

A Comparative Simulation Study for Multiple Traffic Scheduling Algorithms over GPRS

Maurizio D'Arienzo[‡], Antonio Pescapè[‡], Rajiv Chakravorty[†], Giorgio Ventre[‡]

[‡]Computer Science Department
University of Napoli "Federico II"
Via Claudio 21 - 80125 Napoli
{maudarie, pescape, giorgio}@unina.it

[†]University of Cambridge Computer Laboratory
William Gates Bldg., JJ Thompson Avenue
Cambridge CB2 0FD. U.K.
{rajiv.chakravorty}@cl.cam.ac.uk

ABSTRACT

General Packet Radio Service (GPRS) - a bearer service to GSM - has been deployed worldwide, and is widely considered a technology precursor to the evolving third generation (3G) wireless networks. The general conception has been that while users will be exposed to faster wide-area wireless data access, experience gained from GPRS could well prove useful for 3G, and also for systems beyond 3G deployment.

In this paper, we present a comprehensive simulation study for different traffic scheduling algorithms for Quality of Service in General Packet Radio Service (GPRS) at the IP level. We first study the correlation between GSM and GPRS users, and show how a dynamic channel allocation scheme between GSM-GPRS can give substantially better performance than the static ones.

We then extend our study by taking into account users' requirements for different QoS profiles, based on seven different scheduling algorithms in GPRS. By simulating traffic related to an ATIS (Advanced Travelers Information System) at the IP level, we show how traffic scheduling algorithms perform by taking into account different performance parameters such as the average traffic, average waiting time in the scheduler, packet loss probabilities in the scheduler based on static and dynamic channel allocation schemes, packet priorities as well as average throughput per-GPRS user. The study gives a comparative analysis for various scheduling algorithms (network designers can benefit from this study, and by extending this to several other scenarios).

Keywords: Wireless and mobile communication networks, performance evaluation, IP and GPRS networks, discrete-event simulation, scheduling algorithms, ATIS.

1. INTRODUCTION

GPRS is a standard from the European Telecommunications Standards Institute (ETSI) on packet data in GSM systems [1]. By adding GPRS functionality to GSM network, TCs (Telecom Operators) can give their subscribers resource-efficient wireless access to external Internet protocol-based networks, such as the Internet and corporate intranets. The base of GPRS is to provide a packet-switched service in a GSM network. As impressively demonstrated by the Internet, packet-switched networks are more efficient in the use of resources for bursty data applications and they provide more flexibility.

This paper describes a discrete-event simulator for GPRS at IP level. The simulator is developed in Matlab environment. The simulator focuses on the communication over the radio interface because this is one of the most crucial aspects of GPRS operation. In fact, the air interface mainly determines the

performance of GPRS. We studied the correlation between GSM and GPRS users with both a static and dynamic channel allocation scheme. The basic DCS (Dynamic Channel Stealing) [2] concept is to temporarily assign traffic channels dedicated to circuit-switched connections usually unused because of statistical traffic fluctuations. This can be done at no expense in terms of radio resource and with no impact on circuit-switched services performance where the channel allocation to packet-switched services is only permitted for idle traffic channels. Stolen channels are immediately released when requested by the circuit-switched service. After a performance study of static and dynamic channel allocation scheme we consider users with different QoS profiles and we exploit seven different scheduling algorithms in order to analyze the performances. We show experimental results and a comparative study in terms of average carried traffic and packet loss probabilities.

Since in an end-to-end path the wireless link is typically the bottleneck, and due to the advanced traffic asymmetry, the simulator focuses on resource contention in the downlink (i.e., the path BSC -> BTS -> MS) of the radio interface. Because of that advanced traffic asymmetry, the amount of uplink traffic, e.g. the traffic induced by acknowledgments, is assumed to be negligible. The functionality of the GPRS core network is not included and the arrival stream of packets is modeled at the IP layer. Other works are present in literature on this relevant aspect. One of these [3] has inspired this work. As far as the workload (simulated traffic) in the simulation environment we considered the characterization of the GPRS traffic for an ATIS (Advanced Travelers Information System) system that models the operative framework we can find in these systems. We also analyze the fundamental parameters (packets loss probability, throughput, channels occupation). Therefore, the simulation environment is characterized with respect to mobile users and data traffic on both urban and suburban situation. The reason of this choice is that nowadays it's possible to receive various types of information on mobile devices. It's possible both to phone and to send e-mails and browsing web pages as well. Moreover, with a third generation mobile device it's possible to contact a system for traffic information too. A system that manages information concerning the traffic and all about pertaining to a "customer of the road" is known as ATIS. An ATIS system is an articulate and heterogeneous collector of information, resources and device for accounting and delivering of traffic information. An ATIS system assists travelers for foresees, planning, analysis, safety and efficiency of the travel. The ATIS applications provide a shared resources for an efficient analysis on the mobility and data integration. The ATIS architecture deals with various traffic sources: traffic circulation reports, road events, probes, maps, etc. The actor of the system are motorist owning mobile devices and users with PC (at home, office,

