Energy efficiency, latency and reliability trade-offs in M2M uplink scheduling

Moses K Torkudzor | Stefan Schwarz | Jamal-Deen Abdulai | Markus Rupp

1 Department of Computer Engineering and Computer Science, University of Ghana, PMB, Legon, Accra, Ghana
2 CD-Lab for Dependable Wireless Connectivity, Institute of Telecommunications, TU Wien, Vienna, Austria

Abstract
Issues of energy efficiency, latency and reliability have engaged researchers’ attention in wireless communications for a long time due to their importance in the dependable delivery of data over wireless networks. The emergence of M2M communication and IoT has intensified this attention because energy efficiency and latency are critical factors affecting their performance. In this paper, a scheduler is designed by utilising the probability density function of the signal-to-noise ratio of Rayleigh fading channels to define a threshold used for resource allocation. This threshold, combined with the mean SNR of an M2M device, determines whether or not an M2M device is eligible for scheduling, given its instantaneous channel conditions. The trade-off between energy efficiency, latency and packet drop of this proposed scheduling strategy is investigated. The performance of the proposed scheduler is compared to round robin, maximum throughput, and proportional fair schedulers. Compared to these standard scheduling strategies, the scheduler provides the advantage of trading off latency for energy efficiency by tuning the threshold parameter.

1 INTRODUCTION

Machine-to-Machine (M2M) communication has recently developed into an important technology generating significant revenues to mobile network operators. In this technology, nodes are equipped with sensor modules to monitor and collect various forms of data such as temperature, humidity, and pressure for onward transmission through communication networks. There are many applications of M2M communication, some of which are: smart metering, intelligent transportation, healthcare, smart city, public safety and consumer devices [1–4].

M2M communication, however, comes with peculiar challenges which are dependent on its typical characteristics, including a massive number of devices, small size data transmission, infrequent traffic patterns, and battery limitation in contrast to human-to-human (H2H) communications. These problems are well-known and are variously discussed [5–11]. Energy efficiency and low latency are very critical considerations in current and future wireless networks. Since the types of wireless services become increasingly diverse as technologies evolve, networks must also be built to meet the corresponding applications and different delay requirements, knowing when and how to trade off tolerable delay for low power [12]. Data transmission methods in wireless networks should be designed to achieve target delay and throughput at reduced power consumption [13]. The authors in [14] conduct an extensive survey of energy efficiency trade-off mechanisms and underscore the fact that trade-off techniques are needed in each protocol layer to achieve energy efficiency. The challenges of reliability, energy efficiency, latency reduction, a massive increase in connection density, and capacity improvement are the main goals of fifth generation (5G) wireless networks [15]. In this paper, we extend our work in [16] which reveals maximum throughput scheduling (MTS) as more energy efficient compared to proportional fair scheduling (PFS) and round robin scheduling (RRS) to a design of a scheduler more energy efficient than MTS. We furthermore investigate the trade-off between energy efficiency, latency and reliability of this novel scheduler.

The main contributions of the paper are:

• The signal-to-noise ratio’s probability density function (PDF) of Rayleigh fading channels is employed to design a scheduler to investigate energy efficiency, latency and reliability trade-offs.
• Threshold parameters are defined in the scheduler which can be tuned to trade off latency for energy efficiency.
• Order statistics are employed to derive a closed-form analytic PDF expression for resource allocation by the scheduler.
• A closed-form analytic equation for expected energy efficiency and latency is also derived. The analytic results are validated using Monte Carlo simulations.

Notation
\( U \) is used to denote a set of M2M users, \( f(\cdot) \) for PDF, \( F(\cdot) \) for cumulative density function (CDF), \( P(\cdot) \) for probability, and \( E(\cdot) \) for expected value.

The rest of the work is organised as follows: Section 2 discusses the related literature. In Section 3, we present the system model of our proposed scheduler. Designing the scheduler is covered in Section 4. In Section 5, we derive the SNR PDF of the scheduled user. Analytic equations for the performance measures are derived in Section 6. We discuss and analyse simulation results in Section 7. Section 8 finally concludes the work.

