IOSR Journal of Mobile Computing & Application (IOSR-JMCA)
e-ISSN: 2394-0050, P-ISSN: 2394-0042.Volume 4, Issue 3 (Jul. - Aug. 2017), PP 01-12
www.iosrjournals.org

Downlink Scheduling Algorithms in LTE Networks: A Survey

Muhammad Aminu Lawal[1], Ibrahim Saidu[2], Aminu Mohammed[3], Yusra Abdullahi Sade[4]

[1, 2]*(Department of Mathematics and Computer Science, Umaru Musa Yar'adua University Katsina, Katsina State, Nigeria.)*
[3]*(Department of Mathematics, Usmanu Danfodio University, Sokoto, Nigeria)*
[4]*(Department of Physics, Umaru Musa Yar'adua University Katsina, Katsina State, Nigeria.)*

Corresponding Author: Muhammad Aminu Lawal

Abstract: *Long Term Evolution (LTE) standard is conceived as an all IP network to achieve higher data rate, low latency, scalable bandwidth, mobility and extended coverage. The network guarantees Quality of Service (QoS) for diverse applications such as VoIP, video and web browsing according to the Third Generation Partnership Project (3GPP) specifications. The Radio Resource Management (RRM) techniques such as packet scheduling algorithm play a vital role in providing such guarantees. Thus, several algorithms have been proposed to allocate bandwidth resources while ensuring QoS to wireless applications. This paper presents a survey of downlink scheduling algorithms. These algorithms are classified into QoS unaware and QoS aware. The operational procedure, strengths and weaknesses of each algorithm are discussed. The comparative analysis of these algorithms is also presented. The analysis provides an insight on open research issues for future research.*

Keywords*: LTE networks, QoS aware, QoS unaware, Radio Resource Management.*

Date of Submission: 16-08-2017 Date of acceptance: 28-08-2017

I. Introduction

Recently, the increase in the level of traffic such as VoIP, Video, Web browsing e. t. c with diverse Quality of Service (QoS) requirements has strained the capability of the existing wireless networks. Report according to Cisco indicates that the level of mobile data traffic has grown exponentially and will continue to increase by 1000 times in the next five years [1]. The continuous growth of these traffics and the need to achieve required QoS of the emerging wireless applications necessitate the industrial and research communities to provide better solutions in wireless communication systems. One of these solutions is the LTE networks introduced by the 3GPP in order to achieve higher data rate, low latency, scalable bandwidth, mobility and extended coverage.

The LTE network adopts Orthogonal Frequency Division Multiple Access (OFDMA) for downlink transmissions. It adopts a scalable radio resource bandwidth of 1.4 MHz to 20 MHz. This radio resource bandwidth is divided into equal sub-channels of 180 KHz each in frequency domain and a Transmission Time Interval (TTI) of 1ms each in time domain. A TTI comprises of two time slots of 0.5 ms each. Thus, a radio resource in time/frequency domain across one time slot in time domain and one sub-channel in frequency domain is termed a Resource Block (RB). A RB is the smallest unit of radio resource that can be allocated to a User Equipment (UE) for data transmission.To efficiently allocate RBs while providing QoS to the downlink flows, radio resource management (RRM) techniques such as packet scheduling algorithm is highly needed. Therefore, several algorithms have been proposed to allocate radio resources while ensuring QoS to wireless application [2] .

In this paper, a survey of the downlink scheduling algorithms is presented. The algorithms are classified into QoS aware and QoS unaware. The operational procedures, strengths and weaknesses of each algorithm are highlighted. The comparative analysis of these algorithms is also discussed in order to provide open research issues for future direction. The remainder of the paper is organized as follows: Section 2, presents an overview of LTE system. Section 3, describes a survey of the downlink scheduling algorithms and comparative analysis. Finally, Section 4 concludes the paper.

II. Overview Of The LTE Networks

The LTE network was designed to surpass the attributes of 3G networks [3] .It targets doubling the spectral efficiency; improving on the bit rate of cell edge users compared to the earlier networks[2]. Table 1. shows a summary of the main LTE performance targets.

Table 1: Main LTE Performance Targets [2] .

Performance Metric	Target
Peak Data Rate	Downlink: 100 Mbps
	Uplink: 50 Mbps
Spectral Efficiency	2 - 4 times better than 3G systems
Cell-Edge Bit-Rate	Increased whilst maintaining same site locations as deployed today
Mobility	Optimized for low mobility up to 15 km/h
	High performance for speed up to 120 km/h
	Maintaining connection up to 350 km/h
Scalable Bandwidth	From 1.4 to 20 MHz
RRM	Enhanced support for end-to-end QoS
	Efficient transmission and operation of higher layer protocols
Service Support	Efficient support of several services (e.g., web-browsing, FTP, video-streaming, VoIP)
	VoIP should be supported with at least a good quality as voice traffic over the UMTS network

III. LTE Network Architecture

The LTE network is built on a flat architecture called the Service Architecture Evolution shown in Fig. 1.The figure consists of the radio access network and the Evolved Packet Core (EPC). The EPC provides the overall control of the UE and establishment of the bearers [4] which consists of Mobility Management Entity (MME), Serving Gateway (SGW), and Packet Data Network Gateway (PGW). The MME controls handover within LTE, user mobility, and UEs paging as well as tracking procedures on connection establishment. The SGW performs routing and forwarding of user data packets between LTE nodes as well as handover management between the LTE and other 3GPP technologies. The PGW connects the LTE network with other IP networks around the globe and provides the UEs access to the internet[2].The radio access network known as the Evolved-Universal Terrestrial Radio Access Network (E-UTRAN) performs all radio related functions [4] ,which comprises of the eNB and the UE. The UE represents the different types of devices used by the users while the eNB performs radio resource management (RRM) functions along with control procedures for the radio interface such as packet scheduling, CAC etc.

