

Article

A Resource Allocation Scheme for Packet Delay Minimization in Multi-Tier Cellular-Based IoT Networks

Jin Li ^{1,2,*} , Wenyang Guan ^{1,2} and Zuoyin Tang ³ 

¹ School of Artificial Intelligence, Beijing Technology and Business University, Beijing 100048, China; 20151005@btbu.edu.cn

² Key Laboratory of Industrial Internet and Big Data, China National Light Industry, Beijing Technology and Business University, Beijing 100048, China

³ School of Computer Science and Digital Technologies, Aston University, Birmingham B4 7ET, UK; z.tang1@aston.ac.uk

* Correspondence: jinli@btbu.edu.cn

Abstract: With advances in Internet of Things (IoT) technologies, billions of devices are becoming connected, which can result in the unprecedented sensing and control of the physical environments. IoT devices have diverse quality of service (QoS) requirements, including data rate, latency, reliability, and energy consumption. Meeting the diverse QoS requirements presents great challenges to existing fifth-generation (5G) cellular networks, especially in unprecedented scenarios in 5G networks, such as connected vehicle networks, where strict data packet latency may be required. The IoT devices with these scenarios have higher requirements on the packet latency in networking, which is essential to the utilization of 5G networks. In this paper, we propose a multi-tier cellular-based IoT network to address this challenge, with a particular focus on meeting application latency requirements. In the multi-tier network, access points (APs) can relay and forward packets from IoT devices or other APs, which can support higher data rates with multi-hops between IoT devices and cellular base stations. However, as multiple-hop relaying may cause additional delay, which is crucial to delay-sensitive applications, we develop new schemes to mitigate the adverse impact. Firstly, we design a traffic-prioritization scheduling scheme to classify packets with different priorities in each AP based on the age of information (AoI). Then, we design different channel-access protocols for the transmission of packets according to their priorities to ensure the QoS in networking and the effective utilization of the limited network resources. A queuing-theory-based theoretical model is proposed to analyze the packet delay for each type of packet at each tier of the multi-tier IoT networks. An optimal algorithm for the distribution of spectrum and power resources is developed to reduce the overall packet delay in a multi-tier way. The numerical results achieved in a two-tier cellular-based IoT network show that the target packet delay for delay-sensitive applications can be achieved without a large cost in terms of traffic fairness.

Keywords: AoI; multiple access protocol; packet transmission delay; resources allocation; cellular-based IoT network; 5G

MSC: 37M99



Citation: Li, J.; Guan, W.; Tang, Z. A Resource Allocation Scheme for Packet Delay Minimization in Multi-Tier Cellular-Based IoT Networks. *Mathematics* **2023**, *11*, 4538. <https://doi.org/10.3390/math11214538>

Academic Editor: Daniel-Ioan Curiac

Received: 1 October 2023

Revised: 24 October 2023

Accepted: 26 October 2023

Published: 3 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the rapid growth of wireless services beyond fifth-generation (5G) technology and the proliferation of innovative applications, a massive number of devices have been deployed to collect and share data to achieve their full potential; this is known as the Internet of Things (IoT) [1,2]. IoT technology plays a key role in modern wireless communication systems. IoT networks are expected to realize features like local decision making and/or remote monitoring and control by utilizing sensory mechanisms in real-time applications by deploying a massive number of devices [3]. IoT technologies, such as smart cities,

vehicular communication, autonomous driving, and remote health care, demand a high amount of data traffic for wireless networks and are growing explosively [2]. In this context, providing reliable wireless access for the massive IoT devices and the diversification of data traffic is essential to the realization of 5G IoT networks and has attracted considerable attention from academia and industry [2,4]. Cellular-based IoT networks are considered the most promising potential solution to provide last-mile connectivity for IoT devices due to their cost-effective deployments and guaranteed services, such as high scalability, diversity, and security, which do not need massive additional infrastructure deployments and significant changes in current cellular networks [2,5].

In cellular-based IoT networks, massive IoT devices initially access the cellular base station (BS) for network connections directly or through IoT gateways by performing a random access (RA) procedure [5]. However, the massive number of IoT devices simultaneously access and uplink the transmission of small data packets and even need to frequently forward the real-time updating status information. Then, preamble collision among IoT devices is inevitable, and as such, improving the access mechanism of existing cellular systems for better quality of service (QoS) with less power consumption of IoT devices is one of the key challenges for cellular-based IoT networks [2,5,6].

In IoT networks, a broad range of IoT services are enabled by massive amounts of IoT devices to monitor environmental factors such as temperature and humidity [4,7]. In this case, IoT devices need to process different data traffic streams for various receivers, where each traffic stream has different performance metrics, such as delay, throughput, or AoI, in characterizing the timeliness [1,2,4]. Then, there is an unavoidable need for designing innovative data-collection mechanisms for cellular-based IoT networks to enhance efficiency and scalability [7]. Due to the delay-tolerant and uplink-preferable characteristics of IoT traffic with small data packets, the main access technology to request the channel resources in the uplink transmission is the contention-based RA [2]. Meanwhile, different IoT services have different QoS requirements based on the properties of the traffic data and functionality. Then, providing access mechanisms with satisfactory network performance for a massive number of IoT devices and the corresponding variable network services is still a key challenge. However, IoT devices repeatedly generate updated information regarding the environmental factors being observed and transmit them to the BSs [3]. Due to the properties of IoT networks, the purpose of frequent updating of the information status is to keep the information as fresh as possible, which means the delay is as minimal as possible. In [8], the authors introduced the idea of the Age of Information (AoI), which measures the freshness of information, and they discussed the possibility of utilizing the efficient design of freshness-aware IoT networks. Then, a series of works utilized the results of [8] by characterizing the temporal mean of AoI or other freshness-related metrics for various variations in queuing models. In [4], the authors applied the concept of AoI to measure the freshness of information at the BSs regarding the random processes monitored by IoT devices, while [1] studied the AoI of K-tier cellular-based IoT networks while considering weighted path loss association policy and fractional power control strategy. However, artificial designs, like utilizing timing advance (TA) information to avoid collisions, can result in unfairness, which is not considered in most existing studies.

In this paper, we propose a multi-tier cellular-based IoT network with a queuing theory-based theoretic model and utilize a traffic-prioritization scheduling scheme to investigate minimum packet transmission delay to meet particular application latency requirements on AoI. Firstly, a multi-tier hierarchical cellular-based IoT network is modeled by utilizing Voronoi tessellation. Secondly, data packets from each IoT device are allocated different priorities based on their AoI in real time. A multiple access protocol is designed for data packets to access the uplink channel and the data packet with the highest priority without the regulatory requirement for listen-before-talk (LBT), while the others have carrier-sense multiple access with a collision-avoidance (CSMA/CA) mechanism. Thirdly, the mean packet-transmission delay is obtained for different traffic packets by analyzing the M/G/1 queue in each AP. Fourthly, a mean packet total delay minimization problem

by optimally allocating network resources for a two-tier cellular-based IoT network is given and solved by utilizing the gradient descent method with the bisection method. The numerical results demonstrate the superior effectiveness of the proposed mechanism and algorithm, which not only achieve the minimum mean packet delay but also give consideration to delay-sensitive traffic based on the AoI metric.

The remaining parts of this paper are structured as follows. In Section 2, we review the related literature. In Section 3, we model the system for a cellular-based IoT network with multi-tier APs and investigate the queuing model in each AP to achieve mean packet delay for packets with different priorities. In Section 4, we present the problem formulation to minimize mean total packet delay in a multi-tier cellular-based IoT network and solve it. In Section 5, we assess the effectiveness of the proposed mechanism by the mean packet delay in a two-tier cellular-based IoT network by showing numerical results, followed by conclusions in Section 6.