2 RELATED WORK

Scheduling and resource allocation are considered as important strategies in wireless communication networks because of their ability to manage the limited radio resources at the access level in a way that maximises the system performance such as latency and energy consumption. A power allocation algorithm for massive M2M communication using Lagrange multipliers is proposed in [17] to decrease signalling overhead. The main objective behind this Lagrange multiplier power allocation algorithm in [17] is to maximise the total energy efficiency of a group of M2M devices, while satisfying the time delay of the M2M devices. Careful scheduling of sensed data can reduce power consumption for M2M nodes thereby increasing the lifetime of the M2M network [18]. In [19], a joint resource allocation and clustering algorithm was designed which maximises system energy efficiency. Uplink scheduling and transmit power control are investigated in [20] to minimise the energy consumption of battery-driven devices deployed in LTE networks using clustering techniques. According to [21], two main challenges affecting the effective operation of low power M2M networks including processing complexity and transmission range constraints are tackled through cooperative communication which employs M2M gateways as relays. However, some emerging research challenges of cooperative M2M communication indicated in [21] are an increase in end-to-end latency, complex scheduling, increased overhead, increased intra and inter network interferences, and the difficulty of clustering process due to the problem of uneven distribution of M2M devices. Maximum energy efficient uplink M2M data packets transmission is investigated in [22] by formulating it into a joint problem of modulation and coding scheme (MCS) selection, resource allocation, and power control. The allocation and scheduling scheme were formulated into a mixed-integer linear fractional programming problem which was reconstructed into a linear form and an optimisation scheme based on the Charnes–Cooper transformation and the Glover linearisation scheme was proposed. Results indicate that low data packets dropping ratios and optimal energy efficiency for a large number of M2M nodes were achieved. The inclusion of MCS in the scheduling and resource allocation in [22] is unique as it caters for modulation and coding aspects that impact this performance measures in wireless communication. The paper, however, did not consider circuit energy consumption. In [23], a jointly delay-optimal multiclass packet scheduler designed at the application server and M2M aggregators are presented. This work in [23] iteratively searches for an optimal fraction of time-sharing between all preemptive priority scheduling policies and uses a sigmoidal function to map delay requirements of classified QoS onto utility functions of packet delay to maximise proportionally fair system utility metric. The iterative approach reduced computational complexity resulting in near-minimal delay jitter for delay-sensitive traffic but at the expense of higher delay jitter for delay-tolerant traffic. It is thus not suitable for delay tolerant devices and also did not account for the effect of this in terms of energy efficiency. An uplink scheduling algorithm is proposed in [24] to solve the problem of the inability of the uplink scheduling to guarantee the transmission of real-time traffic packets within the delay period by establishing an target integer linear programming model using delay constraints of real-time traffic. The problem that the bandwidth of low priority traffic is insufficient and the overall QoS is affected when multi-class traffic is processed is also addressed in [24] by designing a priority adjustment algorithm for user equipment to choose lower priority traffic to be transmitted first when the transmission of higher priority traffic can be safely delayed without violating its delay constraints in a given TTI. When compared with dynamic Hungarian algorithm with modification, [24] shows better performance for fairness, throughput, number of packets delivered by the user having the worst channel conditions, and the delay metrics. In [25], packet dropping caused by delay constraint violation of M2M traffic at energy harvesting gateways is minimised by developing an optimisation framework. In designing this, rate and delay requirements, energy, data causality, and SC-FDMA constraints were taken into account. The optimisation problem was further expressed in terms of its discrete and continuous variables and solved by decomposing it into primal and master problems using generalised Benders decomposition. A low-complexity method was then proposed to solve the primal problem with a heuristic resource block allocation algorithm which performs better than two other benchmarked heuristic methods. However, ignoring of packet dropping at the M2M devices and the assumption of them having reliable energy sources would eventually degrade the performance in the real life situation.

