Aggregated Traffic Models for Real-World Data in the Internet of Things

Miguel López-Benítez, *Senior Member, IEEE,* Chitradeep Majumdar and Shabbir N. Merchant, *Member, IEEE*

***Abstract*—Traffic models play a key role in the analysis, design and simulation of communication networks. The availability of accurate models is essential to investigate the impact of traffic patterns created by the introduction of new services such as those forecasted for the Internet of Things (IoT). The Poisson model has historically been a popular aggregated traffic model and has been extensively used by the IoT research community. However, the Poisson model implicitly assumes an infinite number of traffic sources, which may not be a valid assumption in various plausible application scenarios. The practical conditions under which the Poisson model is valid in the context of IoT have not been fully investigated, in particular under a finite (and possibly reduced) number of traffic sources with random inter-arrival times. In this context, this work derives exact mathematical models for the packet inter-arrival times of aggregated IoT data traffic based on the superposition of a finite number of traffic sources, each of which is modelled based on real-world experimental data from typical IoT sensors (temperature, light and motion). The obtained exact models are used to explore the validity of the Poisson model, showing that it can be extremely inaccurate when a reduced number of traffic sources is considered. Finally, an illustrative example is presented to show the importance of having accurate and realistic models such as those presented in this work.**

***Index Terms*—Internet of Things, aggregated traffic, traffic modelling, Poisson process.**

I. INTRODUCTION

WITH the advent of a myriad of machine-type devices interconnected through the Internet of Things (IoT), communication networks are facing unprecedented challenges to efficiently support the dramatic increase of traffic loads. In order to optimise current and future communication systems for IoT, it is essential to first understand and model the specific traffic patterns generated by IoT data. While a broad range of traffic models have been proposed in the context of IoT in the literature [1], the Poisson process is certainly the most widely adopted model, including standardisation bodies such as 3GPP where this model has been employed to assess the performance of IoT communications over cellular mobile networks [2], [3].

A Poisson process is essentially a renewal process with exponentially distributed inter-arrival times. The Poisson process is one of the oldest traffic models (if not the oldest one), dating back to the early days of landline telephony, where it was used to characterise the arrival of calls to a telephone exchange. The theoretical basis for this model is the Palm-Khintchine theorem [4], which shows that the superposition of several independent processes converges to a Poisson process (irrespective of the statistics of the individual component processes) as the number of superimposed processes tends to infinity.

Poisson processes are fairly common in practical scenarios where the traffic from a large number of independent traffic sources is aggregated. Thus, traffic flows on the main arteries of communication networks are commonly believed to follow a Poisson process (e.g., in IoT cloud servers where data from a large number of IoT devices are collected and aggregated in a data centre). However, the aggregation of individual traffic sources does not always result in a Poisson process. This is particularly true in practical scenarios with a moderate number of sources, such as indoor small cells or indoor Wi-Fi access points serving as gateways for smart-home IoT devices.

The observation above raises the practical question of how many individual traffic sources need to be aggregated such that the assumption of a Poisson process is realistic. This problem has recently received some attention in the context of periodic IoT traffic sources where data packets are generated periodically at deterministic fixed time intervals [5], such as e.g. smart grids, where smart meters periodically report power usage levels to a data centre. However, in event-driven IoT applications, data packets are generated at variable time intervals in response to random events [6] (e.g., whenever a certain metric of interest varies by a predefined quantity or exceeds a threshold). To the best of the authors' knowledge, the conditions under which the Poisson process is a valid aggregated traffic model for IoT sources with random inter-arrival times has not been investigated to the date. Moreover, the optimisation of scenarios where the Poisson model may not be valid claims for new traffic models that can describe accurately the aggregated traffic flows regardless of the number of aggregated traffic sources, including those cases where the number of individual traffic sources is too low for the Poisson process to be a valid aggregated traffic model. In this context, this work fills the existing gap by exploring the validity of the Poisson process as an aggregated traffic model for IoT sources with random inter-arrival times. Leveraging on some recent models for individual traffic sources based on experimental data from real-word IoT devices, this work analytically develops exact aggregated models for an arbitrary finite number of traffic sources. The developed models are exploited to determine the conditions under which the Poisson process is a realistic aggregated traffic model. The main novelty and contribution of this work is a critical analysis of the validity of the Poisson process as an aggregated traffic model

This work was supported by the Royal Society of UK and the Science and Engineering Research Board (SERB) of India under a Royal Society-SERB Newton International Fellowship (ref. NF170943).

