EAI Endorsed Transactions
on Scalable Information Systems

Research Article EAI.EU

NOMA Assisted Energy-Efficient MEC for Environmental Severity Monitoring in Power IoT Networks

Guangmao Li*, Gang Du, Hongbin Wang, Hongling Zhou, Jie Yang, and Zhikai Pang

Guangzhou Power Supply Bureau of Guangdong Power Grid Co., Ltd., Guangzhou, China

Abstract

This paper proposes an energy-efficient mobile edge computing (MEC) scheme that utilizes non-orthogonal multiple access (NOMA) for environmental severity monitoring in power Internet of Things (IoT) networks. The primary objective of the proposed scheme is to optimize energy consumption while ensuring that tasks are completed within their respective latency and meet the reliability constraints. The scheme integrates NOMA's superposition coding with mobile edge computing to improve task offloading efficiency and reduce computational delay. To achieve this, an iterative water-filling (IWF) algorithm is applied to dynamically adjust the power allocation for each task based on the channel conditions and latency requirements. The optimization problem is formulated to minimize the energy consumption while respecting the given constraints, including outage probability and transmission rate. Simulation results demonstrate that the proposed IWF-based scheme significantly outperforms the traditional ones. For instance, under the latency threshold of 10 ms, the IWF scheme reduces the energy consumption by approximately 30% compared to the conventional ones. Morever, even as the latency threshold increases, the IWF scheme consistently maintains a noticeable advantage, achieving up to 20% lower energy consumption compared to other ones.

Keywords: NOMA, MEC networks, IoT networks, performance analysis

Received on 28 March 2025; accepted on 26 June 2025; published on 15 July 2025

doi:10.4108/eetsis.8980

1. Introduction

In the field of wireless communication, non-orthogonal multiple access (NOMA) has received much attention due to its advantages on both access performance and data rate [1, 2]. NOMA has been integrated with satellite communication systems, where the performance of downlink NOMA in low earth orbit (LEO) satellite systems has been analyzed, revealing significant potential in enhancing system throughput and fairness among users [3–5]. In addition, the outage performance of NOMA system has been analyzed, including paired NOMA in high-speed UAV networks, providing insights into its asymptotic behavior and diversity order, which helps understand the limits of NOMA systems in high-mobility scenarios [6–8]. Moreover, the performance of NOMA with different fading channels, such as Nakagami-m, has been extensively studied, demonstrating how different channel conditions impact the system's outage probability and throughput [9, 10]. Additionally, some studies have analyzed the system-level performance of NOMA in downlink scenarios, especially in vehicle-to-vehicle communication systems, suggesting that NOMA can significantly improve the data throughput compared to traditional orthogonal systems [11, 12]. Performance evaluations have also extended to cooperative NOMA systems, where relay selection and user pairing strategies have been used to further enhance the network performance. In further, NOMA has been combined with massive MIMO and intelligent reflecting surfaces (IRS), showing that NOMA can be integrated with advanced technologies to achieve superior energy efficiency and higher throughput in both uplink and downlink systems [13, 14]. These studies indicate the NOMA's potential to enhance

*Corresponding author. Email: GuangmaoLi@hotmail.com

the wireless communication by providing flexible, efficient, and high-performance solutions in modern network architectures, particularly in dense environments and next-generation 5G systems.

On the other hand, mobile edge computing (MEC) have also gained much attention, particularly focusing on various network applications, optimization techniques, and new technologies that enhance MEC systems [15–17]. The resource allocation, task offloading, and energy efficiency in MEC networks have been investigated to improve the overall performance of edge computing systems [18–20]. In particular, MEC's integration with 5G and beyond has led to significant progress in improving network efficiency by reducing latency, enhancing data throughput, and minimizing the energy consumption [21, 22]. Moreover, MEC can enable real-time computing by distributing computational tasks closer to the end-users, which significantly reduces the latency and enhances the user experience [23, 24]. The performance of MEC has been analyzed under various networking conditions, such as the impact of network congestion, interference, and mobility, with a focus on optimizing the offloading decisions between devices and edge servers [25, 26]. In addition, efficient task offloading algorithms have been proposed for MEC networks, through managing the trade-off between computational resources and energy consumption, and handling the dynamic nature of the wireless communication environment [27]. Moreover, MEC-enabled IoT and Industrial IoT (IIoT) systems have been investigated, with a particular focus on QoS-aware offloading and real-time applications. In further, multi-input multi-output (MIMO) and artificial intelligence (AI)-based techniques have been applied into MEC networks, helping to further optimize the resource management, increase the computational efficiency, and lower the operational costs [27, 28].