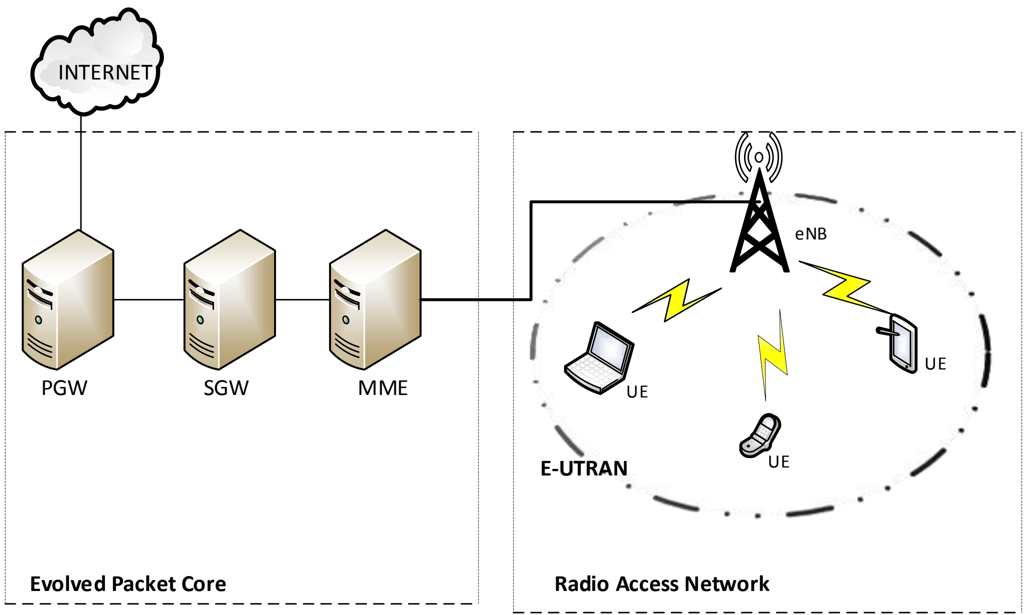

Figure 1 The Service Architecture Evolution of LTE Network.

IV. Quality Of Service (QOS) And Evolved Packet System (EPS) Bearers

The LTE's QoS structure is conceived to grant an end-to-end QoS support [5]. Towards this objective, the LTE permits flow differentiation based on the QoS requirements. These QoS requirements are managed by radio bearers which are classified into two: default and dedicated. The default bearer which corresponds to non

Guaranteed Bit Rate (non-GBR) is created at the beginning of every connection. It does not grant bit rate guarantees and remains until the end of the connection. The dedicated bearer which represents either GBR or non-GBR is created every time a new service is issued[2]. Every bearer has an associated QoS class identifiers (QCI) shown in Table 2.

Table 2: Standardized QoS Class Identifiers (QCI) for LTE [4].

QCI	Resource Type	Priority	Packet Delay Budget (ms)	Packet Loss Rate	Example Service
1	GBR	2	100	10^{-2}	Conversational voice
2	GBR	4	150	10^{-3}	Conversational video (live streaming)
3	GBR	5	300	10^{-6}	Non-Conversational video (buffered streaming)
4	GBR	3	50	10^{-3}	Real time gaming
5	Non-GBR	1	100	10^{-6}	IMS signaling
6	Non-GBR	7	100	10^{-3}	Voice, video (live streaming), interactive gaming
7	Non-GBR	6	300	10^{-6}	Video (buffered streaming)
8	Non-GBR	8	300	10^{-6}	TCP based (e.g., WWW, e-mail),
9	Non-GBR	9	300	10^{-6}	chat, FTP, P2P file Sharing

V. Air Interface

The LTE physical layer employs OFDMA and SC-FDMA as the radio spectrum access method in the downlink and uplink, respectively. Both OFDMA and SC-FDMA permit multiple accesses by allocating sub-carriers to every user. The OFDMA utilizes the sub-carriers within the whole spectrum; it offers high scalability and robustness as well as simple equalization to prevent time-frequency selective nature of radio channel fading. The SC-FDMA exploits only the adjacent sub-carriers; it is employed at the uplink to improve power efficiency of user equipment since they are mostly battery dependent[2].

VI. Resource Management

In LTE network the radio resources are shared to users in a time/frequency domain as shown in Fig. 2. The time domain is divided into frames; every frame is made up of 10 successive TTIs and each TTI lasts for 1ms. In addition, every TTI consists of two time slots with duration of 0.5ms. In the frequency domain, the entire bandwidth is partitioned in to sub-channels of 180 KHz each. Therefore, a time/frequency radio resource ranging across one time slots in the time domain and one sub-channel in frequency domain is known as a resource block (RB). A RB is the minimum radio resource unit that can be allocated to user equipment for data transmission. While number of Resource Blocks (RBs) corresponds to the configuration of the system bandwidth e.g. 6, 12, 25, 50, 75, or 100 RBs corresponds to 1.25 MHz, 2.5 MHz, 5 MHz,10 MHz, 15 MHz, or 20 MHz , respectively [6].

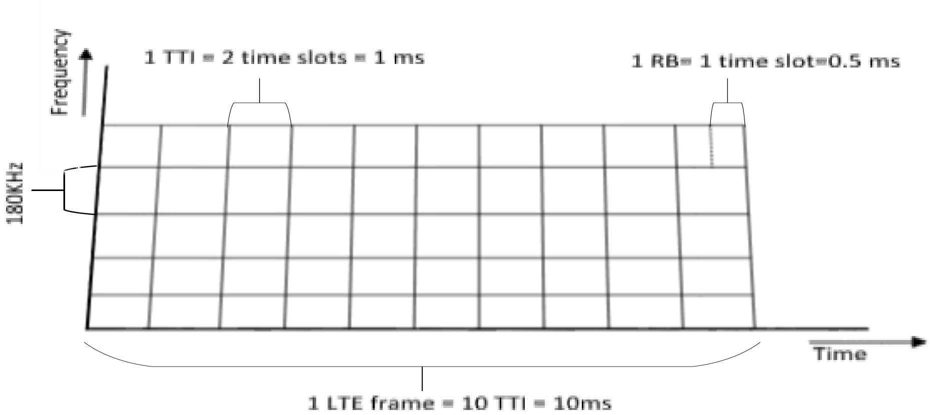

Figure 2 Radio Resources in Time/Frequency Domain.

VII. Model Of Packet Scheduler

The model shown in Fig. 3 consists of the UE and the eNB where the packet scheduler and the Adaptive Modulation and Coding (AMC) module are located. The UE sends the calculated Channel Quality Indicator (CQI) report to the eNB based on the channel condition. The packet scheduler uses CQI report to make decisions and fills up RB allocation "mask". The Adaptive Modulation and Coding (AMC) module chooses the optimum Modulation and Coding Scheme (MCS) for transmission of the scheduled users. The Physical Downlink Control Channel (PDCCH) conveys user information, RB allocation and the chosen MCS to the UE. The UE decodes the PDCCH payload to check whether it is scheduled so that it can access the right PDSCH payload. These series of operations are repeated at each TTI [2].