M. López-Benítez is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, L69 3GJ, United Kingdom, and also with the ARIES Research Centre, Antonio de Nebrija University, 28040 Madrid, Spain. Email: m.lopez-benitez@liverpool.ac.uk.

C. Majumdar is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, L69 3GJ, United Kingdom. Email: chitradeep.majumdar@liverpool.ac.uk.

S. N. Merchant is with the Department of Electrical Engineering, IIT Bombay, Mumbai, 400076, India. Email: merchant@ee.iitb.ac.in.

Digital Object Identifier 10.1109/LWC.2020.XXXXXXX.

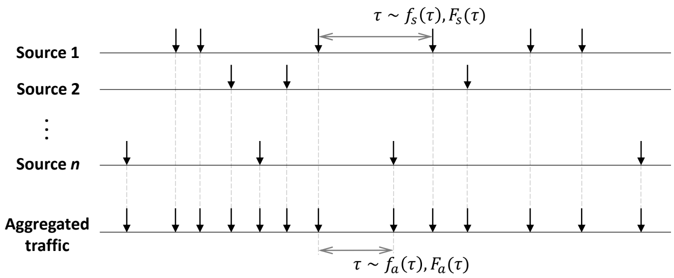

Fig. 1. System model considered in this work.

in practical IoT scenarios and the proposal of an alternative modelling approach that can provide a significantly improved accuracy in those cases where the Poisson model is not valid.

The rest of this work is organised as follows. First, Section II presents the system model and formulates the problem addressed in this work. Exact aggregated models for IoT data sources with random inter-arrival times are then derived in Section III based on source traffic models available in the literature from empiric data. The conditions under which the Poisson process is a valid aggregated model are investigated in Section IV. Finally, Section V concludes this work.

II. System Model and Problem Formulation

Let n be the number of individual IoT traffic sources, each of which generates data at random intervals τ as shown in Fig. 1. The random inter-arrival times τ of the n traffic sources are assumed to be mutually independent and identically distributed according to a certain Probability Density Function (PDF) and Cumulative Distribution Function (CDF), denoted by $f_s(\tau)$ and $F_s(\tau)$, respectively. These inter-arrival times can follow any arbitrary distribution, which is not necessarily (and, indeed, in many practical cases is not) exponential [7].

The data generated by the n sources are multiplexed at some point in the network, thus leading to a single aggregated traffic flow with random inter-arrival times τ distributed according to a certain PDF, $f_a(\tau)$, and a corresponding CDF, $F_a(\tau)$, respectively. The arrival time instants of the aggregated flow can be obtained by rearranging the arrival time instants of each individual traffic source into a single increasing sequence.

The objective of this work is twofold. First, to obtain closed-form expressions for $f_a(\tau)$ and $F_a(\tau)$, for some known $f_s(\tau)$ and $F_s(\tau)$, as a function of the number of aggregated traffic sources n. Second, to exploit such expressions to determine (based on results from real-world IoT data) the minimum number of sources for which the aggregated traffic can be accurately modelled as a Poisson process. Such threshold will also determine the operation conditions under which the models developed in this work are the only realistic models.

III. Aggregated Traffic Models

This section provides closed-form expressions for the aggregated distributions $f_a(\tau)$, $F_a(\tau)$ for some particular source distributions $f_s(\tau)$, $F_s(\tau)$. The analysis is first carried out for exponentially distributed traffic sources and later on conducted for other distributions that, according to [7], characterise more accurately the traffic generated by real-world IoT devices.

The mathematical analysis presented in this section benefits from the analytical result obtained in [8, eq. (31)], by virtue of which the distribution of interest can be obtained as follows:

$$f_a(\tau) = -\frac{\partial}{\partial \tau}\Bigg\{ \underbrace{[1 - F_s(\tau)] \left(\frac{1}{m_s} \int_{\tau}^{\infty} [1 - F_s(z)]\, dz \right)^{n-1}}_{=1-F_a(\tau)} \Bigg\} \tag{1}$$

where m_s is the source's mean inter-arrival time (i.e., mean time between successive packets generated by a single source). Notice that $F_a(\tau)$ can be obtained from the term inside the braces as indicated in (1) and its derivation leads to $f_a(\tau)$.