Motivated by the above literature review, this paper introduces an energy-efficient mobile edge computing framework that leverages NOMA for environmental severity monitoring in power IoT networks. The proposed scheme aims to optimize the energy usage while ensuring timely completion of tasks within their latency threshold and meeting the reliability requirements. By integrating NOMA's superposition coding with MEC, the proposed scheme enhances the task offloading efficiency and reduces the computational latency. To achieve the optimal performance, an iterative water-filling (IWF) algorithm is used based on dynamically adjusting the power allocation for each task in response to varying channel conditions and latency constraints. The optimization problem is carefully formulated to minimize the energy consumption while ensuring compliance with key constraints, such as outage probability and transmission rate. Simulation results reveal that the IWF-based scheme significantly outperforms traditional ones. Specifically, when subjected to the latency threshold of 10 ms, the proposed IWF scheme achieves a 30% reduction in the energy consumption compared to conventional ones. Moreover, as the latency threshold increases, the IWF scheme consistently delivers superior performance, maintaining up to a 20% reduction in the energy consumption relative to other schemes.

2. System Model

In this work, we study a scenario involving a single user offloading tasks using non-orthogonal multiple access. In this model, an end-user device, which could be a wearable device, sensor, or similar IoT-based gadget, offloads computational tasks to multiple base stations (BSs) to reduce the computational latency and optimize performance. In this scenario, each BS is connected to a dedicated edge server, enabling enhanced computational capabilities for end-user devices. Additionally, all transceivers are assumed to operate with a single antenna configuration.

To formally characterize the offloading process, let K represent the total number of computational tasks, each of which is allocated to a designated base station (BS) for execution. The j-th task is characterized by a tuple of parameters,

$$(L_j^{tot}, S_j^{tot}, D_j, \alpha_j, c_j), \tag{1}$$

where L_j^{tot} represents the total number of Nats contained in the task, indicating the size of the data to be processed, S_j^{tot} denotes the deadline by which both the transmission and the computation must be completed, D_j refers to the number of CPU cycles required to compute a single data for the j-th task, α_j is the target outage probability for task execution reliability, and c_j is the device-to-base-station distance for the assigned task.

Additionally, we adopt a composite wireless channel model that simultaneously captures large-scale path loss and small-scale Rayleigh fading effects. The resulting channel gain between the local device and the j-th BS is,

$$|g_j|^2 = |h_j|^2 c_j^{-\epsilon}, \tag{2}$$

where ϵ denotes the path loss exponent, while $h_j \sim CN(0,1)$ represents the small-scale Rayleigh fading coefficient. We assume that the transmitter has access only to statistical channel state information (CSI), relying on large-scale channel characteristics to describe the quality of the communication link.

Note that the channel gain affects the achievable data rate at which the messages will be transmitted. Instead of transmitting at an instantaneous achievable rate, a predetermined target data rate is used. This

target rate is selected to ensure that the transmission remains reliable under varying channel conditions. To evaluate the transmission reliability, we can use outage probability, defined as the probability that the instantaneous achievable data rate is insufficient to meet the target rate, resulting in a communication failure.

In addition, the offloading and transmission strategy in this setup should be designed to minimize the latency and ensure that the tasks are executed within the given threshold, while meeting the required reliability constraints and optimizing the overall communication and computation efficiency. Moreover, unequal transmission durations are observed across tasks, primarily driven by differences in data size and time sensitivity. Consequently, the signal components transmitted during each phase of the transmission process will change, reflecting these differences.

In this work, non-orthogonal multiple access is applied, wherein task signals are integrated through the use of superposition coding (SC) multiplex several task streams in a single transmission, improving the spectral efficiency. However, in certain phases, where only a single task stream is present, for example, the scenario is reduced to an orthogonal multiple access (OMA). Therefore, the transmission mechanism in this case is a hybrid form of NOMA, combining the characteristics of both pure NOMA and OMA strategies, as described in previous works such as [7] and [17].

Let S_s denote the duration of the s-th time block. Additionally, define $\pi(k)$ as the index of the task assigned to the k-th decoding order in the successive interference cancellation process. The SIC process ensures that the signal corresponding to task $\pi(k)$ is successfully decoded before the signals of subsequent tasks with indices $l > k$ can be decoded. During the s-th time block, the end-user device transmits the following signal,

$$T^s = \sum_{k=1, L_S(\pi(k))\geq s} \sqrt{O_s(\pi(k)) T_{\pi(k)}(k)}, \tag{3}$$

where $T_{\pi(k)}(k)$ denotes the unit-power signal associated with task $\pi(k)$. The variable $O_s(\pi(k))$ specifies the transmit power allocated to this task, while $L_S(\pi(k))$ indicates the number of time blocks assigned for its transmission.