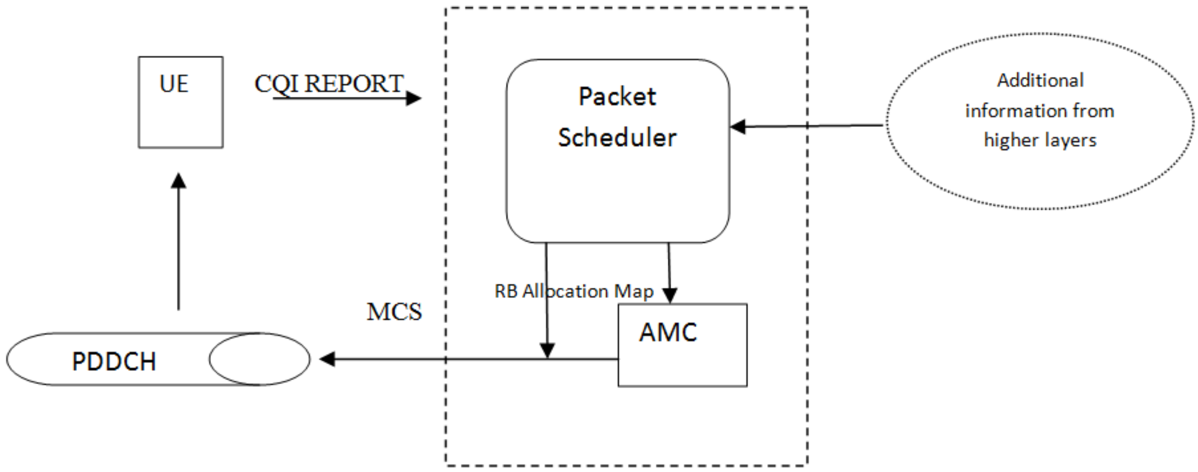

Figure 3 A model of Packet Scheduler.

VIII. QOS Unaware Algorithms

QoS unaware scheduling algorithms provides only the throughput among users without consideration to the QoS requirements such as delay constraints and packet loss rate and hence these algorithms are unsuitable for wireless multimedia traffics [2]. These algorithms are reviewed as follows:

In [7a], a Maximum Throughput (MT) algorithm was proposed to improve the system spectral efficiency. The MT serves users with best Channel Quality Identifier (CQI).It achieves maximum throughput and hence improves spectral efficiency but leads to unfairness because users under bad channel conditions are starved.

In [7b], a Blind Equal Throughput (BET) algorithm was proposed to provide fairness. The BET serves users with equal throughput irrespective of their channel conditions. It uses the past average throughput as a factor which governs its allocation. The algorithm achieves high level of fairness but suffers poor spectral

A Proportional Fair (PF) algorithm [8] was proposed to address the problem of both the unfairness and the system spectral efficiency in [7]. The PF serves users according to ratio of achievable instantaneous throughput and time averaged throughput. It achieves fair share of resources to users and improve spectral efficiency but real-time applications have poor QoS because delay constraints of real- time applications are ignored.

In [9], a Generalized Proportional Fair (GPF) algorithm was propose to regulate the trade-off between spectral efficiency and fairness for best effort traffic. The GPF introduces two weighting factors: a and b to adjust the effect of allocation policy of achievable instantaneous throughput and time averaged throughput. A conventional PF is achieved when a=b=1. The algorithm is skewed to either BET when a = 0 and b = 1 or MT when a=1 and b = 0. The algorithm provides a higher spectral efficiency or higher level of fairness depending on how the weighting factors are set but fails to adapt the weighting factors in a running system [10].

A Delay-Based Weighted Proportional Fair (DBWPF) algorithm [11] was proposed to achieve delay fairness and implementation rate fairness. The DBWPF algorithm uses a weighted average delay based on PF to distribute resources to the users with non-empty buffers. The algorithm achieves the delay fairness, implementation rate fairness as well as approximate throughput and throughput fairness. However, the algorithm experience poor throughput when users with higher average delay under heavy bursty traffic are considered.

An Optimal and Data Rate Guaranteed Radio Resource Allocation algorithm [12] was proposed to meet the minimum data rate requirement of each user. The algorithm classifies users into priority and non-priority. The priority users are assigned resources first and then the non-priority users when the remaining resources are available. It also uses data rate to assign order of RB allocation and introduces margin as well as optimality to prevent waste of RB under given condition. The algorithm achieves an efficient utilization of resources and guarantees rate requirement for large number of users but starves the non-priority users when resources are insufficient.

IX. QoS Aware Algorithms

QoS aware scheduling algorithms consider QoS requirements of users based on the traffic characteristics such as delay constraints and packet loss rates. These algorithms are reviewed based on the operational procedures, strengths and weaknesses as follows:

In [13], a Modified-Largest Weighted Delay First-Virtual Token (M-LWDF-VT) algorithm was proposed to improve the QoS of real-time services. The M-LWDF-VT algorithm combines M-LWDF with a token mechanism to ensure not only the delay but also the minimum throughput to flows. The algorithm enhances throughput in real-time services. However, it starves non-real time services because most resources are allocated to video flows [14][15].

In [14], a downlink scheduling algorithm was proposed to enhance interclass fairness. The algorithm modifies the M-LWDF [16] and the M-LWDF-VT [13] algorithms by considering packet delay and queue size of each flow. The algorithm enhances QoS parameters of diverse class of traffics. However, the algorithm increases PLR of VoIP traffic under large number of UEs.

A novel delay based scheduling algorithm [17] was proposed to improve the throughput of video traffic. The algorithm derives a metric based on some of the properties of the LOG RULE and the MT algorithms in order to allocate resources. The algorithm enhances throughput and Packet Loss Rate (PLR).However, it starves non real time flows are starved due to degradation policy when the video traffic is high.

In [18], an Exponential Earliest Deadline First (EXP-EDF) algorithm was proposed to guarantee QoS of real-time applications. The EXP-EDF algorithm utilizes the characteristics of multi user diversity, packet deadline and difference in the channel quality of transmission to allocate resources to flows. It provides QoS guarantees for real-time flows but starves non real-time traffic due to high priority given to real-time traffic.

In [19], a Modified-Earliest Deadline First-Proportional Fair (M-EDF-PF) algorithm was proposed to improve QoS of video and VoIP. The M-EDF-PF algorithm employs the EDF algorithm to schedule flows with closest expiration deadline and the PF algorithm to fulfill throughput as well as guarantee fairness among flows. The algorithm introduces adjustable factors for flexible resource allocation to real-time services. The algorithm improves the QoS requirements of real-time services but experiences delay under large number of users.

A Deadline based scheduler algorithm [20] was proposed to enhance performance and fairness in allocation of radio resources. The algorithm assigns a deadline for each flow that has a packet queued at the eNB. The deadlines are computed based on maximum delay of the flow and the Head of Line (HOL) delay of the flow. The algorithm allocates RBs based on metrics computed from the deadline, average transmitted data, and spectral efficiency. It improves performance in terms of packet loss rate and delay as well as throughput for video flows but increase delay under large number of UEs.