A. Extended Exponential (Poisson) Model

The aggregated distributions are first obtained for the Poisson model, assuming that the source inter-arrival times are exponentially distributed. An *extended* version of the exponential distribution given by $F_s(\tau) = 1 - e^{-\lambda(\tau-\mu)}$ $(\tau > \mu \geq 0)$ is here considered, where $\lambda > 0$ is the arrival rate (inverse scale parameter) and μ is the (not necessarily zero) minimum inter-arrival time (location parameter). The commonly used exponential distribution, $F_s(\tau) = 1 - e^{-\lambda\tau}$ $(\tau > 0)$, can be obtained as a particular case for $\mu = 0$. The introduction of this minimum value in the model allows for an increased flexibility and thus enables a more accurate characterisation in practical scenarios where the actual source inter-arrival times might be constrained by a non-zero lower bound.

Notice that the inter-arrival times of the aggregated traffic flow can take any value within the interval $[0, \infty)$ regardless of the minimum source inter-arrival time μ. As a result, (1) needs to be evaluated over the whole interval $[0, \infty)$. However, $F_s(\tau) = 0$ for $\tau < \mu$ and $F_s(\tau) \neq 0$ for $\tau > \mu$, which requires a separate evaluation of (1) for each case. Introducing the extended exponential distribution into (1) leads to:

$$1 - F_a(\tau) = \begin{cases} \left(\frac{1}{m_s} \left[\int_{\tau}^{\mu} dz + \int_{\mu}^{\infty} e^{-\lambda(z-\mu)} dz \right] \right)^{n-1}, & \tau < \mu \\ e^{-\lambda(\tau-\mu)} \left(\frac{1}{m_s} \int_{\tau}^{\infty} e^{-\lambda(z-\mu)} dz \right)^{n-1}, & \tau > \mu \end{cases}$$

where $m_s = \mu + 1/\lambda$, which yields the following CDF:

$$F_a(\tau) = \begin{cases} 1 - \left(1 - \frac{\lambda\tau}{1+\lambda\mu} \right)^{n-1}, & \tau \leq \mu \quad \text{(2a)} \\ 1 - \frac{e^{-\lambda n(\tau-\mu)}}{(1+\lambda\mu)^{n-1}}, & \tau \geq \mu \quad \text{(2b)} \end{cases}$$

and its derivation provides the associated PDF:

$$f_a(\tau) = \begin{cases} \frac{\lambda(n-1)}{1+\lambda\mu} \left(1 - \frac{\lambda\tau}{1+\lambda\mu} \right)^{n-2}, & \tau < \mu \quad \text{(3a)} \\ \frac{\lambda n e^{-\lambda n(\tau-\mu)}}{(1+\lambda\mu)^{n-1}}, & \tau > \mu \quad \text{(3b)} \end{cases}$$

Notice that the PDF has a discontinuity at $\tau = \mu$ and as a result the case $\tau = \mu$ is not explicitly included in (3).

For $\mu = 0$, (2) reduces to $F_a(\tau) = 1 - e^{-\lambda n\tau}$ and (3) reduces to $f_a(\tau) = \lambda n e^{-\lambda n\tau}$, showing that the superposition of n (standard) Poisson processes with rate λ results in another (standard) Poisson process with rate λn as expected.

B. Weibull Model

The Weibull distribution has been reported in [7] to be an accurate model for the inter-arrival times of the data packets generated by temperature and light intensity sensors that send a new report (i.e., a new data packet) every time the measured temperature or light intensity differs from the last reported value by a certain predefined amount (i.e., based on differential reporting). Temperature and light sensors are commonly found in IoT scenarios and are considered in this subsection.

The Weibull distribution (lower-bounded by μ) is given by $F_s(\tau) = 1 - \exp(-[(\tau-\mu)/\lambda]^\alpha)$ $(\tau > \mu \geq 0)$, with mean $m_s = \mu + \lambda\Gamma(1+1/\alpha)$, where $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ is the gamma function. The shape parameter $\alpha > 0$ confers the Weibull distribution an increased flexibility to fit empirical data compared to the exponential distribution, which can in fact be obtained as a particular case (concretely, for $\alpha = 1$, the Weibull distribution becomes an exponential with rate $1/\lambda$).

