

A Design of Packet Scheduling Algorithm to Enhance QoS in High-Speed Downlink Packet Access (HSDPA) Core Network

Sohail Ahmed¹, Mubashar Ali²
Department of Computer Science
University of Lahore
Gujrat, Pakistan

Bilal Akbar⁵
Department of Mechanical Engineering
Mirpur University of Science and Technology
AJK, Pakistan

Abdullah Baz³
Dept. of Computer Engineering
College of Computer and Information Systems
Umm Al-Qura University
Makkah, Saudi Arabia

Imran Ali Khan⁶
Department of Computer Science
COMSATS University Islamabad
Abbottabad Campus
Pakistan

Hosam Alhakami⁴
Dept. of Computer Science
College of Computer and Information Systems
Umm Al-Qura University, Makkah, Saudi Arabia

Adeel Ahmed⁷, Muhammad Junaid⁸
Department of Information Technology
The University of Haripur
KPK, Pakistan

Abstract—Voice over Internet Protocol (VOIP) in an efficient manner is a basic requirement of modern era. The real time and non-real traffics demand customized communication provisioning to get guarantee of service. For this we proposed a user fulfillment design for facilitating packets switching in 3G cellular network to insure provisioning of QoS (quality of service) in DiffServ (Differentiated Services) Network. To enhance QoS for real time traffic by reducing delay, packet loss and jitter, we proposed Low latency queuing (LLQ) algorithm. In this paper, we focused on packet scheduling, Diffserv and QoS classes mapping into Universal Mobile telecommunication System (UMTS) classes and buffering. To associate different types of real time multimedia traffic, the QoS provisioning mechanism used different code points of Diffserv. The new idea in LLQ is to map the video and voice traffics against two separate queues and used priority queuing in Low latency queuing for voice traffic. The results got from reproductions shows that proposed calculation meets the QoS prerequisites.

Keywords—Packet Scheduling; Classification; DiffServ; LLQ; EURANE

I. INTRODUCTION

In this era, the transmission of voice as a packet over IP network in an optimized way is relatively core research area. The people are massively turning to VoIP technology because it provides the facility of cost effective and free call. This popularity is demanding the quality of service for real-time voice and video services. The 3G wireless network has the ability to handle different real-time traffic like video, Voice, and other non-real time traffic applications. For the purpose of analyzing and design 3G network, the main issue is to get

required QoS level during transmitting real-time data packets on wireless network. The ability to provide real-time services with the guaranty of QoS is the main potential of HSDPA [1]. There are two different mechanisms standardized by IETF to provide quality of service in current IP networks, like integrated Services (InterServ) [21] and Diffserv [22]. InterServ has complexity and scalability problems while DiffServ is a simple mechanism and can be implemented in HSDPA with simple policy management [2]. In Packet's IP header the DiffServ attach a code points known as Differentiated Services code points (DSCP). At the network boundaries, different traffics are classified with Per Hop Behaviors (PHBs) using different DSCP values [6,11]. The implementation of PHB in routers has got noteworthy attention because there is no any specific implementation mechanism defined by PHB definition standards.

The architecture and concept of QoS for HSDPA network which is defined in 3GPP concentrates only on QoS of signaling between Gateway GPRS Support Node (GGSN) and user entity [3]. It is quite difficult to handle the procedure of QoS for transportation of packets. To get provisioning of end-to-end QoS in HSDPA, there should be a QoS mechanism for user data transportation. There also need to map the class of IP network traffic into class of UMTS network.

In previous work [12,14] to get QoS in end-to-end HSDPA Network, focuses on customized mapping of video and voice communication but do not pay attention to other traffic classes of HSDPA. The scheduling algorithm Weighted Fair Queue (WFQ) and Priority Queue (PQ) have combined in the paper [9], resulting the video conferencing delay was reduced but

the voice traffic delay was increased. For provisioning of QoS in HSDPA network, a project of SEACORN [16] has contributed into implementation and development of Radio Resource Management (RRM).

Enhanced UMTS Radio Access Network- Extension (EURANE), which is the extension of UMTS for Network Simulator (NS-2), is an important contribution to SEACORN project [13, 16]. We selected EURANE tool for simulation of UMTS QoS scenario [23]. The proposed work covered following main sections: introduced the concept of video and voice telephony mapping to different classes of QoS, implementation idea for LLQ scheduler, strategies of packet treatment for HSDPA core network and analyzed the results in a large simulation environment. The fairness in traffic handling is an important feature of efficient scheduler [15]. Our concentration is also to provide fairness in traffics and handles the matter that how real time traffic affected by non-real time traffic in several scenarios.

