Mitigating the Massive Access Problem in the Internet of Things

Erol Gelenbe[1,2,3,4](✉), Mert Nakıp[1], Dariusz Marek[5], and Tadeusz Czachorski[1]

[1] Institute of Theoretical and Applied Informatics Polish Academy of Sciences, 44-100 Gliwice, Poland
{seg,mnakip,tadek}@iitis.pl
[2] Yaşar University, Bornova, İzmir, Turkey
[3] CNRS Lab. I3S Université, Côte d'Azur, Nice, France
[4] CNRS Lab. Abraham de Moivre, London, UK
[5] Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, 44-100 Gliwice, Poland
dariusz.marek@polsl.pl

Abstract. The traffic from the large number of IoT devices connected to the IoT is a source of congestion known as the Massive Access Problem (MAP), that results in packet losses, delays and missed deadlines for real-time data. This paper reviews the literature on MAP and summarizes recent results on two approaches that have been designed to mitigate MAP. One approach is based on randomizing the packet arrival instants to IoT gateways within a given time interval that is chosen so that packet arrivals do not exceed their deadlines, but also so that they do not constitute bulk arrivals. The second approach is a novel traffic shaping policy named the Quasi-Deterministic-Transmission-Policy (QDTP) which has been proved to drastically reduce queue formation at the receiving gateway by delaying packet departures from the IoT devices in a judicious manner. Both analytical and experimental results are summarized, that describe the benefits of these techniques.

Keywords: Internet of Things (IoT) · IoT Gateways · Massive Access Problem · Diffusion approximation · Trace driven simulations · Quasi-Deterministic Transmission Policy

1 Introduction

The number of Internet of Things (IoT) devices is increasing rapidly with the increasing needs of smart cities, healthcare applications, autonomous systems, and smart vehicles [6,7,12,40], causing the overload of communication channels and gateways [27]. This results in the Massive Access Problem (MAP) where high latency and queue lengths can lead to packet loss and deadline violations. In addition, congestion can lead to increased energy consumption at IoT devices and gateways due to repeated requests for access and increased processing times

E. Gelenbe et al. (Eds.): EuroCybersec 2021, CCIS 1596, pp. 118–132, 2022.
https://doi.org/10.1007/978-3-031-09357-9_10

[2,15], thus contributing to the worldwide increase in energy consumption for ICT [14].

Thus, substantial work over the last several years [3,11,28–30,34,35,39,46–51,55,56] attempts to solve MAP in various ways.

In this paper, we first briefly review methodologies and results with regard to reactive or proactive (predictive) solutions that can mitigate MAP. Then, we summarize two recent research avenues: Randomization of Transmission Times [36] and a novel traffic shaping policy – the Quasi-Deterministic-Transmission-Policy (QDTP) [18,19,25]. We illustrate the results these approaches offer via analytical techniques and trace driven simulations using a publicly available dataset of up to 6400 IoT devices [1] with different deadline constraints.

The remainder of this paper is organized as follows. Section 2 reviews recent research focusing on MAP. Section 3 summarizes two recent studies on MAP based on analytical and experimental results. Finally in Sect. 4 our main conclusions are presented.

2 Review of Prior Work on MAP

This section reviews the prior work on MAP in two categories as reactive solutions and predictive/proactive solutions. Early research addressed MAP by reducing congestion through adaptive Random Re-Routing (RRR) [20–22] which improves the QoS of a sensor network by dynamically changing packet routes when congestion is detected. In related work [42], an information theoretic technique selectively reduces the amount of traffic by increasing transmission efficiency, and improves the QoS in sensor networks.

More recent work has proposed solutions to MAP, assuming that IoT traffic is generated at random, and using approaches mostly based on Access Class Barring (ACB). In [35] ACB is enhanced by using Markov chains to model the status of preambles and to forecast active devices, while other work [30] developed the recursive ACB algorithm based on instantaneous detection of idle preambles. Recent work [28] also developed recursive ACB which adapts the probability that a device sends a preamble based on estimating the number of active IoT devices. The performance of ACB has been analyzed under different parameters for Machine-to-Machine (M2M) communications [55], and enhanced by Reinforcement Learning (RL) to select its parameter (i.e. barring rate) with respect to network conditions [56]. In [29] deep RL techniques are proposed to maximize the number of devices that successfully access the medium without collisions.

In [34] an access scheme for M2M communication that clusters machine-type devices according to their requirements and locations is used. The Non-Orthogonal Multiple Access (NOMA) based technique is presented for networks with a massive number of devices in [51]. Moreover, in order to address MAP, in [3] a collision detection based random access technique is developed, and in [49] a hybrid technique that combines slotted-Carrier-Sense Multiple Access with Collision Avoidance (CSMA/CA) and Time-Division Multiple Access (TDMA) schemes are suggested.

2.1 Proactive Solutions

IoT devices are quite simple and cannot easily be coordinated for timing or scheduling [24] based on distributed control [8,10]. Thus recent experimental results [37,38] show that machine learning techniques can be used to predict IoT traffic generated by individual devices, and other work [11,39,46–48,50] designs proactive/predictive access schemes that determine the transmission times from IoT devices based on such predictions to mitigate the MAP.