2. Related Literature

In [9], a queuing-theory-based model that allows for cross-layered optimization was proposed to investigate the possibility of leveraging multi-RAT to reduce transmission delay without losing the requisite QoS and maintaining the freshness of the received information via the AoI metric in the multi-hop IoT networks. Meanwhile, previous research [10] proposed a three-dimensional resource allocation algorithm for maximizing the total throughput in a cognitive radio IoT network with simultaneous wireless information and power transfer. In [11], the authors maximized the energy efficiency and spectral efficiency of the space-air-ground IoT networks by optimizing subchannel selection, power control, and UAV position deployment. Refs. [12–14] investigated the resource allocation in cellular IoT networks, which revealed the rising attention of researchers given to next-generation cellular networks and IoT techniques. The authors in [15] maximized the energy efficiency in a UAV-based network that collects information from smart devices, sensor devices, and IoT devices by optimizing UAV trajectory, power allocation, and time slot assignment. However, as the most distinguishing feature, diverse QoS requirements, especially the strict data packet latency at difficulty scenarios with multi-hop transmissions, are not comprehensively investigated in the existing literature, which requires a more elaborately designed optimization mechanism based on the AoI metric, and this need motivates this paper. In [16], authors studied the resource allocation optimization mechanism to minimize mean packet transmission delay in a three-dimensional cellular network with multi-layer UAV networks. Furthermore, in [17], data packets were classified into incumbent packets and relayed packets based on the data source locations. Packet delays for various packet classes at each layer of a multi-layer UAV network were investigated, and minimum total packet delay was achieved by optimally allocating spectrum and power resources among layers of the UAV network. However, all the relayed packets in each UAV for relaying to the macrocell base station were allocated with the same priority to be served in the queuing model, which is less efficient in utilizing limited radio resources with guaranteed network performance, especially for the packets that already experienced large delay that accumulated at the previous relaying UAV layers. Meanwhile, multiple-hop relaying causes a delay penalty to packets, which is fatal to delay-sensitive packets in life-limited UAV networks. Then, it is essential to investigate and guarantee packet delay performance for all the relayed packets in the hierarchical multi-tier networks based on the AoI metric, which are the main contributions and innovations of this paper. Furthermore, we extend the technologies into a different but more general scenario, which is the cellular-based IoT networks, and it can be extended to the scenarios for heterogeneous data with different delay requirements based on the preference of the network designer.

3. System Model

In this paper, we investigate a cellular-based IoT network with multi-tier APs, which is denoted by a graph $\mathbf{G}(\{\mathbf{M}, \mathbf{A}\}, \mathbf{D})$. In \mathbf{G} , $\mathbf{M} = \{m_1, \dots, m_M\}$ is the set of M MBSs,

$\mathbf{A} = \{a_1, \dots, a_A\}$ is the set of A APs, and $\mathbf{D} = \{d_1, \dots, d_D\}$ is the set of D IoT devices. The nodes in \mathbf{M} , \mathbf{A} and \mathbf{D} follow Poisson point process (PPP) distribution with intensities δ_m , δ_a , and δ_d . The IoT devices transmit data packets to the corresponding MBS through several AP relays due to the limited transmission range in IoT networks. Then, as shown in Figure 1, each AP tier in the network consists of the APs with the same number of transmission counts from the MBS.

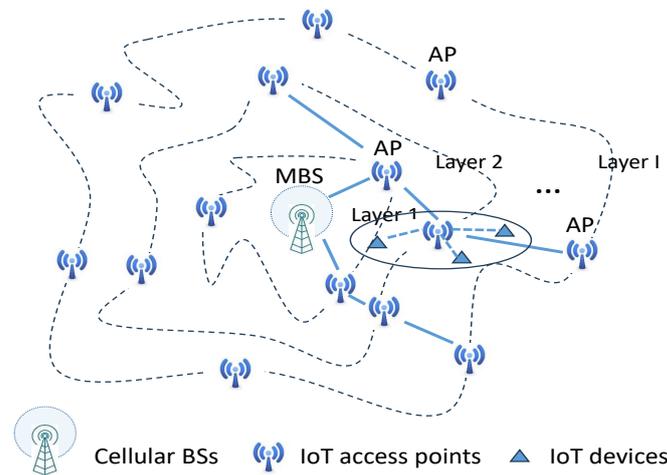


Figure 1. Multi-tier cellular-based IoT network model.

3.1. Channel Model

In cellular-based IoT networks, the communication environment is complex, and it is impossible to achieve ideal LOS communication channels for all the IoT devices, even with the utilization of multi-tier APs to extend network coverage and improve network connectivity. To achieve reliable analysis results for general communication scenarios, this paper models propagation channels by considering LOS and non-LOS components along with their occurrence probabilities separately in (1). The LOS connection probability is decided by the communication environment, height, and density of surrounding buildings, and the elevation angle between the IoT devices and APs along locations of them [18]. Depending on the LOS or non-LOS connection between the IoT device and AP, the received power $p_{jj'}^r$ at node j' from node j is shown as

$$p_{jj'}^r = \begin{cases} p_{jj'}^t g_{jj'} d_{jj'}^{-\alpha}, & \text{LOS link,} \\ \eta p_{jj'}^t g_{jj'} d_{jj'}^{-\alpha}, & \text{Non-LOS link,} \end{cases} \quad (1)$$

where η is an additional attenuation factor due to the non-LOS connection and $\forall j, j' \in \mathbf{M} \cup \mathbf{A} \cup \mathbf{D}, j \neq j'$.

This paper utilizes the Rayleigh fading model in non-LOS paths for both desired signal and interference to achieve reliable analysis results in general communication scenarios, which leads to a random channel power gain following an exponential distribution with unit mean. Then the received power at node j' from node j , $p_{jj'}^r$, is shown as [16]

$$p_{jj'}^r = p_{jj'}^t g_{jj'} d_{jj'}^{-\alpha}, \quad \forall j, j' \in \mathbf{M} \cup \mathbf{A} \cup \mathbf{D}, j \neq j' \quad (2)$$

where power transmitted from node j to node j' is the $p_{jj'}^t$. The $g_{jj'}$ is a random variable that follows an exponential distribution with mean $\frac{1}{h}$ to account for the multi-path fading. Path loss exponent is α . $d_{jj'} = \sqrt{|x_j - x_{j'}|^2 + |y_j - y_{j'}|^2 + |z_j - z_{j'}|^2}$ is the distance between node j and node j' , where (x_j, y_j, z_j) and $(x_{j'}, y_{j'}, z_{j'})$ denote locations of node j and node j' in the Voronoi tessellation [19].

The transmission rate $C_{jj'}$ from node j to node j' is

$$C_{jj'} = b_{jj'} \log_2 \left(1 + \frac{p_{jj'}^r}{I_c + N_0} \right), \tag{3}$$

where the channel bandwidth between j and j' is $b_{jj'}$ with the co-channel interference and white Gaussian noise are I_c and N_0 .

3.2. Multi-Tier Cellular-Based IoT Network Model

In a cellular-based IoT network, the expected number of APs covered by one MBS can be given in the Poisson–Voronoi (PV) system, as shown in [19,20]

$$E[N_a^m] = \delta_a 2\pi \int_0^\infty l_m e^{-\delta_m \pi l_m^2} dl_m = \frac{\delta_a}{\delta_m}, \tag{4}$$

where the radius of MBS cell coverage is l_m . By utilizing the same methodology, the expected number of IoT devices in one AP coverage $E[N_a^d] = \frac{\delta_a}{\delta_a}$.

With modifications in (4), the expected total length of connections between APs and MBS is as shown as

$$E[L_a^m] = \delta_a 2\pi \int_0^\infty l_m^2 e^{-\delta_m \pi l_m^2} dl_m = \frac{\delta_a}{2\delta_m^{\frac{3}{2}}}, \tag{5}$$

and expected total length of connections from device to AP is $E[L_d^a] = \frac{\delta_a}{2\delta_a^{\frac{3}{2}}}$.

The expected average length of connections between APs and MBS can be turned into the radius of MBS cell coverage, which can be given by utilizing the properties of Voronoi properties mentioned above:

$$E[l_m] = \frac{E[L_a^m]}{E[N_a^m]} = \frac{\frac{\delta_a}{2\delta_m^{\frac{3}{2}}}}{\frac{\delta_a}{\delta_m}} = \frac{1}{2\delta_m^{\frac{1}{2}}}. \tag{6}$$

Then, the mean number of tiers of APs covered by one MBS is

$$I = \left\lceil \frac{E[l_m]}{r_{min}} \right\rceil, \tag{7}$$

where r_{min} is the distance between neighboring tiers of APs, which is determined by the minimum transmission range needed to avoid co-channel interference.