In Section II, we described the experimental setup where focused on basic components of HSDPA. The proposed packet scheduling algorithm is briefly discussed using flow chart in Section III. The test scenarios and results are illustrated in Section IV. The conclusion is described in Section V.

II. EXPERIMENTAL SETUP

There are three basic components in architecture of HSDPA system, which are known as terminal equipment (TE) or user equipment (UE), UTRAN and UMTS core network. The functionality of core network depends on different routers. In our experimental setup as shown in Fig. 1, there are core router, edge routers, and four application servers. From external network, application servers sent data packet to the edge router. The edge router is a combination of Egress and Ingress routers. Through these routers, according to the type of application, the data packets are assigned a pre-defined DiffServ Code point (DSCP) and sent to GGSN (core router). The GGSN router differentiates each flow of IP data packet according to its DSCP value (The DSCP value remains same as assigned from external network). After this, these IP data packets are transmitted to Serving GPRS support Node (SGSN) router with specific scheduling and queuing scheme.

After receiving packets SGSN router transmits these packets to Radio Network Controller (RNC), where the packets of IP data are transformed into the RLC SDU. On next stage the DCH is used with acknowledge mode. The maximum transmission time for RLC layers are unlimited. This configuration needs some customization due to some limitation of EURANE [23]. The remaining end to end SDU

losses are due to dropping of a packet in bottleneck of queue's [5, 12], which is over flow of queue for all other traffic or early dropping of a packet from real-time traffic. On another side, the end to end delay is much more uncertain due to unlimited retransmission of RLC PDU. As defined in the specification of UMTS [17,18] there are four classes for QoS in packet domain. These are streaming, conversational, Interactive and background classes. The classes are categorized according to delay sensitivity factor. The conversational and streaming classes represent real time application. So, these are more sensitive to delay [8,10]. On the other hand, background and interactive classes are less delay sensitive, and transmit Packets without any restricted delay requirement. The mappings of DSCP, according to their delay sensitive class are done in the edge router.

A. Mapping of Diffserv into UMTS QoS Parameters

To get end to end QoS provisioning, a layered architecture is defined by 3GPP standards. To realize QoS requirements for a specific network, a Bearer Service (BS) assigned functionality have implemented from the perspective of the source to the destination of service. It also has all parameters to facilitate provisioning of QoS. The configuration of traffic parameters is shown in Table I. The Bearer service of UMTS classifies the QoS, and also introduces a level of service facilitated to BS user. To restrict the specific level of delay in real-time traffic, mapping of QoS between UMTS services and IP DiffServ is very necessary [14]. The service class expedited forwarding (EF) per hop behavior (PHB) is defined as less delay, less packet loss, and low jitter. In this class, the traffic is treated with high priority. If EF traffic's arrival rate crosses the defined limit then it dropped in advance. The conversational class's traffic i.e. VOIP is comparatively very less sensitive for packet loss. So, the conversational class is much suitable for EF DiffServ class and these mapping of QoS is shown in Table I. Another class is Assured Forwarding (AF PHB) which is specified by IETF [10], provides the guarantee of packet delivery within the boundary conditioned of user rate. If there is congestion on the link, and traffic crosses the limit of arrival rate, then it is a high probability that the packets shall drop. There are four dropping classes with three levels of dropping precedence. These all levels are implemented in "Assured Forwarding PHB" for delivery of data packet [19]. Each class is specified with different dropping levels and buffering configuration, according to the sharing of bandwidth. The streaming class is less sensitive to delay and supported by AF PHB class. The interactive class has no any boundary condition and reliability requirement, so mapped this class with AF31. The last class is known as background class and mapped with the best effort class, which is a default class.