Other work [31,45,48] has developed proactive access schemes for Machine-to-Human (M2H) or Human-to-Machine (H2M) traffic, while earlier research [43,44] has focused on Human-to-Human (H2H) traffic. In [43] a schedule-based protocol an expert system is used to determine schedules that minimize delay and maximize channel utilization. In [44], forecasts of data rates of individual applications are used to schedule channel scheduling, and network load has also been balanced based on the forecast of the total load of all machine-type devices [31]. To mitigate the latency bottleneck due to the contention in optical or wireless networks [48], the proactive allocation of bandwidth to transmissions of packet bursts based on Artificial Neural Network (ANN) forecasts is studied, while in [45] the prediction of network throughput for better Quality of Experience (QoE) is investigated.

Fast Uplink Grant (FUG) is one of the predictive access schemes presented in 3GPP Release 14 to provide predetermined uplink allocations for IoT devices [4,5,52]. In [50], IoT packet transmissions have been modelled with a binary Markov process, while in [11] a FUG allocation technique was developed by combining Support Vector Machines (SVM) and Long-Short Term Memory (LSTM) neural networks.

Another trend [39,46,47] on predictive access schemes addresses MAP by using Joint Forecasting Scheduling (JFS), and in [39] JFS was proposed to schedule transmissions based on forecasts of generation times and sizes of bursts. JFS can be recursively enhanced with a Multi-Scale Algorithm (MSA), where the performance of JFS and the length of scheduling horizon, are significantly increased [46] (over 96% throughput) for 6400 devices with variable latency constraints. However, the computational requirements of MSA are very high and in [47] a scheduling heuristic is used to determine JFS transmission times for using multiple frequency channels.

Most recently the Randomization of Generation Times (RGT) preprocessing algorithm [36] is shown to significantly improve the performance of scheduling heuristics with very low computational cost for large numbers of IoT devices. Also, the Quasi-Deterministic Transmission Policy (QDTP) traffic shaping approach [19,25] has been shown, using queueing theory, diffusion approximations [16] and trace driven simulations, to mitigate the MAP by drastically reducing the waiting time at gateways. This research has shown that RGT and QDTP can mitigate MAP with very low computational requirements for up to 6400 IoT devices. In the remainder of this article, we shall outline how these two avenues of recent research can alleviate MAP.

3 Mitigating MAP Using Queueing Theory and Diffusion Approximations

We will summarize together the results of two recent articles [19,36], which offer solutions to reduce MAP. To this end, we first present the analysis in [19] of the probability that the deadline of an IoT packet is met, providing a basis for access policies in IoT networks. Then we review the RGT [36] and QDTP [19] algorithms and their performance.

In [19], the collection of IoT devices that generate packets, the communication channel, and the receiving gateway are represented as two cascaded queues [9, 32,41,53]:

- The first queue translates the generation instant at the IoT device for the j-th packet r_j, into its transmission instant t_j, and
- The second queue starts with the transmission instants t_j that feeds directly into the IoT gateway where the j-th packet is served in FIFO (First-In-First-Out) order with a service duration p_j.
- Note that in this case, the transmission delay within the communication channel is taken to be zero, i.e. it is assumed to be small enough to be as compared to p_j and to the durations between the other successive instants, so that the packet leaving the IoT device at time t_j arrives at the gateway at the same time instant.

Let the IoT device generats traffic in bursts of bits that are sent at the same time instant, where burst j is generated at discrete time slot r_j and should be recieved by d_j. That is, we assume that there is a deadline Δ_j for each burst j beyond which j is of "no value". Thus, the burst j (which can also be considered as a packet) must arrive at the receiver gateway by $d_j = r_j + \Delta_j$. Furthermore, the packets of various IoT devices are processed in time ordered fashion in First-In-First-Out (FIFO) order, and p_j be the "service time" during which the receiving gateway is occupied by packet j.

3.1 The Probability of Meeting Deadlines

Let the j-th packet sent from any of the IoT devices, enumerated in time order (i.e. the $j - th$ packet is generated before the $j + 1$-th packet), be transmitted from its IoT device exactly at the instant r_j when it is generated. Then the time spent between the generation time r_j and the time when it starts being processed at the gateway, i.e. its total waiting time, is denoted by W_j, and is given by Lindley's recursive equation [32,53]:

$$W_{j+1} = [W_j + p_j - r_{j+1} + r_j]^+ \ , \ j = 0, 1, 2, \ . \tag{1}$$

where $r_0 = 0$. Note that the conventional notation $[X]^+$, for a real number X, means that $[X]^+ = 0$ when $X < 0$, and $[X]^+ = X$ if $X \geq 0$.

Assuming that the generation times coincide with the transmission times, if the sequence generation and service times, and deadlines $\{r_j,\ p_j,\ \Delta_j\}_{j\geq 0}$

constitute a stationary random process, the probability Π_j that the packet j does not meet its deadline is given by:

$$\Pi_j = Prob[R_j = W_j + p_j > r_j + \Delta_j],\ and\ \Pi = \lim_{j\to\infty} \Pi_j, \tag{2}$$

where R_j is known as the response time.

