Optimizing User-Level Packet Scheduling Performance through Optimal CQI-Based Resource Allocation in LTE

Erdenebayar Lamjav\(^1\), Otgonbayar Bataa\(^1\), Chuluunbandi Naimannaran\(^2\)

\(^1\)Department of Communications, School of Information and Communication Technology, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia
\(^2\)School of Information and Communication Technology, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia

e-mail: erdenebayar.l@must.edu.mn

Abstract

Our study delves into the exploration of Dynamic allocation of resources in the downlink of LTE-A networks based on OFDMA. It comprehensively investigates user-level packet scheduling performance, aiming to optimize the efficiency of resource allocation. To achieve this, we employ a Traffic Differentiator stage that effectively separate packet queues deriving from attached users into separate service queues according to their corresponding service categories. Within each service category, users are then prioritized by taking into account their unique QoS requirements and the prevailing wireless channel conditions. This prioritization is accomplished through the utilization of the innovative SPSSA technique. We propose the PITDSA algorithm in the TD Scheduler stage, which diligently allocates the appropriate amount of radio resources to all type of services, while optimal distribution of the remaining resources to background services with efficiency. In the FD scheduler stage, we introduce an optimal CQI selection algorithm that optimizes packet scheduling by effectively leveraging FD multi-user diversity.

Keywords: Scheduling algorithms, Radio resource allocation, Cellular networks, Quality of Service.

1. INTRODUCTION

The Long-Term Evolution (LTE) network consistently confronts the demanding task of satisfying the quality of service (QoS), which is carried by scheduling information. The scheduling information includes crucial details such as the allocated PRBs through which users’ data will be transmitted. To ensure reliable and seamless communication, the emergency information of all scheduled users is transmitted based on a First Come First Serve
(FCFS) rule. Various algorithms, proposed by researchers, have to enhance the effectively of 4G resources by integrating traffic considerations, which is include system throughput and ensuring fairness among users during the decision-making process for radio resource scheduling. LTE networks employ advanced transmission techniques such as Orthogonal Frequency Division Multiple Access (OFDMA) for the Downlink (DL) and Single Carrier Frequency Division Multiple Access (SC-FDMA) for the Uplink (UL) [1]. Downlink transmission has peak data rate of 100 Mbps, whereas the UL transmission achieves a commendable 50 Mbps, both within a 20 MHz bandwidth and an impressive spectrum efficiency of 1.5bps/Hz [2]. LTE networks have already been successfully commercialized in countries like the US and the UK, signaling their widespread adoption and effectiveness. In a continuous effort to enhance LTE capabilities, the development of LTE-Advanced (LTE-A) was undertaken by 3GPP under Release 12. This ambitious endeavor aims to push the boundaries further, targeting an exceptional peak data rate of up to 3 Gbps for the DL and 1.5 Gbps for the UL, coupled with a remarkable spectrum efficiency of 30 bps/Hz within a 20 MHz bandwidth.

OFDMA represents an Orthogonal Frequency Division Multiplexing (OFDM) system, incorporating a unique approach to data transmission. The fundamental principle of OFDMA lies in its ability to assign dedicated sub-channels for data transmission from different users, ensuring seamless connectivity. To further optimize the transmission process, a subset of sub-carriers, referred to as individualized OFDMA channels, is allocated to each user. This strategic allocation ensures that every user has exclusive access to a designated traffic channel at any given moment.

![LTE content transmission structure](image)

LTE-A networks, also known as next-generation LTE-A networks, have been specifically engineered to provide support for a wide spectrum of multimedia applications. These applications encompass an extensive range of functionalities, including but not limited to voice telephony, internet browsing, interactive gaming, video messaging, email services, and more. Key factors contributing to this complexity include the scarcity of radio resources, the highly dynamic and fluctuating channel conditions, and the persistent resource contention issues encountered among numerous users. Paramount considerations such as average packet delay, Packet Error Rate (PER), and minimum throughput further
2. RELATED WORKS