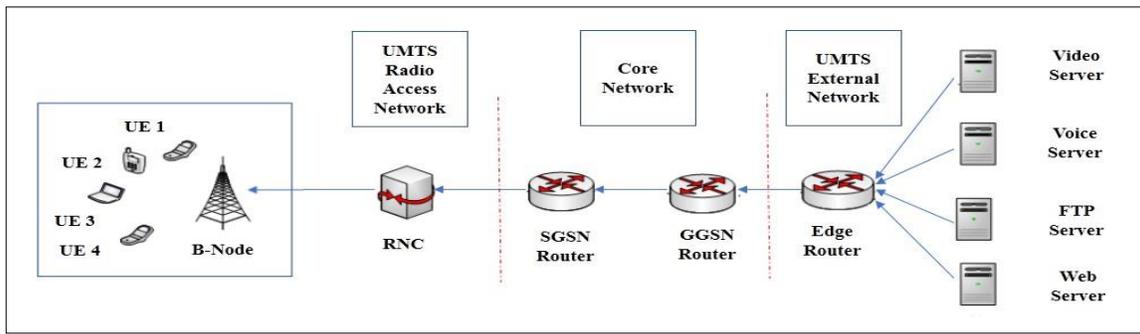


Fig. 1. Functional Blocks in HSDPA Network.

TABLE I. TRAFFIC MODEL AND ARAMETERS

Traffic Types	UMTS Service Class	DSCP Value	DiffServ DSCP	Size of IP Packet	Transport layer Protocol	Holding time distribution	Traffic Source model	Bottleneck Bandwidth
Video Traffic	Streaming	10 12	AF11 AF12	160	UDP	Exp	Exp on/off	1000
Voice Traffic	Conversational	46	EF	120	UDP	Exp	Exp on/off	
FTP Traffic	Background	0	BE	480	TCP	Pareto Distributed	Pareto on/off	
HTTP Traffic	Interactive	18 20	AF21 AF22	240	TCP	Log-normal	Pareto on/off	

For marking the packets, we used Improved Time Sliding Window Three Color Marker (ItswTCM) in which three colors red, yellow and green is assigned to different packet for given different level of priority [4, 22]. These assigning of colors depends on packet arrival rate. There are two types of rates that are Peak Target Rate (PTR) and Committed Target Rate (CTR) [7]. If the packet's throughput meets the PTR then ItswTCM will assign red color to that packet with high priority. If the throughput is between PTR and CTR then assign yellow color and if throughput less then CTR then assigns green color to that packet that's mean low priority traffic. The marking of DSCP is activated by a function at the edge router, which is known as policer. Its means different IP Packets will handle with different virtual / physical queue and treated accordingly. After marking and queuing mechanism the packet forwarded to the scheduling section. The scheduler improves the efficiency of the end to end system using mechanism of sharing common resources between different classes [20]. The task of scheduling algorithm is to manage available core network resources between authorized user's groups. In congestion the system demands queuing and scheduling mechanisms. The scheduling mechanism will decide the queue to handle first. If the link is not congested, then all arrived packets shall transmit without any delay. The delay limits for each type of services are shown in Table II.

TABLE II. HSDPA QOS REQUIREMENTS FOR EACH SERVICE CLASS

Service Type	Streami ng	Conversational	Background	Interactive
Network Delay	$\leq 260\text{ms}$	$< 100\text{ms}$	Not Applicable	Not Applicable
Loss Rate	$\leq 10^{-3}$	$\leq 10^{-3}$	$\leq 10^{-2}$	$\leq 10^{-2}$

III. PROPOSED LLQ PACKET SCHEDULING ALGORITHM

One of the major attentions in this paper is the functions of buffering and scheduling. These mechanisms are deal with packets. The buffering and scheduling functions are the main part of this paper. These functions are used to control the packets. The buffering and scheduling algorithm involve each time either a packet received or sent. Whether an incoming packet should accept or not, this will be decided by buffering and policing algorithm. An implemented diagram of these functions is shown in Fig. 2.

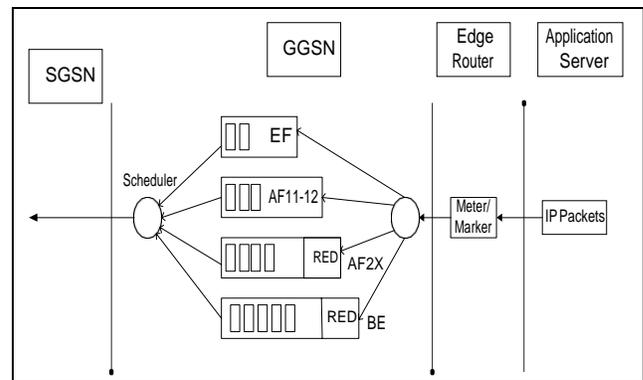


Fig. 2. Buffering and Scheduling Algorithm Implementation.