Since the focus of the work in [19,36] is on selecting the transmission instant of each packet j, denoted by t_j, to minimize Π_j under each traffic load of the network, W_j will be replaced by a total end-to-end delay V_j to each packet, where D_j is a scheduling delay imposed to each successive packet, and V_j includes the delay at the IoT device plus the delay at the gateway:

$$V_j = W_j^a + W_j^b,\ j \geq 0,\ V_0 = W_0^a = W_0^b = 0,\ and \tag{3}$$

$$W_{j+1}^a = [W_j^a + D_j - (r_{j+1} - r_j)]^+,\ t_j = r_j + W_j^a, \tag{4}$$

$$W_{j+1}^b = [W_j^b + p_j - (t_{j+1} - t_j)]^+. \tag{5}$$

In other words, we impose an initial scheduling delay D_j to each successive packet, and then consider the resulting effect on the transmission instant t_j and on the resulting delay at the IoT device W_j^a followed by the delay at the gateway W_j^b. These matters are analyzed in detail in [25], wth a resulting effect on the probability of missing the deadlines:

$$\Pi_j^* = Prob[R_j^* = V_j + p_j > r_j + \Delta_j],\ and\ \Pi^* = \lim_{j\to\infty} \Pi_j, \tag{6}$$

when $\{r_j,\ p_j,\ D_j,\ \Delta_j\}_{j\geq 0}$ constitute a stationary random process.

3.2 Interarrival and Service Time Statistics

Next, in order to compute the interarrival and service time statistics, the approach in [19] assumes that the p_j's of all packets are independent random variables with the same distribution whose mean is $E[P]$ and its SCV is C_B^2. Moreover, λ denotes the interarrival rate of packets, such that $\lambda = E[r_{j+1} - r_j])^{-1}$. It has also been assumed that the value of λ increases with the number of IoT devices M that are connected to the IoT gateway, and the system will operate under variable λ (or M) but under stable conditions, i.e. $\lambda E[P] < 1$. Also, let C_A^2 denote the SCV of interarrival times of packets. Then, C_B^2 and C_A^2 are respectively defined as

$$C_B^2 = \frac{E[P^2]}{(E[P])^2} - 1,\ \text{and}\ C_A^2 = \frac{E[(r_{j+1} - r_j)^2]}{(E[r_{j+1} - r_j])^2} - 1. \tag{7}$$

3.3 Using the Diffusion Approximation:

Last, the diffusion approximation [13,33] has been used in order to determine the probability that $R_j \leq \Delta_j$, denoted by $F_R(\Delta)$, where $\Delta_j = \Delta$ which is constant for all packets of all IoT devices. Then, the probability of missing deadline $\Pi = 1 - F_R(\Delta)$.

Subsequently, using the diffusion approximations [19] one obtains the probability density function of the response time and the probability of missing deadline as

$$f_R(t) = \int_0^\infty \frac{x}{\sqrt{2\pi\alpha t^3}} e^{-\frac{(x+\beta t)^2}{2\alpha t}} f(x)dx, \text{ then } \Pi = 1 - \int_0^\Delta f_R(\tau)d\tau, \quad (8)$$

where $\beta = \lambda - \mu$, $\alpha = \lambda C_A^2 + \mu C_B^2$, $\mu = 1/E[P]$, and $f(x)$ is yielded by the diffusion model [13].

3.4 Numerical Results Concerning the Diffusion Analysis

Now, the results concerning a publicly available IoT traffic dataset [1] present how Π varies with C_A^2 and Δ.

First, Fig. 1 displays $log_{10}(\Pi)$ for $\lambda = 0.8$ and $C_B^2 = 1$ and different values of Δ and C_A^2. Note that the minimum and maximum values of C_A^2 for the traffic in the dataset [1] which are 1.6 and 2.18, and the approximate value of C_A^2 for the uniform distribution which will be used for the randomization policy in Sect. 3.5, are shown as vertical bars. The results in this figure show that Π increases with C_A^2 and decreases with Δ when λ remains constant.

Then, Fig. 2 presents the values of Π against increasing number of devices M as well as the corresponding λ and C_A^2. The results show that the measured value of Π, which is the fraction of packets that do not meet their deadlines, increases as M or corresponding λ increases.

In summary, the results in Fig. 1 and Fig. 2 from [19] show that reducing C_A^2 significantly increases the probability that any IoT traffic packet meets its deadline.

3.5 Randomization of Data Generation Times (RGT)

While the results in Sect. 2 show that fast and computationally inexpensive heuristic algorithms are promising for MAP, the work in [36] develops the RGT preprocessing algorithm which relieves the traffic load by distributing the generation times of packets over a scheduling window of duration of T_{sch}.

In RGT, r_j is updated by adding to it an offset which is a uniformly distributed random variable (Recall that $C_A^2 = 1/3$ for uniform distribution which is shown to be significantly lower than the minimum value in the dataset) as

$$r_j^{new} \leftarrow r_j + U[\Delta_j - S_j] \quad (9)$$

where S_j is a *safety delay* that limits the upper bound of r_j^{new} such that $0 \leq S_j \leq \Delta_j$ which indicates the maximum randomization and $S_j = \Delta_j$ indicates that there is no randomization.

