Accurate Modelling of IoT Data Traffic Based on Weighted Sum of Distributions

Chitradeep Majumdar*, Miguel López-Benítez*† and S. N. Merchant‡
*Department of Electrical Engineering and Electronics, University of Liverpool, United Kingdom
†ARIES Research Centre, Antonio de Nebrija University, Spain
‡ Department of Electrical Engineering, IIT Bombay, India
email: Chitradeep.Majumdar@liverpool.ac.uk, M.Lopez-Benitez@liverpool.ac.uk, merchant@ee.iitb.ac.in

***Abstract*—This work proposes a novel mathematical approach to accurately model data traffic for the Internet of Things (IoT). Most of the conventional results on statistical data traffic models for IoT are based on the underlying assumption that the data generation follows standard Poisson or Exponential distribution which lacks experimental validation. However, in some of the use case applications a single statistical distribution is not adequate to provide the best fit for the inter-arrival time of the data packets generation. Based on the real data collected for over 10 weeks using our customized experimental IoT prototype for smart home application, in this paper we have established this very fact, citing barometric air pressure as an example. The statistical distribution of the inter-arrival time between the data packets for a specified barometric pressure fluctuation threshold is initially determined by approximating the best-fit with a set of standard classical distributions. The goodness-of-fit with the empirical data is numerically quantified using Kolmogorov-Smirnov (KS) Test. Furthermore, it is observed that any single standard distribution is unable to provide a good fit which is at least less than 10%. Therefore, a novel weighted distribution scheme is proposed that could provide an acceptable fit. The weighing factor including the location, scaling and weighing parameters of the best fitting distribution are estimated and analyzed. The distribution parameters are finally expressed as a function of the differential pressure value that can be used for different theoretical analysis and network optimization.**

***Index Terms*—KS-Kolmogorov Smirnov Test, IoT-Internet-of-Things, D2D-Device-to-Device, 3GPP, Traffic Modelling.**

I. Introduction

Next generation cellular communication standard under development like 5G or LTE-m is based on the fundamental premise that it has to support not just human centric communication but also machine type communications like IoT [1], [2]. This implies the future cellular standard must be empowered to handle high volume of data generated by these devices. At the same time it has to be designed taking into account factors like scalability, robustness, energy and spectrum efficiency, reliability and latency of the overall network. IoT is the state of the art paradigm for such short range machine type communication and extensive research is currently under way to achieve an efficient convergence of the IoT paradigm with emerging 5G standard [3], [4]. Novel protocol stacks like LoRaWAN are developed to be integrated with IoT applications targeted towards industry, the concept of Industrial-IoT (IIoT) [5].

This work focusses primarily on the estimation of accurate Data Traffic Modelling (DTM) of specific IoT application. An accurate prediction of the DTM is essential to develop an efficient and optimised network. A detailed classification of the data traffic types for machine-to-machine (M2M) communication including structured mathematical modelling of the data traffic is explained in [6]. Based on such traffic modelling, efficient Medium Access Control (MAC) schemes are proposed in [7] for random access satellite channel. In [8] two classes of M2M traffic generated from smart metering and vehicular applications are investigated where efficient data aggregation scheme is proposed to mitigate the signaling overhead. Machine learning based concepts like variant advanced forms of feature selection to segregate IoT sensor data from the multimedia data is proposed in [9], which is particulary very useful in the industrial environments. So far most of the proposed concepts are theoretical lacking experimental validation. To overcome this challenge, we selected an IoT based smart home application system as a user case scenario to investigate the key challenges associated with DTM based on real data. The motivation to select smart home application is due to its relatively simplistic implementation constrains and flexibility to work both with point source and aggregate source data traffic. Furthermore, there are few literatures available like [10], [11] where the authors have successfully addressed the challenges of convergence IoT based smart home applications with LTE-A and 5G standard. To this end we developed a customized integrated Raspberry Pi assisted IoT compliant smart home module and tried to thoroughly investigate the aspects of IoT data traffic modelling based on real data collected for over 10 weeks [12]. In the presence of a large number of IoT nodes and aggregate data traffic scenario, the Palm-Khintchine theorem holds true, which assumes that the data traffic is generated following a Poisson distribution [13]. However, in practical simple smart home scenario IoT nodes are limited while data generation for specific application is bursty and intermittent. Therefore, it is more realistic to model the data traffic as a point source and estimate the DTM based on this specific assumption. We have established in our previous work in [14] that the assumption of Poissoness for single source data is not valid. Detailed mathematical framework is proposed and numerically quantified in [14] to estimate the appropriate distributions for different applications based on real data collected using our home IoT prototype. However, for some special application like the reported barometric air pressure data it was not possible to find a single standard statistical distribution that could provide the best-fit. The main contribution of this paper is the real data dependant proposed mathematical framework based on sum of weighted distributions that could provide best-fit for applications where a single distribution is not adequate to capture and quantify the distribution of the