When a packet is received at edge router, it measures the arriving rate using meter function. As discussed in mapping to UMTS class section red, yellow and green colors are assigned to packet using ItswTCM. On the bases of color, specified DSCP is assigned to different packets. Now the red, yellow and green color marked packets are queued in EF, AF and BE queues respectively. After this the GGSN router uses LLQ scheduling algorithm to dequeue packets from these queues.

To solve delay issue in real time application a strict priority queue is integrated with Class-Based Weighted Fair Queue (CBWFQ). First of all, the scheduler will check the EF queue which reserved for high priority voice packet. If there are packets then dequeue this using priority queue algorithm. After dequeuing all packets from there, the scheduler will check the AF queue, which has less priority than EF Queue. The scheduler will dequeue 3 quanta of 30 bytes each from AF queue. After serving packets, the scheduler will serve BE queue which has less priority than AF class. The scheduler now dequeue 2 quanta of 30 bytes each from BE queue and move onto the first AF queue. This loop will be continuing in a CBWFQ manner. While serving queue, if any packet arrives in EF queue, the scheduler will pause the serving and switch onto the EF queue. After serving all packets from there it will resume again. After scheduling, these dequeued packets are given to SGSN router. In this scenario, there is a bottleneck of downlink streaming in an outgoing link between GGSN and SGSN. So, the target of our design is to improve the link consumption in this bottleneck. For this improvement, we have to minimize the blocking rate of a session and enhance the throughput of the link while end to end QoS specification for UMTS classes should maintain. The enqueueing and marking function is illustrated in Algorithm 1 and dequeuing function is illustrated in Algorithm 2.

ALGORITHM 1

```

Input → Packet "g"
If ((No. of total bytes P < Size of EF Queue) && (Packet Rate H > CTR)) Red
color mark → g
Enqueue g → EF Queue
Else if ((No. of Total bytes Q < Size of AF queue) && (CTR < H < PTR))
Yellow color mark → g
Enqueue g → AF Queue
Else if ((No. of Total bytes R < Size of BE Queue) && (H < CTR))
Green color mark → g
Enqueue g → BE Queue
Else
Packet dropped from queue by RED
    
```

ALGORITHM 2.

```

If (EF Queue P! = Null)
EF Queue → Dequeue using PQ
Else if (AF Queue Q! = Null) && (Packet weight AF Queue < Defined
Weight))
AF Queue → Dequeue using CBWFQ
Else if (BE Queue R! = Null) && (Packet weight BE Queue < Defined
Weight))
BE Queue → Dequeue using CBWFQ
    
```

The LLQ (Low latency Queuing) algorithm is very flexible because we can easily modify the related importance of each class, to solve delay issue in real time and non-real time traffics. This mechanism is also suitable for jitter sensitive and delay sensitive traffic as well. LLQ received real-time traffic and sent it to SGSN router without any delay. If PQ has no any packet the scheduler executes CBWFQ for remaining traffic. This mechanism provides a guaranty of bandwidth to non-real time traffic and avoids the starvation issue. This algorithm is also represented in a flow chart form in Fig. 3. In this paper, the main target is to achieve the best utilization of bandwidth in bottleneck link while keeping packet loss of IP data and E2E delay within predefined limit, as depicted in Table III. So, the core parameters for calculating the performance are IP data packet loss, delay, jitter and

bandwidth of bottleneck link. The scenario of simulation is shown in Fig. 2 where investigated the parameters, which are described in Table III.