At the receiving base station, the received signal corresponding to task $\pi(k)$ is given by,

$$x^s_{\pi(k)} = g_{\pi(k)} \sum_{k=1, L_S(\pi(k))\geq s} \sqrt{O_s(\pi(k)) T_{\pi(k)}(k)} + l_{\pi(k)}, \tag{4}$$

where $g_{\pi(k)}$ represents the channel gain associated with task $\pi(k)$, and $l_{\pi(k)}$ is the additive white Gaussian noise (AWGN) introduced in the communication channel, with variance β^2.

The achievable data rate for task $\pi(m)$ in the SIC process, where $m \geq k$, is given by,

$$D^s_{\pi(m)\to\pi(k)} = \ln\left(1 + \frac{|g_{\pi(m)}|^2 O_s(\pi(k))}{|g_{\pi(m)}|^2 \sum_{n\in\mathcal{N}^s_{\mathrm{IF}}(k)} O_s(\pi(n)) + \beta^2}\right), \tag{5}$$

where $\mathcal{N}^s_{\mathrm{IF}}(k)$ represents the set of task indices corresponding to the interfering signals for task $\pi(k)$. This set is defined as,

$$\mathcal{N}^s_{\mathrm{IF}}(k) = \{n \mid n > k, L_S(\pi(n)) \geq s\}, \tag{6}$$

where task $\pi(1)$ experiences substantial interference caused by concurrent transmissions of tasks $\pi(2)$ and $\pi(3)$ during the initial time slot. Thus, we have,

$$\mathcal{N}^1_{\mathrm{IF}}(1) = \{2, 3\}. \tag{7}$$

The communication outage occurs when the transmission signal $T^s_{\pi(k)}$ for task $\pi(k)$ cannot be decoded correctly at the receiver. The outage probability for the task $\pi(m)$ at the s-th base station is calculated as:

$$O^{\pi(m)}_{\mathrm{out},s} = 1 - \Pr\left\{\bigcap_{k=1, L_S(\pi(k))\geq s} D^s_{\pi(m)\to\pi(k)} \geq Q^s_{\pi(k)}\right\}, \tag{8}$$

where $Q^s_{\pi(k)}$ is the target rate for task $\pi(k)$ during the s-th time block.

3. Proposed Offloading for MEC Networks

As discussed earlier, efficient rate allocation across distinct transmission phases plays a critical role in overall system performance, particularly when the tasks have varying urgency levels. The primary objective of the optimization is to minimize the total energy consumption, subject to the latency and outage constraints. Given the signal superposition mechanism inherent in non-orthogonal multiple access, the transmission rate, decoding order, and signal power should be jointly optimized. This leads to the following optimization,

$$P_0: \min_{\pi,Q,O} I = \sum_{k=1}^{K} \sum_{s=1}^{L_S(\pi(k))} O_s(\pi(k)) S_s, \tag{9a}$$

$$\text{s.t.} \quad \sum_{s=1}^{L_S(\pi(k))} Q_s(\pi(k)) S_s A = L_{\mathrm{mec}}(\pi(k)), \quad \forall k, \tag{9b}$$

$$O^{\pi(k)}_{\mathrm{out},s} \leq \alpha_{\pi(k)}, \quad \forall k, s, \tag{9c}$$

$$Q^{\pi(k)}_s \geq 0, \quad \forall k, s. \tag{9d}$$

Here, the objective function I minimizes the total energy consumption. The constraint (9b) ensures that the data transmission of each task should meet the given deadline. The constraint (9c) enforces the outage probability, ensuring that each task is decoded successfully under the reliability requirements. The constraint (9d) ensures that the data rate is non-negative.

The successive interference cancellation order plays a crucial role in determining the order in which signals are decoded. The decoding process must account for both channel gains and outage requirements. The order in which signals are decoded influences the energy consumption because higher-level signals can cause interference, thereby affecting the decoding of lower-level signals. In this work, we assume that the SIC order follows the optimal decoding order proposed in [13], which ensures that the total energy consumption is minimized. Specifically, the decoding order is given by,

$$\zeta_{\pi^*(1)} \geq \zeta_{\pi^*(2)} \geq \cdots \geq \zeta_{\pi^*(K)}, \tag{10}$$

where the optimal SIC order $\pi^*(k)$ satisfies the following relation,

$$\zeta_{\pi^*(k)} = \frac{c_{\pi^*(k)}^{-\epsilon}}{\ln(1-\alpha_{\pi^*(k)})}. \tag{11}$$

This order minimizes the total energy consumption while adhering to the constraints imposed by the system.

