Modeling of Traffic Flows in Internet of Things Using Renewal Approximation

Florian Wamser, Phuoc Tran-Gia, Stefan Geißler and Tobias Hossfeld

Abstract This paper proposes a versatile approach to model aggregated traffic flows in the Internet of Things (IoT) using renewal approximation. The modeled traffic originates from a large number of sources or devices consisting of a set of sensors mixed with classical elastic random traffic modeled as Poisson arrival process. The work shows the exact derivation in the simple case for periodic sensors. It shows further results in the mixed case with periodic sensors and a background process. The renewal approximation allows to derive the required number of sensors such that the aggregated traffic can be approximated as Poisson process.

Keywords: Internet of Things, Traffic Modeling, Renewal Approximation

1 Introduction

Internet of Things (IoT) is a growing area in mobile communication applications [1, 2]. It is expected that millions of devices will be found on the networks in the near future, each sending independently or via gateways over the mobile network. In such a scenario, IoT devices encompass all types of physical nodes or objects that are connected to the Internet to receive and respond to requests, or to store data. IoT devices can be subdivided into (1) stand-alone devices with independent Internet connectivity, (2) device groups that communicate in an aggregated manner with servers through Internet gateways on the Internet, or (3) devices that communicate with one another based on direct peer-to-peer connections.

A typical situation is the aggregation of traffic streams from many independent sensors to an Internet gateway as described in [3, 4]. This class of sensors send data at periodically time-fixed intervals to store measurements or to request input data and

Florian Wamser, Phuoc Tran-Gia, Stefan Geißler, Tobias Hoßfeld
University of Würzburg, Institute of Computer Science, Germany,
e-mail: {florian.wamser|trangia|stefan.geissler|hossfeld}@informatik.uni-wuerzburg.de

updates. The time-fixed intervals result from mechanisms to conserve power or from continuous measurements at specific time intervals such as in Smart Grids [5, 4]. As in the classic Internet, this traffic overlaps with background traffic from various other sources, which can be considered independent due to the large number of sources.

In previous work, aggregated IoT traffic was modeled for periodic traffic with a fixed sending period of the individual sensor node intervals [3, 4, 6]. We extend this idea with an additional component, a random and heterogeneous background traffic, as described in [4]. To this end, a closed-form expression for the approximation of the inter-arrival time distribution of the superposition of different arrival processes of IoT devices is derived in this work using renewal approximation. We consider a class of sensors that send data at consecutive, time-fixed intervals combined with a continuous-time Markov arrival stream in form of a background Poisson random process. We provide a detailed derivation for the approximation of the inter-arrival time distribution based on the renewal approximation including an exact determination of the coefficient of variation, which can be well approximated by a Poisson process such that the statistical differences are below a threshold ϵ.

The rest of the paper is structured as follows. After this introduction, in Section 2 related work is discussed. In Section 3, the aggregated IoT traffic is modeled for an Internet gateway. We introduce the used notation and definitions and provide a detailed description of the approach. Further on, two cases are described in detail in Section 3.1 and 3.2. After showing numerical results in Section 4, we conclude the work in Section 5.

2 Related Work

The superposition of a number of deterministic flows is a subject of various papers in the last decades, especially during the development of Asynchronous Transfer Mode (ATM) technology [6, 7]. The resulting process of n deterministic flows, each of rate $1/T$, is a periodic non-renewal process of the same period T [8, 9]. Assuming now that the traffic sources are independent, e.g. due to a very large number of sources, one can model the superposition of deterministic point processes as a Poisson process as limiting case [8]. There are papers [10, 11] that discuss the renewal assumption that holds true in the Poisson case and does not hold for deterministic processes when the number is small. In this paper we apply the renewal approximation and check whether and when it is valid. Further work on traffic modeling can be found in e.g. [12, 4, 3]. Directly linked to our work are the works of Metzger *et al.* [4] and Hoßfeld *et al.* [3]. They both refer to the same modeling context as the present work. These papers are pioneer in this area, facing the same problem, and providing basic ideas for IoT traffic flows modeling. Our work is based on these approaches and complements them with further definitions. Our approach is different in that we apply the renewal approximation to derive a closed form. In [4] a comprehensive list of IoT traffic models is given, showing how important periodic traffic characteristics are in the IoT environment.