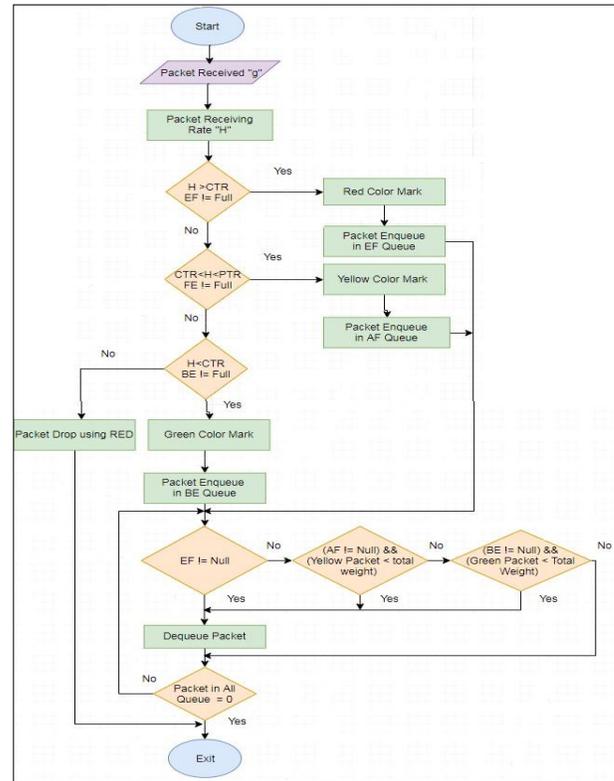


Fig. 3. Proposed Marking and Scheduling Algorithm.

TABLE III. PARAMETERS OF NETWORK

Parameters	Values
Active UE number	20
Quantum size	30 bytes
Fast Power Control	Ideal
DCH bandwidth	384Kbps
Mobility model	No
Radio link RLC PDU error	Uniform, mean=0.01

IV. TEST SCENARIO AND RESULTS

This investigation is performed using different experiments of simulation, where we implemented LLQ, PQ, and WRR scheduling algorithm and then evaluated it for different link congestion scenario. The traffic load in the link is started from 10 % and increase gradually to 150 %. The conversational class sent 20 % traffic from whole generated traffics. The streaming class sent 70 %, background class 7 % and interactive class sent 3 % of overall traffic. The voice traffic using PQ scheduler assigned highest level of priority, while least level of priority is specified for background class. The weight in scheduler is representing the use of output bandwidth in percentage i.e. 20 % weight for conversational, 70% weight for streaming, 7 % weight for background and 3 % weight for interactive class. There are different traffic types for the experiment, as depicted in Table I.

It is clear from Fig. 4 that in case of Weighted Round Robin (WRR) algorithm, voice delay is increasing rapidly at 100% load because there is no priority for voice. In case of PQ algorithm voice, traffic has priority but other traffic is facing starvation of resource. In case of previous LLQ delay of voice.

Traffic is increasing due to use of WFQ treatment for each packet. Our proposed LLQ has less delay in voice traffic due to implementing CBWFQ who treats the whole traffic class instead of each packet. We can see that, the delay is less than 100ms and also remains less in high congestion.

For video streaming as depicted in Fig. 5, all schedulers show almost same behavior at 75% load because the traffic for video streaming is very high so, queuing of packets is required. The WRR has a maximum delay than other algorithms. PQ has a starvation problem, so our modified LLQ has less delay than WRR but almost equal to the previous LLQ.

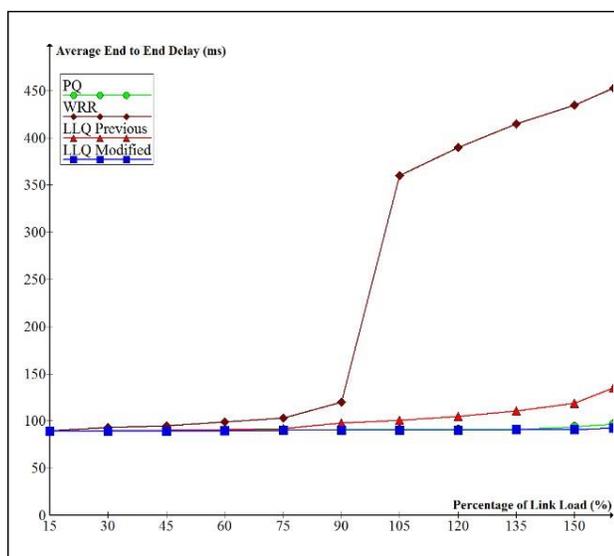


Fig. 4. The Average Delay Invoice Class.