Introducing the Weibull model into (1) leads to the following CDF for the inter-arrival times of the aggregated stream:

$$F_a(\tau) = \begin{cases} 1 - \left(1 - \dfrac{\tau}{\mu + \lambda\Gamma\left(1+\frac{1}{\alpha}\right)}\right)^{n-1}, & \tau \leq \mu \quad \text{(4a)} \\ 1 - \exp\left(-\left[\dfrac{\tau-\mu}{\lambda}\right]^\alpha\right) \times & \\ \quad \times \left(\dfrac{\lambda\Gamma\left(\frac{1}{\alpha}, \left[\frac{\tau-\mu}{\lambda}\right]^\alpha\right)}{\alpha\left[\mu + \lambda\Gamma(1+\frac{1}{\alpha})\right]}\right)^{n-1}, & \tau \geq \mu \quad \text{(4b)} \end{cases}$$

and its derivation provides the associated PDF:

$$f_a(\tau) = \begin{cases} \frac{n-1}{\mu+\lambda\Gamma\left(1+\frac{1}{\alpha}\right)} \left(1 - \frac{\tau}{\mu+\lambda\Gamma\left(1+\frac{1}{\alpha}\right)}\right)^{n-2}, & \tau < \mu \quad \text{(5a)} \\ \frac{\alpha}{\lambda}\left(\frac{\tau-\mu}{\lambda}\right)^{\alpha-1} \exp\left(-\left[\frac{\tau-\mu}{\lambda}\right]^\alpha\right) \times & \\ \quad \times \left(\frac{\lambda\Gamma\left(\frac{1}{\alpha}, \left[\frac{\tau-\mu}{\lambda}\right]^\alpha\right)}{\alpha\left[\mu+\lambda\Gamma(1+\frac{1}{\alpha})\right]}\right)^{n-1} + & \\ \quad + \frac{n-1}{\mu+\lambda\Gamma(1+\frac{1}{\alpha})} \exp\left(-2\left[\frac{\tau-\mu}{\lambda}\right]^\alpha\right) \times & \\ \quad \times \left(\frac{\lambda\Gamma\left(\frac{1}{\alpha}, \left[\frac{\tau-\mu}{\lambda}\right]^\alpha\right)}{\alpha\left[\mu+\lambda\Gamma(1+\frac{1}{\alpha})\right]}\right)^{n-2}, & \tau > \mu \quad \text{(5b)} \end{cases}$$

where $\Gamma(s,z) = \int_z^\infty t^{s-1}e^{-t}dt$ represents the upper incomplete gamma function. Notice that the PDF has a discontinuity at $\tau = \mu$ and as a result the case $\tau = \mu$ is not explicitly included in (5). It can be shown that, when $\alpha = 1$, the expressions in (4) and (5) reduce to (2) and (3), respectively, with rate parameter $1/\lambda$.

C. Generalised Pareto Model

The Generalised Pareto distribution has been shown in [7] to be an accurate model for the inter-arrival times of the data packets generated by motion sensors, which are another type of sensors commonly found in practical IoT applications.

The CDF of the generalised Pareto distribution is given by $F_s(\tau) = 1 - [1 + \alpha(\tau-\mu)/\lambda]^{-1/\alpha}$ $(\tau > \mu \geq 0)$, with mean $m_s = \mu + \lambda/(1-\alpha)$, where $\mu > 0$, $\lambda > 0$ and $\alpha > 0$ are the location, scale and shape parameters, respectively. This distribution is also more flexible than the exponential distribution, which can also be obtained as a particular case (concretely, for $\alpha = 0$, the generalised Pareto distribution becomes an exponential distribution with rate $1/\lambda$).

Introducing the generalised Pareto distribution into (1) leads to the following CDF for the inter-arrival times of the aggregated traffic stream:

$$F_a(\tau) = \begin{cases} 1 - \left(1 - \dfrac{(1-\alpha)\tau}{\lambda + \mu(1-\alpha)}\right)^{n-1}, & \tau \leq \mu \quad \text{(6a)} \\ 1 - \left(\dfrac{\lambda + \alpha(\tau-\mu)}{\lambda + \mu(1-\alpha)}\right)^{n-1} \times & \\ \quad \times \left(1 + \dfrac{\alpha(\tau-\mu)}{\lambda}\right)^{-\frac{n}{\alpha}}, & \tau \geq \mu \quad \text{(6b)} \end{cases}$$