3 Modeling Aggregated Traffic Flows

The modeling of aggregated traffic flows in the IoT environment relies on a fundamental consideration and definition of the arrival processes of the individual sources. As described in [4], the predominant consideration of these traffic flows in the literature is the Poisson arrival process. This is in contrast to the work from ATM times [6], which specifies the need for more detailed consideration of periodic traffic. In the following, periodic traffic flows are defined and modeled in detail. We use a description consistent with [4], from which we derive the distribution function and its moments. The latter serves to answer the question of how many devices aggregated traffic flows can be approximated as Poisson process.

Variables and Notation As the resulting processes of flows in IoT environments are generally point processes, we employ a renewal approximation technique to derive the inter-arrival distribution function of the resulting flow, assuming it follows renewal input process properties [13]. The main steps of the renewal approximation used in the analysis are:

1. Consider an independent outside observer looking at the process at an independent point in time.
2. Derive the residual time distribution of the resulting process, i.e. the interval from observation instance until the next arrival to occur.
3. With the assumption that the resulting process is a renewal process we then derive the inter-arrival distribution function out of the forward residual time.

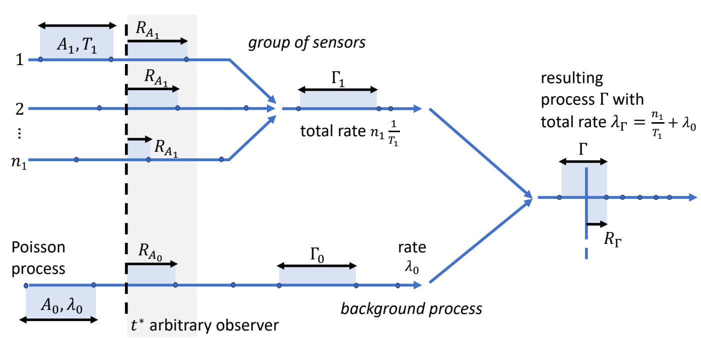


Fig. 1: General model with n_1 deterministic sensor sources plus Poisson source.

In the following, we consider the scenario described in Fig. 1. There is a group of different sensors. Each device sends periodically messages with period T_1. There are n_1 devices in this class. A node k starts randomly at time $t_{1,k} \in [0; T_1]$ and thus, the sending times are $t_{1,k} + z \cdot T_1$ with $z \in \mathbb{N}$. We denote the distribution function for this process from the sensor nodes as $A_1(t)$, respectively $a_1(t)$ as density distribution function. The traffic pattern for this group is repeated after period T which is the least common multiple of the sending periods for this class. The resulting stream originating from the group of sensors is described with Γ_1 as inter-arrival time of the resulting process. We model a background random Poisson process as Γ_0 and rate λ_0. The final aggregated process of the sensors and the background process is denoted as Γ with a total rate of $\lambda_\Gamma = n_1/T_1 + \lambda_0$.

Table 1: Notation

T_1	$\triangleq$	sending period of a sensor; without loss of generality, we assume $T_1 \in \mathbb{N}$
n_1	$\triangleq$	number of devices within the group of sensors
A_1	$\triangleq$	inter-arrival time of one input process $A_1(t)$, $a_1(t)$
$R_{A_1}(t)$	$\triangleq$	distribution function of residual time $R_{A_1}(t) = P(R_{A_1} \leq t)$
$r_{A_1}(t)$	$\triangleq$	density function of residual time
$R^C_{A_1}$	$\triangleq$	complementary distribution function of residual time $R^C_{A_1} = P(R_{A_1} > t)$
Γ	$\triangleq$	inter-arrival time of the resulting process
R_Γ	$\triangleq$	residual time of the resulting process

Residual Time Distribution The residual time or the forward recurrence time is the time between any random observation time until the next arrival. We consider a random observer looking at the process, the interval to the next observed arrival is denoted by the random variable R_Γ. The residual time until the next message arrival is the minimum of the residual time of participating processes:

$$R_\Gamma = \min(\underbrace{R_{A_1}, R_{A_1}, ..., R_{A_1}}_{n_1 \text{ times}}, R_{A_0}). \tag{1}$$

This leads to the complementary cumulative distribution function (CCDF) of the resulting process

$$P(R_\Gamma > t) = 1 - R_\Gamma(t) = \underbrace{P(R_{A_1} > t) \cdot P(R_{A_1} > t) \cdot \ldots \cdot P(R_{A_1} > t)}_{n_1 \text{ times}} \cdot P(R_{A_0} > t)\,. \tag{2}$$