The Weighted-Fairness algorithm, proposed in [3], determines user selection by considering a comprehensive approach that incorporates both instantaneous downlink Signal-to-Noise Ratio (SNR) values and user priority weights. This novel algorithm intelligently allocates radio resources by taking account of not only channel conditions but also the importance and specific requirements of each user. The algorithm, known as Dynamic Proportional Allocation (DPA) was introduced [4]. DPA takes a proactive approach to address fairness concerns by dynamically adapting the allocation of packet transmission time based on changing network conditions and user demands. An Adaptive Modulation and Coding and Time Division Multiplexing (AMC/TDM) system prompted the development of the Dynamic Quality-of-Service (DQoS) algorithm [5, 6]. DQoS introduces an framework that considers both the QoS requirements of multimedia applications and wireless channels. Research has been conducted in system-level simulations of PHY resource allocation within LTE networks [7-9], load model [10] and multi-user allocation [11]. Literature surveys and reviews on the relationship between throughput and fairness are limited. While many studies discuss the overall QoS performance of well-known scheduling schemes [12][13][14][15][16].

The field of scheduling algorithms has attracted considerable research attention. Habaebi et al. (2013) conducted a comparative analysis of scheduling algorithms, including RR, Best CQI (Channel Quality Indicator), and PF schedulers, were validated for their performance in terms of throughput and block error rate (BER) using a MATLAB-based system-level simulator [17]. Capozzi et al. (2013) provided an extensive survey on downlink packet allocation strategies in LTE networks, focusing on QoS provision [18]. Abu-Ali et al. (2014) presented a detailed survey on LTE uplink schedulers, categorizing them into best-effort, QoS-based, and power-optimizing schedulers [19]. Zavyalova (2015) developed a simulation model to evaluate RR and Best CQI scheduling algorithms' system throughput under varying load intensities in LTE [20]. Shams, Abied & Hossain (2016) compared small cell networks (ScNet) and heterogeneous networks (HetNet) in terms of average UE throughput, cell edge throughput, and spectral efficiency for downlink performance metrics [21]. Subramanian, Sandrasegaran & Kong (2016) compared PS algorithms, including PF, MLWDF, and EXP/PF, using metrics such as throughput, PLR, delay, and fairness [22]. Mahdi, Ali Yahiya & Kirici (2019) investigated variable packet scheduling algorithms such as WFQ, PQ, and FIFO, optimizing QoS for real-time and non-real-time applications [23]. Thienthong et al. (2019) examined schedulers while considering the system’s integration of cell range expansion (CRE) and almost blank subframe (ABS) mechanisms [24].

3. RESEARCH METHODOLOGY

The efficient scheduling of packets is fundamental to the effective allocation of radio resources, serving as a critical factor in meeting the varied QoS requirements of users. Traditional packet scheduling algorithms, such as Randomized Priority Assignment (RPA), Spectrum-Aware Throughput Optimization (SATO), and Dynamic Fairness Maximization (DFM), take a comprehensive approach by optimizing system-level performance, including fairness and spectral efficiency. Several fundamental metrics are employed for evaluating the efficiency of resource allocation, encompassing system throughput, fairness in
throughput distribution among users, and the average PER between Real-Time (RT) and Non-Real-Time (NRT) users.

Our approach adopts a cross-layer design that uses the interactions between different layers to facilitate intelligent scheduling decisions, as illustrated in Figure 2. The scheduling process encompasses the Traffic Differentiator, the Time Domain (TD) Scheduler, and the Frequency Domain (FD) Scheduler. By integrating information from diverse layers, our methodology enables a network dynamics and facilitates optimized resource allocation. Through our research methodology, which capitalizes on a cross-layer design and leverages information from multiple layers, we proposed advanced scheduling algorithms designed to enhance resource allocation, maximize QoS provisioning, and optimize network performance. By taking account of the unique requirements and dynamics across different layers, our methodology signifies a substantial step toward the realization of robust and efficient wireless communication systems.