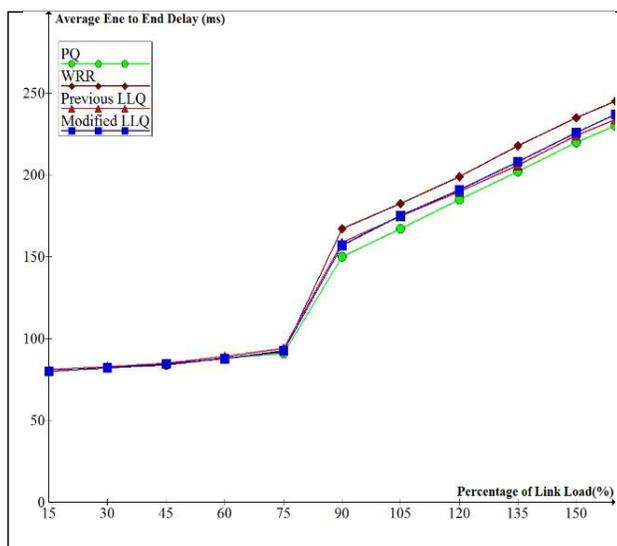


Fig. 5. The Average Delay in Video Streaming Class.

Fig. 6 depicts average jitter in conversational class. In case of voice class, the WRR has maximum jitter because the size of the voice packet is smaller than other traffic's packet. Our modified LLQ algorithm has good result than other.

Fig. 7 depicts average jitter in video streaming traffic. In this case, the WRR has maximum jitter and priority queue has minimum jitter. But our LLQ has more jitter than the previous LLQ because CBWFQ treated the video whole class traffic instead of individual packet.

Fig. 8 depicts packet loss rate for voice conversational class. We can see that the performance of WRR is greatly affected by link and the rate of packet loss remains within specified limit while using PQ and LLQ. The Packet loss occurred due to queuing of the packet, but we are giving priority to voice traffic so there is no queuing in our algorithm.

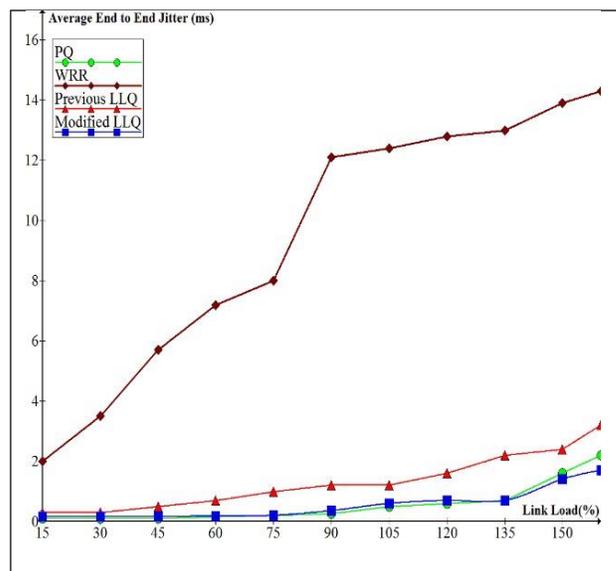


Fig. 6. The Average Jitter in Voice Conversational Class.

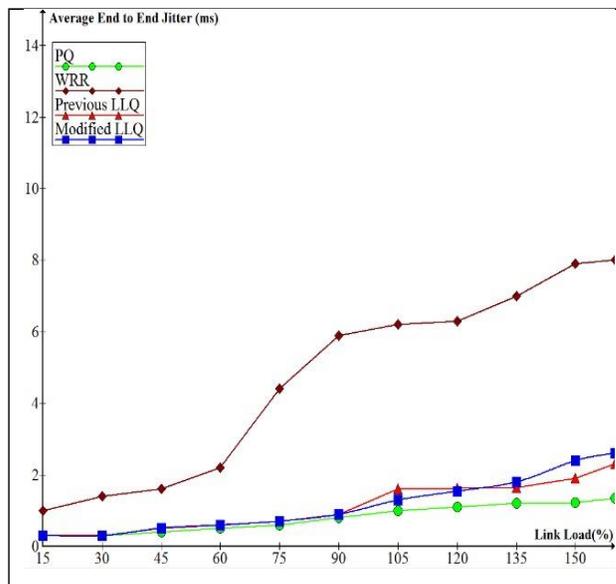


Fig. 7. The Average Jitter in Video Streaming Class.