We obtain subsequently the distribution function of the residual time of the resulting process. In assuming the resulting process to be a renewal process, we can use the basic result of renewal theory $r(t) = \frac{1}{E[\Gamma]}(1 - \Gamma(t))$ to derive the inter-arrival time distribution with Eq. (2)

$$\Gamma(t) = 1 - E[\Gamma] \cdot r(t) = 1 - E[\Gamma] \cdot \frac{d}{dt} R_\Gamma(t)\,. \tag{3}$$

It is obvious that the aggregated stream of deterministic traffic processes is non-renewal. However, with the superposition of a very large number of processes, like in IoT environments with huge sets of sensors, the inter-arrival time occurs in microscopic scale compared to the periodicity of a single participating process. We expect that the resulting process is "more renewal" with a growing number of superimposed processes. In this paper, we investigate under which conditions the results using renewal approximation is accurate enough for practical use in IoT systems and try to quantify accuracy of the renewal approximation. In the following we consider two consecutive cases and model their properties.

3.1 nD: Deterministic Case for a Group of Periodic Sensors

This case outlines the aggregation of n_1 deterministic flows solely to an aggregated stream of IoT traffic, in our case Γ_1 see Fig. 1. In the IoT context this model is employed to describe a (large) number of measurement data flows from a set of sensors. This process was also often used to model ATM traffic flows on aggregated cell patterns [6], where it is often denoted as nD.

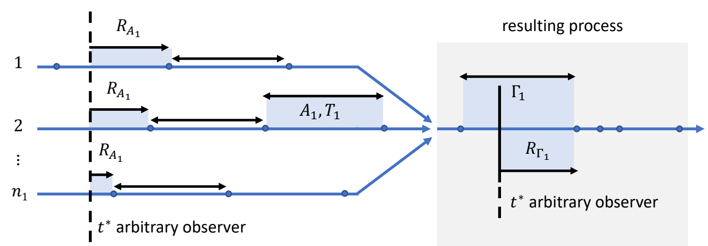


Fig. 2: Model Case 1 with arrival processes according to A_1, fixed T_1, and n_1 streams.

This basic model is depicted in Fig. 2 with an arbitrary observer at t^*. Each of the input processes, e.g. to represent traffic emitting from a sensor, is a deterministic process with inter-arrival time A_1 with distance T_1 and flow rate $\frac{1}{T_1}$. The corresponding CDF $R_{A_1}(t)$ and the probability density (PDF) of the recurrence time of A_1 are:

$$R_{A_1}(t) = P(R_{A_1} \leq t) = \begin{cases} 0 & \text{for } t < 0 \\ t/T_1 & \text{for } 0 \leq t \leq T_1 \\ 1 & \text{for } t > T_1 \end{cases}, \tag{4}$$

$$\frac{d}{dt} R_{A_1}(t) = r_{A_1}(t) = \begin{cases} 1/T_1 & \text{for } 0 \leq t \leq T_1 \\ 0 & \text{otherwise} \end{cases}. \tag{5}$$

With Γ_1 denoting the random variable of the inter-arrival time of the resulting process with corresponding residual time R_{Γ_1}, we obtain the residual time distribution function (of the superposition Γ_1 of n_1 deterministic flows) with density as

$$P(R_{\Gamma_1} \leq t) = R_{\Gamma_1}(t) = \begin{cases} 0 & t < 0 \\ 1 - (1 - \frac{1}{T_1} t)^{n_1} & 0 \leq t < T_1 \\ 1 & t \geq T_1 \end{cases} , \tag{6}$$

$$\frac{d}{dt} R_{\Gamma_1}(t) = r_{\Gamma_1}(t) = \begin{cases} \frac{n_1}{T_1}(1 - \frac{1}{T_1} t)^{n_1 - 1} & 0 \leq t < T_1 \\ 0 & \text{otherwise} \end{cases} . \tag{7}$$

Assuming the renewal property for the resulting process, we arrive at

$$\Gamma_1(t) = 1 - \frac{T_1}{n_1} r_{\Gamma_1}(t) = \begin{cases} 0 & t < 0 \\ 1 - (1 - \frac{1}{T_1} t)^{n_1 - 1} & 0 \leq t < T_1 \\ 1 & t \geq T_1 \end{cases} . \tag{8}$$