After sorting the tasks according to the optimal SIC order, the signal power allocation for each task at the s-th time block is given by,

$$O_s^k = \rho_k^s \left(\zeta_k + \sum_{n \in \mathcal{N}_{\mathrm{IF}}(k)} \zeta_n \rho_n^s \prod_{l \in \mathcal{N}_{\mathrm{IF}}(k,n)} f^{Q_l^s} \right), \tag{12}$$

where $\mathcal{N}_{\mathrm{IF}}(k)$ is the set of interfering tasks for task $\pi(k)$, and Q_l^s denotes the transmission rate allocated to the l-th task during the s-th time block.

The problem formulation P_1, which minimizes the total energy under the constraints from (9b) and (9d), is convex. This convexity can be proven using the Karush-Kuhn-Tucker (KKT) conditions, which are instrumental in finding the optimal rate allocation solution. Specifically, it suffices to show that the equation in (9a) is convex. This is because the expressions in (9b) and (9d) are affine functions, which are known to maintain convexity. Specifically, we need to compute the partial derivative of the objective function with respect to Q_i^F to determine the behavior of the function with respect to this parameter. The goal is to show that the derivative of the objective function with respect to the transmission rate Q_i^F equals zero for $k > i$, where k is an index representing the decoding order of the signals.

As to the derivation of the partial derivative, we begin by considering the derivative expression that needs to be computed,

$$\frac{\partial I}{\partial Q_i^F} = \left(\sum_{k=1}^{K} L_S(k) \geq s \right) \frac{\partial O_i^S}{\partial Q_i^F} S_s = 0. \tag{13}$$

This result arises because the transmission data rate associated with signals that have a higher decoding order do not affect the transmit power of those with a lower decoding order. Therefore, the derivative is set to zero for $k > i$.

When $k = i$, we have,

$$\frac{\partial O_i^S}{\partial Q_i^F} = f^{Q_i^F} \left(\zeta_i + \sum_{n \in N_i} \zeta_n \rho_n^S \prod_{l \in N_i(k,n)} f^{Q_i^F} \right). \tag{14}$$

For the case when $k = i - 1$, we can compute the derivative as,

$$\frac{\partial O_i^S}{\partial Q_i^F} = \left(f^{Q_i^F - 1} - 1 \right) \left(\zeta_i f^{Q_i^F} + \sum_{n \in N_i} \zeta_n \left(f^{Q_i^F} - 1 \right) \prod_{l \in N_i(k,n)} f^{Q_i^F} \right), \tag{15}$$

which is simplified to

$$\frac{\partial O_i^S}{\partial Q_i^F} = f^{Q_i^F - 1} \left(\zeta_i + \sum_{n \in N_i} \zeta_n \left(f^{Q_i^F} - 1 \right) \prod_{l \in N_i(k,n)} f^{Q_i^F} \right). \tag{16}$$

Next, combining the results from both cases, we have,

$$\frac{\partial O_i^S}{\partial Q_i^F} + \frac{\partial O_i^S}{\partial Q_i^F} = f^{Q_i^F - 1} f^{Q_i^F} \left(\zeta_i + \sum_{n \in N_i} \zeta_n \left(f^{Q_i^F} - 1 \right) \prod_{l \in N_i(k,n)} f^{Q_i^F} \right). \tag{17}$$

Thus, the final expression for the partial derivative is,

$$\frac{\partial O_i^S}{\partial Q_i^F} + \frac{\partial O_i^S}{\partial Q_i^F} = f^{Q_i^F - 1} \left(\zeta_i + \sum_{n \in N_i} \zeta_n \left(f^{Q_i^F} - 1 \right) \prod_{l \in N_i(k,n)} f^{Q_i^F} \right). \tag{18}$$

By applying the same method from $i = K$ to $i = 1$, we can have,

$$\frac{\partial I}{\partial Q_i^F} = \prod_{s=1}^{i} f^{Q_i^F} \left(\zeta_i + \sum_{n \in N_i} \zeta_n \left(f^{Q_i^F} - 1 \right) \right) \prod_{l \in N_i(k,n)} f^{Q_i^F} S_s. \tag{19}$$

Thus, the required derivative has been calculated for the transmission power with respect to the transmission

rate Q_i^F. This derivative is crucial for proving the convexity of the objective function. The KKT conditions for the optimization problem are given by:

$$\frac{\partial I}{\partial Q_k^s} - \eta_k S_s \geq 0, \quad \forall k, \tag{20}$$

$$\left(\frac{\partial I}{\partial Q_k^s} - \eta_k S_s\right) Q_k^s = 0, \quad \forall k, \tag{21}$$

$$\sum_{s=1}^{L_S(\pi(k))} Q_k^s S_s A - L_{\text{mec}}(\pi(k)) = 0, \tag{22}$$

where η_k is the dual variable associated with the constraint (9b). The Lagrange multipliers η_k should satisfy the complementary slackness condition, which ensures the optimality of the solution. The derived value of Q_k^s follows a water-filling method, given by,

$$Q_k^s = (\ln \eta_k - \ln b_k^s)^+, \tag{23}$$

where b_k^s is a function of the signal power and rate allocation of other tasks, and z^+ represents the positive part of z.