Other technologies being explored to reduce latency and energy consumption are computational task offloading, resource provisioning, and fog networking whose computing environment connects resources together to improve life through usage of computation resources at the network edge. This idea of fog computing is therefore to drive resource computation and storage towards the end-user to solve the problem of long-delay links between the cloud data center and end-user equipment. A comprehensive review of machine learning and
stochastic-based computation offloading mechanisms in mobile edge computing environments are respectively compared in [26] and [27] under the performance metrics, case studies, utilised techniques, evaluation tools, advantages, and weaknesses. The performance metrics considered in [26, 27] are energy consumption, latency, QoS, quality of experience, response time, and cost. Optimisation of task scheduling in a fog computing environment is a challenging issue because it is considered an NP-hard problem [28]. In order to solve this problem, a task scheduling algorithm using the moth-flame optimisation algorithm which assigns an optimal set of tasks to fog nodes is presented in [28] which minimises total tasks execution. In fog networking, however, computational offloading of tasks for the delay-sensitive devices is done at the expense of additional transmission latency. This trade-off issue is formulated in [29] as a delay-sensitive data offloading problem considering the local task execution and transmission latencies to jointly optimise the computing and communication resources in the fog node. An offloading scheme in fog computing to reduce the weighted sum of energy consumption and total delay for task processing per end-user is proposed in [30]. This was achieved by formulating and solving a non-convex optimisation problem using semidefinite relaxation. The authors in [30] however ignored the cost due to local energy consumption to execute and offload a task at the fog node and remote cloud which will not present a comprehensive energy consumption model. A workload clustering-based resource provisioning mechanism presented in [31] for executing cloud-based applications reduces delay, cost, and energy consumption. Summary of scheduling algorithms under the performance metrics, technique, evaluation tools, advantages, and disadvantages is presented in Table 1.

### 3 | SYSTEM MODEL

We consider in this paper an uplink multi-user system in a single cell scenario as shown in Figure 1. We assume a clustering scenario where \( n \) M2M devices (cluster members) communicate through an M2M gateway (cluster head) to the base station. However, only the communication between a single cluster head and its cluster members is investigated in this paper.

The SNR \( \gamma_u(t) \), of an M2M user \( u \) at a given time instance \( t \) is:

\[
\gamma_u(t) = \frac{P_u |b_u(t)|^2}{\sigma_u^2}
\]

(1)

where \( P_u, b_u(t) \) and \( \sigma_u^2 \) are the transmit power, the channel coefficient between the M2M device, the M2M gateway, and the

### Table 1: Summary of scheduling algorithm

| Paper | Metric | Technique | Advantages | Disadvantages | Evaluation Tool |
|-------|--------|-----------|------------|---------------|----------------|
| [22]  | Energy efficiency. Bandwidth utilisation. Drop rate. | Charnes–Cooper transformation and Glover linearisation | Low data packets dropping ratio | Less energy efficient due combination of scheduling and resource allocation | Lp solve Toolbox in MATLAB |
| [23]  | Delay, jitter | Distributed and Iterative optimisation | Low computational complexity | Higher delay jitter for delay-tolerant traffic | unknown |
| [24]  | Fairness, throughput, delay, packet drop | Linear programming | Improved bandwith utilisation, Improve low priority traffic QoS | Slightly larger delay, Priority given to users with poor channel will degrade throughput and energy efficiency | OMNET++ |
| [25]  | Packet drop | Generalised Benders decomposition-heuristic | Low-complexity Slow convergence of algorithm due to NP hardness | unknown |
| [27]  | Execution time, task transfer time | Moth-flame optimisation | Less total execution time consumption | Does not address energy consumption and communication cost | iFogSim |
| [28]  | Task execution delay, transmission delay | Convex optimisation, quadratically constraint quadratic programming (QCQP) | Minimum end-to-end latency Multi-task Scenario with different delay deadline for each task was not considered | Unknown |
| [29]  | Energy consumption, task processing delay | Semidefinite relaxation | Reduced computational load, consume less energy | Ignores latency | MATLAB |
We define the scheduling SNR threshold of user $u$ as follows:

$$k_u = \bar{\gamma}_u T. \quad (6)$$

Let the subset of M2M users that have at the current time instant packets in their packet buffer to transmit be $U_1 = \{1, ..., n\}$. Out of these users, the subset $U_2 \subseteq U_1$ defines the M2M users whose current SNR is at least $k_u$, $U_2 = \{u \in U_1 | \gamma_u(t) \geq k_u\}$. The M2M user, $u$, with the maximum SNR from the subset $U_2$ is then scheduled:

$$u = \text{argmax}_{u \in U_2} \gamma_u(t). \quad (7)$$

Devices which do not meet the threshold condition at time instant $t$ have the probability of meeting it at another instance.