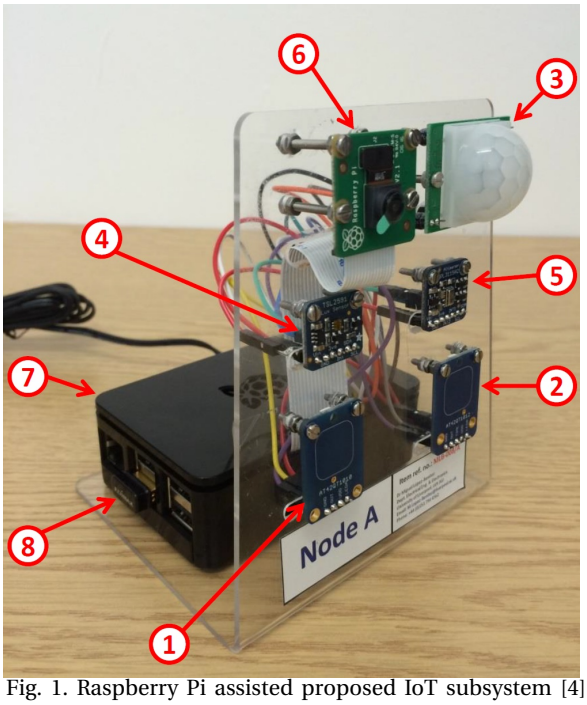

Fig. 1. Raspberry Pi assisted proposed IoT subsystem [4]

inter-arrival time of data packet generation.

The remainder of this paper is as follows. Section II contains the description of the experimental setup and the methodology. Section III describes the techniques used towards statistical fitting of the inter-arrival duration of the data packets for the differential air pressure data. Section IV contains the analysis of the empirical data and Section V contains the conclusion.

II. Experimental Platform and Methodology

A. Experimental Setup

Detailed description of the individual components of our custom made IoT smart home sensor node [12] is shown in Fig. 1. The sensor node is equipped with six sensor modules (Element 1 to 6). Element 7 labelled in Fig. 1 is the Raspberry Pi minicomputer to which all the sensor elements are connected. The Raspberry Pi module is connected to the central processing unit through USB WiFi adapter (Element 8). Details can be found in [12].

The data traffic analysis for the temperature, luminous intensity and motion sensor data is already investigated in our previous work [14]. Therefore, in this paper we investigate the outcome of the barometric pressure data obtained using Element 5 of the sensor node and propose a mathematical framework to determine the best-fitting distribution.

B. Differential Data Reporting scheme and Inter-arrival Time estimation

The data traffic generated by the IoT system is strongly dependant on the reporting strategy of the respective application. The reporting strategy can be periodic in nature [13] where the data packet is generated to be reported at every predetermined time interval. However, this work is based on a differential reporting strategy where the to be reported data packets are generated only when the sensed physical parameter fluctuates by a specified threshold denoted as ΔD. The elapsed time durations between the parameter fluctuation exceeding ΔD is a non-deterministic parameter and therefore a random variable. Therefore, the inter-arrival time between the reported data packet generation is a random variable. The key objective of this work is to determine the statistical distribution of this inter-arrival time which is a random variable because the data traffic is generated in accordance with this estimated distribution.

In this work we have adopted the strategy to record the parameter data at every 200 ms interval for over 10 weeks. The recorded data is stored at the CPU back-end. Subsequently during post processing, the time durations between the two parameter values whenever it exceeded the parameter threshold ΔD is determined and this process is repeated over the entire recorded data set to form a vector of the inter-arrival time X. With changing ΔD values every time a new vector X is generated for the overall analysis. Since the physical parameter in this case is the barometric air pressure therefore, ΔD is now renamed to be as ΔP. The ΔP is varied from 50 Pa to 400 Pa where Pa is Pascals (unit of air pressure).

III. Statistical fitting of the Inter-Arrival Time

The observed inter-arrival times from the captured experimental data were processed to calculate the ECDF and then it was used to fit the distribution models. The inter-arrival times are random variable as the time duration among the parameter fluctuations are non-deterministic. One of the most standard procedure to estimate the Empirical Cumulative Distribution Function (ECDF) of the random variable. It is the distribution related to the empirical measure of the sample. Upon estimation of the ECDF, next step is to determine the statistical distribution which would provide best fit to the ECDF. The shaping and scaling parameters of the best fitted distribution is subsequently determined using either Method-of-Moments (MoM) or Maximum Likelihood Estimation (MLE) technique. The best fit is quantified in terms of the Kolmogorov-Smirnov distance denoted as D_{KS} which is the maximum of the absolute difference value between the empirical CDF and the fitted statistical CDF estimated at each sample point and expressed as:

$$D_{KS} = \max_{X}\left\{abs\left(F_X^{ECDF}(x) - F_X^{fitted}(x)\right)\right\} \quad (1)$$

Seven standard statistical distributions are considered for our exhaustive investigation. They are exponential, generalized exponential, Pareto, generalized Pareto, gamma, lognormal and Weibull distributions as per [15].