As discussed above, in general, the resulting process is non-renewal. During an interval of length T_1, there are exactly n_1 arrivals, which form a periodic pattern depending on the starting constellation of the flows. Thus, microscopically, the process is periodic, with infinite number of possible patterns. If n_1 is sufficiently large, from microscopic views, during a time interval sufficiently smaller than T_1, the inter-arrival process appears more random and a renewal process approximation appears more sensible. In the IoT context this model is employed to describe a (sufficiently large) number of measurement data flows from a set of sensors. We expect that if n_1 becomes large enough, the resulting process will quickly approach Poisson. We try here to compute this limit analytically.

From Eq. (8), we can assess the accuracy of the renewal approximation in more detail. The variance and the coefficient of variation of the resulting process are:

$$Var[\Gamma_1] = \frac{T_1^2 (n_1 - 1)}{n_1^2 (n_1 + 1)}, \qquad c_{\Gamma_1} = \frac{\sqrt{Var[\Gamma_1]}}{E[\Gamma_1]} = \sqrt{\frac{n_1 - 1}{n_1 + 1}} . \tag{9}$$

It can be seen from this expression that the coefficient of variation of the resulting flow just depends on the number n_1 of aggregated flows, not from the inter-arrival distance T_1. Furthermore for the case of one flow, $c_{\Gamma_1} = 0$ as expected for a deterministic process. For the limiting case $n_1 \to \infty$, we obtain $c_{\Gamma_1} \to 1$, which corresponds to the Markovian property. The resulting process approaches a Poisson process.

If we set a threshold $c_{\Gamma_{95\%}} = 0.95$ to answer the question, how many flows we need to deliver for a process with 95 % of the randomness of a Poisson stream, we arrive at $n_1 = 19.51$, i.e. with just $n_1 = 20$ flows, the superposition on average can already approximated with a Poisson process with more than 95% accuracy.

3.2 $nD + M$: Mixed Case with Periodic Sensors and a Background Process

The derivation of the case with an additional Poisson background traffic is analogous to the basic case above. This case is shown in Fig. 1. There are n_1 sensors periodically sending data. The sending period is T_1. The nodes start randomly within $[0, T_1]$. There is also background traffic with rate λ_0 with inter-arrival times A_0 following a negative-exponential distribution function: $A_0(t) = 1 - e^{-\lambda_0 t}$. The residual time R_0 has the same expression as for A_0.

The residual time of a sensor is $R_{A_1} \sim U(0, T_1)$ with CDF $R_{A_1}(t) = t/T_1$ for $0 \le t \le T_1$. The residual time for the aggregated traffic is $R_\Gamma = min(R_{A_1}, R_{A_1}, ..., R_{A_1}, R_{A_0})$. With $R_0 = A_0$, the CDF is given with

$$R_{\Gamma_1}(t) = \begin{cases} 0 & t < 0 \\ 1 - (1 - \frac{1}{T_1}t)^{n_1} \cdot e^{-\lambda_0 t} & 0 \le t < T_1 \\ 1 & t \ge T_1 \end{cases} . \tag{10}$$

The interarrival time distribution $\Gamma(t)$ can be derived using Eq. (3) with $E[\Gamma] = \frac{T_1}{n_1 + \lambda_0 T_1}$. The CDF for $\Gamma(t)$ for $0 \le t < T_1$ derives to

$$\Gamma(t) = 1 - \frac{T_1 e^{-\lambda_0 t} \left(1 - \frac{t}{T_1}\right)^{n_1} (n_1 + \lambda_0(T_1 - t))}{(T_1 - t)(n_1 + \lambda_0 T_1)} . \tag{11}$$

The coefficient of variation c_Γ can be derived in this case with standard mathematical tools analogous to the result of the basic case described above. The coefficient of variation is shown and explained below in the result section.

4 Numerical Results

In order to substantiate and validate our results, we compare the results obtained by renewal approximation with (i) an event-by-event simulation of an exact point process and (ii) results from previous works [4]. The simulation randomly generates a sufficiently long point process according to the given properties, over which the same statistical measures can be derived after many iterations as obtained analytically. For smart city use cases, typical sensor periods are 1 h and 12 h as e.g. proposed by 3GPP, see [4] for an overview on IoT traffic models. Hence, we assume $T_1 = 1$ h and $T_2 = 12$ h.