To solve the optimization problem, we employ the iterative water-filling algorithm (IWF), which alternates the adjustment of data rates Q_k^s and the waterline η_k. In each iteration, the power allocation and data rate calculations are updated until convergence, where the energy consumption is minimized, and the transmission data rates should satisfy the given constraints. As to the convergence and complexity, the majority of studies on iterative water-filling (IWF) primarily concentrate on optimizing power allocation to maximize the overall sum-rate, while this work emphasizes minimizing energy consumption. The total energy consumption decreases in each iteration, as it is non-increasing with respect to the transmission rates associated with other tasks. The complexity of the proposed algorithm is linear in the number of tasks K, making it suitable for large-scale task offloading problems. Consequently, the IWF-based rate allocation approach effectively scales to scenarios involving numerous offloading tasks, all while preserving computational efficiency. The whole procedure of the proposed IWF scheme can be found in Algorithm 1.

Algorithm 1 Transmission data rate calculation for power IoT networks

Require: Data size $L_{\text{mec}}(\pi(k))$, task information ζ_k, time duration S_s
Ensure: Transmission data rates Q_k^s
1: **Initialize:**
2: Sort the tasks, and initialize the data rates $Q_k^s = 0$
3: **for** $k = 1, \ldots, K$ **do**
4: **for** $s = 1, \ldots, L_S(k)$ **do**
5: Calculate b_k^s for $s = 1, \ldots, L_S(k)$
6: Update the waterline η_k and data rates Q_k^s for each s and the water-filling method
7: **end for**
8: **end for**
9: **Energy Calculation:**
10: Calculate the total energy I
11: **Convergence:**
12: Repeat steps 2 to 4 until the energy I converges
return Transmission data rates Q_k^s

4. Simulation Results and Discussions

In this part, some simulations are presented to validate the proposed scheme for the NOMA based MEC networks. The path loss exponent is set to 3, and the channel experiences Rayleigh fading with zero-mean unit-variance complex Gaussian coefficients. The distances between the user and the base stations vary for different tasks and are used to calculate large-scale fading. The data size for each task is defined in Nats, and the latency threshold is varied between 0.65 seconds and 0.95 seconds to evaluate performance under different time constraints. The outage probability targets are set between 0.01 and 0.05 to reflect different reliability levels. The noise power is normalized, and the transmission scheme combines both NOMA and OMA depending on the number of active streams in each time block. Transmission occurs in multiple unequal time blocks due to task-specific deadlines and data sizes. The decoding order of tasks in successive interference cancellation is determined by a reliability-aware metric based on the path loss and outage requirement. The proposed IWF-based algorithm jointly optimizes transmission power, data rates, and decoding order, and iteratively adjusts parameters until convergence. These settings collectively reflect a realistic and scalable NOMA-MEC environment designed to evaluate energy-efficient offloading strategies under latency and reliability constraints.

Fig. 1 and Table I illustrate the energy consumption of several schemes as a function of the latency threshold, where the latency threshold varies from 0.65s to 0.95s. From this figure and table, we can find that the proposed IWF scheme consistently outperforms conventional ones, particularly at lower latency thresholds. For instance, at the latency threshold of 10 ms, the proposed IWF scheme reduces energy consumption by approximately 30% compared to the conventional ones, showcasing its superior ability to optimize the power usage while meeting the stringent latency requirement.

Table 1. Numerical energy consumption of several schemes versus the latency threshold.

Latency threshold	IWF-based (Optimal order)	Uniform rate (Optimal order)	Uniform rate (Random order)
0.65	**0.6796**	0.7627	0.8269
0.70	**0.5031**	0.5564	0.6047
0.73	**0.4454**	0.4844	0.5300
0.78	**0.4192**	0.4483	0.4940
0.82	**0.4081**	0.4266	0.4726
0.86	**0.4000**	0.4119	0.4598
0.90	**0.3933**	0.4010	0.4493
0.94	**0.3891**	0.3924	0.4446