informative centers, etc...). The big innovation in the last years has been the use of wireless devices as well as the use of traffic information on these devices. Thanks to real time traffic information the quality and efficiency of a travel can be increased. In this paper the workload is related to an ATIS system, but achieved results can be easily extended to others scenario where the traffic model is different. Now we are focusing on ATIS systems because actually there is a particular attention both to life quality and ecological problems. This work would be a little example where ICT solutions can help people.

The paper is organized in 8 sections. After this introduction, next section presents the reference scenario where ATIS environment and GPRS network are explained. The third section presents a short view on related works whereas in the section 4 we present the characterization of GPRS traffic for an ATIS system. In this section the simulation model is presented while results related to the static and dynamic allocation are presented in Section 5. The main issues related to seven implemented scheduling algorithms are presented in Section 6 while the experimental results are showed and analyzed in the Section 7. Finally, Section 8 provides some concluding remarks. In order to increase the readability of this paper an appendix (section 11) with some output graphics is reported.

2. REFERENCE SCENARIO

An ATIS system provides real-time information making easy the decisions about travels (time of departure, choosed path and all the other information about the other critical situation concerning viability and atmospheric environment). Information can be transmitted to the various customers using several device like mobile phones or vehicular navigation systems, but usually transmission over wireless networks are currently very critical. The reason of the big increase of ATIS applications is the ongoing market of new network technologies (2.5G and 3G networks). The services offered by GSM are circuit switched and the max bit-rate is 9600 bit/s. GPRS instead has been thought and designed in order to provide solutions to these restrictions and to offer a packet switched service. In the initial phase the GPRS will allow a bit rate of 60 Kbit/s approximately whereas, for the successive phases, a further speed increment will be possible. Thanks to this rate, phone mobile users will be enabled to e-mail, browsing, downloading and e-commerce applications. Moreover, since a GPRS user is constantly connected to the network even if he does not transmit anything, it is necessary a per-byte user accounting. The traffic is generated from several applications characterized by different requirements of Quality of Service (QoS).

As far as GSM, a physical channel is permanently allocated for a particular user during the entire call. In contrast, GPRS allocates channels only when data packets are sent or received, and they are released after the transmission. For bursty traffic this results in a much more efficient usage of the scarce radio resource. Stemming from this principle, multiple users can share one physical channel. GPRS allows a single mobile station to transmit on multiple time slots of the same TDMA (Time Division Multiple Access) frame. This results in a very flexible channel allocation: one to eight time slots per TDMA frame can be allocated to one mobile station. On the other hand a time slot can be assigned temporarily to a mobile station, so that one to eight mobile stations can use one time slot. GPRS includes the functionality to increase or decrease the amount of radio resources allocated to GPRS on a dynamic basis. The PDCHs (Packet Data Channels) are taken from the common pool of all available channels in the cell. The mapping of physical channels to either packet-switched (GPRS) or circuit-switched (conventional GSM) services can be performed statically or

dynamically ("capacity on demand"), depending on the current traffic load. A load supervision procedure monitors the load of the PDCHs in the cell. According to the current demand, the number of channels allocated for GPRS can be changed. Physical channels not currently in use by conventional GSM can be allocated as PDCHs in order to increase the quality of service for GPRS. When there is a resource demand for services with higher priority, e.g. GSM voice calls, PDCHs can be de-allocated. Because of the poor wireless channel capacity, aggressive admission control will likely be employed to fully utilize the wireless link. Therefore GPRS subscribers can choose their own QoS profile consisting of priority class, delay class, reliability class, peak throughput class and mean throughput class. For a detailed description of the GPRS network architecture and the GPRS Radio Interface, we refer to [4], [5], [6] [7] while for QoS profiles proposed by the ETSI to [1].