In Fig. 3, the coefficient of variation (CoV) is depicted for the deterministic case, where the CoV is shown as function of the number of sensor nodes. It can be seen when the coefficient of variation reaches certain values to justify the approximation by a Poisson process, e.g. $c_{\Gamma_1} = 0.95$ or $c_{\Gamma_1} = 1$. Here, only the aggregated periodic traffic of n_1 nodes is considered with period T_1. Since the simulation generates a point process with a specific deterministic pattern for each run, it is possible to plot the CoV as a distribution over all appearing instances. With this, and in addition to the analytic result, Fig. 3 shows (1) the coefficient of variation according to the renewal approximation, (2) the mean of the coefficient of variation from the simulation of 1000 random superpositions, (3) the quantiles of this simulation, and (4) the fitted result as specified by paper [4] with $\bar{C} = 1 - \frac{1}{n}$.

Both the empirically fitted formula from [4] and the value of the renewal approximation are close to the values of the simulation of many instances of the exact process. Furthermore, all curves run together with a large number of sensor nodes n_1. Nevertheless, the quantiles show that the mean values conceal the extreme cases. With a small number of devices, e.g $n_1 = 20$, the mean increases to 1, but the 5 % and 95 % quantiles are still more than 20 % away from the mean value. Overall, the renewal approximation can be used as a simple closed-form expression if one considers a high number of nodes and also takes into account the quantiles, which show that there are some highly variable occurrence of arrivals in individual cases.

For the numerical results of the mixed case with deterministic arrivals and a Poisson background arrival process, we consider a scenario with $\beta = 10\,\%$ of background traffic. The aggregated periodic traffic leads to an arrival rate n_1/T_1. Hence, the arrival rate of the background traffic is $\lambda_0 = \beta\lambda_\Gamma = \beta(n_1/T_1 + \lambda_0)$ which leads to $\lambda_0 = \frac{\beta}{1-\beta}\frac{n_1}{T_1}$.

Fig. 4a shows a comparison between a single simulation run of $nD + M$, Poisson process, and renewal approximation. All results are plotted for $n = 3, 10, 50$ nodes and $T_1 = 1\,\mathrm{h}$ with a ratio of $\beta = 0.1$ of Poisson background traffic. It shows the convergence of both the simulation runs to the renewal approach and the convergence of all approaches with a large number of nodes. With a larger number of devices, the

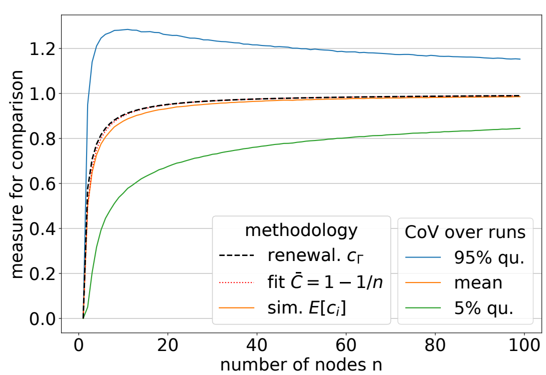
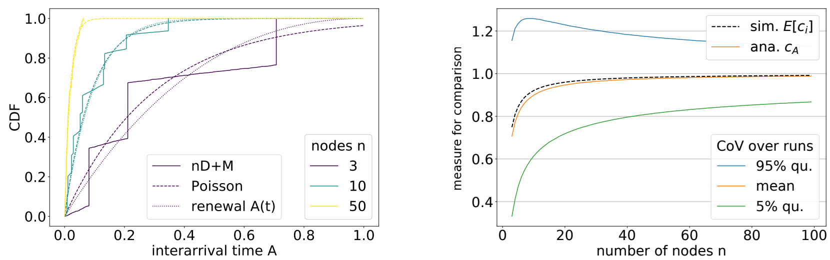


Fig. 3: Comparison of coefficients of variation: simulation, results from [4] and renewal approximation for different number of sensor nodes.

single simulation run curve loses its steps, i.e., visually the convergence to the curve of the renewal approximation can be viewed in this figure.