In [21], a RB Preserver Scheduler algorithm was proposed to provide QoS for real time flows. The algorithm allocates resources in two levels. The first level utilizes the LTE frame characteristics of combining several sub-frames in order to assure the user's QoS. The second level employs PF_MAX, Delay Rule (DR), EXP/PF and Weighted Delay (WD) algorithms to allocate Resource Blocks (RBs) to real time flows in order to satisfy user's QoS. In addition, the level also uses PF to assign RBs to non real-time flows in order to achieve fairness. The algorithm ensures the QoS requirements of RT flows. However, the algorithm is unfair to non real-time flows because higher ratio of RBs is allocated to real-time flows.

In [22], a Delay Prioritized Scheduling (DPS) algorithm was proposed to maximize system throughput and satisfy QoS requirements of video streaming applications. The DPS algorithm uses packet delay information and instantaneous channel conditions in making scheduling decisions. It enhances system throughput and maintains lower average system delay as well as a fair resource allocation. However, the algorithm fails to optimize system throughput and PLR performance in heterogeneous real-time traffic environment due to its inability to differentiate the QoS level of diverse real-time traffics[23].

A QoS-Aware scheduling algorithm [24] was proposed to limit the resources used by the real-time traffic. The algorithm employs a Time Domain (TD) and Frequency Domain (FD) scheduler. The TD scheduler classifies bearers into Guaranteed Bit Rate (GBR) and non- Guaranteed Bit Rate (non GBR). It further chooses users requesting resources based on QoS requirements and current status of the channel. The FD utilizes the token bucket to identify GBR bearers that can be assigned the resources and also uses M-LWDF and PF to distribute the resources to GBR and non GBR, respectively. The algorithm satisfies the QoS requirements and improves the cell performance. However, it violates the delay budget under high traffic due to token bucket used to limit GBR traffic.

In [25], a QoS aware scheduling algorithm was proposed to optimize video delivery quality. It utilize the available channel rate of a user on a given resource block, packet delay constraint of the video application and the historical average data rate of each user on a given resource block to dynamically allocate resources using Weighted Round Robin algorithm (WRR) .The algorithm further employs a cross layer framework using a mean square error (MSE) between the received pixels and original pixels of the video frames as a distortion metric [26] to improve user perceived video quality. The algorithm enhances the radio resource allocation as well as the user perceived video quality for end users but unfair because VoIP and non-real time applications are ignored.

A Two- level downlink scheduling algorithm [27] was proposed to guarantee bounded delays of real-time applications. The first level employs a discrete-time linear control theory to compute amount of data to be transmitted in a frame for each service flow. While the second level uses PF to allocate resource blocks (RBs) to real-time flows in each TTI and utilizes the leftover RBs for the best effort service. The algorithm improves network performance and Quality of Experience (QoE), but suffers unfairness problem because non-real time applications are scheduled if and only if real-time flows are satisfied. It also suffers low resource utilization because spectral efficiency is neglected and hence degrades performance when the system load is above a certain utilization threshold [28] .

In [29], a Frame Level Scheduler-Advanced (FLSA) algorithm was proposed to prioritize real-time flows. The FLSA allocates radio resources to users in three levels. The first level employs quota of data formula in [27] to estimate the amount of data a real-time flow transmits in each TTI. The second and third level uses M-LWDF algorithm to distribute RBs to real-time flows and allocate the remaining RBs to real-time and best effort, respectively. The algorithm achieves a better resource allocation for real-time flows but starves non real-time flows under high real-time flows.

A Packet Prediction mechanism (PPM) algorithm [30] was proposed to support QoS of real-time applications. The mechanism uses three stages to allocate resources. The first stage considers a user with the best CQIs. The second stage utilizes the prediction mechanism to determine packets that may fail to meet the delay requirements based on the queue status using a virtual queue. While the third stage employs a cut-in process to re-arrange the transmission order and ignores packets that fail to meet delay demand. It achieves an acceptable enhancement in terms of goodput and invalid packet rate. But the algorithm suffers packet drop of VoIP due to large demand of video traffic [6] and fails to classify real-time applications according to priorities.

In [16], a Modified-Largest Weighted Delay First (M-LWDF) algorithm was proposed to guarantee Quality of Service (QoS) for real time applications. The M-LWDF serves users based on the channel condition and the queue status of each user. The algorithm increases the number of users supported with the QoS and provides minimum throughput guarantees. However, it starves non-real time application and fails to provide the QoS requirements according to 3GPP specifications[15].

In [31], an Exponential rule (EXP RULE) and a Logarithmic rule (LOG RULE) was proposed to provide QoS to heterogeneous traffic. The algorithms serve users using equation 3a and 3b respectively, according to the channel condition and queue status. These algorithms guarantee the QoS requirements with bounded but algorithms fail to conform to the QoS requirements of 3GPP specifications [15].

In[15], a scheduling framework for the downlink of LTE system was proposed to satisfy the QoS requirements as defined by the QoS architecture of 3GPP specifications. The framework classifies flows into GBR bearers and non-GBR bearers according QCI values provided in [32][33]. It uses a delay-dependent scheme obtained using a sigmoid function [34] combined with a rate shaping function to schedule GBR bearers and a utility maximization scheme based on [35] to schedule non-GBR bearers. The framework integrates the schemes as a novel algorithm to determine relative priority among QCIs that fail to meet the Packet Delay Budget (PDB) for all bearers. It achieves the QoS requirements based on 3GPP specifications and provides an improved spectral efficiency but non-GBR may experience Packet drop during congestion.

A service-differentiated downlink flow scheduling (S-DFS) algorithm [6] was proposed to satisfy the QoS demand of GBR flows and ensure throughput fairness of non-GBR flows. It allocates resources to flows in four phases: In the first phase, it uses a CQI and QCI as the basis for allocating the RBs to User Equipment (UEs). In the second phase, packets that have not met the delay requirement in the current TTI are dropped in order to trim the queue and predict flows that will encounter packets drop in the subsequent TTI. In the third phase, reallocatable RBs i.e. the RBs not utilized in the first phase due to multiple allocations can be further allocated to other flows in the next phase. In the fourth phase, re-assignment of RBs obtained from the third phase is executed based on the flows that may experience packet drop in the second phase; the flows compete for these RBs based on the queue status and the QCI. The algorithm improves the cell spectral efficiency, alleviate the dropping ratio of VoIP flows, reduce the average delay of video flows, and also keep higher CBR data throughput. However, the algorithm experiences low video throughput under high number of UEs. Also, it incurs high scheduling complexity due to metric base flow selection per TTI [28] .

In [36], an Adaptive Exponential/Proportional Fair (EXP/PF) algorithm was proposed to guarantee QoS requirements of real-time applications and assures throughput demand of non-real time applications. The EXP/PF employs two schemes: EXP rule and PF. The EXP allocates resources to real-time applications while the PF distributes resources to the non-real time applications. It provides QoS to real-time applications and assures a satisfactory system throughput. However, it is unfair to non real-time application under large number of users [37] .