IV. Modelling of the Empirical Barometric Pressure Data

Table I shows the KS distance for the distributions for change in atmospheric pressure parameter ΔP from 50 to 400 Pascals. At ΔP= 50 Pa, gamma distribution under MoM estimate gives the best fit with KS value of 0.0831. From ΔP= 100, 150 and 200 Pa, the generalized exponential distribution provides best fit with KS values at 0.1145, 0.1519 and 0.1670 respectively. At ΔP= 300 and 400 Pa, Weibull distribution provides best fit with KS values at 0.1900 and 0.1608. Under generalized-exponential distribution, it can

Table I. KS Distances of different statistical distributions- Differential Pascals Reporting (in pascals) for 10 weeks duration

Distributions	ΔP= 50 Pa		100 Pa		150 Pa		200 Pa		300 Pa		400 Pa	
	MoM	ML	MoM	ML	MoM	ML	MoM	ML	MoM	ML	MoM	ML
Exp	0.1039	0.1039	0.1312	0.1312	0.2008	0.2008	0.2737	0.2737	0.3561	0.3561	0.4449	0.4449
G.Exp	0.0862	***0.0862***	0.1145	***0.1145***	0.1519	***0.1519***	0.1670	***0.1670***	0.2250	***0.2250***	0.2353	***0.2353***
Pareto	0.8905	0.4162	0.8695	0.4304	0.8153	0.3806	0.7748	0.3230	0.7217	0.2716	0.6673	0.2343
G.Pareto	0.1019	0.1033	0.1263	0.1336	0.1948	0.2023	0.2595	0.2710	0.3407	0.3557	0.4273	0.4195
Log-N	0.1524	0.2446	0.1705	0.3122	0.2388	0.2933	0.3094	0.2531	0.3886	0.2169	0.4684	0.1817
Gamma	***0.0831***	0.0845	0.1154	0.1360	0.1556	0.1904	0.1749	0.2261	0.2323	0.2847	0.2492	0.3291
Weibull	0.0911	0.1437	0.1184	0.2113	0.1675	0.2291	0.2043	0.2040	0.2811	***0.1900***	0.3316	***0.1608***

be observed that at ΔP= 50, 300 and 400 Pa, the KS values are 0.0862, 0.2250 and 0.2353 marked underlined. Likewise previous cases, these KS distances are not too far from the minimum KS distance estimated for the specific ΔP values under gamma and Weibull distribution. Therefore, from Table I, generalized-exponential distribution could be considered as best fit for the change in atmospheric pressure data. However, it is clear that other than ΔP= 50 Pa where the KS distance under generalized exponential distribution is 0.0862, for all the rest ΔP values the KS distances are way above 10% which usually cannot be considered in general as a favourable fit.

Therefore, to address this challenge in this case a single distribution with different shaping and scaling parameter is considered and weighted by a factor less than 1. For example if $F_X(x,\lambda_1,\alpha_1)$ and $F_X(x,\lambda_2,\alpha_2)$ are two similar distributions with different parameters, then the best fitting CDF is expected to be achieved with the weighted combination of these two distributions with $(0<\gamma<1)$

$$F_X^{fit}(x_i) = \gamma F_X(x_i,\lambda_1,\alpha_1) + (1-\gamma) F_X(x_i,\lambda_2,\alpha_2). \quad (2)$$

Since generalized exponential distribution provided a comparatively better fit relative to other distributions over the range of ΔP values, in this work sums of weighted exponential and generalized-exponential distributions are considered. In case of exponential distribution only two scaling parameters λ_1 and λ_2 and the weighting factor γ are considered as exponential distribution does not have a shaping parameter. The optimal values of the parameters which would maximize the log-likelihood is considered as the best fit under that particular distribution. The probability distribution function is now also a function of the parameters λ_1 and λ_2 from [6]. The probability density function is now estimated to be as

$$f_X^{fit}(x_i) = \frac{\partial F_X(x_i,\lambda_1,\alpha_1,\lambda_2,\alpha_2)}{\partial x_i}. \quad (3)$$