and its derivation provides the associated PDF:

$$f_a(\tau) = \begin{cases} \frac{(n-1)(1-\alpha)}{\lambda+\mu(1-\alpha)}\left(1 - \frac{(1-\alpha)\tau}{\lambda+\mu(1-\alpha)}\right)^{n-2}, & \tau < \mu \quad \text{(7a)} \\ \frac{n}{\lambda}\left(\frac{\lambda+\alpha(\tau-\mu)}{\lambda+\mu(1-\alpha)}\right)^{n-1}\left(1 + \frac{\alpha(\tau-\mu)}{\lambda}\right)^{-\left(\frac{n}{\alpha}+1\right)} - & \\ \quad - \frac{\alpha(n-1)}{\lambda+\mu(1-\alpha)}\left(\frac{\lambda+\alpha(\tau-\mu)}{\lambda+\mu(1-\alpha)}\right)^{n-2} \times & \\ \quad \times \left(1 + \frac{\alpha(\tau-\mu)}{\lambda}\right)^{-\frac{n}{\alpha}}, & \tau > \mu \quad \text{(7b)} \end{cases}$$

Notice that the PDF has a discontinuity at $\tau = \mu$ and as a result the case $\tau = \mu$ is not explicitly included in (7). It can be shown that, when $\alpha = 0$, the expressions in (6) and (7) reduce to (2) and (3), respectively, with rate parameter $1/\lambda$.

IV. Numerical Results

The validity of the Poisson process as an aggregated traffic model is determined by evaluating the deviation of the analytical results in Section III-A from those in Sections III-B and III-C in terms of the Kolmogorov-Smirnov (KS) distance, which is defined as the maximum absolute difference between two CDFs and in the context of this work is evaluated as:

$$D_{KS} = \max_\tau \left|F_a^{\mathrm{P}}(\tau) - F_a^{\mathrm{W,GP}}(\tau)\right| \quad (8)$$

where $F_a^{\mathrm{P}}(\tau)$ is (2) and $F_a^{\mathrm{W,GP}}(\tau)$ is either (4) or (6). The KS distance can be used to test the equality of two continuous distributions. The lower the KS distance, the more similar the distributions are (with $D_{KS} = 0$ indicating identical distributions). By noting that the true distribution of empirical data can be accurately modelled by $F_a^{\mathrm{W,GP}}(\tau)$ as shown in [7], the KS distance in (8) can be used as a metric to quantify how far the Poisson process model $F_a^{\mathrm{P}}(\tau)$ is from real data and thus its accuracy and practical validity. The parameters of the Weibull and generalised Pareto distributions are set based on [7], where these models were fitted to real experimental data from temperature/light and motion sensors, respectively. For a fair comparison, the parameters of the Poisson model are set to reproduce the same minimum and mean inter-arrival times (i.e., $\mu_{\mathrm{P}} = \mu_{\mathrm{W}}$ and $\lambda_{\mathrm{P}} = [\lambda_{\mathrm{W}} \cdot \Gamma(1 + 1/\alpha_{\mathrm{W}})]^{-1}$ when comparing to the Weibull model, while $\mu_{\mathrm{P}} = \mu_{\mathrm{GP}}$ and $\lambda_{\mathrm{P}} = (1-\alpha_{\mathrm{GP}})/\lambda_{\mathrm{GP}}$ for the generalised Pareto model).

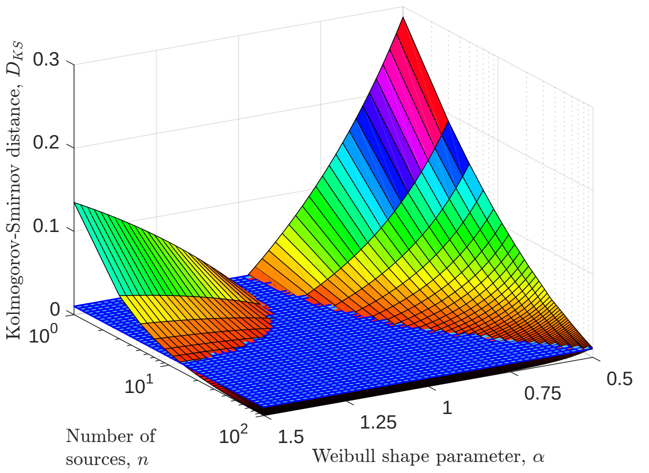

Fig. 2. Deviation of the Poisson process model from the Weibull model. Weibull model parameters: $\mu = 5 \cdot 10^3$ s, $\lambda = 2 \cdot 10^5$ s (based on [7]).