In [38], a two level scheduling algorithm was proposed to provide fair resource distribution. The upper level uses a game theory and Shapley value concept to provide a fair distribution of resources. The lower level employs EXP/PF to enhance the level of fairness and throughput among real-time and non-real time flows. The algorithm provides an efficient resource distribution but increases PLR under large number of users.

A utility based resource allocation with delay scheduler (U-DELAY) [39] was proposed to support both the real and non-real time applications. The U-DELAY works in two phases; the first phase employs a cooperative game theory and a sigmoid utility function in allocation of resources to different class of services while second phase employs the packet delay budget to decide which flow is transmitted within a class. It achieves an acceptable performance in terms of QoS requirements, fairness, and throughput as well as user satisfaction. However, the algorithm suffers poor spectral efficiency due to its failure to consider interference. Furthermore, the system delay increases under heavy network load.

In [40], a Delay Scheduler (DS) coupled Throughput Fairness (TF) resource allocation algorithm was proposed to assure QoS requirement of real-time applications and throughput fairness for non- real time applications. The algorithm categorizes real-time and non-real-time applications into urgent and non-urgent, respectively. The urgent are applications with high delay constraint usually the real-time. The non-urgent are real time applications with lower delay constraint and non-real-time applications. It first serves urgent with higher scheduling priorities and then non urgent with equal opportunities to access the wireless resources. The algorithm enhances the QoS of real-time application as well as throughput fairness for non-real-time applications. However, it experiences a rise in packet drop due to limited resources and delay constraints of QoS under large number of user arrivals.

A Service Based Scheduler (SBS) algorithm [41] was proposed to satisfy QoS requirements of real-time services. The algorithm utilizes the service requirements to create T_{max} and T_{min} lists for real-time and non-real-time flows, respectively. It then allocates RBs to the T_{max} list irrespective of their CQI and to the T_{min} list based on the good channel condition if the current TTI has additional resources. The algorithm enhances the system performance in terms of fairness, delay, PLR and average goodput. However, it wastes RBs when real-time receives allocation irrespective of its CQI. In addition; it starves non real-time flows when real-time flows are large.

In [42], a downlink scheduling algorithm was proposed to enhance QoS of multimedia service along with use of power saving mechanism at the User Equipment (UE). The algorithm employs opportunistic scheduling to determine the UEs priorities and allocates resources based on channel condition, average throughput, UE buffer status, Discontinuous Reception (DRX) status, delay and GBR / non GBR. It improves throughput fairness and minimizes packet delay under power saving environment. In addition, the algorithm also efficiently allocates resources under non saving environment. However, it starves the non-GBR when GBR experience data rate lower than the defined GBR.

A Rate-Level-Based Scheduling (RLBS) Algorithm [23] was proposed to support the diverse traffic in the downlink. The RBLS algorithm utilizes radio channel condition, packet delay information and Guaranteed Bit Rate (GBR) information to prioritize UEs requests. It serves the GBR first because they have a higher priority than the non GBR. The RBLS algorithm achieves a good trade-off between satisfaction of QoS requirements and fairness as well as enhances Packet Loss Ratio (PLR). However, it starves non-GBR due to limited resources under large number of users. Moreover, it applies packet delay upper bound to non-GBR traffic that is delay tolerant [15].

In [43], a QoS aware MAC scheduler algorithm was proposed to distinguish between the diverse QoS classes and their demands. The algorithm employs Time Domain (TD) and Frequency Domain (FD) schedulers to distribute resources to users. The TD scheduler classifies incoming packets into GBR and a non GBR flow, which corresponds to five QoS classes based on the QCI indices. The FD starts scheduling with the GBR flows and then with non GBR. The algorithm confirms that it is possible to assure diverse QoS requirements. However, it leads to the starvation of non GBR flows when resources are insufficient. In addition, the algorithm fails to increase the user's priority when the HOL approaches the upper bound [15].

A Highway and Rate Prediction Downlink Scheduler (HRPDS) [44] to improve QoS in real-time flows. The HRPDS employs a Predicted Bit Rate (PBR), initialization and update algorithms to allocate resources. The PBR algorithm decides the required Resource Blocks (RBs) for real-time and non-real time flows. The initialization algorithm initializes lines of highway to allocate resources if the RBs are sufficient for both real-time and non-real time flows without waiting in the buffer and revises the allocation values according to the percentage of each flow if the RBs are insufficient. Finally, the update algorithm updates the real-time flows approaching expiration in the buffer in order to provide priority. The scheduler guarantees an improved trade-off between bounded delay and low packet loss rate (PLR) but non-real-time flows suffer starvation due to its degradation under heavy real-time flows.

In [28], a novel three level LTE downlink scheduling framework (TLS) was proposed to assure pre-specified QoS requirements of Real-time Variable Bit Rate (RT VBR) flows. The TLS allocates RBs to active flows in three levels: Super Frame (SF) level, Frame level and TTI level in time domain. The SF level uses active RT VBR flows as input to estimate the amount of data to be transmitted in each flow in order to ensure QoS requirements and subsequent SF are computed using overload handling scheme. The Frame level employs a low overhead technique to distribute RBs of the subsequent frames to RT flows, instantaneously incoming RT data burst and non RT flows at each frame boundary within a super-frame. The TTI level utilizes scheduling matrix obtained from the Frame level scheduler to allocate RBs to flows. The framework improves QoS requirements and reduces computational overheads. However, the algorithm has poor spectral efficiency and goodput due to heavy prioritization of RT over NRT.

In [45], a Dynamic Multi-traffic scheduler algorithm was proposed to find trade-off between throughput and fairness. The algorithm classifies traffic based on QCI and employs dynamic bandwidth properties to distribute RBs according to size of traffic in service classes. It uses RR, BQCI, PF and opportunistic algorithms to schedule users. The algorithm enhances the throughput and fairness index. However, it starves non real-time traffic due to degradation during network congestion.

In [46], a Delay Aware Resource Block Management Scheduling algorithm was proposed to satisfy QoS requirements of RT traffic. The algorithm employs the RB ratio and the delay threshold (critical value) to allocate resources to Real-time (RT) and Non Real-time (NRT). It improves the average packet delay, system throughput and packet loss ratio. However, the algorithm leads to starvation of NRT traffic due to NRT degradation to admit RT traffic in critical zone.