5 | Derivation of the SNR PDF of the Scheduled User

In the following, we derive the PDF of the SNR of the user that has been selected by the scheduler. According to (6), this PDF is determined by the maximum of the SNRs of the users in $U_2$ that meet the threshold condition. Hence, the PDF can be obtained from the order statistics of the users in $U_2$. To simplify the derivation, we assume in the following that the average SNR $\bar{\gamma}$ of all users is equal, i.e. $\bar{\gamma}_u = \bar{\gamma}$, $\forall u$ and therefore, the threshold $k_u = k, \forall u$. We furthermore assume that all $u$ users, i.e. $U_1 = \{1, ..., n\}$ have data to transmit. Assuming Rayleigh fading, the conditional PDF of the SNR of the users that satisfy the threshold condition is:

$$f(\gamma(t) | \gamma(t) \geq k) = \frac{\lambda e^{-\lambda \gamma(t)}}{C}, k \leq \gamma(t) \leq \infty. \quad (8)$$

Here, $\lambda$ is a shape parameter computed as:

$$\lambda = 1/\bar{\gamma} \quad (9)$$

and $C$ is a normalising constant which is given as:

$$C = \int_{k}^{\infty} \lambda e^{-\lambda \gamma} d\gamma = e^{-\lambda k}. \quad (10)$$

Substituting Equation (10) into Equation (8) gives:

$$f(\gamma(t) | \gamma(t) \geq k) = \lambda e^{-\lambda \gamma(t)-k}, k \leq \gamma(t) \leq \infty. \quad (11)$$

This implies that the CDF of the SNR of the users that satisfy the threshold condition is:

$$F(\gamma(t) | \gamma(t) \geq k) = 1 - e^{-\lambda \gamma(t)-k}. \quad (12)$$

We use order statistics to derive a closed-form expression for the SNR statistics of the user selected by our scheduler. The algorithm for closed-form solution of SNR PDF for the scheduled user is presented in Algorithm 1. The number of M2M users with packets to transmit whose current SNR is at
In this section, we derive a closed-form analytic equation for the energy efficiency. Steps for deriving these closed solutions are respectively shown in Algorithm 1 and 2. We consider the same assumption $\bar{\gamma} = \bar{\gamma}$ for all users as in Section 5 for the same purpose of simplifying the derivations.

The expected energy efficiency of our scheduler, which schedules the M2M device with largest SNR out of those M2M devices that surpass the required SNR threshold of $\gamma_{i}(t) \geq k$, is:

$$E(EE_{sc}) = \frac{B}{P_{t} + P_{e}} \int_{k}^{\infty} \log_{2}(1 + \gamma_{\max}) \cdot f(\gamma_{\max})d\gamma_{\max}$$

(16)

This means that the expected Energy Efficiency, $E(EE_{sc})$, of M2M user $u$ is:

$$E(EE_{sc}) = E(EE_{sc}) \cdot P(\gamma_{sc}).$$

(17)

In Equation (17), $P(\gamma_{sc})$ as derived in Equation (18), is the unconditional probability of scheduling an M2M device.

### 6.2 Latency analytic equation

To estimate the expected latency analytic equation, we need to know the average number of the LTE frames $N_{f}$ required to transmit a single packet and the probability $P(\gamma_{sc})$ that a user is scheduled. To this end, we first calculate the conditional probability that a user is selected by the scheduler when it has already passed the first scheduling stage, i.e. $\gamma_{i}(t) \geq k$. In this case, when there are $\bar{n}$ other users in the second scheduling stage, the considered user is scheduled with probability $1/(\bar{n} + 1)$, since all users follow the same SNR statistics. The conditional probability is thus obtained as:

$$P(\gamma_{sc}|\gamma_{i}(t) \geq k) = \sum_{j=0}^{\bar{n}} \frac{1}{\bar{n} + 1} \left( \sum_{\gamma_{i}(t) \geq k} \right) \cdot \left( 1 - \gamma_{i}(t) \right)^{\bar{n} - \bar{n}} \cdot f(\gamma_{\max}, t)$$