By taking the same procedure as shown in (4), mean number of APs at the i th tier of AP network is represented as [20]

$$E[N_i^a] = \frac{\delta_a}{\delta_m} (e^{-\delta_m \pi (i r_{min})^2} - e^{-\delta_m \pi ((i+1) r_{min})^2}), \tag{8}$$

where $i \in \{1, \dots, (I - 1)\}$.

3.3. Queueing Model in Multi-Tier Cellular-Based IoT Network

As shown in the multi-tier cellular-based IoT network model, each AP at the i th tier receives packets not only directly from the IoT device served by it but also from APs at the $(i + 1)$ th tier for relaying. To utilize the limited resources efficiently while considering the QoS requirement of the delay-sensitive IoT devices and the corresponding applications, we utilize different access protocols for different data traffic streams in the queueing model. The priority of each data packet is shown as

$$pri_q = 1 - \frac{AoI_q}{Delay_{thr}^q}, \tag{9}$$

where $q \in \{1, \dots, Q\}$ are the data packets in the waiting list to access the channel. The data packet that has the highest priority is named q_1 , and the others with different sequence numbers from $\{2, \dots, Q\}$, to indicate the priorities.

Data packet q_1 with the highest priority in the queueing model immediately occupies the channel as soon as the channel is sensed to be idle without the LBT procedure, while the other packets follow the CSMA/CA mechanism in the same batch. After the packet q_1 is successfully transmitted, the data packet with the highest priority among all the remaining data packets that not have been successfully transmitted yet follows the CSMA/CA mechanism in the next sensing process. In this paper, we name the data packet q_1 with the highest priority as a packet with priority and the other packets as packets without priority.

For the data received from the APs at the $(i + 1)$ th tier, the priority of each packet is decided by the location of the AP that the generating device is attached to, which is because the packet generated by the device attached to the APs far from the i th tier already experienced delay, which makes it more emergent to give priority to it to satisfy the QoS requirement by utilizing a shorter contention window size for these packets than that of the packets near the i th tier.

3.4. Mean Packet Delay of Packet with Priority

3.4.1. Without Packets in the Channel

The packet service time of packets with priority is determined by the status of the packets without priority. Under the average fraction of time that packets without priority do not occupy the channel, packets with priority in the queueing instantly occupy the channel, and the packet service time in this case is as shown as [17]

$$S_{s,case1,1} = S_{s,o}, \tag{10}$$

where there is no packet with priority occupying the channel at the moment. $S_{s,o}$ is the channel occupancy time in service of packets with priority, and that for packets without priority is $S_{r,o}$. $S_{s,o}$ and $S_{r,o}$ follow the exponential distribution with $E[S_{s,o}] = \frac{1}{\mu_s}$ and $E[S_{r,o}] = \frac{1}{\mu_r}$, respectively. Packet arrival rates for packets with priority and packets without priority are λ_s and λ_r , which follow Poisson distribution, respectively.

On the other hand, the packet service time of packets with priority with existing packets with priority occupying the channel is given by

$$S_{s,case1,2} = R_{s,busy} + S_{s,o}, \tag{11}$$

where $R_{s,busy}$ is the residual busy period of packets with priority, whose Laplace–Stieltjes transform (LST) is as shown as [21]

$$\tilde{R}_{s,busy}(s^*) = \frac{\mu_s}{s^* + \lambda_s - \lambda_s \tilde{B}_s(s^*) + \mu_s} (\tilde{B}_s(s^*))^{\lambda_s E[S_s]}. \tag{12}$$

The LSTs of B_s based on the $S_{s,case1,1}$ are expressed as follows:

$$\tilde{B}_s(s^*) = \frac{\mu_s}{s^* + \lambda_s - \lambda_s \tilde{B}_s(s^*) + \mu_s}, \tag{13}$$

where $\tilde{B}_s(s^*)$ is a root of quadratic equation.

The mean packet service time of packets with priority in this case is

$$\begin{aligned} S_{s,case1} &= (1 - \lambda_s E[S_{s,o}]) S_{s,case1,1} + \lambda_s E[S_{s,o}] S_{s,case1,2} \\ &= \lambda_s E[S_{s,o}] R_{s,busy} + S_{s,o}. \end{aligned} \tag{14}$$

3.4.2. With Packets in the Channel

In addition, under the average fraction of time that packets without priority occupy the channel, packets with priority wait until the channel is released and occupy the channel as soon as the channel is sensed to be idle. The packet service time in this case is given by

$$S_{s,case2} = R_{r,o} + S_{s,o}, \tag{15}$$

where $R_{r,o}$ is the residual service time of $S_{r,o}$. It is worth mentioning that this case is based on the fact that it is impossible to have packets with priority occupying the channel under the assumptions in this case.

Then, based on the fundamental theorem, the packet service time of packets with priority is shown as

$$\begin{aligned} S_s &= (1 - \lambda_r E[S_{r,o}])S_{s,case1} + \lambda_r E[S_{r,o}]S_{s,case2} \\ &\stackrel{(a)}{=} \frac{\lambda_s}{\mu_s} S_{s,o} + S_{s,o} - \frac{\lambda_r}{\mu_r} \frac{\lambda_s}{\mu_s} S_{s,o} + \frac{\lambda_r}{\mu_r} S_{r,o} \\ &= \frac{\lambda_r}{\mu_r} S_{r,o} + S_{s,o}, \end{aligned} \tag{16}$$

where (a) is achieved by utilizing the memoryless property of exponential distribution.

The LST of S_s can be given as

$$\begin{aligned} \tilde{S}_s(s^*) &= \tilde{S}_{s,o}(s^*) \tilde{S}_{r,o}\left(\frac{\lambda_r}{\mu_r} s^*\right) \tilde{R}_{s,busy}\left(\frac{\lambda_s}{\mu_s} \left(1 + \frac{\lambda_r}{\mu_r}\right) s^*\right) \\ &= \frac{\mu_s}{s^* + \mu_s} \frac{\mu_r}{\frac{\lambda_r}{\mu_r} s^* + \mu_r} \tilde{R}_{s,busy}\left(\frac{\lambda_s}{\mu_s} \left(1 + \frac{\lambda_r}{\mu_r}\right) s^*\right) \\ &= \left(\frac{\mu_s}{s + \mu_s}\right) \left(\frac{\frac{\mu_r^2}{\lambda_r}}{s + \frac{\mu_r^2}{\lambda_r}}\right). \end{aligned} \tag{17}$$

3.5. Mean Packet Delay of Packets without Priority

3.5.1. Without Packets with Priority in the Channel

By taking the analysis procedure as shown in the previous section for packets with priority, the packet service time of packets without priority at the i th tier of the multi-tier cellular-based IoT network can be achieved based on the status of packets with priority. In the case that packets with priority do not take the channel, the packet service time of packets without priority is [17]

$$S_{r,case1,1} = S_{DIFS} + S_{back} + S_{r,o}, \tag{18}$$

where S_{DIFS} is the time for successful completion of the distributed coordination function (DCF) inter-frame space (DIFS) without any packets with priority arriving. We also assume that there are no packets without priority occupying the channel in this case at the moment. However, if the channel is occupied during the duration of the DIFS, a new DIFS restarts after the channel is idle again, which can be given as a summation of stopped time for the time between the first DIFS starts and the last DIFS duration $T_{stop,DIFS}$ with a busy period of packets with priority B_s . Then the S_{DIFS} analyzed above for packets without priority with CSMA/CA mechanism is [22], as shown in (19).

$$\begin{aligned} S_{DIFS} &= T_{DIFS} + \sum_{k=1}^{\infty} ((T_{stop,DIFS,s} + B_s) Pr_{k,arr,DIFS}^s + (T_{stop,DIFS,I} + B_I) Pr_{k,arr,DIFS}^I \\ &\quad + (T_{stop,DIFS,(I-1)} + B_{(I-1)}) Pr_{k,arr,DIFS}^{(I-1)} + (T_{stop,DIFS,(I-2)} + B_{(I-2)}) Pr_{k,arr,DIFS}^{(I-2)} \\ &\quad + \dots + (T_{stop,DIFS,(i+3)} + B_{(i+3)}) Pr_{k,arr,DIFS}^{(i+3)} + (T_{stop,DIFS,(i+2)} + B_{(i+2)}) Pr_{k,arr,DIFS}^{(i+2)}). \end{aligned} \tag{19}$$

In (19), the $Pr_{k,arr,DIFS}^{(i+2)}$ is

$$Pr_{k,arr,DIFS}^{(i+2)} = (1 - Pr_{k,arr,DIFS}^s)(1 - Pr_{k,DIFS,I})(1 - Pr_{k,DIFS,(I-1)}) \dots (1 - Pr_{k,DIFS,(i+3)})Pr_{k,DIFS,(i+2)}, \tag{20}$$

and $Pr_{k,arr,DIFS}$ is the probability that there are arriving packets within the k th DIFS attempt duration.