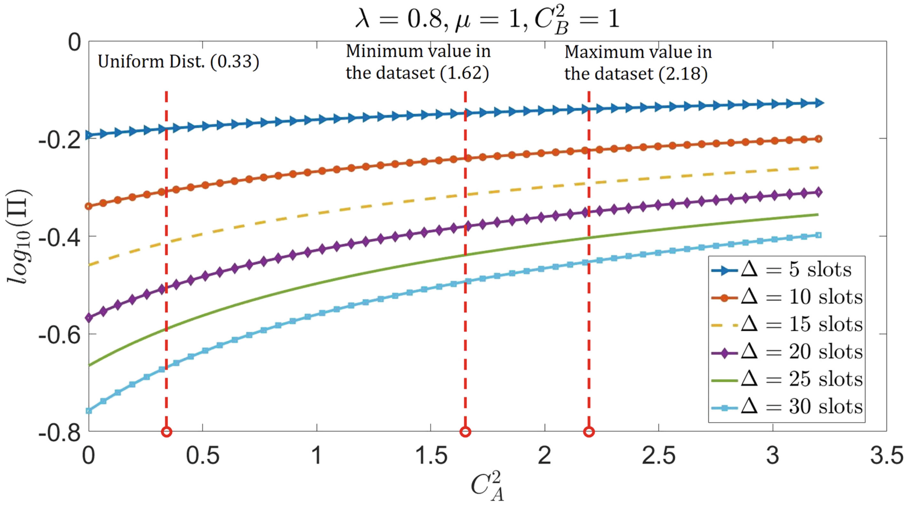

Fig. 1. We show the probability of missing the deadline (y-axis) in logarithmic scale (to the base ten), estimated with the diffusion approximation. We see that it increases significantly as C_A^2 increases and when the deadline Δ measured in slots decreases, for a fixed but high value of the arrival rate $\lambda = 0.8$. The average service rate μ and the SCV of service time C_B^2 are both fixed to 1.

Then, in [36] the value of S_j was selected by using queueing theory [17,23] as:

$$S_j \approx min[\ p_j + \frac{E[P]}{\frac{1}{\rho} - 1},\ \Delta_j\ ], \tag{10}$$

where $E[P]$ is the average processing time, and $\rho = \lambda E[P]$.

3.6 Experimental Results Concerning RGT

We now review the performance evaluation of RGT for two known heuristic scheduling algorithms, Priority based on Average Load (PAL) [39] and non-preemptive version of Earliest Deadline First (EDF) [26], where PAL enhanced with RGT is called R-PAL and EDF enhanced with RGT is called R-EDF. The performance evaluation is performed on the publicly available dataset [1] with respect to each of throughput η and fraction of successfully delivered bursts ζ metrics. Also, the performances of R-PAL and R-EDF are compared with the upper bound performances, where $S_j = \gamma \Delta_j$ and exhaustively search for the value of γ in the range $[0, 1]$ with increments of 0.05 to maximize each of η and ζ.

Accordingly, Fig. 3 displays the comparison of R-PAL and R-EDF with the upper bounds of those as well as PAL and EDF heuristics for η and ζ. The results in this figure show that RGT preprocessing significantly improves the throughput performance of each heuristic while the fraction of successfully delivered bursts

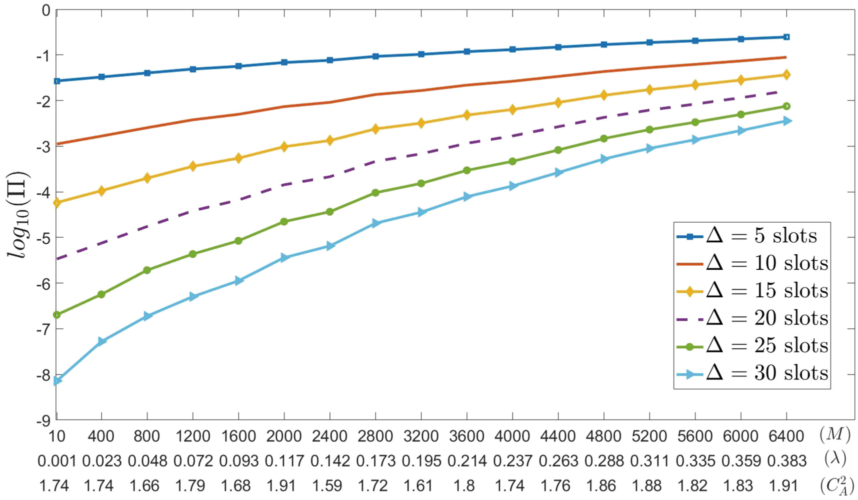

Fig. 2. The probability of missing the deadline (y-axis) in logarithmic scale (to the base ten) estimated with the diffusion approximation, using the traffic statistics of the real dataset of [1], is plotted against the number of IoT devices M (x-axis) that are being used. Note that each value of M corresponds to specific measured values of λ and C_A^2 shown along the $x - axis$.

remains the same. Furthermore, the results in Fig. 3 (top) show that the R-PAL and R-EDF significantly outperform the original versions of the heuristics PAL and EDF for a higher number of devices while both R-PAL and R-EDF are able to achieve almost the same throughput with their upper bounds. On the other hand, in Fig. 3 (bottom), one sees that ζ is almost the same for the enhanced heuristics (R-PAL and R-EDF) and the original heuristics (PAL and EDF).