(18)

In Equation (18), we sum over all possible numbers $\bar{n}$ of other users in scheduling stage 2. Within the sum, we calculate the
Algorithm 3 Algorithm for closed-form latency analytic equation

1: Run Algorithm 2 to find the unconditional probability of scheduling an M2M user by using Equation (19)
2: Determine the average number of transmitted bits per frame duration using Equation (21)
3: Calculate the average number of frames required to transmit a single packet as in Equation (20)
4: The approximate latency of an M2M user is found using Equation (22)

probability that exactly \( \hat{n} \) other users are in scheduling stage 2. Using Bayes’ rule, the unconditional probability of scheduling an M2M user is therefore given as:

\[
P(u_{sc}) = P(y_i(t) \geq k) \cdot P(y_i(t) \geq k) = e^{-\lambda k} \cdot P(y_i(t) \geq k).
\]  

The average number of frames required to transmit a single packet is:

\[
N_f = \frac{n_{\text{bits}}}{\bar{R}_u} \tag{20}
\]

where \( n_{\text{bits}} \) and \( \bar{R}_u \) are the number of bits per packet and the average number of transmitted bits per frame duration \( T_f \) respectively. The average number of transmitted bits per \( T_f \) is:

\[
\bar{R}_u = B \cdot T_f \int_k^{\infty} \log_2(1 + y_{\max}) \cdot f(y_{\max})dy_{\max} \tag{21}
\]

We therefore calculate the approximate latency \( L \) by estimating the number of frames an M2M user has to wait until it is scheduled \( N_f \) times:

\[
E(L) \approx N_f \cdot \frac{1}{P(u_{sc})} \tag{22}
\]

6.3 Algorithm complexity

We now discuss the problem complexity of the three algorithms as stated above. It can be seen from Algorithm 1 that it is linear in the number of users due to the linearity of the equations involved. Though Equations (14) and (15) in Algorithm 1 are complex yet we have them available in closed-form and so complexity is not an issue. Similarly, in Algorithms 1 and 2, we provide every mathematical equations in closed-form and the values can be calculated in linear complexity.

7 SIMULATION RESULTS DISCUSSION AND ANALYSIS

7.1 Performance investigation

Following our model in Section 3, our communication network simulation test bed consists of one M2M gateway with \( n \) M2M nodes connected to the gateway in a cellular environment. We consider two different scenarios for the performance investigation, one in which all M2M devices have the same SNR and a second, in which they have different SNRs. Under each of these scenarios, we use a traffic model similar to [32] as shown in Figure 2 which generates packets randomly at time \( \tau \). After \( \tau \), subsequent packets are generated periodically at \( T \) [ms]. The packets are queued in the buffer and are scheduled when the respective M2M devices meet the scheduling condition discussed in Section 4. If packets are delayed in the buffer beyond the packet generation interval they are dropped making our traffic model delay-sensitive. We compare the performance of our scheduler, called SNR Threshold Scheduler (THS) with RRS, MTS and PFS. In RRS the same priority is given to allocating the same number of resource blocks (RBs) to each user in an ordered manner to maintain fairness. The MTS, on the other hand, allocates a different number of RBs to users based on their respective channel quality to maximise average throughput. The PRS allocates RBs to users with the best relative channel quality by keeping track of the average throughput of users in a past window to balance the maximisation of throughput and fairness. The performance measures considered for the respective scenarios include energy efficiency, latency, and the ratio of packets not successfully transmitted. The packets which are not successfully transmitted are those which were delayed in the buffer beyond the packet generation interval.