2.1 OVERVIEW ON GPRS

As part of the transition towards GPRS, new components have been added in the network subsystem (NSS) to the traditional GSM network. The two new nodes SGSN (Serving GPRS Support Node) and GGSN (Gateway GPRS Support Node) are used for GPRS that will be later upgraded for third generation (3G) Universal Mobile Telecommunication Network (UMTS). The SGSN node acts as a packet switch that performs signalling similar to a mobile switching center (MSC) in GSM, along with cell selection, routing and handovers between different Base Switching Centers (BSCs). It controls the mobile terminal's access to the GPRS network and routes packets to the appropriate BSC.

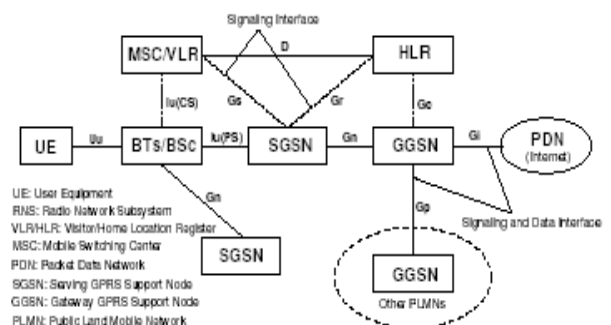


Figure 1: The GSM-GPRS extension

When migrating to "All-IP" UMTS, SGSN will be enhanced to replace the MSC altogether, where it will switch packets to the correct UMTS terrestrial radio network (UTRAN). The GGSN is the last "port of call" that acts as a gateway between the mobile packet routing of GPRS, and the fixed IP routing of the Internet. The MSC/visitor location register (VLR), Home Location Register (HLR), and short message service (SMS) center are functional entities tied to the circuit-switched GSM. To exchange GPRS subscriber information with the SGSN, the HLR is extended by a GPRS register (GR). When a mobile terminal (MT) wishes to use GPRS, it will first attach itself to the network through a signalling procedure. The attach procedure can be performed either when the MT is switched on or when the user wishes to transfer packet data. Depending upon the MT device class, it can connect to either circuit switched or to packet switched services, or both simultaneously [36]. Mobile terminals are classified according to the number of time slots they are capable of operating on simultaneously. For example, many current GPRS devices are classified as '3+1' meaning that

at any given time they can listen to 3 downlink channels (from base station to mobile), but can only transmit on 1 uplink channel to the base station.

A reliable RLC (radio link control) mode ensures that packets are delivered in order, while a selective repeat ARQ (automatic repeat request) coupled with the modulo-128 numbering of data blocks using temporary RLC flow identifiers (TFI) helps to recover from packets received in error. In this scheme, the sender transmits blocks within a window of 64 blocks, and receiver side periodically sends ACK/NACK messages. While every ACK acknowledges all correctly received RLC blocks indicated upto a sequence number (BSN), the NACKs act as a bitmap to selectively request erroneously received RLC data blocks for re-transmission [39]. The sender then just re-transmits the erroneous RLC data blocks further sliding the sending window. However, all this happens at the expense of variable throughput and higher delay due to retransmissions [42]. Radio conditions change with time, and achievable data rates over GPRS can vary, depending on other factors, and to the external environmental interference. Higher interference will lead to higher block error rates over GPRS, and consequently, higher data transfer times. The level of interference is typically specified in channel-to-interference (C/I) ratio of the radio environment. A low C/I (for e.g. < 6-8dB) gives tough radio conditions (high block error rates), a C/I of 13-18dB indicates moderate radio conditions while a high C/I (e.g. > 25dB) gives good channel conditions. GPRS copes with a wide range of radio/channel conditions by making use of 4 different coding schemes (CS-1 TO CS-4) [36] [38], with varying levels of FEC (forward error correction). Most of the currently deployed GPRS networks support only CS-1 and CS-2 [43]- the other two are not used as block error rates would be typically too high for the applications to be useful. While CS-1 is meant for use during tough radio conditions (e.g. < 7-8dB of C/I), CS-2 is particularly useful during tough-to-moderate channel conditions (e.g. < 15-18dB of C/I).