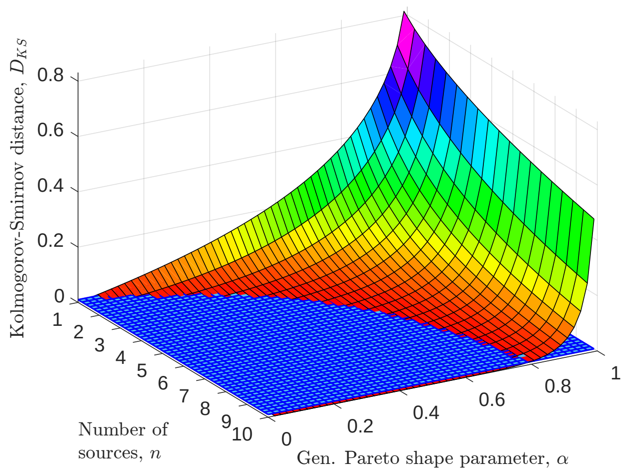

Fig. 3. Deviation of the Poisson model from the generalised Pareto model. Generalised Pareto model parameters: $\mu = 8.8$ s, $\lambda = \mu \cdot \alpha$ (based on [7]).

Fig. 2 shows the KS distance between the Poisson process and Weibull models as a function of the number of sources and the Weibull shape parameter α (the impact of the location and scale parameters is much less significant and therefore is not shown here). A horizontal plane at $D_{KS} = 0.01$ is also shown, which represents a maximum deviation of 1% with respect to the true distribution of the packet inter-arrival times of the aggregated data traffic and is used as a reference to determine the regions where the accuracy of the Poisson model can be considered satisfactory (i.e., $D_{KS} \leq 0.01$). As appreciated, the accuracy of the Poisson model improves (i.e., D_{KS} decreases) as the number of traffic sources increases, which is not surprising since the Poisson model implicitly assumes an infinite number of traffic sources. The minimum number of sources required for the Poisson process to be valid depends on the Weibull shape parameter. Such minimum requirement increases as the value of the shape parameter diverges from $\alpha = 1$ (for which the Weibull distribution becomes exponential as pointed out in Section III-B). Fig. 3 shows the counterpart for the generalised Pareto model; similar comments can be made (in this case, the minimum number of sources required for the Poisson process to be valid increases as the shape parameter diverges from $\alpha = 0$, where the generalised Pareto distribution is equivalent to an exponential distribution as pointed out in Section III-C).

Fig. 4 evaluates the accuracy of the Poisson model for the aggregated traffic of temperature sensors which report when a temperature difference of 4.5 °C is detected (Weibull model) and motion sensors (generalised Pareto). The parameters of the distributions are configured based on the empirical results obtained in [7]: $\mu = 5 \cdot 10^3$ s, $\lambda = 2 \cdot 10^5$ s and $\alpha = 0.7664$ for the Weibull distribution, and $\mu = 8.8$ s, $\lambda = 6.5490$ s and $\alpha = 0.7442$ for the generalised Pareto distribution. As it can be appreciated, a minimum of 8 motion and 50 temperature sensors are required for the Poisson process to be a valid aggregated traffic model. In some scenarios such as indoor small cells or indoor Wi-Fi access points serving as gateways for smart-home IoT devices, the number of IoT devices can be below these thresholds. In such scenarios, the Poisson process would not be a suitable aggregated traffic model.

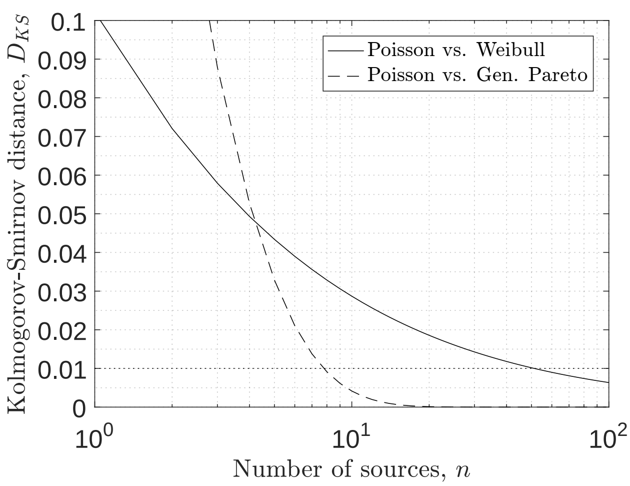

Fig. 4. Accuracy of the Poisson model for the aggregated traffic of temperature sensors (Weibull) and motion sensors (generalised Pareto) [7].