On the basis of this, the coefficient of variation over the number of sensors is shown in Fig. 4b to discuss the approximation using renewal assumption in the mixed case $nD + M$. The black dashed line is the analytic solution using renewal approximation. The numerical results are derived using numerical integration. The solid lines are from simulation runs. We use $T_1 = 1$ h and vary n_1. We keep a constant ratio of background traffic which is again $\beta = 0.1$.



(a) Comparison of $nD + M$ with different number of devices n_1 in the deterministic case for Poisson process, renewal approximation, and a single simulation run.

(b) Coefficients of variation for the aggregated case with background traffic from simulation and renewal approximation on the number of sensor nodes.

Fig. 4: Mixed case $nD + M$, periodic traffic of sensor nodes and background Poisson traffic.

The simulation and analytic solution from the renewal approximation coincide; in fact, they converge for a large number of sensor nodes. Hence, the renewal approximation can also be used here for a large number of nodes. The results in this case, however, again show large distances to the 5 % and 95 % quantiles of the simulation runs, which is also due to the low ratio of background traffic with $\beta = 0.1$. If β increases, the curve from the analytic solution approaches 1 more quickly, which means the process becomes more random and converges faster to a random process where the renewal approximation can be employed.

5 Conclusion

This paper had the objective to describe an aggregated traffic mix of IoT devices with (1) periodic traffic patterns and (2) background traffic using renewal approximation. It is based on the papers [3, 4]. In contrast to them, in this paper a closed-form expression of the approximation of the distribution function for the aggregated traffic mix is derived using renewal approximation. Both the simple case with periodic-sending sensors and the mixed case with Poisson background traffic were calculated. The numerical results demonstrate the consistency of this approach with simulated instances of an exact point processes for a large number of devices. In the analytical form, it is shown in this paper that the coefficient of variation for $n \geq 20$ goes sufficiently against 1 and allows to quantify the required nodes, such that the aggregated traffic can be approximated by a Poisson process.

References

1. Shancang Li, Li Da Xu, and Shanshan Zhao. 5g internet of things: A survey. *Journal of Industrial Information Integration*, 10:1–9, 2018.
2. I-Scoop, 5G and IoT in 2018 and beyond: the mobile broadband future of IoT. https://www.i-scoop.eu/internet- of-things-guide/5g-iot/. Accessed: 2019-04-03.
3. Tobias Hoßfeld, Florian Metzger, and Poul E Heegaard. Traffic modeling for aggregated periodic iot data. In *2018 21st Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*, pages 1–8. IEEE, 2018.
4. Florian Metzger, Tobias Hoßfeld, André Bauer, Samuel Kounev, and Poul E Heegaard. Modeling of aggregated iot traffic and its application to an iot cloud. *Proceedings of the IEEE*, 107(4):679–694, 2019.
5. Rapeepat Ratasuk, Athul Prasad, Zexian Li, Amitava Ghosh, and Mikko Uusitalo. Recent advancements in M2M communications in 4G networks and evolution towards 5G. In *2015 18th International Conference on Intelligence in Next Generation Networks*, pages 52–57. IEEE, 2015.
6. James Roberts and Jorma Virtamo. The superposition of periodic cell arrival streams in an atm multiplexer. *IEEE Transactions on Communications*, 39(2):298–303, 1991.
7. Ottmar Gihr and Phuoc Tran-Gia. A layered description of atm cell traffic streams and correlation analysis. In *IEEE INFCOM'91. The conference on Computer Communications. Tenth Annual Joint Comference of the IEEE Computer and Communications Societies Proceedings*, pages 137–144. IEEE, 1991.
8. David Roxbee Cox and Walter Smith. On the superposition of renewal processes. *Biometrika*, 41(1-2):91–99, 1954.
9. Aleksandr Yakovlevich Khinchin, DM Andrews, and Maurice Henry Quenouille. *Mathematical methods in the theory of queuing*. Courier Corporation, 2013.
10. Susan L Albin. On poisson approximations for superposition arrival processes in queues. *Management Science*, 28(2):126–137, 1982.
11. Ward Whitt. Approximating a point process by a renewal process, i: Two basic methods. *Operations Research*, 30(1):125–147, 1982.
12. Muhammad Asad Arfeen, Krzysztof Pawlikowski, Andreas Willig, and Don McNickle. Internet traffic modelling: from superposition to scaling. *IET networks*, 3(1):30–40, 2014.
13. David Roxbee Cox and Hilton David Miller. The theory of stochastic processes. *Chapman and Hall, New York*, 1965.