Most GPRS network operators insist that CS-1 and CS-2 is a good compromise coding scheme for the moment. Later, when applications would become more error resilient, then it should be possible to use even CS-3 and CS-4. An other difficulty using CS-3 and CS-4 is that it cannot be supported by many GPRS networks, since the 'Abis' and 'Gb' interface (see figure 1) is currently capacity limited. While CS-4 scheme removes FEC correcting capabilities altogether, the CS-2 scheme employs a coding rate of approximately 2:3, to obtain a transmission rate as high as 13.4 Kb/s per GSM time slot [38]. The effective GPRS data rate is slightly less, among other factors, due to protocol header overhead and signalling messages.

Radio resources of a cell are shared between all GPRS and GSM mobile stations located in the cell. Most network operators typically configure the network to give GSM (voice) calls strict priority over GPRS for time slot allocation. The time slots available for GPRS use, known as packet data channels (PDCHs), are then dynamically allocated (using capacity on demand principle) between mobile terminals with data to send or receive. GPRS can multiplex time slots between different users, and can also allow multiple time slots to be used in parallel to increase bandwidth to/from a particular mobile terminal.

When there is contention for GPRS resources, individual PDCHs may be multiplexed between different users.

When this occurs, the specification allows for packets to be prioritised according to various Quality of Service (QoS) levels. A user can request for a desired QoS profile during the packet data protocol (PDP) context activation phase.

GPRS Release 99 defines several QoS parameters to meet the application requirements for different levels of network QoS.

The release offers several benefits when compared to its predecessor (Release 97/98) [40] - such as - BSS aware QoS profile negotiation, MT and GGSN initiated QoS profile (re)negotiation based on application or network requirements, and multiple PDP contexts per PDP address. Further, four distinct GPRS traffic classes are specified: conversational, streaming, interactive and background [36], [40]. Applications that are delay sensitive belong to the conversational class. The conversational class offers strict delay and bandwidth guarantees, while the background class offers neither quantitative nor qualitative guarantees. It can be at best referred to as the best-effort traffic class. Currently, in the 'phase one' of GPRS deployment, operators only support a single best-effort service class [43]. Further information about GPRS network design and operation can be found in [36] [37].

3. RELATED WORK

In our simulated architecture we present an evaluation of several traffic scheduling methods, including FIFO, Priority FIFO, Static Priority Scheduling (SPS), Shortest Job First (SJF), Earliest Deadline First (EDF), Weighted Round Robin (W.R.R.) and Token Bank Leaky Bucket (T.B.L.B.), with the objective to compare the obtained results in a GPRS scenario.

Scheduling is an important aspect in the QoS support over GPRS networks. The study of GPRS scheduling algorithms would like to clear how the most common scheduling algorithms can be adapted to GPRS scenario. Scheduling problems in GPRS are different from scheduling in other packet switched network due to the existence of specific MAC protocol, multislot capability restriction and difference in QoS requirements. While the QoS profiles for a number of GPRS classes has been specified by ETSI, how QoS management is provided by means of traffic scheduling, traffic shaping, and connection admission control, in a GPRS network is an implementation issue that is attracting significant current research interest.

Although there are a lot of scheduling methods or service disciplines [9-19] that have been studied, e.g. First-In-First-Out (FIFO), Static Priority Scheduling (SPS), Virtual Clock, Weighted Fair Queuing (WFQ), Self-Clocked Fair Queuing, Start-Time Fair Queuing (STFQ), Worst-case Fair Weighted Fair Queuing (WF2Q), Earliest Deadline First (EDF), Delay Earliest-Due-Date (Delay-EDD), Jitter Earliest-Due-Date (Jitter-EDD), Stop-and-Go, Weighted Round-Robin (WRR), Deficit Round Robin (DRR), Hierarchical Round Robin (HRR), Rate-Controlled Static Priority (RCSP), and Leave-in-time etc, we are not aware neither of any extensive evaluation in literature concerning traffic scheduling relative to GPRS requirements (predictive and best effort) nor of many simulation studies.

J. Sau and C. Scholefield [29] consider two scheduling algorithms SPS and MED (Modified Earliest Deadline) in order to study both channel utilization and congestion (in order to study the congestion they use for the metric both the queue length and the normalized frame scheduling).

Q. Pang, A. Bigloo, V. C. M. Leung, C. Scholefield [30] study how QoS management is provided by means of traffic scheduling. They carried out simulation results with respect to three scheduling algorithms: FIFO, SPS and EDF. The objective of their study is meeting the delay profile defined for a number of GPRS classes. They focus on the forward link which represents the bottleneck of a typical GPRS data connection. Following the interest on scheduling algorithms other works [31] [32] [33] have carried out various interesting stuff on novel scheduling techniques and comparative study between promising scheduling algorithms. As far as performance analysis of scheduling algorithms over QoS-aware network we cite [34],

whereas with respect to GPRS delay analysis we have studied [35].