To illustrate the potential impact that an inaccurate traffic modelling could have, a practical example is provided based on a scenario where the gateway is programmed to enter a low power mode (idle state) after η seconds of inactivity from the last received packet, and switch back to active mode when a new packet is received. It can be shown that the fraction of time the gateway would remain in idle state, $\varphi \in [0, 1]$, is given by (details on the analytical derivation are provided in the Appendix):

$$\varphi = \frac{n}{m_s} \int_{\eta}^{\infty} (\tau - \eta) \, f_a(\tau) \, d\tau \tag{9}$$

Fig. 5 shows (for the same configuration as Fig. 4) the result of numerically evaluating (9) based on (3), (5) and (7) when $\eta = m_s$ (simulation results are also included for validation). The Poisson model results in a significant underestimation of φ and the corresponding energy savings, with errors of

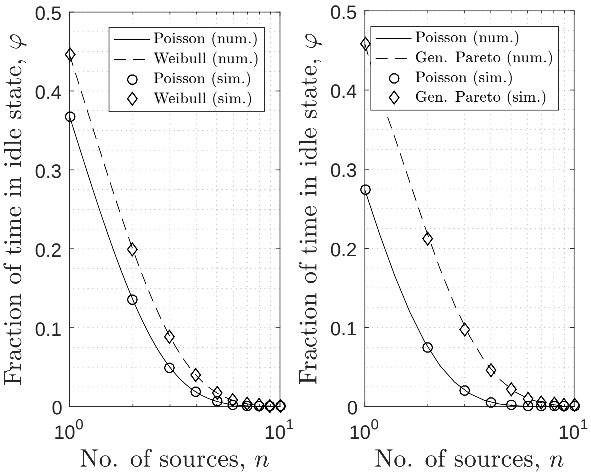

Fig. 5. Fraction of time that an IoT gateway would remain in idle state for the different traffic models as a function of the number of traffic sources [7].

up to 8% for temperature sensors (Weibull model) and 18% for motion sensors (generalised Pareto model). The minimum number of sensors for which the Poisson model would be accurate according to Fig. 4 is not relevant in this example since there would be no energy savings for that number of IoT devices according to Fig. 5. This is an example where the interest is in the aggregation of a low/moderate number of IoT traffic sources and where the Poisson model would not be suitable. In this type of IoT scenarios, the models developed in this work, which are exact for any number of sources, would therefore constitute a more realistic and accurate alternative.

V. Conclusions

The Internet of Things (IoT) will embrace a myriad of heterogeneous devices with radically different traffic patterns than those generated by legacy, human-driven services. The availability of accurate traffic models for all plausible scenarios is therefore essential. In this context, this work has investigated the suitability of the Poisson model as an aggregated traffic model for IoT. The Poisson model assumes an infinite number of traffic sources, an assumption that this work has demonstrated to be invalid in IoT scenarios that involve a low/moderate number of traffic sources, such as indoor small cells or indoor WiFi access points serving as gateways for IoT devices. A more convenient modelling approach for the inter-arrival times of aggregated traffic flows has been presented and illustrated for some typical real-world IoT applications. The proposed modelling approach provides exact results for any arbitrary number of traffic sources and is therefore suitable for those scenarios where the Poisson model fails to provide the required level of accuracy and realism.