Therefore, the Likelihood function becomes

$$L(x,\lambda_1,\lambda_2) = \prod_{i=1}^{N} \{\gamma f_X(x_i,\lambda_1,\alpha_1) + (1-\gamma) f_X(x_i,\lambda_2,\alpha_2)\}, \quad (4)$$

where N is the number of data points. Taking logarithm on both sides we get

$$\log L = \sum_{i=1}^{N} \log\{\gamma f_X(x_i,\lambda_1,\alpha_1) + (1-\gamma) f_X(x_i,\lambda_2,\alpha_2)\}. \quad (5)$$

We set the log-likelihood function as

$$\tilde{L} = \log L(x,\lambda_1,\alpha_1,\lambda_2,\alpha_2,\gamma) \quad (6)$$

Now our objective is to maximize this log likelihood function for the weighted distribution with respect to λ_1 and λ_2. This can be done analytically by solving for the $\frac{\partial \tilde{L}}{\lambda_1} = 0$ and $\frac{\partial \tilde{L}}{\lambda_2} = 0$ for a fixed value of γ. The process can be iterated over a range of γ values ranging between $0<\gamma\leq 1$. The γ and its corresponding optimal λ_1 and λ_2 are selected that would minimize the KS distance with the empirical CDF. However, it is worth to mention that this analytical treatment was possible in the simulations presented for the change in atmospheric pressure parameter ΔP because exponential and generalized exponential distribution provided a good fit. The estimation of the derivative of the Log-Likelihood function for the weighted exponential and generalized exponential distributions are shown in the Appendix of the paper. For other distributions, with increasing degrees of freedom and with the inclusion of the shaping parameter, the analytical calculation becomes highly non-trivial and it may not be possible to obtain a unique maxima for the log-likelihood function. In that case for the numerical analysis, Optimization toolbox from MATLAB can be used to solve the problem.

Upon numerical computation of the KS distances for different pressure values for the sum of weighted exponential distribution, the scaling parameters λ_1, λ_2 and the weighing parameter γ are estimated for all the values from 150 to 500 Pa with an increase of 5 Pa. These values are then interpolated with higher resolution. Subsequently, conventional curve fitting technique is applied using MATLAB curve fitting toolbox to obtain a generalized analytical expression for the parameters as a function of pressure fluctuation ΔP. From simulation, the best analytical fit for λ_1 and λ_2 turns out be generalized Gaussian which is simply sum of weighted Gaussian exponentials while for the weighing parameter γ is Fourier series. Based on these analytically calculated parameter values, the sum of weighted exponential distribution over inter-arrival times X is estimated for a specific ΔP. The KS distance is then measured with respect to the empirical CDF of the X for that given value of ΔP. The values of the scaling and the weighing parameters are expressed as

$$\lambda(\Delta P) = \sum_{i=1}^{K} a_i e^{-\left(\frac{\Delta P - b_i}{c_i}\right)^2} \quad (7)$$

$$\gamma(\Delta P) = a_0 + \sum_{i=1}^{K} \{a_i \cos(i.w) + b_i \sin(i.w)\}, \quad (8)$$

where a_0= 0.3214 and w= 0.01421 are parameters for the Fourier fitting and K = 8 is the number of terms.

In Fig. 2(a) and Fig. 2(b), two of the scaling parameters λ_1 and λ_2 are shown for the weighted exponential distribution for ΔP= 160 to 500 Pa. Fig. 2(c), shows the weighing parameter γ obtained numerically and analytically. Fig. 3 shows the KS distances with the empirical CDF estimated using the shaping parameters obtained numerically and analytically. It could be observed that using sum of weighted exponential distribution for the atmospheric pressure change parameter ΔP, the KS values are less than 10% for almost all throughout the range of ΔP. This outperforms the performance as quantified in the Table I with normal generalized-exponential distribution which is way over 10%. Both scaling parameters are analytically determined in terms of ΔP which follows (7). Table II(a) provides the value of the coefficients for λ_1. However, while estimating λ_2, (7) is required to be scaled by a factor of 10^{-5}. γ could be obtained using (8) in order to have the analytical fit with the numerical λ_2 values. Table II(b) provides the value of the coefficients for λ_2. For the weighing parameter (γ), the coefficients values are shown in Table II(c).

Interestingly, it is intuitive that with the increase in the degree of freedom of the fitting distribution, there could be improvement in the fit. Therefore, the case of weighted exponential distribution is now extended to the case of weighted generalized exponential distribution. As shown in (15), the weighted generalized exponential distribution has five degrees of freedom that includes the two scaling and shaping factors λ_1, λ_2, α_1 and α_2 along with the weighting factor γ. As it could be seen in Fig. 4, the KS distance values obtained numerically for ΔP ranging from 160 to 500 Pa is less than 5% which is considered to be a very good fit. With increase in the degree of freedom, the sensitivity of curve fitting to obtain the analytical expression for the parameters increases. Therefore, marginal fitting error in any of the parameter usually could affect the final KS distance estimation. This is the reason when the parameters for the sum of weighted generalized exponential distribution is estimated analytically and used to calculate the KS distance, it is although lesser than 10% but it is more than 5% for all the ΔP values as shown in Fig. 4.