The LST of $T_{stop,DIFS}$ and B_s are as shown as follows:

$$\begin{aligned} \tilde{T}_{stop,DIFS}(s^*) &= \frac{\lambda_s}{s^* + \lambda_s} \frac{1 - e^{-(s^* + \lambda_s)T_{DIFS}}}{1 - e^{-\lambda_s T_{DIFS}}} \frac{\lambda_I}{s^* + (\lambda_s + \lambda_I)} \frac{1 - e^{-(s^* + (\lambda_s + \lambda_I))T_{DIFS}}}{1 - e^{-\lambda_I T_{DIFS}}} \\ &\dots \frac{\lambda_{i+1}}{s^* + (\lambda_s + \lambda_I + \dots + \lambda_{i+1})} \frac{1 - e^{-(s^* + (\lambda_s + \lambda_I + \dots + \lambda_{i+1}))T_{DIFS}}}{1 - e^{-\lambda_{i+1} T_{DIFS}}}, \end{aligned} \tag{21}$$

and

$$\tilde{B}_s(s^*) = \left(\frac{\mu_s}{s^* + \lambda_s - \lambda_s \tilde{B}_s(s^*) + \mu_s} \right) \left(\frac{\frac{\mu_r^2}{\lambda_r}}{s^* + \lambda_s - \lambda_s \tilde{B}_s(s^*) + \frac{\mu_r^2}{\lambda_r}} \right), \tag{22}$$

where the $\tilde{B}_s(s^*)$ is a root of cubic equation.

Then, the LST of S_{DIFS} is shown as

$$\begin{aligned} \tilde{S}_{DIFS}(s^*) &= \tilde{T}_{DIFS}(s^*) \tilde{T}_{stop,DIFS}((e^{\lambda_s T_{DIFS}} - 1)(1 + Pr_{n-s,DIFS} + Pr_{n-s,DIFS} Pr_{n-(k+1)th,DIFS} + \dots + \\ &Pr_{n-s,DIFS} Pr_{n-(k+1)th,DIFS} \dots Pr_{n-(I-1)th,DIFS})s^*) \tilde{B}_s((e^{\lambda_s T_{DIFS}} - 1)(1 + Pr_{n-s,DIFS} \\ &+ Pr_{n-s,DIFS} Pr_{n-(k+1)th,DIFS} + \dots + Pr_{n-s,DIFS} Pr_{n-(k+1)th,DIFS} \dots Pr_{n-(I-1)th,DIFS})s^*). \end{aligned} \tag{23}$$

S_{back} is the time for the backoff procedure in the CSMA/CA mechanism. After detecting the channel is not occupied for the DIFS duration, the station initiates the transmission only if the channel remains idle for an additional random time duration S_{back} , which is determined by the contention window size and is given by

$$S_{back} = T_{back} + N_{back,s}(B_s + S_{DIFS,s}) + N_{back,i+1}(B_{i+1} + S_{DIFS,i+1}) + \dots + N_{back,I}(B_I + S_{DIFS,I}), \tag{24}$$

where T_{back} is the backoff interval and $N_{back,s}$ is the number of arriving packets with priority during T_{back} . $N_{back,i+1}$ is the number of arriving packets without priority transmitted from devices served by APs at the $(i + 1)$ th tier.

Then, the LST of S_{back} can be given by

$$\begin{aligned} \tilde{S}_{back}(s^*) &= \tilde{T}_{back}(s^*) (\tilde{B}_s(s^*) \tilde{S}_{DIFS}(s^*))^{N_{back,s}} (\tilde{B}_{i+1}(s^*) \\ &\tilde{S}_{DIFS,i+1}(s^*))^{N_{back,i+1}} \dots (\tilde{B}_I(s^*) \tilde{S}_{DIFS,I}(s^*))^{N_{back,I}}. \end{aligned} \tag{25}$$

On the other hand, the packet service time of packets without priority with existing packets without priority occupying the channel is given by

$$S_{r,case1,2} = R_{r,busy} + S_{r,case1,1}, \tag{26}$$

where $R_{r,busy}$ is the residual busy period of packets without priority occupying the channel, whose LST is shown as

$$\tilde{R}_{r,busy}(s^*) = \frac{\mu_{r_o}}{s^* + \lambda_{r_o} - \lambda_{r_o} \tilde{B}_{r_o}(s^*) + \mu_{r_o}} (\tilde{B}_{r_o}(s^*))^{\lambda_{r_o} E[S_{r_o}]}. \tag{27}$$

The LSTs of B_{r_o} are expressed as follows:

$$\tilde{B}_{r_o}(s^*) = \left(\frac{\mu_{r_o}}{s^* + \lambda_{r_o} - \lambda_{r_o} \tilde{B}_{r_o}(s^*) + \mu_{r_o}} \right) \left(\frac{\frac{\mu_r^2}{\lambda_r}}{s^* + \lambda_{r_o} - \lambda_{r_o} \tilde{B}_{r_o}(s^*) + \frac{\mu_r^2}{\lambda_r}} \right), \tag{28}$$

where r_o is the packets without priority occupying the channel. $\tilde{B}_{r_o}(s^*)$ is a root of cubic equation.

Then, the mean packet service time of packets without priority in this case is shown as

$$\begin{aligned} S_{r,case1} &= (1 - \lambda_r E[S_{r,o}]) S_{r,case1,1} + \lambda_r E[S_{r,o}] S_{r,case1,2} \\ &= \lambda_r E[S_{r,o}] R_{r,busy} + S_{r,case1,1}. \end{aligned} \tag{29}$$

3.5.2. With Packets with Priority in the Channel

The packet service time of packets without priority in the case that packets with priority occupy the channel is presented as

$$S_{r,case2} = R_{s,busy} + S_{r,case1,1}. \tag{30}$$

It is also worth mentioning that this case is based on the fact that it is impossible to have packets without priority occupying the channel under the assumptions in this case.

Then based on the fundamental theorem, the packet service time of packets without priority is shown as

$$\begin{aligned} S_r &= (1 - \lambda_s E[S_{s,o}]) S_{r,case1}(T_{back}, N_{back,s}) + \lambda_s E[S_{s,o}] S_{r,case2}(T_{back}, N_{back,s}) \\ &= S_{DIFS} + S_{back}(T_{back}, N_{back,s}) + S_{r,o} + \frac{\lambda_s}{\mu_s} R_{s,busy} + \frac{\lambda_r}{\mu_r} \left(1 - \frac{\lambda_s}{\mu_s} \right) R_{r,busy}. \end{aligned} \tag{31}$$

The LST of S_r is

$$\tilde{S}_r(s^*) = \tilde{S}_{DIFS}(s^*) \tilde{S}_{back}(s^* | T_{back}, N_{back,s}) \tilde{S}_{r,o}(s^*) \tilde{R}_{s,busy} \left(\frac{\lambda_s}{\mu_s} s^* \right) \tilde{R}_{r,busy} \left(\frac{\lambda_r}{\mu_r} \left(1 - \frac{\lambda_s}{\mu_s} \right) s^* \right). \tag{32}$$

4. Problem Formulation

4.1. Mean Packet Arrival Rate

In this paper, the successful transmission probability between the device and AP is β , then the mean packet arrival rate of packets with priority from the device to the serving AP λ_s is

$$\lambda^{uv} = E[N_u^v] C^{uv} \beta, \tag{33}$$

where C^{uv} is the transmission rate from one device to the serving AP.