Fig. 9 depicts packet loss rate for video streaming. The packet size of video streaming is larger than voice traffic so after 90% load the loss rate increases rapidly, but the loss rate of LLQ is more than the previous LLQ. The Table IV contains the average value of throughput for background and interactive class. In high congestion, our proposed LLQ algorithm provides best result of throughput in both traffic classes because of fair distribution of bandwidth. But the PQ provides least throughput in case of both classes.

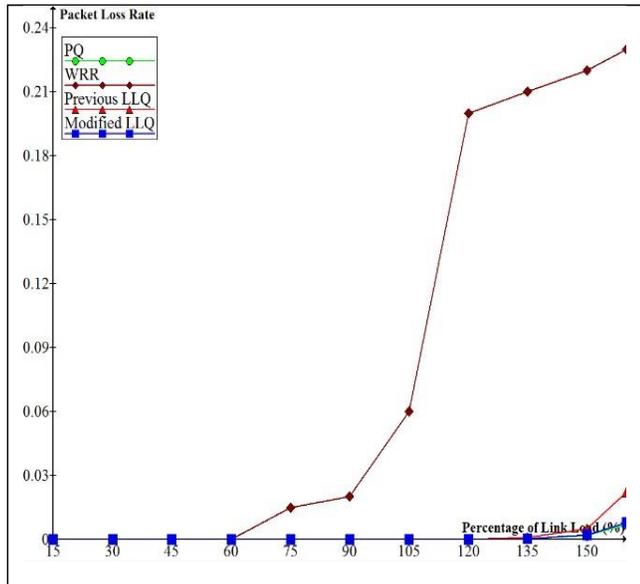


Fig. 8. The Average Packet Loss Rate in Voice Conversational Class.

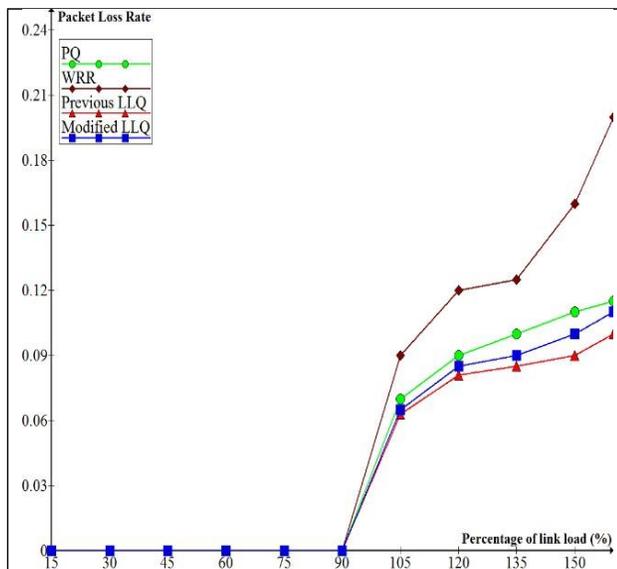


Fig. 9. Average Packet Loss Rate in Voice Streaming Class.

TABLE IV. AVERAGE THROUGHPUT OF NON-REAL TIME TRAFFIC

Traffic Class	Throughput (Kbps)		
	LLQ	PQ	WRR
Background Class	17.23	10.04	16.21
Interactive Class	8.24	4.31	8.57

V. CONCLUSIONS

In this paper, we investigated the quality of service for HSDPA real-time traffic and improved the performance of the end to end network. We proposed an algorithm called LLQ packet scheduling algorithm. The improved algorithm has combined Diffserv to HSDPA quality of service mapping, LLQ scheduling and multiple queuing with optimized parameters. The performance of LLQ scheduling algorithm is evaluated with different scenarios. The simulation results have been proved that this scheduling algorithm reduces the ratio of delay, packet loss, and gets better utilization of link's bottleneck within boundary limit conditions of QoS. The algorithm also gives a fair resource to voice traffic even in highly links congestion and increase the capacity of the overall system, but there is some increased of jitter in traffic of video streaming.

In future, we will combine Diffserv with HSUPA, and reduced the jitter in the traffic of video streaming.

ACKNOWLEDGMENT

We are very thankful to our colleagues provided their expertise and assisted in this research work. This work was funded under grant number 18-COM-1-01-0001 from Department of Computer Science and Engineering, College of Computer and Information Systems, Umm Al-Qura University, Makkah, Saudi Arabia.