For each user, the Traffic Differentiator extracts an array of crucial information from various layers. At the application layer, it captures traffic type specifics, shedding light on the unique characteristics and requirements of the data being transmitted. The network layer contributes QoS-related insights, including parameters like delay budget and minimum required throughput which significantly influence scheduling decisions. Furthermore, the Radio Link Control (RLC) layer supplies queue status information, enabling a real-time assessment of the data waiting to be transmitted. Lastly, the physical layer furnishes channel status information, empowering the system to adapt scheduling strategies based on
prevailing channel conditions. The TD scheduler plays a key role in the process, leveraging QoS measurements, specifically the average PER obtained from the MAC layer. Simultaneously, the FD scheduler operates to map the Physical Resource Blocks (PRBs) in alignment with user priorities and the CQI reports specific to each PRB.

4. PROPOSED RESOURCE ALLOCATION ALGORITHM

4.1. Traffic Differentiator

Within the scope of this study, a novel approach to packet scheduling has been presented, introducing a cutting-edge architecture that effectively manages mixed traffic.

![Traffic Differentiator Stage](image)

Figure 5. Traffic differentiator Stage

The architecture incorporates six distinct queues to accommodate various types of data. Specifically, these queues comprise a RT queue, which caters to emergency message content in addition to RT and NRT queues dedicated to streaming video services. Prioritization is applied hierarchically, with the highest priority assigned to emergency message content, followed by streaming video services. This approach guarantees the orderly and timely delivery of essential information, contributing to enhanced communication reliability. Moreover, Non-Real-Time streaming video services require a minimum throughput guarantee over an extended duration to ensure optimal quality for video streaming purposes.

![Dynamic C-mean clustering algorithm](image)

Figure 6. Dynamic C-mean clustering algorithm.
As part of our research, we introduce two innovative queue sorting algorithms specifically designed for prioritizing emergency messages, real-time and non-real-time services. These algorithms, referred to as Service Priority Specific queue Sorting Algorithms (SPSSA), aim to optimize system performance while ensuring fairness among different service types. To increase efficiency, we employ the $P_{i,k}$ coefficient as a key metric for prioritizing users in the background service queue. The $P_{i,k}$ coefficient enables us to strike a delicate balance between system throughput and fairness considerations. The SPSSA algorithms offer an applied approach to queue sorting, taking account of the specific requirements and priorities associated with all type of services.

The comprehensive functionality of the Traffic Differentiator is illustrated in Figure 6. This module plays a critical role in effectively distinguishing and categorizing mixed traffic into three distinct groups: RT emergency users, NRT streaming users, and their corresponding control information stored in a dedicated queue. To optimize the prioritization of RT users, a dynamic c-mean clustering algorithm is employed. This algorithm takes account of the individual $P_{i,k}$ priority coefficient of each selected RT user, enabling a refined arrangement of priorities. The goal of the prioritization process is to enhance the $P_{i,k}$ value for each user, resulting in improved overall performance. By utilizing the dynamic c-mean clustering algorithm and considering the individual $P_{i,k}$ priority coefficients, the Traffic Differentiator ensures a more efficient and effective distribution of resources. This enables better management and prioritization of RT emergency users, ultimately leading to enhanced $P_{i,k}$ values and an overall superior system performance.

4.2. Time Domain Scheduler

In the TD Scheduler stage, a cutting-edge algorithm known as the Adaptive Resource Allocation Algorithm with Pseudo-Inverse (ARA-PI) is introduced. This revolutionary algorithm revolutionizes resource allocation by leveraging the power of pseudo-inverse computation. By utilizing this advanced mathematical technique, the algorithm dynamically determines the optimal number of RT and NRT users to be scheduled during each Transmission Time Interval (TTI). The primary objective of the ARA-PI algorithm is to achieve a well-balanced allocation of resources across various service types, thereby promoting optimal system-level performance. By incorporating the ARA-PI algorithm in the TD Scheduler stage, the proposed packet scheduling architecture achieves remarkable adaptability and flexibility in resource allocation. This results in a highly efficient and balanced utilization of resources, ultimately leading to superior system-level performance.