3.7 The Quasi-Deterministic Transmission Policy (QDTP)

As the diffusion analysis that is discussed in Sect. 3.4 suggests, minimizing the SCV of interarrival times of traffic packets, C_A^2, reduces the probability of missing deadlines of those packets. Accordingly, in [19] a "Quasi-Deterministic Transmission Policy" (QDTP) is developed to minimize the probability of missing deadlines Π by setting almost all of the intertransmission times to a constant

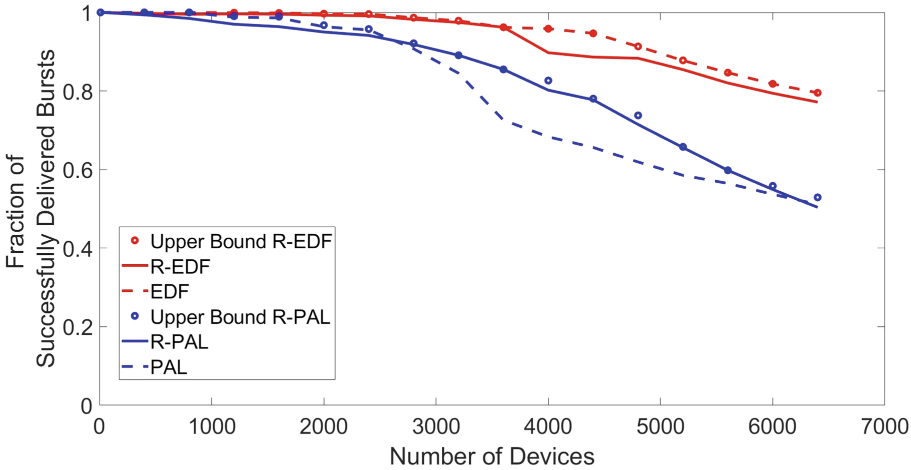

Fig. 3. Comparison of the upper bound R-PAL, R-PAL, PAL and the upper bound R-EDF, R-EDF, EDF algorithms for 12 to 6400 devices with respect to η (top) and ζ (bottom)

D so as to reduce the value of C_A^2. In QDTP whose pseudo-code is presented in Algorithm 1, $D = \frac{1}{\lambda}$, where λ is the interarrival time of burst generation times.

Algorithm 1: Pseudo-code QDTP

```
n = 1;
t_n = a_n;
for n ∈ {2, ..., N} do
    if a_n ≤ t_{n−1} + D then
        t_n = a_{n−1} + D;
    else
        t_n = a_n;
    end
end
```

For the practical application of QDTP, the generation times of the packets must be known in advance, similar to other predictive protocols. While advanced knowledge of the λ value is required, it can be easily calculated based on the creation times of the packages. Also, each IoT device must be informed of the time slot reserved for transmission of a packet created by that device before the start of the reserved slot. For this purpose, information regarding channels reserved for transmission permissions will be sent to devices via the downlink channel. Therefore, the communication channel must be bidirectional or individual IoT devices must also have the ability to detect the channel.

3.8 Experimental Results Concerning QDTP

We now examine the performance evaluation of QDTP on the same dataset [1] used for the evaluation of RGT. On the other hand, the bits are now packetized as IP packets, where the $21 - Byte$ headers (similar to the header size in LoRa-WAN) are followed by the payload, which is the bits in a burst. In addition, considering the effective data transmission rate of LoRa-WAN in free space is $5400\ bit/s$ [54], the transmission time for the average packet length in the dataset is 33.33 ms.

Figure 4 presents the SCV of interarrival times of packets, C_A^2, for both the original generation times in the dataset and the transmission time determined via QDTP. The results in this figure show that C_A^2 is significantly reduced for all values of M when QDTP is used to schedule packet transmissions.

Figure 5 displays the comparison between the performance of QDTP and that of original generation times in the dataset with respect to the base 10 of the (empirically measured) probability Π that the deadline is missed for different values of deadline Δ. In this figure, we see that Π is reduced to practically zero for $\Delta \geq 5$ when QDTP is used. On the other hand, when QDTP is not used (i.e. packets are transmitted when they are generated), Π increases with M and approaches 1.

