \approx 0.13.$$

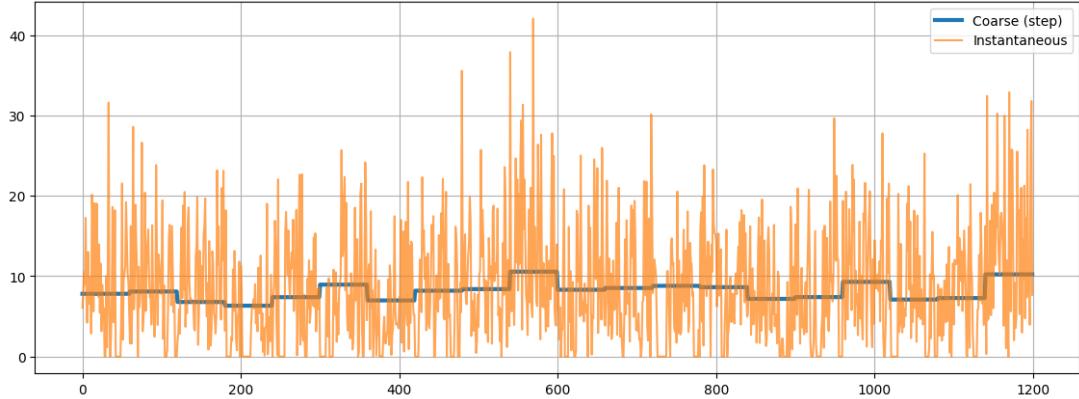
This confirms correct implementation of Eq.(1).

## 4.2 Qualitative Behavior

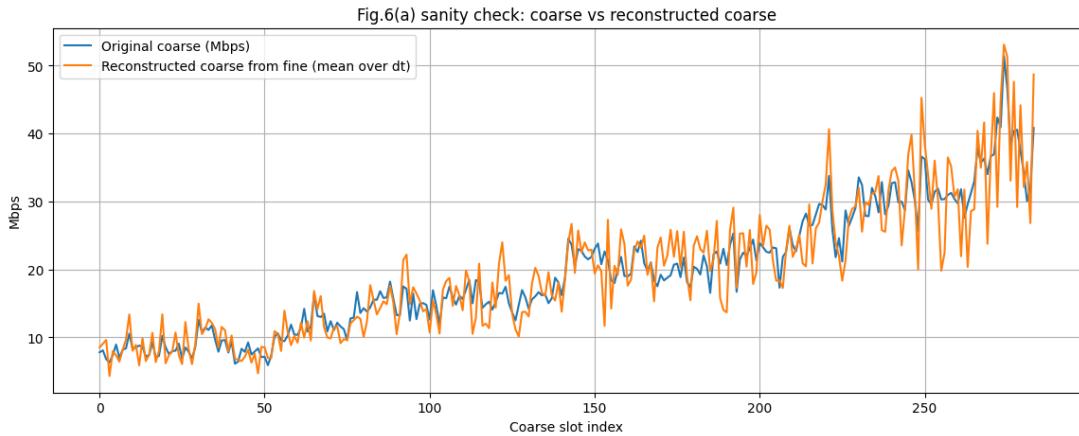
The instantaneous signal:

- Exhibits silent OFF periods,
- Shows burst peaks during ON states,
- Matches the coarse envelope.

This reproduces the qualitative behavior shown in Fig.6(a) of the KTH paper.



**Figure 1:** Instantaneous IPP signal versus coarse staircase signal.



**Figure 2:** Reconstructed coarse signal from fine-grained data.