The simulation parameters are summarised in Table 2. The performance of the different SNR scenario for the energy efficiency, latency and reliability are respectively shown in Figures 3–5. The average SNR in dB for M2M devices in this scenario are 44.77, 43.39, 42.01...7.53, 6.15, 4.77. As seen from
Figure 3, the energy efficiency of our scheduler increases with a growing threshold, outperforming the other schedulers when the threshold is sufficiently large. It can be seen from Figures 3 and 4 that energy efficiency and latency increase respectively with growing threshold while reliability decreases as shown in Figure 5. Thus, THS is trading off energy efficiency for latency and reliability which means that for low latency and high reliability to be achieved, some energy efficiency has to be sacrificed. The increase in latency and decrease in reliability for THS as seen in Figures 4 and 5 is due to the fact that devices have to wait for some time for their SNR to meet the condition for scheduling. Energy is thus wasted by the devices while waiting to achieve the minimum SNR requirement of our scheduler. Therefore, it will be necessary to account for the loss of energy during idle time which has not been given much attention in the literature. As seen from Figure 4, MTS achieves the lowest latency with THS better than PFS and RRS until beyond threshold values of 2.08 and 2.79, respectively. In terms of reliability from Figure 5, the ratio of packet drop rate for THS and MTS are the same until a threshold value of 1 where that of THS increases but is still better than that of PFS and RRS until beyond threshold values of 1.89 and 2.55, respectively. However, a careful look at Figures 3–5 reveals that for better performance to be achieved, our THS scheduler is the choice when operated at lower threshold values.

Considering the scenario with an equal average SNR for all users from Figures 6–8, a similar order of performance for the schedulers as described above is exhibited except for MTS and PFS recording the same performance for energy efficiency and reliability. This behaviour occurs due to the same average SNR for M2M devices resulting in them having equal probability to be scheduled.

Using Bootstrap estimate for the 95 percent confidence intervals of the population mean, we plot the confidence interval of our results to determine whether the differences shown in the figures are statistically significant and if their results are consistent and of less variation. The confidence interval plot which is considered for only different SNR scenario is shown in Figures 3–5 where the regions of the confidence intervals around RRS, MTS, PFS, and THS are indicated in colours of magenta, black, green and yellow respectively. As seen in these figures, the regions of the confidence intervals are very small indicating more consistency and less variation in the results from the schedulers showing that they are statistically significant. Furthermore, we confirmed the Bootstrap estimate with results from paired t-test as summarised in Table 3. The significance level considered for the paired t-test for the statistical evaluation is $p \leq 0.05$. Our null hypothesis is, there is no
difference between THS and the algorithms compared and the alternative hypothesis is, there is a difference between THS and the compared algorithm. As seen in Table 3, the \( p \) value is less than 0.05 in all the cases signifying that the null hypothesis is rejected. It thus shows from Table 3 that there is a significant difference between the proposed THS over the other schedulers. Stated differently, there is improvement in the proposed scheduler over the others. Reliability is not shown in the table because the analysis has to be done under the condition that all packets are successfully transmitted. As such there would not be any packet drop for the analysis of unsuccessful packet transmission. The provided confidence intervals and \( p \)-values agree on the statistical significance of our simulations. To determine whether test results are statistically significant, both, \( p \)-values as well as confidence intervals can be used. Confidence intervals provide likely boundaries for any improvement to aid in determining if a difference is noteworthy. It also provides meaningful estimates because it gives ranges that usually contain the parameter.

### 7.2 Validation of analytical results

We now discuss the validity of THS by fitting the analytic equations and simulated results. We consider the same SNR scenario since we also assume equal SNR in our analysis. The average SNR for M2M devices in this case is 25 dB. To make our simulation set up suitable for the purpose of the fitting, packets are generated for each device at random time instances unlike the periodic time interval described in Figure 2. Figure 9 shows a CDF graph of Equation (15) and the simulated result respectively.
Clearly, it can be seen from Figure 9 that the analytic result is the same as the simulated result. We also plot the 95% confidence interval in Figure 9 which can be seen in yellow colour around the analytical and simulated graphs. The very smaller region on the confidence interval shown here proves that we are highly confident about the validity of our proposed THS.

The graph of Equation (16) for the expected energy efficiency and its corresponding simulated result are shown in Figure 10 which indicates that the two are the same. The latency results from the simulation and its corresponding approximated analytic Equation (22) is presented in Figure 11 which also show a good fitting with slight variation occurring beyond the threshold value of 4.