Both in case of weighted exponential and weighted generalized exponential distributions, the location parameter (μ) for all values of the reported ΔP is interestingly same and equals 0.2 corresponding to 200 ms which is the minimum inter-arrival time. This is due to the fact that in general the normal atmospheric pressure at mean sea level is 1 atm ≈ 105 kPa. For stationary indoor condition the rate of air pressure fluctuation is relatively slower. Therefore considering the time window of our entire data recording which is for 10 weeks, fluctuation parameter threshold ΔP in the order of 1 atm or higher Pascal values was unavailable. Within the scope of available data, analysis of ΔP within an estimated range of 150 to 500 Pa was feasible and for every ΔP value there has been at least one single or multiple instances where the corresponding parameter variation occurred within the minimum data sampling interval.

Table II. Pressure-Scaling Parameter fitting parameter coefficients with weighted exponential distribution

(a) λ_1 parameter

$\mathbf{a_i}$	$\mathbf{b_i}$	$\mathbf{c_i}$
1.867	110.1	3.279
0.496	125.4	11.94
0.1904	155.8	3.046
0.2131	187.2	7.169
0.3252	166.2	17.06
2.816	-712.6	542.3
-0.2017	162.4	51.97
-0.3747	47.13	84.72

(b) λ_2 parameter

$\mathbf{a_i}$	$\mathbf{b_i}$	$\mathbf{c_i}$
2.466	128.3	80.68
0.09805	190.2	2.401
0.1841	220.9	9.798
0.425	251.8	20.27
0.3819	288.6	17.57
0.1497	298.7	4.598
-0.1626	275.9	47.57
3.019	157.9	439.9

(c) γ parameter

$\mathbf{a_i}$	$\mathbf{b_i}$
0.1114	-0.08738
0.06726	0.07031
-0.05857	0.08055
-0.04354	-0.02673
0.01242	-0.04226
0.02905	0.02265
-0.01407	0.0304
-0.01835	-0.01453

V. Conclusion

In this work we have established the fact that for few specific IoT applications the distribution of the data traffic generated by the single source cannot be modelled using a single statistical distribution. For similar scenarios we have proposed a comprehensive mathematical framework to determine the distribution based on sum of weighted single distributions. The parameters of the weighted distribution is determined as expressed as a function of the reported parameter difference threshold. Such analysis could be useful in case of applications where the reported physical parameter fluctuation is highly sensitive and required to be quantified at greater order of resolution. The distribution of the inter-arrival time of data generation could be used to optimise the cellular network that envisages to support large number of similar IoT based applications.

VI. Appendix

A. Likelihood maximization for the weighted exponential distribution

From the case of weighted exponential distribution, from (2) and (3)

$$F_E(x_i, \lambda_1, \lambda_2, \mu, \gamma) = \gamma\left(1 - e^{-\lambda_1(x_i-\mu)}\right) + (1-\gamma)\left(1 - e^{-\lambda_2(x_i-\mu)}\right) \quad (9)$$

Differentiating the CDF $F_E(x_i, \lambda_1, \lambda_2, \mu)$ with respect to x_i, we get the Probability Density Function (PDF)

$$f_E(x_i, \lambda_1, \lambda_2, \mu, \gamma) = \gamma\lambda_1 e^{-\lambda_1(x_i-\mu)} + (1-\gamma)\lambda_2 e^{-\lambda_2(x_i-\mu)} \quad (10)$$

The log-likelihood function turns out to be

$$\tilde{L} = \sum_{i=1}^{N} \log\left\{\gamma\lambda_1 e^{-\lambda_1(x_i-\mu)} + (1-\gamma)\lambda_2 e^{-\lambda_2(x_i-\mu)}\right\}, \quad (11)$$