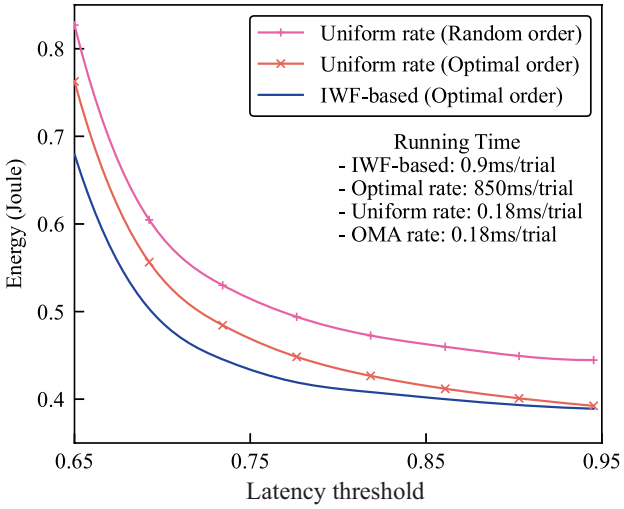

Figure 1. Energy consumption of several schemes versus the latency threshold.

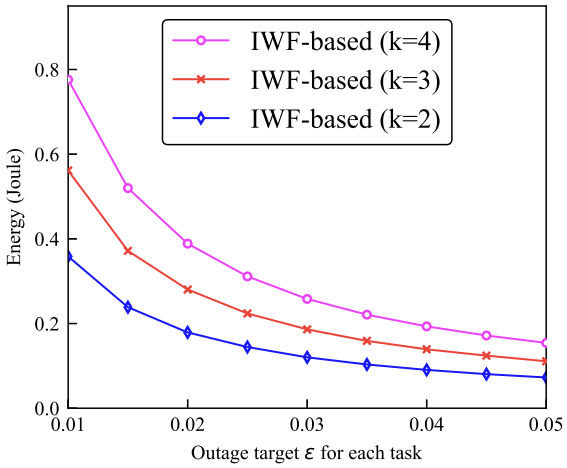

Figure 2. Energy consumption of several schemes versus the outage target.

As the latency threshold increases, the energy consumption gap narrows, but the proposed IWF scheme maintains a noticeable advantage, achieving around 15-20% lower energy consumption even at higher thresholds (e.g., 50 ms). This efficiency stems from the IWF scheme's dynamic power allocation, which iteratively adjusts transmit power based on channel conditions and latency constraint, unlike conventional schemes that rely on static or suboptimal power allocation. The results in this figure and table indicate the IWF scheme's adaptability and effectiveness in balancing the energy efficiency and latency, making it a robust solution for latency-sensitive applications.

Fig. 2 and Table II illustrate the energy consumption performance of several schemes as a function of the outage target ε for each task, where three different configurations of the IWF-based approach with varying values of parameter k, specifically $k = 2$, $k = 3$, and $k = 4$, are investigated. It is evident from this figure and table that the energy consumption of the proposed scheme decreases as the outage target ε increases across all parameters of k, indicating a trade-off between reliability (lower outage) and energy efficiency. In particular, the proposed scheme with $k = 2$ consistently achieves the lowest energy consumption, demonstrating a superior energy efficiency. For instance, when the outage target is set at $\varepsilon = 0.01$, the energy consumption with $k = 2$ is approximately 0.36 Joules, while the values for $k = 3$ and $k = 4$ are about 0.56 Joules and 0.78 Joules, respectively. This gap persists across the range of ε: at $\varepsilon = 0.03$, the energy usage for $k = 2$ drops to about 0.13 Joules, compared to roughly 0.22 Joules for $k = 3$ and 0.28 Joules for $k = 4$. Even at the higher outage target of $\varepsilon = 0.05$, $k = 2$ maintains its advantage with energy consumption below 0.1 Joules, while $k = 3$ and $k = 4$ still consume more energy, at around 0.13 and 0.16 Joules respectively.

Fig. 3 and Table III present the data rate of several schemes as a function of the latency threshold, where the data rate is measured in Nat/s/Hz and the latency threshold ranges from 0.65s to 0.95s. We can find from this figure and table that the OMA-based scheme demonstrates the highest growth in

Table 2. Numerical energy consumption of several schemes versus the outage target.

Energy consumption	IWF-based $k = 2$	IWF-based $k = 3$	IWF-based $k = 4$
0.010	**0.3584**	0.5615	0.7760
0.015	**0.2386**	0.3717	0.5199
0.020	**0.1790**	0.2803	0.3887
0.025	**0.1445**	0.2238	0.3112
0.030	**0.1202**	0.1862	0.2581
0.035	**0.1033**	0.1590	0.2209
0.040	**0.0905**	0.1391	0.1933
0.045	**0.0805**	0.1242	0.1717
0.050	**0.0726**	0.1109	0.1543

Table 3. Numerical data rate of several schemes versus the latency threshold.