Steps from these cited works, our approach is more complete both in terms of simulated scheduling algorithms and results analysis. We work with seven different scheduling algorithms and we introduce four traffic classes (with different priority) and thanks to it we can analyze the different behaviour for each class. We carried out results with respect to PDCH occupation, average time in the scheduler (and in the queue) and finally for packet loss probability.

4. SIMULATION MODEL

Mobile users and data concerning urban and suburban situations are the actors of our scenario. Therefore we have four data classes: Data from traffic probe, Report on accidents or events, Video-Images, Data from ambient probe. These information can be sent to both the main operating central and the user. Furthermore, such a classification of the traffic can be divided (as far as the bandwidth occupation) in text, graphical and multimedia traffic.

In a simulation problem, the choice of the model to be implemented is very important. In this field there are several models in bibliography. These models have in common the following two aspects:

- The GPRS traffic is a little part of the total traffic and the GSM calls use the major part of the radio resources.
- The GPRS flow between the network and the MS (Mobile Station) is strongly asymmetric: the downlink channel (network to MS) is more stressed, especially in a web-browsing context.

The used model covers the various aspects: network model, GSM traffic model and GPRS traffic model.

The deployed architecture is shown in Figure 2 and the software for the simulation implements the BSS (Base Station Subsystem). The BSS consists of two parts: BSC (Base Station Controller) and BTS (Base Transceiver Station). The IP packets arrive to the BSC and are stored in the "Access Queue" that can contain 1000 packets of 1 Kbyte. The scheduling applied to this queue is the subject of this paper. In the simulation we study two different channel allocations:

- Static schema: on the N total logic channels in the cell there is a static division between GSM and GPRS calls, $N_{GSM} + N_{GPRS} = N$.
- Dynamic schema: the N total logic channels are shared between GSM and GPRS with priority on the GSM calls.

In the second case GSM calls are preemptive with respect to a GPRS packet (GPRS packet goes in the "Suspend Queue"). The Suspend Queue has an higher priority than the Access Queue as well as the successive allocation of a PDCH with regard to a GPRS packet.

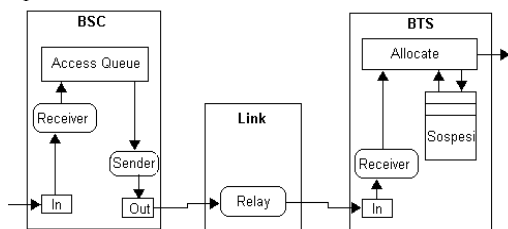


Figure 2: Simulated Network Model

CS-2 is the codec used with one timeslot per user: the data rate is 13.4 kbit/s. An IP packet is fragmented in radio blocks of 416 bit sent onto the radio channel. The radio channels in a cell are 20 ($N=20$): since with TDMA every channel is divided in 8

timeslots, we have a total amount of 160 TCH (Traffic Channel) in the GSM case or PDCH in the GPRS case.

The choice of an appropriate traffic model reflects the customers behaviour of a telecommunication system supplying the results of the TC. This choice enables it for taking decisions about network design. We have decided to model two types of traffic: "Sessions Traffic" and "Single Packets Traffic". In the case of "Sessions Traffic" the user approaches the network in order to receive information "opening" a session: the user requests for packets alternating reception, thinking and transmission phases. In the thinking phase (e.g. the user is reading information for the choice of the travel route) the channel is left idle. The session characteristics are:

- Average time of 3 min (exponential distribution) with minimum of 1 min and maximum of 8 min.
- Average packet length of 1500 byte (exponential distribution) with minimum of 10000 byte and maximum of 30000 byte.
- Reading phase with average time of 20 sec. (exponential distribution) with minimum of 5 sec. and maximum of 40 sec.

Figure 3 shows the case of "Sessions Traffic" with 4 users. The reading time can be used by the system to send packets belonging to other users in order to optimize radio resources occupation. With "Single Packets Traffic" the user receives "light" information from the network: he can receive updated data, images or packets with short messages. This traffic have these characteristics: average packet length of 2500 bytes (exponential distribution) with minimum of 100 bytes and maximum of 3500 bytes.