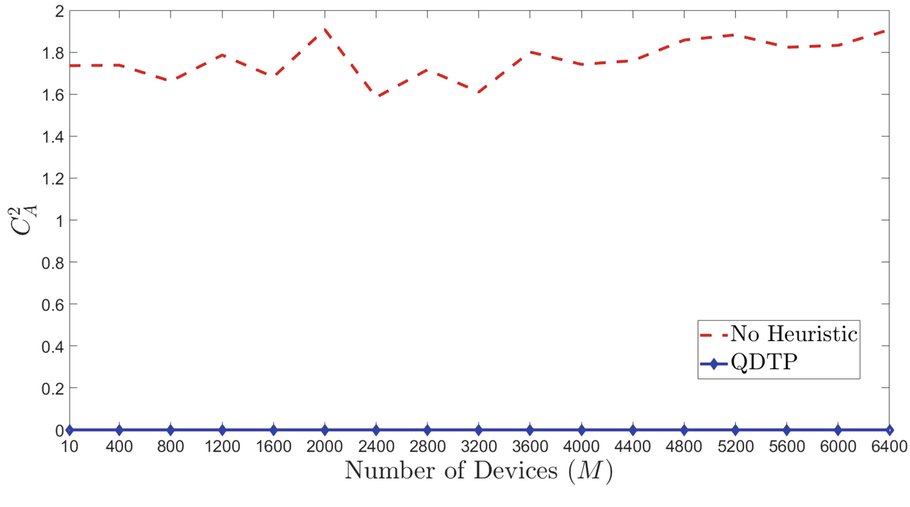

Fig. 4. Measurements of the SCV of interarrival times, both for the raw IoT data from [1], and for the same data using QDTP, for a varying number of active IoT devices M. We observe that QDTP has substantially reduced the empirically measured SCV C_A^2, reducing it to zero for all the distinct numbers of devices M in the dataset of [1].

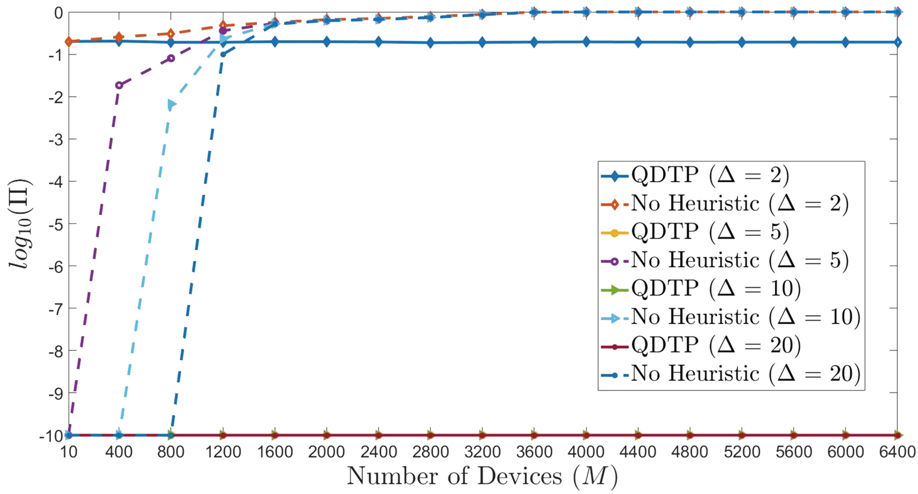

Fig. 5. We compare the logarithm to the base 10 of the (empirically measured) probability Π that the deadline is missed under both the raw dataset of [1] and under the case where the QDTP is used with the same dataset for values of the deadline $\Delta \in \{2, 5, 10, 20\}$.

4 Conclusions

The MAP, which can occur when a massive number of devices attempt the access a single gateway, is one of the most significant challenges for future IoT networks. Due to a high latency during the transmissions of data packets caused by MAP, the deadlines for delay-critical applications can be missed. Much work has been conducted in recent years to address this problem.

This paper reviews the work on reactive and proactive approaches to MAP, showing that methods based on proactive (i.e., predictive) techniques are a highly promising avenue to mitigate MAP. The observations of the most recent research results can be recapitulated with the following remarks:

1. The diffusion analysis proposed in [19] for the probability of missed deadlines in MAP shows that the latency requirements of IoT devices can be met in networks with a massive number of devices by reducing the SCV of the interarrival times of packets.
2. The Randomization of Generation Times (RGT) preprocessing algorithm proposed by [36] significantly improves the performance of fast scheduling heuristics by randomizing generation time of packets with uniform distribution yielding an inter-arrival time SCV of close to 1/3.
3. The Quasi-Deterministic Transmission Policy (QDTP) [25] meets the deadlines of almost all packets for up to 6400 IoT devices by reducing the queue length at IoT gateways to nearly zero, and the SCV of packet inter-arival times is also brought down to nearly zero.

Acknowledgments. This research has been supported by the European Commission H2020 Program through the IoTAC Research and Innovation Action, under Grant Agreement No. 952684.