4.3. Frequency Domain Scheduler

The literature on channel-aware scheduling predominantly assumes the availability of accurate CQI reports from all users to the eNB in each Transmission Time Interval (TTI). This assumption forms the foundation for many existing studies and research works in the field. Recognizing the significance of precise channel information for effective scheduling decisions, the assumption of accurate and readily available CQI reports enables the eNB to make informed choices regarding resource allocation. By leveraging this vital information, the eNB can optimize the scheduling process and allocate resources based on the real-time channel conditions of each user. While acknowledging the importance of accurate CQI reports, it is essential to note that practical implementations may encounter challenges such as measurement errors or delays in report transmission. Therefore, it is important for future studies to investigate and address these potential limitations to ensure the robustness and reliability of channel-aware scheduling algorithms.
The allocation of available resource blocks to users follows a dynamic and iterative process. In each iteration, a single Resource Block (RB) is allocated to the user, optimizing a novel priority function specifically designed for this purpose. This function determines the most suitable RB allocation for the user, considering factors such as channel conditions, service requirements, and system-level performance. By adopting an iterative approach, the allocation process aims to continually refine and enhance the RB assignments, ensuring an optimal utilization of the available resources. This iterative nature allows for adaptive adjustments and fine-tuning based on the dynamic changes in the network environment and user demands. It is worth noting that the proposed priority function is a key component of the allocation process, as it plays a vital role in guiding the selection of RBs.

The research results and experimentation are necessary to evaluate the effectiveness and efficiency of the proposed iterative RB allocation approach, as well as to explore potential enhancements and refinements to the priority function, ensuring its applicability across various network scenarios and deployment settings.

\[
PRF(k) = \frac{RSRP_t(k) \times BLER_t(k) \times P_i(k)}{RSSI_t \times PER_t \times p_i}
\]  

\( RSRP_t(k) \) is the signal strength measurement for user \( k \) is a crucial parameter in wireless communication systems. It provides the quality and reliability of the signal received by the user. \( BLER_t(k) \) is the block error Rate (BLER) for user \( k \) is a significant metric in wireless communication systems that quantifies the reliability of data transmission. It represents the probability of encountering errors or corrupted blocks of data received by user \( k \). \( P_i(k) \) is the service priority for user \( k \) is a crucial aspect of network resource allocation and management. It determines the level of importance assigned to the user \( k \)'s communication needs and influences the scheduling and allocation decisions within the system. \( RSSI_t(k) \) is the total signal strength for a specific Resource Block (RB) plays a significant role in assessing the overall quality of the wireless channel and determining the potential performance of data transmission within that RB. \( PER_t \) is a critical performance metric that quantifies the reliability of data transmission over a specific set of available Resource Blocks (RBs) in a wireless communication system. It represents the probability that a packet sent over the RBs will receive errors at the intended receiver. In addition, \( p_z \) is the concept of maximum priority plays a crucial role in service provisioning and resource allocation in various communication systems. It refers to the highest level of importance or urgency assigned to each service based on specific criteria and requirements.

The aggregate transmit power of an OFDM symbol within a Transmission Time Interval (TTI) can be mathematically represented by the following expression:

\[
RSSI_t = \sum_{r=0}^{R} (a_{k,r} \cdot RSRP_t(k)) \quad \forall r \in S_r^k
\]  

In this context, we introduce our novel algorithm for selecting the optimal CQI.
5. EXPERIMENTAL RESULTS AND ANALYSIS

The Best-CQI scheduler focuses on maximizing of the CQI value within the system, prioritizing superior performance. The Max-Min Throughput algorithm aims to maximize the minimum throughput across the users. As a result, the Round-Robin scheduler emerges as a proportional fair scheduler, distribution of system resources given configuration. These outcomes underline the algorithm's effectiveness and ability to satisfy performance requirements, showcasing its suitability for real-world implementation.