Appendix
Derivation of Equation (9)

The fraction of time that the gateway in the example of Section IV remains in idle state, $\varphi \in [0,1]$, can be obtained as the ratio:

$$\varphi = \frac{\mathbb{E}(\tau_{idl})}{\mathbb{E}(\tau_{idl}) + \mathbb{E}(\tau_{act})} \tag{10}$$

where $\mathbb{E}(\tau_{idl})$ and $\mathbb{E}(\tau_{act})$ are the average times in the idle and active states, respectively. When the inter-arrival time between two successive packets at the gateway is lower than the threshold (i.e., $\tau \leq \eta$), the gateway will not enter the low power mode and therefore its average active time will be given by $\int_0^\eta \tau f_a(\tau|\tau \leq \eta) d\tau$. Otherwise (i.e., when $\tau > \eta$), the gateway will enter the low power mode after η seconds and its active time will be equal to η. Therefore, the average active time is given by:

$$\begin{aligned}\mathbb{E}(\tau_{act}) &= P(\tau \leq \eta) \int_0^\eta \tau f_a(\tau|\tau \leq \eta)\, d\tau + P(\tau > \eta)\, \eta \\ &= F_a(\eta) \frac{\int_0^\eta \tau f_a(\tau)\, d\tau}{F_a(\eta)} + [1 - F_a(\eta)]\, \eta \\ &= \int_0^\eta \tau f_a(\tau)\, d\tau + [1 - F_a(\eta)]\, \eta \end{aligned} \tag{11}$$

When the gateway enters the low power mode (i.e., when the next packet arrives later than η seconds after the last received packet), the average idle time will be given by the quantity $\int_\eta^\infty \tau f_a(\tau|\tau > \eta) d\tau$ minus the active time η. Therefore, the average idle time is given by:

$$\begin{aligned}\mathbb{E}(\tau_{idl}) &= P(\tau > \eta) \left(\int_\eta^\infty \tau f_a(\tau|\tau > \eta)\, d\tau - \eta \right) \\ &= [1 - F_a(\eta)] \left(\frac{\int_\eta^\infty \tau f_a(\tau)\, d\tau}{1 - F_a(\eta)} - \eta \right) \\ &= \int_\eta^\infty \tau f_a(\tau)\, d\tau - [1 - F_a(\eta)]\, \eta \qquad (12) \\ &= \int_\eta^\infty (\tau - \eta) f_a(\tau)\, d\tau \qquad (13) \end{aligned}$$

The denominator of (10) is obtained by adding (11) and (12), which is equal to $\int_0^\infty \tau f_a(\tau)\, d\tau$, i.e., the average inter-arrival time of the aggregated traffic stream at the gateway, which in turn is equal to m_s/n. Finally, the introduction of (13) in the numerator of (10) yields the result shown in (9).

References

[1] M. Laner, N. Nikaein, P. Svoboda, M. Popovic, D. Drajic, and S. Krco, "Traffic models for machine-to-machine (M2M) communications: Types and applications," in *Machine-to-machine communications: Architecture, performance and applications*. Woodhead, 2015, ch. 8, pp. 133–154.

[2] 3rd Generation Partnership Project (3GPP), "Technical Specification Group GSM/EDGE Radio Access Network; GERAN improvements for Machine-Type Communications (MTC) (Release 12)," Tech. Rep. 3GPP TR 43.868, Feb. 2014, v12.1.0.

[3] ——, "Technical Specification Group Radio Access Network; Study on RAN Improvements for Machine-type Communications (Release 11)," Tech. Rep. 3GPP TR 37.868, Sep. 2011, v11.0.0.

[4] D. P. Heyman and M. J. Sobel, *Stochastic models in operations research: Stochastic processes and operating characteristics*. Dover, 2003, vol. 1.

[5] T. Hoßfeld, F. Metzger, and P. E. Heegaard, "Traffic modeling for aggregated periodic IoT data," in *Proc. 21st IEEE Conf. Innovation in Clouds, Internet and Networks (ICIN 2018)*, Feb. 2018, pp. 1–8.

[6] M. López-Benítez, T. D. Drysdale, S. Hadfield, and M. I. Maricar, "Prototype for multidisciplinary research in the context of the Internet of Things," *J. Netw. & Comp. Appl.*, vol. 78, pp. 141–161, Jan. 2017.

[7] C. Majumdar, M. López-Benítez, and S. N. Merchant, "Experimental evaluation of the Poissoness of real sensor data traffic in the Internet of Things," in *Proc. 6th IEEE Annual Consumer Communications & Networking Conference (CCNC 2019)*, Jan. 2019, pp. 1–7.

[8] D. R. Cox and W. L. Smith, "On the superposition of renewal processes," *Biometrika*, vol. 41, no. 1/2, pp. 91–99, Jun. 1954.