8 | CONCLUSION

In this paper, we investigated energy efficiency, latency and reliability trade-offs in M2M communication using various scheduling strategies. We proposed a scheduler that allows to trade off the aforementioned parameters by adjusting a threshold parameter. Results indicate that energy efficiency and latency respectively increase with growing threshold parameter while reliability decreases with it exhibiting a fundamental trade-off between energy efficiency, latency, and reliability. We derived a closed-form CDF/PDF of the SNR of the user that is selected by our scheduler. Based on this expression we analysed the energy efficiency, latency and reliability of our scheduler as a function of the threshold parameter. The analytic equations of the CDF, energy efficiency and approximated latency produced good fitting with their respective simulation results validating our analysis. When compared with RRS, MTS and PFS, our scheduler (THS) is the most energy efficient when selecting a large threshold parameter, but at a cost of higher latency and lower reliability. However, our scheduler is the best choice at lower threshold values for optimised energy efficiency, latency and reliability performance.

ACKNOWLEDGEMENT

The financial support by the Austrian Federal Ministry for Digital and Economic Affairs and the National Foundation for Research, Technology and Development is gratefully acknowledged.
REFERENCES

1. Chen, M., Wan, J., Li, F.: Machine-to-machine communications: Architectures, standards and applications. KSHI Trans. Internet Inf. Syst. 6(2), 480–497 (2012)
2. Zhang, Y. et al.: Home M2M networks: architectures, standards, and QoS improvement. IEEE Commun. Mag. 49(4), 44–52 (2011)
3. Verma, P.K., et al.: Machine-to-Machine (M2M) communications: A survey. Journal of Network and Computer Applications 66, 83–105 (2016)
4. Schwarz, S., Rupp, M.: Society in Motion: Challenges for LTE and Beyond Mobile Communications. IEEE Communications Magazine, Feature Topic: LTE Evolution 54(5), 76–83 (2016)
5. Shafiq, M.Z. et al.: A first look at cellular machine-to-machine traffic: large scale measurement and characterization. ACM SIGMETRICS Performance Evaluation Review 40(1), 65–76 (2012)
6. Abu-Al, N. et al.: Uplink scheduling in LTE and LTE-advanced: Tutorial, survey and evaluation framework. IEEE Communications Surveys & Tutorials 16(3), 1239–1265 (2013)
7. Marwat, S.N.K. et al.: Impact of machine-to-machine (M2M) communications on disaster management in future mobile networks. General Sir John Kotelawala Defence University Annual Symposium. Kotelawala Defence University, Ratmalana (2012)
8. Kumar, A., Abdelhadi, A., Clancy, C.: An Online Delay-Optimal Packet Scheduler for Generic M2M Uplink Traffic. https://www.serf.net/wp-content/uploads/2017/08/TR314_OnlineDelayOptimalPacket.pdf. Accessed on October 2019
9. Liu, R. et al.: M2M-oriented QoS categorization in cellular network. In: 7th International Conference on Wireless Communications, Networking and Mobile Computing, pp. 1–5.IEEE, Piscataway, NJ (2011)
10. Khoshkholgh, M.G. et al.: Modeling and characterization of transmission energy consumption in machine-to-machine communications in cellular networks. In: IEEE Wireless Communications and Networking Conference, pp. 1469–1475.IEEE, Piscataway, NJ (2015)
11. Dhillon, H.S. et al.: On resource allocation for machine-to-machine (M2M) communications in cellular networks. IEEE Globecom Workshops, pp. 1638–1643.IEEE, Piscataway, NJ (2012)
12. Chen, Y. et al.: Fundamental trade-offs on green wireless networks. IEEE Commun. Mag. 49(6), 30–37 (2011)
13. Li, Z., Yin, C., Yue, G.: Delay-bounded power-efficient packet scheduling for uplink systems of LTE. 5th International Conference on Wireless Communications, Networking and Mobile Computing, pp. 1–4.IEEE, Piscataway, NJ (2009)