REFERENCES

- [1] Nadim K. M. Madi, Zurina B. M. Hanapi, M. Othman, S. Subramaniam, "Two-level QoS-aware frame-based downlink resources allocation for RT/NRT services fairness in LTE networks," Telecommunication System, Springer, 2017.
- [2] Ted H. Szymanski, "An ultra-low latency guaranteed rate Internet for cloud services," IEEE/ACM Transaction on Networking, VOL.24, NO.1, 2016.
- [3] J. Khamse-Ashari, G. Kesidis, I. Lambadaris, B. Urgaonkar, Y. Zhao, "Max-Min Fair Scheduling of Variable-Length Packet-Flows to Multiple Servers by Deficit Round-Robin, CISS," IEEE, 2016.
- [4] A. M. Alkharasani, M. Othman, A. Abdullah, K. Y. Lun, "An optimized aggregate marker algorithm for bandwidth fairness improvement in classifying traffic networks," EURASIP J. Wireless Comm. Network. 2016.
- [5] J. Jin, M. Palaniswami, D. Yuan, Y. N. Dong, K. Moessner, "Priority Service Provisioning and Max-Min Fairness: A Utility-Based Flow Control Approach," J. Netw Syst Manage, Springer, 2017.
- [6] C. H. Fu, Y. Y. Yen, "A study on a differentiated service queuing scheme based on transmitted bytes within a UMTS core network gateway," Wireless Personal Communication. J. Springer, 2013.
- [7] S. Przyluck, D. Czerwinski, "Priority-Aware Packet Pre-marking for DiffServ Architecture Based on H.264/SVC Video Stream Structure," Wireless Personal Commun. Springer, Oct. 2016.
- [8] Tychogiorgos, G. Leung, "Optimization-based resource allocation in communication networks," Comput. Netw. 66, 32-45, 2014.
- [9] D. Raghupathikumar and. BommannaRaja, "combined low latency and weighted fair queue-based scheduling of an input-queued switch," Jsc, 2014.
- [10] P. Li, S. Guo, S. Yu, A.V. Vasilakos, "An opportunistic feeding and routing protocol for reliable multicast with pipelined network coding, INFOCOM, 2012 Proceedings IEEE, pp. 100-108, 2012.
- [11] R. Deng, J. Yang, C. Liu, "A delay prediction-based content and channel aware packet scheduler for real-time video over HSDPA," IEEE, 2013.
- [12] S. Karamchati, S. Rawat, V. Varma, "A novel architecture to enhance quality of service in IP networks," ICOIN, IEEE, 2017.

- [13] R. Sharma, N. Kumar, S. Talabattula, "Performance of New Dynamic Benefit-Weighted Scheduling Scheme in DiffServ Networks," ICACCI, IEEE, 2014.
- [14] Alreshoodi, M. Woods, J. Musa, "QoE-enabled transport optimisation scheme for real-time SVC video delivery," CSNDSP, 2014.
- [15] Paolo Valente," Reducing the execution time of fair- queueing packet schedulers," Computer Communications, Elsevier, 2014.
- [16] SEACORN Project, Network Simulator, <http://www.isi.edu/nsnam/ns>, 2012.
- [17] TS 23.107 Release 7, Quality of Service, concept, and architecture, 3GPP, September 2007.
- [18] N. Shetty, G. Schwartz, and J. Walrand," Internet QoS and regulations," IEEE/ACM Trans. Netw., vol. 18, no. 6, pp. 1725–1737, Dec. 2010.
- [19] J. Heinanen, F. Baker, W. Weiss, J. Wroclawski, "Assured forwarding PHB group," RFC 2597, 1999.
- [20] Li. Xiuqin, Y. Luoheng, Lan Julong, "A DiffServ Supporting and Distributed Dynamic Dual Round Robin Scheduling Algorithm on DS-CICQ," ICMPE, Elsevier, 2012.
- [21] R. Braden, D. Clark, S. Shenker, "Integrated services in the Internet architecture: an overview," RFC 1633, 1994.
- [22] K. Nichols, S. Blake, F. Baker, D. Black, "Definition of the differentiated services field (DS Field) in the IPv4 and IPv6 headers," RFC 2474, 1998.
- [23] EURANE website, <http://www.ti-mc.nl/eurane/>, SEACORN project, 2004.