In Figure 7, we analyze the Cumulative Distribution Function (CDF) of throughput for following two methods: the Best CQI method and the Optimal CQI. We aim to determine which method performs better and provide reasoning for this assessment. The Best CQI method exhibits a smoothly increasing CDF, starting at 0.4022 and gradually reaching a peak of 0.9738. This method's CDF demonstrates a continuous improvement in throughput, which suggests a consistent and favorable performance across the entire range of throughput values.

On the other hand, the Optimal CQI 2 method shows a somewhat erratic CDF pattern. While it attains a slightly higher peak CDF of 0.9905, the progression is marked by discrete steps, indicating variations in throughput performance at specific points. If the objective is
to achieve a consistently smooth and gradual improvement in throughput, the Best CQI method appears to be the more suitable choice due to its continuous CDF trend. Conversely, if the goal is to prioritize higher peak throughput values, the Optimal CQI method may be preferred, even though its progression is less smooth.

In Figure 8, we examine the CDF of spectrum efficiency for the following two methods: the maxminTP method and the Optimal CQI method. The goal is to determine which method performs better and provide an explanation for this assessment. The maxminTP method demonstrates a gradual increase in its CDF, starting at 0.2614 and reaching a peak of 1. This method showcases a steady improvement in spectrum efficiency, indicating consistent performance across the entire spectrum efficiency range. The CDF curve is relatively smooth to suggest a reliable and stable spectrum allocation strategy.

The Optimal CQI method exhibits a similar trend, but with some noticeable differences. While it starts slightly higher at 0.3295 and also reaches a peak of 1, its CDF curve displays some fluctuations and irregularities. These fluctuations imply variations in spectrum efficiency performance at specific points along the spectrum. If the aim is to achieve a consistently smooth and steady improvement in spectrum efficiency, the maxminTP method appears to be the most suitable choice due to its continuous and stable CDF trend. In contrast, if the goal is to prioritize slightly higher than initial spectrum efficiency values despite some fluctuations, the Optimal CQI method may be preferred. However, based on the consistency and stability of the CDF curve, the maximum method may be considered more favorable for most scenarios.

In Figure 9, we examine the CDF of SINR for the following two methods: the Best CQI and the Optimal CQI methods. The goal is to determine which method performs better and provide an explanation for this assessment. The Best CQI method demonstrates a gradual increase in its CDF, starting at 0.2614 and reaching a peak of 1. This method showcases a steady improvement in SINR, indicating consistent performance across the entire SINR range. The CDF curve is relatively smooth to suggest a reliable and stable SINR allocation strategy.

The Optimal CQI method exhibits a similar trend, but with some noticeable differences. While it starts slightly higher at 0.3295 and also reaches a peak of 1, its CDF curve displays some fluctuations and irregularities. These fluctuations imply variations in SINR performance at specific points along the spectrum. If the aim is to achieve a consistently smooth and steady improvement in SINR, the maxminTP method appears to be the most suitable choice due to its continuous and stable CDF trend. In contrast, if the goal is to prioritize slightly higher than initial SINR values despite some fluctuations, the Optimal CQI method may be preferred. However, based on the consistency and stability of the CDF curve, the maximum method may be considered more favorable for most scenarios.
In Figure 9, we analyze the CDF of Signal-to-Interference-plus-Noise Ratio (SINR) for the following two methods: the Best CQI method and the Optimal CQI method. The objective is to determine which method performs better and explain the assessment based on numerical results. The Best CQI method starts with a SINR value of 0.0094 and gradually increases, reaching a maximum SINR of 26.3537. The corresponding CDF values increase steadily from 0.1986 to 1. This indicates that the best CQI method provides a consistent improvement in SINR across the entire range of values, with no abrupt changes or irregularities in the CDF curve. On the other hand, the Optimal CQI method starts with a higher initial SINR value of -0.0065 and gradually increases to a maximum SINR of 24.3053. The corresponding CDF values follow a smooth upward trend from 0.1381 to 1. While the Optimal CQI method offers a slightly higher initial SINR value, the overall performance is similar to the Best CQI method, characterized by a consistent and continuous increase in SINR. Both methods demonstrate stable and reliable performance in improving SINR.