In [47], a Buffer aware adaptive resource allocation algorithm was proposed to maximize the throughput, reduce the Bit Loss Rate (BLR) and fairness. It employs a user priority determination algorithm to determine the priority value of each user based on remaining life time or the queue overflow probability by considering buffer fullness, data arrival rate, CQI feedback from UEs. In addition, it also uses an online measurement-based algorithm to avoid buffer overflow and ensure QoS requirements considering the queue priorities. The algorithm improves system throughput and fairness as well as minimizes average Bit Loss Rate (BLR).However, the algorithm operates only under perfect channel condition and thus unsuitable to be used in real life scenario due to varying channel condition.

Table 3: Comparative analysis of Downlink Scheduling Algorithms.

S/No	Scheduling Algorithm	Type	Parameters used	Strength	Weakness
1	MT.	QoS Unaware	CQI	Maximum throughput. Improves spectral efficiency.	Starves users with bad channel conditions.
2	BET.	"	Data rate	High level of fairness.	Poor spectral efficiency.
3	PF.	"	CQI and Data rate	Fair distribution of resources. Improve spectral efficiency.	Fails to consider QoS requirements.
4	GPF.	"	CQI, Data rate and weighting factors	Higher spectral efficiency or higher level of fairness .	Fails to adapt the weighting factors in a running system.
5	DBWPF.	"	Buffer status, weighted average delay, CQI and Data rate	Approximate throughput and throughput fairness Improves delay fairness and implementation rate fairness.	Fails to achieve throughput fairness.
6	Optimal and Data Rate Guaranteed Radio Resource Allocation algorithm.	QoS Unaware	Data rate and CQI	Efficient utilization of resources Guarantees rate requirement for large number of users.	Starves non-priority users.
7	M-LWDF-VT.	QoS Aware	CQI, Data rate, Head-of-line token delay, Target delay and PLR	Enhances throughput in real-time services.	Starves non real-time traffic.
8	A Downlink Resource Allocation Algorithm For Multi-Traffic Classes.	"	Queue size, Head-of-line delay, Data rate and CQI	Enhances QoS parameters of diverse class of traffics	Increase PLR of VoIP traffic.
9	A novel delay based scheduling algorithm for video traffic.	"	HOL and Data rate	Enhances throughput Minimizes Packet Loss Rate (PLR).	Starves NRT flows.
10	EXP-EDF.	"	Packet delay budget, HOL and CQI	Provides QoS guarantees for real-time flows.	Starves non real-time traffic.

11	M-EDF-PF.	"	Packet delay budget ,HOL, Data rate and CQI	Improves the QoS requirements of real-time services.	Experience high delay.
12	Deadline based scheduler algorithm.	"	Packet delay budget, HOL, Data rate and CQI	Improves packet loss rate, delay and throughput for video flows.	Increase delay under large number of UEs.
13	RB Preserver Scheduler Algorithm.	"	HOL, Data rate, and Packet delay budget	Ensures the QoS requirements of RT flows.	Unfair to NRT flows.
14	DPS.	"	HOL, CQI and PDB	Enhances system throughput Reduces lower average system delay Fair resource allocation.	Fails to optimize Throughput and PLR.
15	A QoS-Aware Downlink Packet Scheduler Using Token Bucket Algorithm.	"	CQI, QCI, HOL and Data rate.	Satisfies the QoS requirements. Improves the cell performance.	Violates the delay budget.
16	QoS aware LTE OFDMA scheduling algorithm .	QoS Aware	CQI, Packet Delay Budget and Data rate.	Provides smooth video delivery.	Ignores non real-time and VoIP.
17	Two- level downlink scheduling algorithm.	"	CQI, Data rate, PDB and Buffer Status	Improves network performance and Quality of Experience (QoE),	Unfairness to non real-time applications. Low resource utilization
18	A Frame Level Scheduler-Advanced (FLSA) algorithm.	"	CQI, Data rate, Head-of-line packet delay, Target delay , PLR and Buffer status	Better resource allocation for real-time flows.	Starves non real-time flows.
19	PPM.	"	CQI, Data rate, PDB and Buffer status.	Enhances goodput and invalid packet rate.	Suffers VoIP PLR.
20	A Buffer aware adaptive resource allocation algorithm.	"	Buffer status, CQI, Data rate and PDB.	Improves system throughput, fairness and minimizes average Bit Loss Rate (BLR).	Operates only under perfect condition.
21	M-LWDF.	QoS Aware	CQI, Data rate, Head-of-line packet delay, Target delay and PLR	Increases the number of users supported with QoS. Guarantees minimum throughput and flow isolation.	Starves non-real time application. Non conformance to the 3GPP QoS specifications.
22	EXP and LOG rule.	"	CQI, , Data rate , Head-of-line delay (HOL) and Target delay	Guarantee QoS requirements with bounded delay.	Fails to fulfill the QoS requirements of 3GPP specifications .
23	3GPP QoS-based scheduling framework.	"	PDB, HOL, QCI and CQI.	QoS requirements based on 3GPP specifications. Improved spectral efficiency	Non-GBR may experience Packet drop.
24	S-DFS.	"	QCI, CQI, HOL and Buffer Status	Higher transmission efficiency. Decreases both dropping ratio and delay of GBR packets. Increases data throughput of non GBR flows.	High scheduling complexity. Low video throughput.
25	Adaptive EXP/PF.	"	CQI, Data rate, Head-of-line packet delay, Target delay and Target PLR	Provides QoS to real-time applications. Assures a satisfactory system throughput.	Unfair under large number of users .
26	Resource Allocation Algorithm for Downlink using shapley value.	QoS Aware	HOL, CQI, Data rate and Target delay	Provides an efficient resource distribution.	Increases PLR under large number of UEs.
27	U-DELAY.	"	PDB,HOL and CQI	Acceptable performance in terms of QoS requirements, fairness, throughput and user satisfaction.	Poor spectral efficiency. High system delay.
28	Delay Scheduler (DS) coupled Throughput Fairness (TF) resource allocation	"	Data rate, HOL and CQI.	Enhances the QoS of real-time application and throughput fairness for non-real-time applications.	Suffers high PLR.

	algorithm.				
29	Service Based Scheduler (SBS) algorithm.	"	QCI, CQI, Data rate, Packet size and Theoretical departure time	Enhances fairness, delay, PLR and average goodput.	Wastes RBs. Starves non real-time traffic.
30	QoS-Aware LTE Downlink Scheduler for VoIP with Power saving.	"	PDB,HOL.CQI and Data rate	Improves throughput and fairness. Minimizes packet delay Efficiently allocates resources	Starves non- GBR.
31	RLBS.	"	HOL,PDB and CQI	Achieves a good trade-off between satisfaction of QoS requirements and fairness. Enhances Packet Loss Ratio (PLR) performance.	Starves non- GBR.
32	QoS aware MAC scheduler algorithm.	"	QCI,CQI and Data rate.	Assure diverse QoS requirements.	Starves non GBR flows.
33	HRPDS.	QoS Aware	HOL, Data rate and CQI.	Guarantee an improved trade-off between bounded delay and low packet loss rate (PLR)	Starves NRT flows
34	TLS.	"	CQI, HOL, PDB, and Data rate.	Improves QoS requirements, goodput and spectral efficiency.	Poor spectral efficiency and goodput.
35	A Dynamic Multi-traffic scheduler algorithm.	"	CQI, QCI and Data rate.	Enhances the throughput and fairness index.	Starves NRT traffic.
36	Delay Aware Resource Block Management Scheduling algorithm.	"	HOL and Data rate.	Improves average packet delay, the system throughput and the packet loss ratio.	Starves NRT traffic