Then, the mean packet arrival rate in any AP at the i th tier of a multi-tier cellular-based IoT network is given by

$$\lambda_i^v = \lambda_s + \lambda_r = \lambda_s + \frac{E[N_{(i+1)}^v]}{E[N_i^v]} \mu_{(i+1)}^v. \tag{34}$$

If i equals I , then $\lambda_i^v = \lambda^{uv}$. The $\mu_{(i+1)}^v$ is the mean service rate for packets in any AP at the $(i + 1)$ th tier of the network, which is also the packet transmission rate of packets without priority from any AP at the $(i + 1)$ th tier to APs at the i th tier, $C_{(i+1)}^{vv}$.

4.2. Mean Packet Transmission Delay

If we utilize the M/G/1 queue in each AP, the mean waiting time of packets in a random AP at the i th tier $E[W_i^v|N_i^v, \mathbf{b}, \mathbf{p}]$ with given N_i^v, \mathbf{b} and \mathbf{p} is shown in the following [23]:

$$E[W_i^v|N_i^v, \mathbf{b}, \mathbf{p}] = E[L_i^v]E[S_i^v|N_i^v, \mathbf{b}, \mathbf{p}] + \rho_i^v E[R_i^v|N_i^v, \mathbf{b}, \mathbf{p}], \tag{35}$$

where $E[L_i^v]$ is the number of packets in any AP at the i th tier and $\rho_i^v = \frac{\lambda_i^v}{\mu_i^v}$ is utility ratio. $E[S_i^v|N_i^v, \mathbf{b}, \mathbf{p}]$ and $E[R_i^v|N_i^v, \mathbf{b}, \mathbf{p}]$ are mean service time and residual service time, respectively. The \mathbf{b} and \mathbf{p} are spectrum and power allocation ratios for each tier in a multi-tier cellular-based IoT network. At each tier, spectrum and power are uniformly allocated to APs. The condition that $\rho_i^v \leq 1$ must be satisfied to ensure the stability of the queue system. For convenience, the superscript v is removed for all parameters without any ambiguity.

With Little’s law and the relationship between service time and residual service time, which is

$$E[R_i|N_i, \mathbf{b}, \mathbf{p}] = E[S_i^2|N_i, \mathbf{b}, \mathbf{p}]/2E[S_i|N_i, \mathbf{b}, \mathbf{p}], \tag{36}$$

the mean waiting time in (35) can be also represented as

$$E[W_i|N_i, \mathbf{b}, \mathbf{p}] = \frac{\rho_i E[S_i^2|N_i, \mathbf{b}, \mathbf{p}]}{2(1 - \rho_i)E[S_i|N_i, \mathbf{b}, \mathbf{p}]} \tag{37}$$

Then, the mean packet transmission delay in a random AP at the i th tier, $E[D_i|N_i, \mathbf{b}, \mathbf{p}]$, is given by

$$E[D_i|N_i, \mathbf{b}, \mathbf{p}] = E[S_i|N_i, \mathbf{b}, \mathbf{p}] + \frac{\rho_i E[S_i^2|N_i, \mathbf{b}, \mathbf{p}]}{2(1 - \rho_i)E[S_i|N_i, \mathbf{b}, \mathbf{p}]} \tag{38}$$

$E[S_i|N_i, \mathbf{b}, \mathbf{p}]$ and $E[S_i^2|N_i, \mathbf{b}, \mathbf{p}]$ are the first and second moments of conditional packet service time and can be obtained by applying the property of LST, which can be given as

$$E[S_i|N_i, \mathbf{b}, \mathbf{p}] = -\tilde{S}_i'(0|N_i, \mathbf{b}, \mathbf{p}) \tag{39}$$

and

$$E[S_i^2|N_i, \mathbf{b}, \mathbf{p}] = (-1)^2 \tilde{S}_i''(0|N_i, \mathbf{b}, \mathbf{p}). \tag{40}$$

4.3. Channel Occupancy Time in Service

In this paper, the packet transmission time from the i th tier to the $(i - 1)$ th tier is defined as the packet service time S_i , which is $E[S_i] = 1/C_i$ for per unit length packet. The achievable packet transmission rate in any AP at the i th tier is shown as $C_i = \frac{b_i}{E[N_i]} \log_2(1 + \frac{g(p_i/E[N_i])r_{min}^{-\alpha}}{I_{c_i} + N_0})$, where b_i and p_i are the total spectrum and power resources allocated to APs at the i th tier. Then the CDF of S_i can be represented as

$$F_{S_i}(t|N_i, \mathbf{b}, \mathbf{p}) = Pr[S_i < t|N_i, \mathbf{b}, \mathbf{p}] = Pr[g > \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)(I_{c_i} + N_0)}{(p_i/E[N_i])r_{min}^{-\alpha}}]. \tag{41}$$

With the fact that $g \sim \exp(h)$ [24], we can obtain

$$Pr[g > \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)(I_{c_i} + N_0)}{(p_i/E[N_i])r_{min}^{-\alpha}}] = E_{In_i}[g > \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)(In_i + N_0)}{(p_i/E[N_i])r_{min}^{-\alpha}} | In_i] \tag{42}$$

$$\stackrel{(b)}{=} e^{-h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)N_0}{(p_i/E[N_i])r_{min}^{-\alpha}}} L_{I_{c_i}}(h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}}),$$

where the proof for (b) is given in Appendix A.

Then, the PDF of service time S_i at the i th tier can be achieved:

$$f_{S_i}(t|N_i, \mathbf{b}, \mathbf{p}) = \frac{d}{dt} F_{S_i}(t|N_i, \mathbf{b}, \mathbf{p})$$

$$= \frac{d}{dt} Pr[g > \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)(In_i + N_0)}{(p_i/E[N_i])r_{min}^{-\alpha}}] \tag{43}$$

$$= (e^{-h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)N_0}{(p_i/E[N_i])r_{min}^{-\alpha}}})' L_{I_{c_i}}(h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}}) + e^{-h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)N_0}{(p_i/E[N_i])r_{min}^{-\alpha}}} L'_{I_{c_i}}(h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}})$$

where $(e^{-h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)N_0}{(p_i/E[N_i])r_{min}^{-\alpha}}})'$ and $L'_{I_{c_i}}(h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}})$ can be achieved by derivation operations.

$$(e^{-h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)N_0}{(p_i/E[N_i])r_{min}^{-\alpha}}})' = e^{-h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)N_0}{(p_i/E[N_i])r_{min}^{-\alpha}}} \frac{-hN_0}{(p_i/E[N_i])r_{min}^{-\alpha}} 2^{\frac{E[N_i]}{b_i t}} (\ln 2) \frac{E[N_i]}{b_i} (-t^{-2}) \tag{44}$$

and $L'_{I_{c_i}}(h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}})$ is shown as (45).

$$L'_{I_{c_i}}(h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}}) = \exp(-2\pi\delta_v) \int_{r_{min}}^{\infty} \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])} \tau d\tau (-2\pi\delta_v)$$

$$\frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])} + (\frac{\tau}{r_{min}})^{\alpha} \tag{45}$$

$$\int_{r_{min}}^{\infty} \frac{(2^{\frac{E[N_i]}{b_i t}} - 1) \frac{E[N_i]}{p_i} 2^{\frac{E[N_i]}{b_i t}} (\ln 2) \frac{E[N_i]}{b_i} (t^{-2}) - (2^{\frac{E[N_i]}{b_i t}} - 1) \frac{E[N_i]}{p_i} 2^{\frac{E[N_i]}{b_i t}} (\ln 2) \frac{E[N_i]}{b_i} (t^{-2}) (\frac{2^{\frac{E[N_i]}{b_i t}} - 1}{(p_i/E[N_i])} + (\frac{\tau}{r_{min}})^{\alpha})}{(\frac{2^{\frac{E[N_i]}{b_i t}} - 1}{(p_i/E[N_i])} + (\frac{\tau}{r_{min}})^{\alpha})^2} \tau d\tau.$$

With a conditional PDF of the service time, we can obtain the mean packet transmission delay in a random AP at the i th tier.