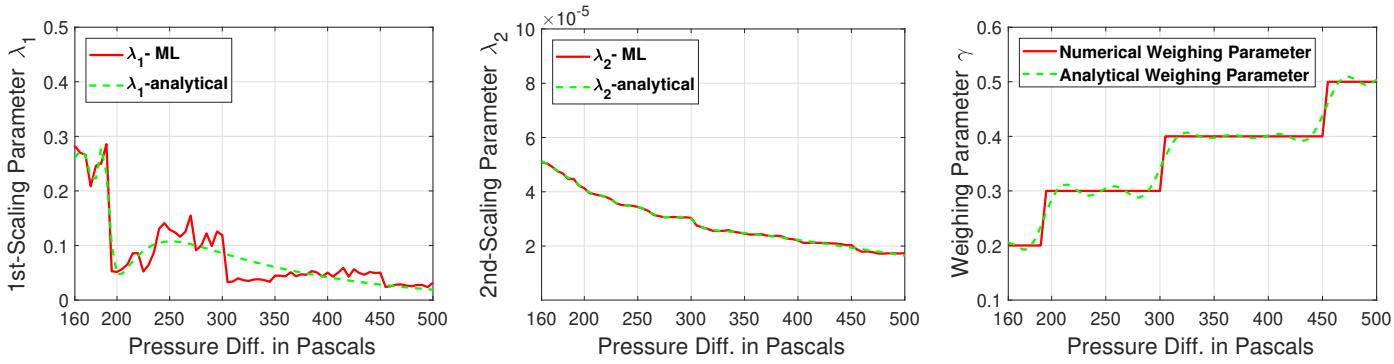

Fig. 2. a. Scaling Parameter-1 (λ_1) vs pressure difference b. scaling Parameter-2 (λ_2) vs pressure difference c. numerically and analytically computed weighing parameter (γ) vs Pressure Difference, all for weighted exponential distribution

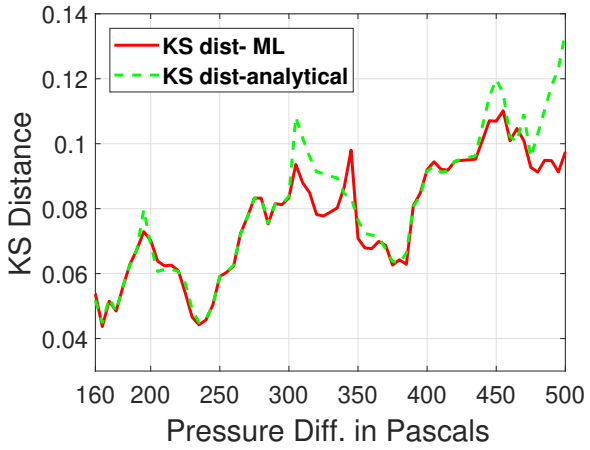

Fig. 3. Numerically and analytically computed KS distances vs reported pressure difference (ΔP), for weighted exponential distribution

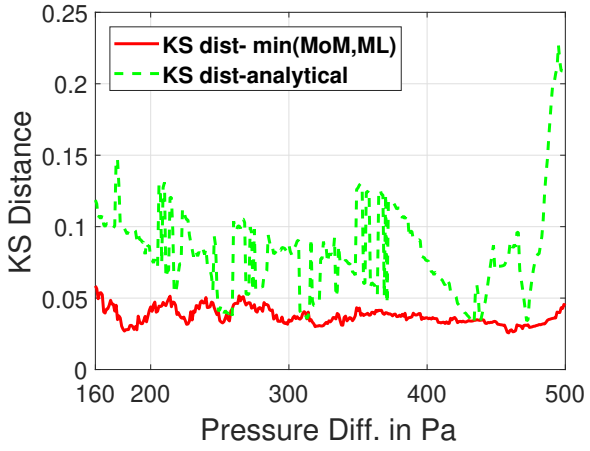

Fig. 4. Numerically and analytically computed KS distances vs reported pressure difference (ΔP), for weighted generalized exponential distribution

where N is the total number of inter-arrival durations for a specific parameter input difference ΔP.

In order to maximize the Log-Likelihood function, $\tilde{L}$ is differentiated with respect to the two scaling parameters λ_1 and λ_2 and their partial derivatives are set to zero to find the solution. Taking partial derivative of (11) we obtain,

$$\frac{\partial \tilde{L}}{\partial \lambda_1} = \sum_{i=1}^{N} \frac{\gamma e^{-\lambda_1 (x_i-\mu)} - \gamma \lambda_1 (x_i-\mu) e^{-\lambda_1 (x_i-\mu)}}{\{\gamma \lambda_1 e^{-\lambda_1 (x_i-\mu)} + (1-\gamma)\lambda_2 e^{-\lambda_2 (x_i-\mu)}\}}, \tag{12}$$

$$\frac{\partial \tilde{L}}{\partial \lambda_2} = \sum_{i=1}^{N} \frac{(1-\gamma) e^{-\lambda_2 (x_i-\mu)} - (1-\gamma)\lambda_2 (x_i-\mu) e^{-\lambda_2 (x_i-\mu)}}{\{\gamma \lambda_1 e^{-\lambda_1 (x_i-\mu)} + (1-\gamma)\lambda_2 e^{-\lambda_2 (x_i-\mu)}\}}. \tag{13}$$