X. Comparative Analysis

Table1.presents a comparative analysis of the various downlink scheduling algorithms in terms of the parameters used in resource allocation, strengths and weaknesses. The algorithms are classified into QoS aware or QoS unaware. The QoS unaware algorithms are proposed to ensure throughput by utilizing buffer status, data rate and CQI. The algorithms in [7] utilize data rate or CQI in allocation decision but lead to unfairness and poor spectral efficiency while algorithms in [8], [9], [11], [12] employ CQI and data rate to improve fairness and spectral efficiency but ignores QoS requirements and thus unsuitable for wireless multimedia traffic [2].

The QoS aware algorithms utilize data rate, CQI, buffer status, delay constraints (HOL and PDB) and QCI in order to guarantee QoS to multimedia applications. The algorithms in [13], [14], [16]–[22], [24], [25], [48], [49], [30] focus only on real time application but unfair to non-real time applications. While the algorithms in [15], [23], [28], [6], [31], [36], [38], [40], [41], [43]–[46], [39], [42] consider both the real-time and non-real time traffics in order to address the unfairness problem but the algorithms in [23], [28], [41], [43]–[45], [42] suffer starvation due to degradation policy applied when resources are insufficient and in [15], [6], [36], [40], [46], [39] experience low throughput, high PLR, or poor spectral efficiency under high network load. In addition, majority of these algorithms incur high scheduling complexity due to metric base flow selection per TTI [36] .

XI. Conclusion

In this paper, we presented a survey of scheduling algorithms proposed in recent literature, aiming at distributing available RBs to users in the downlink direction of LTE networks. The algorithms are classified into QoS aware and QoS unaware, the former uses data rate, CQI, buffer status, delay constraints (HOL and PDB) and QCI to ensure meeting the QoS requirements and the later utilizes buffer status, data rate and CQI to assure throughput. The way each algorithm operates as well as the advantages and the disadvantages are also discussed. Furthermore, comparative analysis has been provided. The analysis indicates that the QoS unaware algorithms are in appropriate for wireless multimedia traffic because the QoS requirements are ignored while the QoS aware algorithms either fail to consider or starve the non real-time traffic. Also, the analysis shows that the algorithms experience low throughput, high PLR, or poor spectral efficiency under high network load.

References

[1] Cisco, "Cisco Visual Networking Index,VNI Forecast and Methodology, 2014 – 2019.," 2015.

[2] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: Key design issues and a survey," *IEEE Commun. Surv. Tutorials*, vol. 15, no. 2, pp. 678–700, 2013.

[3] D. Erik, P. Stefan, S. Johan, and P. Beming, *HSPA and LTE for Mobile Broadband*. 2007.

[4] Alcatel-Luncet, "The LTE Network Architecture, White Paper."

[5] N. Abu-ali, A. M. Taha, S. Member, M. Salah, H. Hassanein, and S. Member, "Uplink Scheduling in LTE and LTE-Advanced : Tutorial , Survey and Evaluation Framework," *IEEE Commun. Surv. Tutor.*, vol. 16, no. 3, pp. 1239–1265, 2014.

[6] Y. C. Wang and S. Y. Hsieh, "Service-differentiated downlink flow scheduling to support QoS in long term evolution," *Comput. Networks*, vol. 94, pp. 344–359, 2016.

[7] J. Puttonen, N. Kolehmainen, T. Henttonen, M. Moisio, M. Rinne, and P. Kela, "Dynamic Packet scheduling Performance in UTRAN long term evolution downlink," *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC*, pp. 308–313, 2008.

[8] H. J. Kushner and P. A. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions," *IEEE Trans. Wirel. Commun.*, vol. 3, no. 4, pp. 1250–1259, 2004.

[9] C. Wengerter, J. Ohlhorst, A. Golitschek, E. Von Elbwart, and P. Mcom, "Fairness and Throughput Analysis for Generalized Proportional Fair Frequency Scheduling in OFDMA," in *IEEE Veh. Tech. Conf.,VTC-Spring*, 2005, vol. 3, no. 2, pp. 1903– 1907.

[10] M. Proebster, C. M. Mueller, and H. Bakker, "Adaptive fairness control for a proportional fair LTE scheduler," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, 2010, pp. 1504–1509.

[11] S. Liu, C. Zhang, Y. Zhou, and Y. Zhang, "Delay-Based Weighted Proportional Fair Algorithm for LTE Downlink Packet Scheduling," *Wirel. Pers. Commun.*, pp. 1955–1965, 2015.

[12] B. Rebekka, S. Sudheep, and B. Malarkodi, "An Optimal and Priority Based Rate Guaranteed Radio Resource Allocation Scheme for LTE Downlink," *Wirel. Pers. Commun.*, vol. 83, no. 3, pp. 1643–1661, 2015.

[13] M. Iturralde, T. A. Yahiya, A. Wei, and A. Beylot, "Performance Study of Multimedia Services Using Virtual Token Mechanism for Resource Allocation in LTE Networks," in *IEEE Vehicular Technology Conference (VTC)*, 2011, pp. 1–5.

[14] M. M. Nasralla and M. G. Martini, "A downlink scheduling approach for balancing QoS in LTE wireless networks," *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC*, pp. 1571–1575, 2013.

[15] P. Ameigeiras *et al.*, "3GPP QoS-based scheduling framework for LTE," *EURASIP J. Wirel. Commun. Netw.*, vol. 78, 2016.

[16] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Commun. Mag.*, vol. 39, no. 2, pp. 150–153, 2001.

[17] B. Saiede and M. D. Farzanegan, "A novel delay based scheduling algorithm for video traffic in LTE," in *2nd International Conference on Knowledge based Engineering Innovation (KBEI)*, 2015.

[18] H. M. J. Magalhães and P. R. Guardieiro, "A Downlink Scheduling based on Earliest Deadline First Discipline for Real-Time Traffic in LTE Networks," in *Vehicular International Workshop on Telecommunication*, 2013.

[19] B. Liu, H. Tian, and L. Xu, "An efficient downlink packet scheduling algorithm for real time traffics in LTE systems," *2013 IEEE 10th Consum. Commun. Netw. Conf. CCNC 2013*, pp. 364–369, 2013.