4.4. Total Packet Delay Minimization Problem

In multi-tier cellular-based IoT networks, it is essential to investigate mean total packet delay for packets due to the delay penalty to packets caused by the multi-hopping transmission, which is the summation of served delay and relayed delay experienced at each tier, and can be shown as follows:

$$D_t = D_s + D_r = D_s + D_{r,(i-1)} + D_{r,(i-2)} + \dots + D_{r,1}. \tag{46}$$

We present a cellular-based IoT network with APs in two tiers to demonstrate the effectiveness of the analysis results and then formulate a mean total packet delay minimization

problem for packets transmitted from devices served by APs at the second tier of the IoT network. Then, the optimization problem with objective function and the corresponding constraints is shown as

$$\begin{aligned}
 & \min_{\phi, \zeta} E[D_t](\phi, \zeta) \\
 & \text{s.t.} \quad \sum_{i=1}^2 b_i = B, \quad \sum_{i=1}^2 p_i = P, \\
 & \quad \rho_{r,1} \leq 1, \rho_{s,i} \leq 1 \quad i \in \{1, 2\}, \\
 & \quad 0 \leq \phi, \zeta \leq 1,
 \end{aligned} \tag{47}$$

where $E[D_t] = E[D_{s,2}] + E[D_{r,1}]$. The total packet delay, $E[D_t]$, includes delay experienced as served packet at the second tier $E[D_{s,2}]$ and relayed packet at the first tier $E[D_{r,1}]$. The ratios of spectrum and power allocated to APs at the first tier are ϕ and ζ , as $\phi = \frac{b_1}{B}$ and $\zeta = \frac{p_1}{P}$, respectively. The P and B are the total power and spectrum for the two tiers of APs, respectively.

The proposed algorithm for total packet transmission delay minimization via resources optimal allocation in the cellular-based IoT networks with multi-tier APs is exhibited in Algorithm 1 [25].

Algorithm 1: Proposed algorithm for total packet transmission delay minimization.

Algorithm input: $\mathbf{G}(\{\mathbf{M}, \mathbf{A}\}, \mathbf{D})$, $\delta_m, \delta_a, \delta_d$, tep size θ , tolerance $\epsilon > 0$.

Initialize ϕ, ζ with $\rho_i(\phi, \zeta) \leq 1$ and $0 \leq \phi, \zeta \leq 1$

begin

Obtain minimum $\sum_{i=1}^I E[D_i](\phi|\zeta)$ and the corresponding ϕ^* by using the Gradient descent method with step size θ .

if $\|\nabla \sum_{i=1}^I E[D_i](\zeta|\phi^*)\|_2 \leq \epsilon$, **then**

| $\zeta^* = \zeta$;
| Break;

else

| Update ζ by utilizing the Bisection method and return to **Step 3**;

end if

end

Algorithm output: ϕ^*, ζ^* (Optimal spectrum and power allocation ratio).

5. Numerical Results

In this paper, we present a cellular-based IoT network with two-tier APs in one MBS cell coverage with dimensions $1 \text{ km} \times 1 \text{ km} \times 1 \text{ km}$. The distance between two tiers of APs is 50 m. In the simulation, δ_m, δ_a and δ_d are 1, 1000, and 10,000 per km^2 , respectively. The total spectrum and power for APs in the network are 1000 MHz and 100 W, respectively. The step size θ to update ϕ and ζ is 0.05. The system parameter T_{DIFS} is $36 \times 10^{-6} \text{ s}$, and the maximum T_{back} is $72 \times 10^{-6} \text{ s}$ by taking slot time as $9 \times 10^{-6} \text{ s}$ with the contention window size as 8.

Figure 2 presents variation trends of the total packet transmission delay $E[D_t]$ as spectrum allocation ratio ϕ for different fixed power allocation ratios ζ . Meanwhile, the minimum $E[D_t]$ for each ζ reveals the optimal value of ϕ , which is ϕ^* as shown with the vertical dotted lines. Based on the facts shown in the figure, we can observe that the value of ϕ^* is decreasing as the ζ keeps increasing, which is due to the fact that more transmission power can compensate the shortage of spectrum resource to achieve the same minimum total packet delay.

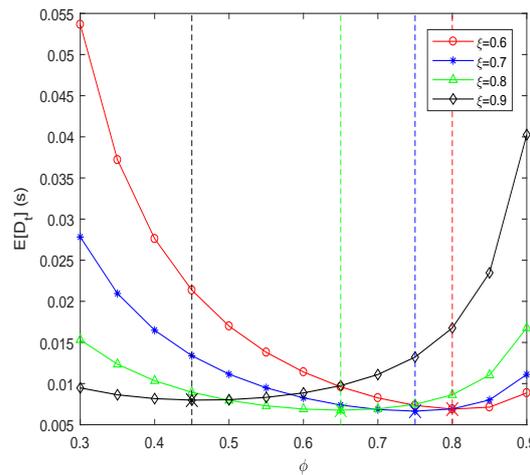


Figure 2. Total packet transmission delay $E[D_t]$ as a function of ϕ .

Figure 3 shows the variation trend of the total packet transmission delay $E[D_t]$ as power allocation ratio ζ for various fixed values of spectrum allocation ratio ϕ . Also, the minimum $E[D_t]$ for each ϕ reveals the optimal value of ζ , which is ζ^* as shown with the vertical dotted lines. The value of ζ^* is decreasing as the ϕ keeps increasing, which reflects the same conclusion as shown in Figure 2 and cross-proves the correctness of the proposed scheme and algorithm. On the other hand, it presents a performance comparison of the proposed mechanism and the spectrum resource allocation with the Monte Carlo simulation. As shown in the figure, the proposed mechanism obtains better minimum mean total packet delay with ϕ^* than the ϕ is any other values obtained by using the Monte Carlo simulation. Meanwhile, the performance comparison of the proposed scheme with the spectrum allocation based on the round-robin scheduling mechanism with $\phi = \phi_{rr}$ by the total packet transmission delay is given in Figure 4. In the round-robin scheduling mechanism, each tier in the multi-tier cellular-based IoT network alternatively utilizes all the network resources, which results in a single-hop transmission network and accumulates all the network traffic. It also shows superior performance than the round-robin scheduling in the figure.

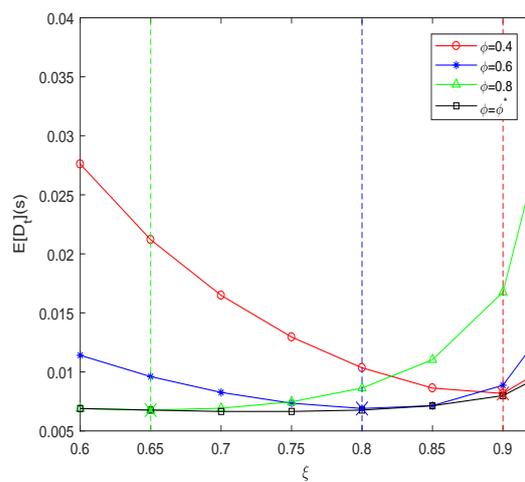


Figure 3. Total packet transmission delay $E[D_t]$ as a function of ζ .