Latency threshold	IWF-based R_2	OMA R_2	Uniform R_2
0.65	**0.3787**	0.3787	0.303
0.70	**0.4215**	0.4215	0.303
0.73	**0.4977**	0.4977	0.303
0.78	**0.61**	0.61	0.303
0.82	**0.61**	0.72	0.303
0.86	**0.61**	0.94	0.303
0.90	**0.61**	1.36	0.303
0.94	**0.61**	2.44	0.303

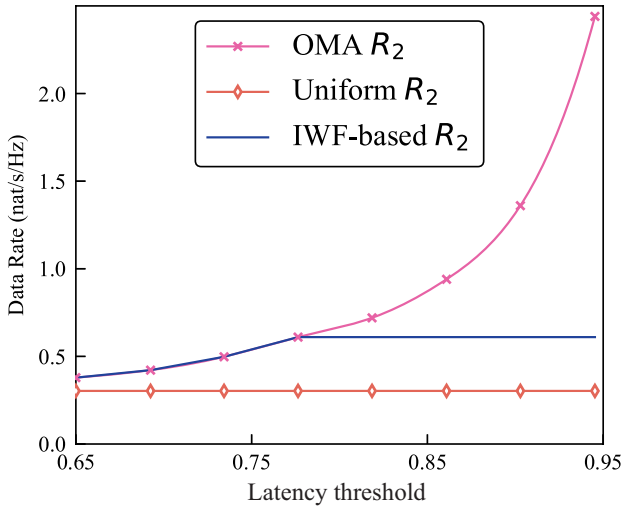

Figure 3. Data rate of several schemes versus the latency threshold.

data rate as the latency threshold increases, rapidly escalating after the latency threshold exceeds around 0.85s. In contrast, the Uniform-based scheme shows a much slower increase, with the data rate remaining relatively constant at approximately 0.5 Nat/s/Hz across the entire latency range. Moreover, the proposed IWF-based scheme maintains a consistent and higher data rate compared to the Uniform-based scheme, remaining flat at around 0.5 Nat/s/Hz for lower latency thresholds but demonstrating a modest rise as the latency threshold increases. Specifically, at the latency threshold of 0.65s, both OMA-based and IWF-based schemes start at a similar value of around 0.5 Nat/s/Hz, while the Uniform-based scheme remains consistently lower at approximately 0.45 Nat/s/Hz. As the latency threshold increases to 0.85s, the OMA-based scheme reaches about 2.0 Nat/s/Hz, while the IWF-based scheme remains stable below 1.0 Nat/s/Hz, indicating a more moderate response to latency relaxation. Overall, the IWF-based scheme offers a stable data rate that does not escalate dramatically with increased latency, distinguishing it from the OMA-based approach, which benefits significantly from relaxed latency thresholds, making the proposed IWF-based scheme a more stable, yet less latency-sensitive solution compared to the aggressive scaling seen in the OMA-based scheme.

5. Conclusions

This paper presented an energy-efficient MEC framework that utilized NOMA for environmental severity monitoring in power IoT networks. The primary goal of the proposed scheme was to optimize energy consumption while ensuring that tasks were completed within their respective deadlines and met reliability requirements. By integrating NOMA's superposition coding with MEC, the scheme improved task offloading efficiency and reduced computational latency. To achieve this, an IWF algorithm was employed to dynamically

adjust the power allocation for each task, accounting for varying channel conditions and latency constraints. The optimization problem was formulated to minimize energy consumption while adhering to essential constraints, including outage probability and transmission rate. Simulation results demonstrated that the IWF-based scheme significantly outperformed traditional schemes. Specifically, under the strict latency of 10ms, the IWF scheme reduced energy consumption by approximately 30% compared to conventional ones. Furthermore, even as the latency threshold increased, the IWF scheme consistently maintained a notable advantage, achieving up to 20% lower energy consumption compared to other schemes. These results underscored the IWF-based scheme's effectiveness in balancing energy efficiency with latency-sensitive performance in power IoT networks.

References

[1] S. Islam, N. Avazov, and O. Dobre, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 3, pp. 1471–1492, 2016.

[2] A. Akbar, S. Jangsher, and F. Bhatti, "NOMA and 5G emerging technologies: A survey on issues and solution techniques," *Computer Networks*, vol. 184, p. 107538, 2021.

[3] S. Islam, J. Zeng, and R. Liu, "Investigation on evolving single-carrier NOMA into multi-carrier NOMA in 5G," *IEEE Access*, vol. 8, pp. 218 462–218 473, 2018.

[4] Y. Liu, W. Yi, and Z. Ding, "Evolution of NOMA toward next generation multiple access (NGMA) for 6g," *IEEE Wireless Communications*, vol. 29, no. 6, pp. 16–23, 2022.