References

1. IoT Traffic Generation Pattern Dataset, January 2021. https://www.kaggle.com/tubitak1001118e277/iot-traffic-generation-patterns
2. Abdelrahman, O.H., Gelenbe, E.: A diffusion model for energy harvesting sensor nodes. In: 2016 IEEE 24th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS), pp. 154–158. IEEE (2016)
3. Alavikia, Z., Ghasemi, A.: Collision-aware resource access scheme for LTE-based machine-to-machine communications. IEEE Trans. Veh. Technol. **67**(5), 4683–4688 (2018)
4. Ali, S., Rajatheva, N., Saad, W.: Fast uplink grant for machine type communications: challenges and opportunities. IEEE Commun. Mag. **57**(3), 97–103 (2019)
5. Astely, D., et al.: LTE release 14 outlook. IEEE Commun. Mag. **54**(6), 44–49 (2016)
6. Augusto-Gonzalez, J., et al.: From internet of threats to internet of things: a cyber security architecture for smart homes. In: 2019 IEEE 24th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), pp. 1–6. IEEE (2019)
7. Bello, O., Zeadally, S.: Toward efficient smartification of the internet of things (IoT) services. Future Gener. Comput. Syst. **92**, 663–673 (2019)
8. Chesnais, A., Gelenbe, E., Mitrani, I.: On the modeling of parallel access to shared data. Commun. ACM **26**(3), 196–202 (1983)
9. Cox, D.R., Miller, H.D.: The Theory of Stochastic Processes. Chapman and Hall, London (1965)
10. Du, J., Gelenbe, E., Jiang, C., Zhang, H., Ren, Y.: Contract design for traffic offloading and resource allocation in heterogeneous ultra-dense networks. IEEE J. Sel. Areas Commun. **35**(11), 2457–2467 (2017)
11. Eldeeb, E., Shehab, M., Alves, H.: A learning-based fast uplink grant for massive IoT via support vector machines and long short-term memory. IEEE Internet Things J. (2021)
12. Frötscher, A., Monschiebl, B., Drosou, A., Gelenbe, E., Reed, M.J., Al-Naday, M.: Improve cybersecurity of c-its road side infrastructure installations: the serIoT-secure and safe IoT approach. In: 2019 IEEE International Conference on Connected Vehicles and Expo (ICCVE), pp. 1–5. IEEE (2019)
13. Gelenbe, E.: On approximate computer system models. J. ACM (JACM) **22**(2), 261–269 (1975)
14. Gelenbe, E., Caseau, Y.: The impact of information technology on energy consumption and carbon emissions. Ubiquity **2015**(June), 1–15 (2015)
15. Gelenbe, E., Ceran, E.T.: Energy packet networks with energy harvesting. IEEE Access **4**, 1321–1331 (2016). https://doi.org/10.1109/ACCESS.2016.2545340
16. Gelenbe, E., Mang, X., Feng, Y.: A diffusion cell loss estimate for ATM with multiclass bursty traffic. In: ATM 1995. IAICT, pp. 233–248. Springer, Boston (1996). https://doi.org/10.1007/978-0-387-35068-4_13

17. Gelenbe, E., Mitrani, I.: Analysis and Synthesis of Computer Systems, 2nd Edition. World Scientific Ltd. & Imperial College Press, London (2010). https://doi.org/10.1142/p643
18. Gelenbe, E., Nakip, M., Czachorski, T.: Improving massive access to an IoT gateway. Submitted for publication (2022)
19. Gelenbe, E., Nakip, M., Marek, D., Czachórski, T.: Diffusion analysis improves scalability of IoT networks to mitigate the massive access problem. In: 29th International Symposium on the Modelling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS) (2021). (in Press)
20. Gelenbe, E., Ngai, E.: Adaptive random re-routing for differentiated QoS in sensor networks. Comput. J. **53**(7), 1052–1061 (2010)
21. Gelenbe, E., Ngai, E., Yadav, P.: Routing of high-priority packets in wireless sensor networks. In: IEEE Second International Conference on Computer and Network Technology, IEEE (2010)
22. Gelenbe, E., Ngai, E.C.H.: Adaptive QoS routing for significant events in wireless sensor networks. In: 2008 5th IEEE International Conference on Mobile Ad Hoc and Sensor Systems, pp. 410–415. IEEE (2008)
23. Gelenbe, E., Pujolle, G.: Introduction to Networks of Queues. Wiley, Chichester (1998)
24. Gelenbe, E., Sevcik, K.: Analysis of update synchronization for multiple copy data bases. IEEE Trans. Comput. **10**, 737–747 (1979)
25. Gelenbe, E., Sigman, K.: IoT traffic shaping and the massive access problem. In: ICC 2022: IEEE International Conference on Communications, pp. 1–6. IEEE, May 2022
26. George, L., Rivierre, N., Spuri, M.: Preemptive and non-preemptive real-time uniprocessor scheduling (1996)
27. Ghavimi, F., Chen, H.H.: M2M communications in 3GPP LTE/LTE-A networks: architectures, service requirements, challenges, and applications. IEEE Commun. Surv. Tutorials **17**(2), 525–549 (2015)
28. Jang, H.S., Jin, H., Jung, B.C., Quek, T.Q.: Resource-optimized recursive access class barring for bursty traffic in cellular IoT networks. IEEE Internet Things J. (2021)
29. Jiang, N., Deng, Y., Nallanathan, A., Yuan, J.: A decoupled learning strategy for massive access optimization in cellular IoT networks. IEEE J. Sel. Areas Commun. **39**(3), 668–685 (2020)
30. Jin, H., Toor, W.T., Jung, B.C., Seo, J.B.: Recursive pseudo-Bayesian access class barring for M2M communications in LTE systems. IEEE Trans. Veh. Technol. **66**(9), 8595–8599 (2017)
31. Kim, H.-Y., Kim, J.-M.: A load balancing scheme based on deep-learning in IoT. Cluster Comput. **20**(1), 873–878 (2016). https://doi.org/10.1007/s10586-016-0667-5
32. Kleinrock, L.: Queueing Systems: Computer Applications. Wiley, Hoboken (1976)
33. Kobayashi, H.: Application of the diffusion approximation to queueing networks i: equilibrium queue distributions. J. ACM (JACM) **21**(2), 316–328 (1974)
34. Liang, L., Xu, L., Cao, B., Jia, Y.: A cluster-based congestion-mitigating access scheme for massive M2M communications in internet of things. IEEE Internet Things J. **5**(3), 2200–2211 (2018)
35. Liu, J., Song, L., et al.: A novel congestion reduction scheme for massive machine-to-machine communication. IEEE Access **5**, 18765–18777 (2017)