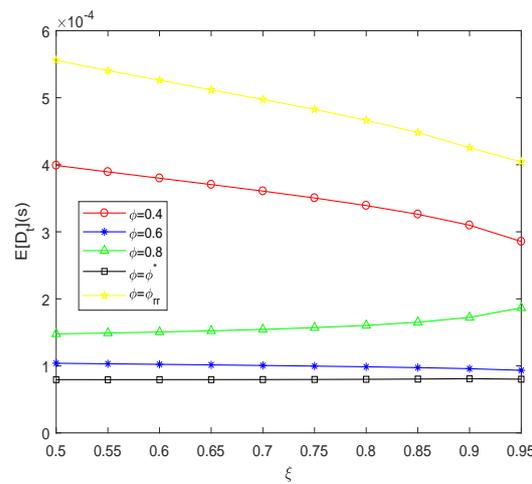


Figure 4. Performance comparison for the proposed scheme.

Based on the results shown in Figures 2 and 3, Figure 5 describes the value of optimal ζ and optimal ϕ , which collaboratively reveal the minimum $E[D_t]$ as shown in the third dimension in the figure. In the figure, the value of ζ^* is decreasing with ϕ increasing, and the value of ϕ^* is decreasing when ζ is increasing, as shown in Figures 2 and 3. With the condition mentioned above that $\rho_i \leq 1$ must be satisfied, ζ cannot be less than 0.5 in Figures 3 and 5.

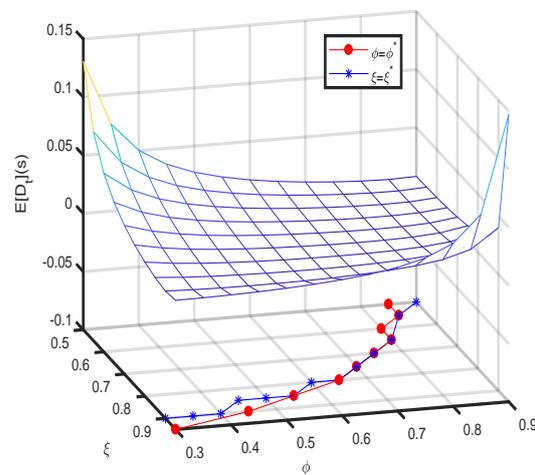


Figure 5. The optimal ϕ and ζ for the minimum $E[D_t]$.

With Figure 5 as a basis, Figure 6 describes the mean delay $E[D_{s,1}]$ at the first tier of the network experienced by the packets transmitted from devices served by APs at the first tier of the multi-tier cellular-based IoT network. The optimal values of ϕ , ϕ^* , and the optimal value of ζ , ζ^* , are constant at the maximum values to achieve the minimum $E[D_{s,1}]$, as shown in Figure 6, on account of more spectrum and power resources leading to the smaller delay to packets from devices attached to the APs in the first tier with constant packets arrival rate and higher priority than the other packets.

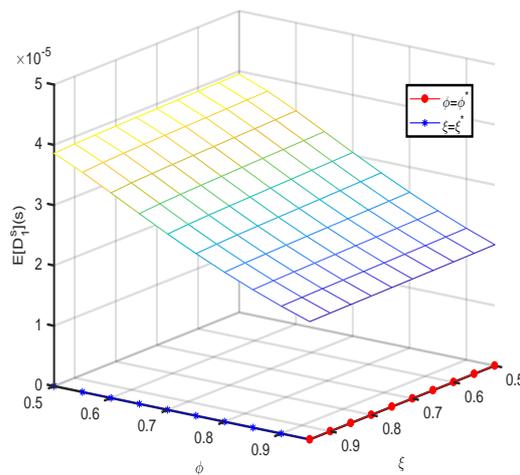


Figure 6. The optimal ϕ and ξ for the minimum $E[D_{s,1}]$.

On the other hand, Figure 7 exhibits the mean delay $E[D_{s,2}]$ at the second tier of the network experienced by the packets transmitted from devices served by APs at the second tier of multi-tier cellular-based IoT networks. The optimal value of ϕ , ϕ^* , and the optimal value of ξ , ξ^* , is constant at the minimum values to obtain the minimum $E[D_{s,2}]$, as shown in Figure 7, for more spectrum and power resources.

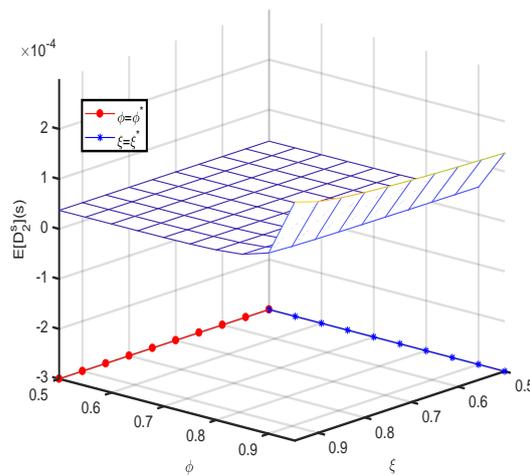


Figure 7. The optimal ϕ and ξ for the minimum $E[D_{s,2}]$.

Figure 8 shows the mean delay $E[D_{r,1}]$ at the first tier of the network experienced by the packets transmitted from devices served by APs at the second tier of the multi-tier cellular-based IoT networks for relaying. The optimal value of ϕ , ϕ^* , is constant at the maximum value, which is due to the relaying delay being the dominant factor in mean total packet delay and needs enough resources to guarantee the QoS performance due to the large AoI. The optimal value of ξ , ξ^* decreases from the maximum value to the minimum value as ϕ increases because the maximum power resource is needed to obtain the minimum mean relaying delay when the spectrum resource cannot be guaranteed, as the spectrum increases until it reaches the value where the minimum mean total packet delay can be guaranteed with any ratio of power resource allocation, which changes to minimum value sharply, and the remaining part is allocated to the second tier of the IoT network. It also shows that the mean total packet delay is more sensitive to ξ compared to ϕ , because the mean delay for relaying is more sensitive to ξ , as shown.

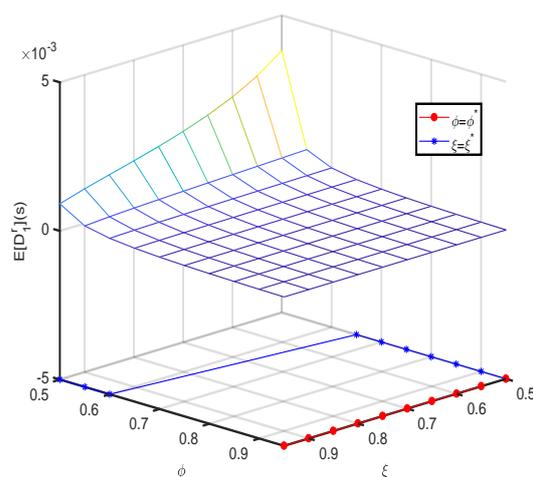


Figure 8. The optimal ϕ and ζ for the minimum $E[D_{r,1}]$.

6. Conclusions

In this paper, we proposed a multi-tier cellular-based IoT network with a traffic prioritization scheduling scheme based on the AoI metric to satisfy strict data packet latency requirements in 5G IoT networks. Our mathematical framework based on the queueing theory models the time-domain behaviors of data packets with different channel access protocols to overcome delay penalties and ensure fairness for packets with large AoI metrics. Then, we minimized the mean total packet delay by optimally distributing spectrum and power resources among multiple tiers in IoT networks. Numerical results demonstrated the effectiveness of the proposed mechanism and algorithm in a two-tier cellular-based IoT network, which achieved the minimum total packet transmission delay to meet strict latency requirements. Moreover, the optimal allocation of spectrum and power resources minimized delays of data packets experienced at each tier in the multi-tier IoT networks, which realized monitoring and guaranteeing the packet latency as a requirement in real-time. Future studies will consider a more general and practical case in which forwarded packets have different priorities based on the location of the transmitting device. Our work contributes to the development of IoT networks by addressing packet delay penalties during multi-hopping transmission and improving fairness for different kinds of data packets.