As explained in the earlier Section IV, the weighting parameter γ can be treated both as discrete ranging between $\gamma = \{0, 0.1, 0.2, ...0.9\}$ and continuous value ranging between $\gamma = [0, 1]$. In discrete case an exhaustive search technique can be applied and for each value of γ, it corresponding optimal λ_1 and λ_2 can be estimated which would maximize the Log-Likelihood function. However, in case of γ as continuous it can be included as one of the decision variable along with λ_1 and λ_2 and the corresponding maxima of the Log-Likelihood function can be determined. In that case,

$$\frac{\partial \tilde{L}}{\partial \gamma} = \sum_{i=1}^{N} \frac{\left\{\lambda_1 e^{-\lambda_1 (x_i-\mu)} - \gamma \lambda_2 e^{-\lambda_2 (x_i-\mu)}\right\}}{\{\gamma \lambda_1 e^{-\lambda_1 (x_i-\mu)} + (1-\gamma)\lambda_2 e^{-\lambda_2 (x_i-\mu)}\}} \tag{14}$$

The maximum of the Log-Likelihood function $\tilde{L}$ can now be easily obtained by solving for $\frac{\partial \tilde{L}}{\partial \lambda_1} = 0$, $\frac{\partial \tilde{L}}{\partial \lambda_2} = 0$ and $\frac{\partial \tilde{L}}{\partial \gamma} = 0$.

B. Likelihood maximization for the weighted generalized-exponential distribution

The CDF of the weighted generalized exponential distribution

$$F_{GE} = \gamma \left[1 - e^{-\lambda_1 (x_i-\mu)}\right]^{\alpha_1} + (1-\gamma)\left[1 - e^{-\lambda_2 (x_i-\mu)}\right]^{\alpha_2}. \tag{15}$$

The PDF is given by

$$f_{GE} = \gamma \lambda_1 \alpha_1 e^{-\lambda_1 (x_i-\mu)}\left[1 - e^{-\lambda_1 (x_i-\mu)}\right]^{\alpha_1 - 1} + (1-\gamma)\lambda_2 \alpha_2 e^{-\lambda_2 (x_i-\mu)}\left[1 - e^{-\lambda_2 (x_i-\mu)}\right]^{\alpha_2 - 1} \tag{16}$$

The Log-Likelihood Function is given by

$$\tilde{L} = \sum_{i=1}^{N} \log\Big\{\gamma\lambda_1\alpha_1 e^{-\lambda_1(x_i-\mu)}\left[1-e^{-\lambda_1(x_i-\mu)}\right]^{\alpha_1-1} + \big(1-\gamma\big)\lambda_2\alpha_2 e^{-\lambda_2(x_i-\mu)}\left[1-e^{-\lambda_2(x_i-\mu)}\right]^{\alpha_2-1}\Big\} \quad (17)$$

Differentiating $\tilde{L}$ with respect to λ_1, λ_2, α_1 and α_2 we get,

$$\frac{\partial\tilde{L}}{\partial\lambda_1} = \sum_{i=1}^{N}\frac{\gamma\alpha_1\left(\zeta_i^{\lambda_1(1)}+\zeta_i^{\lambda_1(2)}\right)}{D_i}, \quad (18)$$

where $\zeta_i^{\lambda_1(1)}$, $\zeta_i^{\lambda_1(2)}$ and D_i are dummy variables as shown in the numerator and denominator of (17) calculated to be as

$$\zeta_i^{\lambda_1(1)} = \left[1-e^{-\lambda_1(x_i-\mu)}\right]^{(\alpha_1-1)}\left\{e^{-\lambda_1(x_i-\mu)} - \lambda_1\left(x_i-\mu\right)e^{-\lambda_1(x_i-\mu)}\right\}, \quad (19)$$

$$\zeta_i^{\lambda_1(2)} = \lambda_1 e^{-\lambda_1(x_i-\mu)}\left\{(\alpha_1-1)\left[1-e^{-\lambda_1(x_i-\mu)}\right]^{\alpha_1-2}\left(x_i-\mu\right)e^{-\lambda_1(x_i-\mu)}\right\} \quad (20)$$

and

$$D_i = \Big\{\gamma\lambda_1\alpha_1 e^{-\lambda_1(x_i-\mu)}\left[1-e^{-\lambda_1(x_i-\mu)}\right]^{\alpha_1-1} + \big(1-\gamma\big)\lambda_2\alpha_2 e^{-\lambda_2(x_i-\mu)}\left[1-e^{-\lambda_2(x_i-\mu)}\right]^{\alpha_2-1}\Big\}. \quad (21)$$