[5] J. Zeng, T. Lv, and R. Liu, "Investigation on evolving single-carrier NOMA into multi-carrier NOMA in 5G," *IEEE Access*, vol. 6, pp. 72 285–72 297, 2018.

[6] F. Zhou, Y. Wu, and Y. Liang, "Is NOMA efficient in multi-antenna networks? a critical look at next generation multiple access techniques," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 62–70, 2021.

[7] M. Vaezi, Y. Liu, and O. Dobre, "Cooperative NOMA: State of the art, key techniques, and open challenges," *IEEE Network*, vol. 34, no. 6, pp. 42–48, 2020.

[8] M. Liaqat, K. Noordin, and T. A. Latef, "Power-domain non orthogonal multiple access (PD-NOMA) in cooperative networks: an overview," *Wireless Networks*, vol. 26, no. 5, pp. 2741–2755, 2020.

[9] B. Makki, A. Behravan, and O. Dobre, "A comprehensive survey on cooperative NOMA networks," *IEEE Access*, vol. 7, pp. 173 532–173 546, 2019.

[10] S. Zhang, Y. Liu, and Z. Ding, "Key techniques in NOMA: Resource allocation and optimization," *IEEE Communications Magazine*, vol. 59, no. 8, pp. 14–22, 2021.

[11] B. Clerckx, Y. Mao, and R. Schober, "Is NOMA efficient in multi-antenna networks? a critical look at next generation multiple access techniques," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 297–308, 2021.

[12] M. Ali, Y. Liu, and E. Hossain, "Non-orthogonal multiple access (NOMA) in cellular uplink and downlink: Challenges and enabling techniques," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 45–51, 2019.

[13] M. Abd-Elnaby, G. Sedhom, and E. El-Rabaie, "NOMA for 5G and beyond: Literature review and novel trends," *Wireless Networks*, vol. 29, pp. 3597–3610, 2023.

[14] H. Tabassum, M. Ali, and E. Hossain, "Non-orthogonal multiple access (NOMA) in cellular uplink and downlink: Challenges and enabling techniques," *arXiv preprint arXiv:1608.05783*, 2016.

[15] Y. Hu, M. Patel, and D. Sabella, "Mobile edge computing-a key technology towards 5G," *ETSI White Paper*, 2015.

[16] Y. Mao, C. You, and J. Zhang, "Mobile edge computing: Survey and research outlook," *arXiv preprint arXiv:1701.01090*, 2017.

[17] H. Li, G. Shou, and Y. Hu, "Mobile edge computing: Progress and challenges," *IEEE Access*, vol. 4, pp. 2957–2969, 2016.

[18] M. Yagoub, A. Jarray, and L. Haibeh, "A survey on mobile edge computing infrastructure: Design, resource management, and optimization approaches," *IEEE Access*, vol. 10, pp. 5095–5113, 2022.

[19] T. Taleb, A. Ksentini, and M. Iqbal, "Mobile edge computing potential in making cities smarter," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 12–18, 2017.

[20] S. Wang, J. Xu, and N. Zhang, "A survey on service migration in mobile edge computing," *IEEE Access*, vol. 6, pp. 28 787–28 798, 2018.

[21] X. Sun and N. Ansari, "EdgeIoT: Mobile edge computing for the Internet of Things," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 1–6, 2016.

[22] T. Tran, A. Hajisami, and P. Pandey, "Collaborative mobile edge computing in 5G networks: New paradigms, scenarios, and challenges," *IEEE Communications Surveys and Tutorials*, vol. 19, no. 3, pp. 1834–1852, 2017.

[23] R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges," *Future Generation Computer Systems*, vol. 79, pp. 366–378, 2018.

[24] E. Ahmed and M. Rehmani, "Mobile edge computing: Opportunities, solutions, and challenges," *Future Generation Computer Systems*, vol. 79, pp. 440–448, 2017.

[25] S. Wang, C. Hsu, and J. Xu, "Edge server placement in mobile edge computing," *Future Generation Computer Systems*, vol. 81, pp. 149–156, 2018.

[26] Y. Zhang, A. Taherkordi, and I. Yaqoob, "Mobile edge computing: A survey," *IEEE Internet of Things Journal*, vol. 4, no. 3, pp. 539–550, 2017.

[27] W. Khan, E. Ahmed, and S. Hakak, "Edge computing: A survey," *Future Generation Computer Systems*, vol. 94, pp. 17–35, 2019.

[28] R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing: A survey on architecture and computation offloading," *Future Generation Computer Systems*, vol. 87, pp. 62–80, 2018.