Author Contributions: Conceptualization, J.L.; methodology, J.L.; software, W.G.; validation, Z.T.; formal analysis, J.L. and Z.T.; investigation, W.G.; resources, Z.T.; data curation, W.G.; writing—original draft preparation, J.L.; writing—review and editing, Z.T.; visualization, W.G.; supervision, Z.T.; project administration, J.L.; funding acquisition, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Beijing Municipal Education Commission (KM202310011002) and the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 824019.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1

The mathematical proof for (b) in Equation (42) is shown as follows:

$$Pr[g > \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)(I n_i + N_0)}{(p_i/E[N_i])r_{min}^{-\alpha}}] = 1 - Pr[g \leq \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)(I n_i + N_0)}{(p_i/E[N_i])r_{min}^{-\alpha}}] \tag{A1}$$

$$= exp(-\mu \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)(I n_i + N_0)}{(p_i/E[N_i])r_{min}^{-\alpha}}).$$

$$L_{Ic_i}(h \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}}) = E_{\Phi}[\prod_{i \in \Phi} \frac{1}{1 + \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}} \tau^{-\alpha}}]$$

$$= E_{\Phi,g}[\prod_{i \in \Phi} E_g[exp(\mu \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}})]] \tag{A2}$$

$$\stackrel{(c)}{=} E_{\Phi}[\prod_{i \in \Phi} \frac{1}{1 + \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}} \tau^{-\alpha}}]$$

$$= exp(-2\pi\delta_v \int_{r_{min}}^{\infty} \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{\frac{(p_i/E[N_i])}{(2^{\frac{E[N_i]}{b_i t}} - 1)} + (\frac{\tau}{r_{min}})^{\alpha}} \tau d\tau)$$

with $r_{min}^2 \rightarrow \tau$ and Φ is interference set for communication. The proof for (c) is given in Appendix B.

Appendix B

The mathematical proof for (c) in Appendix A is shown as follows:

$$E_g[exp(\mu \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}})] = \int exp(\mu \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}}) \mu exp(-\mu g) d_g \tag{A3}$$

$$= \frac{1}{1 + \frac{(2^{\frac{E[N_i]}{b_i t}} - 1)}{(p_i/E[N_i])r_{min}^{-\alpha}} \nu^{-\alpha}}.$$

References

1. Sun, H.; Shen, H.; Yang, H.H.; Wang, X.; Liu, J.; Li, S.; Quek, T.Q.S. Impacts of User Association and Power Control on AoI in Multi-Tier Cellular-Based IoT Networks. *IEEE Wirel. Commun. Lett.* **2022**, *11*, 1196–1200. [CrossRef]
2. Jiang, N.; Deng, Y.; Nallanathan, A.; Kang, X.; Quek, T.Q.S. Analyzing Random Access Collisions in Massive IoT Networks. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 6853–6870. [CrossRef]
3. Mankar, P.D.; Chen, Z.; Abd-Elmagid, M.A.; Pappas, N.; Dhillon, H.S. Throughput and Age of Information in a Cellular-Based IoT Network. *IEEE Trans. Wirel. Commun.* **2021**, *20*, 8248–8263. [CrossRef]
4. Park, J.; Kim, S.; Youn, J.; Ahn, S.; Cho, S. Low-Complexity Data Collection Scheme for UAV Sink Nodes in Cellular IoT Networks. *IEEE Trans. Veh. Technol.* **2021**, *70*, 4865–4879. [CrossRef]
5. Park, E.; Kim, T.; Han, Y. An Efficient Time-Shifted Random Access Scheme for Cellular-Based IoT Networks. *IEEE Commun. Lett.* **2019**, *23*, 522–525. [CrossRef]
6. Jiang, N.; Deng, Y.; Kang, X.; Nallanathan, A. Random Access Analysis for Massive IoT Networks Under a New Spatio-Temporal Model: A Stochastic Geometry Approach. *IEEE Trans. Commun.* **2018**, *66*, 5788–5803. [CrossRef]

7. Ebrahimi, D.; Sharafeddine, S.; Ho, P.-H.; Assi, C. UAV-Aided Projection-Based Compressive Data Gathering in Wireless Sensor Networks. *IEEE Internet Things J.* **2019**, *6*, 1893–1905. [[CrossRef](#)]
8. Abd-Elmagid, M.A.; Pappas, N.; Dhillon, H.S. On the Role of Age of Information in the Internet of Things. *IEEE Commun. Mag.* **2019**, *57*, 72–77. [[CrossRef](#)]
9. Cheikh, I.; Sabir, E.; Aouami, R.; Roy, S.; Sadik, M. Data Freshness and End-to-End Delay in Cross-Layer Two-Tier Linear IoT Networks. *Sensors* **2022**, *22*, 9455. [[CrossRef](#)]
10. Fu, S.; Jiang, D. Multi-Dimensional Resource Allocation for Throughput Maximization in CRIoT with SWIPT. *Energies* **2023**, *16*, 4767. [[CrossRef](#)]
11. Zhao, J.; Mei, Y.; Gao, X.; Yang, J.; Shang, J. Multi-Objective Optimization for EE-SE Tradeoff in Space-Air-Ground Internet of Things Networks. *Electronics* **2023**, *12*, 2585. [[CrossRef](#)]
12. Ma, B.; Xu, D.; Ren, X.; Wang, Y.; Liu, J. Resource Allocation of UAV-Assisted IoT Node Secure Communication System. *Signals* **2023**, *4*, 591–603. [[CrossRef](#)]
13. Ramamoorthy, P.; Sanobar, S.; Di Nunzio, L.; Cardarilli, G.C. Sustainable Power Consumption for Variance-Based Integration Model in Cellular 6G-IoT System. *Sustainability* **2023**, *15*, 12696. [[CrossRef](#)]
14. Kim, T.; Chae, S.; Lim, J.-T.; Bang, I. Resource-Efficient Parallelized Random Access for Reliable Connection Establishment in Cellular IoT Networks. *Sensors* **2023**, *23*, 3819. [[CrossRef](#)] [[PubMed](#)]
15. Tung, T.V.; An, T.T.; Lee, B.M. Joint Resource and Trajectory Optimization for Energy Efficiency Maximization in UAV-Based Networks. *Mathematics* **2022**, *10*, 3840. [[CrossRef](#)]
16. Li, J.; Han, Y. Optimal resource allocation for packet delay minimization in Multi-layer UAV networks. *IEEE Commun. Lett.* **2017**, *21*, 580–583. [[CrossRef](#)]
17. Li, J.; Han, Y. A traffic service scheme for delay minimization in Multi-Layer UAV networks. *IEEE Trans. Veh. Technol.* **2018**, *67*, 5500–5504. [[CrossRef](#)]
18. Mozaffari, M.; Saad, W.; Bennis, M.; Debbah, M. Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 3949–3963. [[CrossRef](#)]
19. Foss, S.G.; Zuyev, S.A. On a Voronoi Aggregative Process Related to a Bivariate Poisson Process. *Adv. Appl. Probab.* **1996**, *28*, 965–981. [[CrossRef](#)]
20. Jin, Z.; Kim, D.; Cho, J.; Lee, B. An Analysis on Optimal Cluster Ratio in Cluster-Based Wireless Sensor Networks. *IEEE Sens. J.* **2015**, *15*, 6413–6423. [[CrossRef](#)]
21. Wang, X.; Quek, T.Q.S.; Sheng, M.; Li, J. On the coexistence of Wi-Fi and LTE-U in unlicensed spectrum. In Proceedings of the 2017 IEEE International Conference on Communications (ICC), Paris, France, 21–25 May 2017.
22. Song, Y.; Sung, K.W.; Han, Y. Impact of packet arrivals on Wi-Fi and cellular system sharing unlicensed spectrum. *IEEE Trans. Veh. Technol.* **2016**, *65*, 10204–10208. [[CrossRef](#)]
23. Adan, I.; Resing, J. *Queueing Theory*; Department of Mathematics and Computer Science, Eindhoven University of Technology: Eindhoven, The Netherlands, 2001.
24. Renzo, M.D.; Lam, T.T.; Zappone, A.; Debbah, M. A Tractable Closed-Form Expression of the Coverage Probability in Poisson Cellular Networks. *IEEE Wirel. Commun. Lett.* **2018**, *8*, 249–252. [[CrossRef](#)]
25. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.