14. Mahapatra, R. et al.: Energy efficiency tradeoff mechanism towards wireless green communication: A survey. IEEE Commun. Surveys & Tutorials 18(1), 686–705 (2015)
15. Parvez, I., et al.: A survey on low latency towards 5G: RAN, core network and caching solutions. IEEE Communications Surveys & Tutorials 20, 3098–3130 (2018)
16. Torkudzor, M.K. et al.: Energy efficiency analysis of scheduling in M2M communication over Rayleigh fading channels. WSA 2019, 23rd International ITG Workshop on Smart Antennas, pp. 1–6.IEEE, Piscataway, NJ (2019)
17. Li, N., et al.: Energy-efficient power allocation for massive M2M communication over LTE-A cellular uplink. EURASIP Journal on Wireless Communications and Networking 2018, 1–11 (2018)
18. Fu, H.-L. et al.: Energy-efficient reporting mechanisms for multi-type real-time monitoring in machine-to-machine communications networks. In: Proceedings IEEE INFOCOM, pp. 136–144.IEEE, Piscataway, NJ (2012)
19. Chai, R., et al.: Energy Efficiency Optimization-Based Joint Resource Allocation and Clustering Algorithm for M2M Communication Systems. IEEE Access 7, 168507–168519 (2019)
20. Azari, A., Miao, G.: Lifetime-aware scheduling and power control for M2M communications in LTE networks. In: IEEE 81st Vehicular Technology Conference, pp. 1–5.IEEE, Piscataway, NJ (2015)
21. Raymon, J.W., Olwal, T.O., Kurien, A.M.: Cooperative communications in machine to machine (M2M): Solutions, challenges and future work. IEEE Access 6, 9750–9766 (2018)
22. Li, Q. et al.: An energy efficient uplink scheduling and resource allocation for M2M communications in SC-FDMA based LTE-A networks. Mobile Networks and Applications 1–12 (2019)
23. Kumar, A., Abdelhadi, A., Clancy, C.: A delay optimal multiclass packet scheduler for general M2M uplink. IEEE Systems Journal 13(4), 3815–3826 (2019)
24. Yang, C., He, J.: Research on scheduling algorithms in uplink of LTE system. Mater. Sci. Eng. 782(3), 032097 (2020)
25. Nobar, S.K. et al.: Uplink resource allocation in energy harvesting cellular network with H2H/M2M coexistence. IEEE Transactions on Wireless Communications 19(8), 5101–5116, (2020)
26. Shakarami, A., Gholbasi-Arani, M., Shahlidinjead, A.: A survey on the computation offloading approaches in mobile edge computing: A machine learning-based perspective. Computer Networks 182, 107496 (2020)
27. Shakarami, A. et al.: A survey on the computation offloading approaches in mobile edge/cloud computing environment: a stochastic-based perspective. Journal of Grid Computing 18, 639–671 (2020)
28. Gholbasi-Arani, M. et al.: An efficient task scheduling approach using moth-flame optimization algorithm for cyber-physical system applications in fog computing. Transactions on Emerging Telecommunications Technologies 31(2), e3770 (2020)
29. Mukherjee, M. et al.: Joint task offloading and resource allocation for delay-sensitive fog networks. In: 2019 IEEE International Conference on Communications, pp. 1–7.IEEE, Piscataway, NJ (2019)
30. Mukherjee, M. et al.: Computation offloading strategy in heterogeneous fog computing with energy and delay constraints. In: 2020 IEEE International Conference on Communications, pp. 1–5.IEEE, Piscataway, NJ (2020)
31. Gholbasi-Arani, M.: A workload clustering based resource provisioning mechanism using biogeography based optimization technique in the cloud based systems. Soft Computing 25, 3813–3830 (2020)
32. Safiulin, I. et al.: Latency and resource utilization analysis for V2X communication over LTE MBSFN transmission. In: 20th International ITG Workshop on Smart Antennas, pp. 1–6.IEEE, Piscataway, NJ (2016)

How to cite this article: Torkudzor MK, Schwarz S, Abdulai Jamal-Deen, Rupp M. Energy efficiency, latency and reliability trade-offs in M2M uplink scheduling. JET Commun. 2021;1–10. https://doi.org/10.1049/cmu2.12201