[20] P. J. Ashidani and P. R. Guardieiro, "Downlink Scheduler Based on Deadlines for LTE Networks," in *International workshop on telecommunications (IWT)*, 2013.

[21] S. Ghassan, A. Ali, M. D. Baba, and M. A. Mansor, "RB Preserver Downlink Scheduling Algorithm for Real-Time Multimedia Services in LTE Networks," in *Computer Application and Industrial Electronics (ISCAIE)*, 2015, pp. 21–26.

[22] S. Kumbesan, A. H. M. Ramli, and B. Riyaj, "Delay-Prioritized Scheduling (DPS) for Real Time Traffic in 3GPP LTE System," in *IEEE Wireless Communication and Networking Conference (WCNC)*, 2010, pp. 1–6.

[23] X. Wu, X. Han, and X. Lin, "QoS Oriented Heterogeneous Traffic Scheduling in LTE Downlink," in *IEEE International Conference On. Communications (ICC)*, 2015, pp. 3088–3093.

[24] Y. Farzaneh and A. Mardi, "A QoS-Aware Downlink Packet Scheduler Using Token Bucket Algorithm for LTE Systems," in *22nd Iranian Conference on Electrical Engineering (ICEE 2014)*, 2014, pp. 1775–1780.

[25] H. Luo, S. Ci, D. Wu, J. Wu, and H. Tang, "Quality-driven cross-layer optimized video delivery over LTE," *IEEE Commun. Mag.*, vol. 48, no. 2, pp. 102–109, 2010.

[26] R. Zhang, S. L. Regunathan, and K. Rose, "Video Coding with Optimal Inter/Intra-Mode Switching for Packet Loss Resilience," *IEEE JSAC*, vol. vol. 18, no. no.6, pp. 966–976, 2000.

[27] G. Piro *et al.*, "Two-level downlink scheduling for real-time multimedia services in LTE networks," *IEEE Trans. Multimed.*, vol. 13, no. 5, pp. 1056–1065, 2011.

[28] S. Kumar, A. Sarkar, S. Sriram, and A. Sur, "A three level LTE downlink scheduling framework for RT VBR traffic," *Computer Networks*, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2015.08.027.

[29] E. Skondras, A. Michalas, A. Sgora, and D. D. Vergados, "A downlink scheduler supporting real time services in LTE cellular networks," in *Information, Intelligence, Systems and Application(IISA)*, 2015, pp. 1–6.

[30] W. K. Lai and C. L. Tang, "QoS-aware downlink packet scheduling for LTE networks," *Comput. Networks*, vol. 57, no. 7, pp. 1689–1698, 2013.

[31] B. Sadiq, R. Madan, and A. Sampath, "Downlink scheduling for multiclass traffic in LTE," *Eurasip J. Wirel. Commun. Netw.*, vol. 2009, 2009.

[32] 3GPP, "3 GPP Tech. Specif. Group Radio Access Network - Physical Channel and Modulation (Release 8), 3GPP TS 36.211," 2011.

[33] 3GPP, "TS 23.203 V8.11.0. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Policy and Charging Control architecture (Rel. 8)," 2010.

[34] P. Ameigeiras, Y. Wang, J. Navarro-Ortiz, P. Mogensen, and J. Lopez-Soler, "Traffic models impact on OFDMA systems design. 2012, 61 (2012)," *EURASIP J. Wirel. Commun. Netw.*, no. 61, 2012.

[35] H. L. H. Yin, "An efficient multiuser loading algorithm for OFDM-based broadband wireless systems," in *in IEEE Global Communications Conference (GLOBECOM'00)*, 2000, pp. 103–107.

[36] J. Rhee, J. Holtzman, and D. Kim, "Scheduling of real/non-real time services: adaptive EXP/PF algorithm," *Conf. 2003. VTC* , pp. 462–466, 2003.

[37] H. A. M. Ramli, R. Basukala, K. Sandrasegaran, and R. Patachaianand, "Performance of well known packet scheduling algorithms in the downlink 3GPP LTE system," *Proc. - MICC 2009 2009 IEEE 9th Malaysia Int. Conf. Commun. with a Spec. Work. Digit. TV Contents*, no. December, pp. 815–820, 2009.

[38] M. Iturralde, T. A. Yahiya, A. Wei, and A.-L. Beylot, "Resource allocation using Shapley value in LTE networks," in *International Symposium on Personal, Indoor and Mobile Radio Communications*, 2011, pp. 31–35.

[39] S. Ali and M. Zeeshan, "A utility based resource allocation scheme with delay scheduler for LTE service-class support," *IEEE Wirel. Commun. Netw. Conf. WCNC*, pp. 1450–1455, 2012.
[40] C. Wang and Y. C. Huang, "Delay-scheduler coupled throughput-fairness resource allocation algorithm in the long-term evolution wireless networks," *IET Commun.*, vol. 8, no. 17, pp. 3105–3112, 2014.
[41] S. F. Sulthana and R. Nakkeeran, "Performance Analysis of Service Based Scheduler in LTE OFDMA System," *Wirel. Pers. Commun.*, pp. 841–854, 2015.
[42] M. Sajid Mushtaq, A. Shahid, and S. Fowler, "QoS -aware LTE Downlink Scheduler for VoIP with Power asaving," *2012 IEEE International conference on Computational Science and Engineering*. 2012.
[43] Y. Zaki, T. Weerawardane, C. Görg, and A. Timm-Giel, "Multi-QoS-aware fair scheduling for LTE," *IEEE Veh. Technol. Conf.*, 2011.
[44] S. Hendaoui, Z. Nawel, and S. Tabbane, "Highway and Rate Prediction Downlink Scheduler (HRPDS) for LTE-A Real Time Applications," in *13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, 2016.
[45] K. G. Sundari, R. Gunasundari, and K. Jayanthi, "Prioritized Dynamic Multi-traffic Scheduler For Downlink in LTE," in *IEEE Sponsored Second International Conference On Electronics And Communication Systems (ICECS)*, 2015, pp. 1546–1550.
[46] K. Korn, J. Aphirak, and A. Phonphoem, "Delay-Aware with Resource Block Management Scheduling Algorithm in LTE," in *International Computer Science and Engineering Conference (ICSEC)*, 2015.
[47] R. Zhu and J. Yang, "Buffer-aware adaptive resource allocation scheme in LTE transmission systems," *EURASIP J. Wirel. Commun. Netw.*, vol. 176, 2015.

Muhammad Aminu Lawal. "Downlink Scheduling Algorithms in LTE Networks: A Survey." IOSR Journal of Mobile Computing & Application (IOSR-JMCA) , vol. 4, no. 3, 2017, pp. 01-12