![Figure 10. SINR vs Throughput](image)

In Figure 10, we compare the Throughput and SINR performance of two scheduling methods: the Round Robin method and the Optimal CQI method. To determine which method is the best, we evaluate their respective throughput values against SINR ratios. The Round Robin method starts with a SINR of approximately 8.57 and a corresponding Throughput of 8.66. As the SINR gradually increases, the Throughput also increases, reaching a maximum of 57.94 when the SINR is around 20.37. The Round Robin method shows a consistent increment in Throughput as SINR improves.

On the other hand, the Optimal CQI method begins with a higher initial SINR of about 13.06 but a lower Throughput of 4.87. As SINR increases, the Throughput also rises, peaking at 50.04 when the SINR reaches approximately 16.94. The Optimal CQI method exhibits a more pronounced increase in Throughput as SINR improves, compared to the Round Robin method. The Optimal CQI method achieves higher Throughput values for similar SINR indicating a more efficient resource allocation strategy. However, the choice between these methods should consider the specific system requirements, as higher Throughput may come at the expense of other performance factors, such as fairness or system complexity. In summary, based on the numerical results and the trade-off between SINR and Throughput, the Optimal CQI method appears to be the more efficient choice for resource allocation, as it achieves higher Throughput values for similar SINR levels compared to the Round Robin method.
6. CONCLUSIONS

The Optimal CQI selection algorithm was compared to existing algorithms, and extensive testing was conducted in diverse scenarios to mimic real-world systems. The results demonstrate enhanced QoS for real-time and non-real-time services while striking a favorable balance between user-level and system-level performance. Although significant progress has been made in LTE downlink scheduling, further research opportunities remain. The proposed scheduling algorithm represents an initial step towards achieving a trade-off between throughput and fairness, and future investigations can focus on optimizing the throughput of this algorithm. Depending on the specific objectives of the scheduling algorithm, enhancements can improve throughput, fairness, or both.

In summary, the Best CQI method show us smooth improvement in throughput with a peak CDF of 0.9738, while the Optimal CQI method achieves a slightly higher peak CDF of 0.9905 but with variations, making the Best CQI method make good consistent throughput performance. In spectrum efficiency, the maxminTP method makes good consistent and stable spectrum allocation. In SINR, with the Best CQI method starting at a lower initial SINR but achieving a slightly higher peak SINR. The Optimal CQI method achieves higher throughput for similar SINR levels compared to the Round Robin method, suggesting more efficient resource allocation.

Compared to the suboptimal algorithms, the optimal algorithms demonstrate slightly better performance. It is worth noting that the latter algorithms offer an acceptable level of performance while providing additional advantages in terms of effectiveness and practicality. In practice, suboptimal methods are highly valuable as they can be readily implemented in real systems, making them more feasible and efficient. Therefore, despite their slight compromise in performance, suboptimal algorithms remain a favorable choice for real-world applications.

REFERENCES

BIOGRAPHIES

Erdenebayar Lamjav is a telecommunication engineer (2006), master (2011) and doctor in Radio communication Engineer (2018) with the Mongolian University of Science and Technology. He is the coordinator of Multimedia and Computer Vision Laboratory. He is currently an associate professor at MUST. His areas of interest are: Information and Communication Technology, DSP, Multimedia Technology.
Otgonbayar Bataa is a radio communication engineer (1978), master (1995) and doctor in Communication Engineer (1996) with the Moscow Institute of Telecommunication. She is currently a consultant professor at MUST. Her areas of interest are: Advanced Communication, Mobile IPTV, Mobile and Wireless access. Dr. Bataa is member of the IEEE.

Chuluunbandi Naimannaran is a measurement electronics engineer (1995), master (1997) and doctor in Radio communication Engineer (2011) with the Mongolian University of Science and Technology. He is currently a director at School of Information and Communication Technology of MUST. His areas of interest are: RF engineering, Microwave, Antenna Design.