36. Nakip, M., Gelenbe, E.: Randomization of data generation times improves performance of predictive IoT networks. In: 2021 IEEE World Forum on Internet of Things (WF-IoT) (2021). (in Press)
37. Nakip, M., Gül, B.C., Rodoplu, V., Güzeliş, C.: Comparative study of forecasting schemes for IoT device traffic in machine-to-machine communication. In: Proceedings of the 2019 4th International Conference on Cloud Computing and Internet of Things, pp. 102–109 (2019)
38. Nakip, M., Karakayali, K., Güzeliş, C., Rodoplu, V.: An end-to-end trainable feature selection-forecasting architecture targeted at the internet of things. IEEE Access **9**, 1–1 (2021). https://doi.org/10.1109/ACCESS.2021.3092228
39. Nakip, M., Rodoplu, V., Güzeliş, C., Eliiyi, D.T.: Joint forecasting-scheduling for the internet of things. In: 2019 IEEE Global Conference on Internet of Things (GCIoT), pp. 1–7. IEEE (2019)
40. Natsiavas, P., et al.: Developing an infrastructure for secure patient summary exchange in the EU context: lessons learned from the konfido project. Health Inf. J. **27**(2), 14604582211021460 (2021)
41. Newell, G.F.: Applications of Queueing Theory. Chapman and Hall, London, June 1971
42. Ngai, E.C.H., Gelenbe, E., Humber, G.: Information-aware traffic reduction for wireless sensor networks. In: 2009 IEEE 34th Conference on Local Computer Networks, pp. 451–458. IEEE (2009)
43. Petkov, V., Obraczka, K.: The case for using traffic forecasting in schedule-based channel access. In: 2011 IEEE Consumer Communications and Networking Conference (CCNC), pp. 208–212. IEEE (2011)
44. Petkov, V., Obraczka, K.: Collision-free medium access based on traffic forecasting. In: 2012 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), pp. 1–9. IEEE (2012)
45. Raca, D., et al.: On leveraging machine and deep learning for throughput prediction in cellular networks: design, performance, and challenges. IEEE Commun. Mag. **58**(3), 11–17 (2020)
46. Rodoplu, V., Nakıp, M., Eliiyi, D.T., Güzelis, C.: A multi-scale algorithm for joint forecasting-scheduling to solve the massive access problem of IoT. IEEE Internet Things J. (2020)
47. Rodoplu, V., Nakip, M., Qorbanian, R., Eliiyi, D.T.: Multi-channel joint forecasting-scheduling for the internet of things. IEEE Access **8**, 217324–217354 (2020)
48. Ruan, L., Dias, M.P.I., Wong, E.: Machine learning-based bandwidth prediction for low-latency H2M applications. IEEE Internet Things J. **6**(2), 3743–3752 (2019)
49. Shahin, N., Ali, R., Kim, Y.T.: Hybrid slotted-CSMA/CA-TDMA for efficient massive registration of IoT devices. IEEE Access **6**, 18366–18382 (2018)
50. Shehab, M., Hagelskjær, A.K., Kalør, A.E., Popovski, P., Alves, H.: Traffic prediction based fast uplink grant for massive IoT. In: 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1–6. IEEE (2020)
51. Shirvanimoghaddam, M., Dohler, M., Johnson, S.J.: Massive non-orthogonal multiple access for cellular IoT: potentials and limitations. IEEE Commun. Mag. **55**(9), 55–61 (2017)
52. Soltanmohammadi, E., Ghavami, K., Naraghi-Pour, M.: A survey of traffic issues in machine-to-machine communications over LTE. IEEE Internet Things J. **3**(6), 865–884 (2016)

53. Takács, L.: Introduction to the Theory of Queues. Oxford University Press, Oxford (1962)
54. Tarab, H.: Real time performance testing of LoRa-LPWAN based environmental monitoring UAV system. University of Windsor, Electronic Theses and Dissertations. 7578 (2018). https://scholar.uwindsor.ca/etd/7578
55. Tello-Oquendo, L., et al.: Performance analysis and optimal access class barring parameter configuration in LTE-A networks with massive M2M traffic. IEEE Trans. Veh. Technol. **67**(4), 3505–3520 (2018)
56. Tello-Oquendo, L., Pacheco-Paramo, D., Pla, V., Martinez-Bauset, J.: Reinforcement learning-based ACB in LTE-A networks for handling massive M2M and H2H communications. In: 2018 IEEE International Conference on Communications (ICC), pp. 1–7. IEEE (2018)