Similarly, for the shaping parameter α_1, the derivative turns out to be,

$$\frac{\partial\tilde{L}}{\partial\alpha_1} = \sum_{i=1}^{N}\frac{\gamma\lambda_1 e^{-\lambda_1(x_i-\mu)}\left(\epsilon_i^{\alpha_1(1)}+\epsilon_i^{\alpha_2(2)}\right)}{D_i}, \quad (22)$$

where

$$\epsilon_i^{\alpha_1(1)} = \left[1-e^{-\lambda_1(x_i-\mu)}\right]^{\alpha_1-1} \quad (23)$$

$$\epsilon_i^{\alpha_1(2)} = \alpha_1\left[1-e^{-\lambda_1(x_i-\mu)}\right]^{\alpha_1-1}\log\left\{1-e^{-\lambda_1(x_i-\mu)}\right\}. \quad (24)$$

Based on the formulations as shown in (18) and (22), $\frac{\partial\tilde{L}}{\partial\lambda_2}$ and $\frac{\partial\tilde{L}}{\partial\alpha_2}$ can be easily calculated just by replacing γ as $\big(1-\gamma\big)$, λ_1 as λ_2 and α_1 as α_2 respectively. Once the partial derivatives of the Log-Likelihood function with respect to all the scaling (λ_1,λ_2) and shaping (α_1,α_2) parameters are obtained, it can be equated to zero to find the values of the parameters which would maximize the log-likelihood function.

VII. Acknowledgement

This work was jointly funded by the Royal Society of the United Kingdom and the Science and Engineering Research Board (SERB) of India under a Royal Society-SERB Newton International Fellowship (grant reference no. NF170943).

References

[1] A. Rico-Alvarino et al., "An overview of 3GPP enhancements on machine to machine communications," *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 14-21, Jun. 2016.

[2] S. Andreev et al., "Understanding the IoT connectivity landscape: a contemporary M2M radio technology roadmap," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 32-40, Sep. 2015.

[3] M. R. Palattella et al., "Internet of Things in the 5G Era: Enablers, Architecture, and Business Models,"*IEEE J. Sel. Areas in Commun.*, vol. 34, no. 3, pp. 510-527, Mar. 2016.

[4] D. Zhang, Z. Zhou, S. Mumtaz, J. Rodriguez and T. Sato, "One Integrated Energy Efficiency Proposal for 5G IoT Communications," *IEEE Internet of Things J.*, vol. 3, no. 6, pp. 1346-1354, Dec. 2016.

[5] J. Navarro-Ortiz, S. Sendra, P. Ameigeiras and J. M. Lopez-Soler, "Integration of LoRaWAN and 4G/5G for the Industrial Internet of Things," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 60-67, Feb. 2018.

[6] 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," edited by M. Dolher and C. Antón-Haro, Woodhead Publishing, 2015.

[7] M. Bacco, P. Cassarà, M. Colucci and A. Gotta, "Modeling Reliable M2M/IoT Traffic Over Random Access Satellite Links in Non-Saturated Conditions," *IEEE J. Sel. Areas in Commun.*, vol. 36, no. 5, pp. 1042-1051, May 2018.

[8] N. Kouzayha, M. Jaber and Z. Dawy, "Measurement-Based Signaling Management Strategies for Cellular IoT," *IEEE Internet of Things J.*, vol. 4, no. 5, pp. 1434-1444, Oct. 2017.

[9] S. Egea, A. Rego Mañez, B. Carro, A. Sánchez-Esguevillas and J. Lloret, "Intelligent IoT Traffic Classification Using Novel Search Strategy for Fast-Based-Correlation Feature Selection in Industrial Environments," *IEEE Internet of Things J.*, vol. 5, no. 3, pp. 1616-1624, Jun. 2018.

[10] A. Javed, H. Larijani and A. Wixted, "Improving Energy Consumption of a Commercial Building with IoT and Machine Learning," IT Professional, vol. 20, no. 5, pp. 30-38, Oct. 2018.

[11] C. Vallati, A. Virdis, E. Mingozzi and G. Stea, "Mobile-Edge Computing Come Home Connecting things in future smart homes using LTE device-to-device communications," *IEEE Consumer Elec. Mag.*, vol. 5, no. 4, pp. 77-83, Oct. 2016.

[12] 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. and Comp. Appl.*, vol. 78, pp. 146-161, Jan. 2017.

[13] T. Hoßfeld, F. Metzger and P. E. Heegaard, "Traffic modeling for aggregated periodic IoT data," in Proc. *Innov. in Clouds, Internet and Netw. and Workshops (ICIN)*, pp. 1-8, Dec. 2018.

[14] 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. *IEEE Consumer Commun. and Netw. Conf.(CCNC)*, pp. 1-7, Jan. 2019.

[15] M. López-Benítez and F. Casadevall, "Time-Dimension Models of Spectrum Usage for the Analysis, Design, and Simulation of Cognitive Radio Networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2091-2104, Jun. 2013.

[16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Numerical Recipes: The Art of Scientific Computing", 3rd ed. Cambridge, U.K.: Cambridge Univ. Press, 2007