Both in the case of "Sessions Traffic" and "Single Packets Traffic" the arrivals of the GPRS calls are modelled with a Poisson process with a variable λ parameter. In particular the λ parameter is used to model the arrival of GSM/GPRS traffic: we used two percentages: GSM equal to 90% and 95% and, therefore, GPRS equal to 10% and 5%.

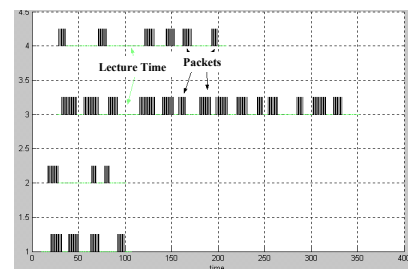


Figure 3 : Characterization of GPRS Sessions

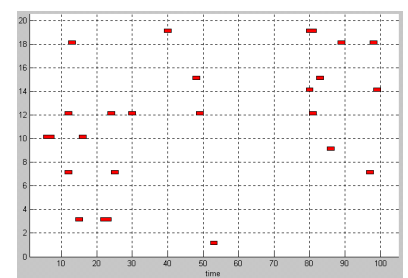


Figure 4 : Characterization of GPRS Packets

Total traffic (sessions and packets) is divided in four priority classes: higher priority is the number 0 whereas the lower priority is the number 4 (this class is the best effort class). The total GPRS traffic (λ_{GPRS}), is divided in: 20% class 1; 20% class 2; 20% class 3; 40% class 4. The two traffic scenario (Figure 5 and Figure 6) are:

1. Urban Context (Type 1): a typical user of this context demands mostly textual information, like “Where is the pharmacy on duty nearby...?” and therefore the response is a little text message. In our model this scenario has been implemented setting the percentage of customers with light traffic to 80% of the total GPRS traffic.
2. Suburban Context (Type 2): a typical user of this context demands mostly graphical/multimedia information. In this case the user has a device with the possibility to visualize images.

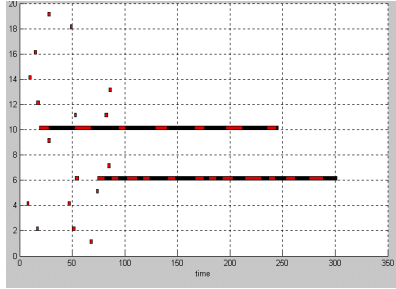


Figure 5 : Type 1 Traffic

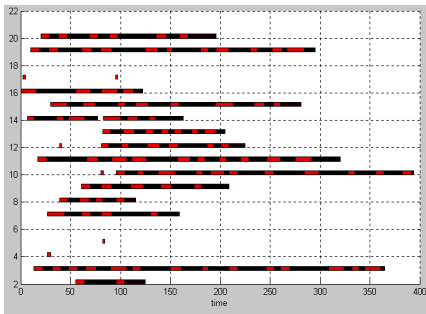


Figure 6 : Type 2 Traffic

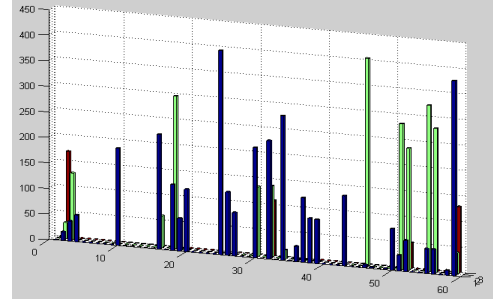


Figure 7 : Characterization of GSM Traffic

The GSM calls (Figure 7) arrivals are modelled as a Poisson process having a λ_{GSM} parameter and their average time is 120 sec. A GSM call has priority on the GPRS traffic (this situation is realistic because a TC uses a time-based billing system for the GSM calls and a volume-based billing system for the GPRS traffic).

The software carried out for the simulation has been developed in Matlab. The modules developed for the simulation are shown in Figure 8. “Channel Allocation” receives the GSM calls from “Traffic Generator” whereas the GPRS traffic comes from the “Access Queue” module only when radio resources are available. If all available channels are busy when a GSM call arrives the routine chooses a channel used by a GPRS packet and releases it: this channel is then allocated to the GSM call. The GSM call is refused only if all the channels are used by GSM users. The de-scheduled GPRS packets are put in “Suspend Queue” and they have priority in the allocation of new channels. Table 1 shows all parameters of the simulation.

Model	Parameter	Value
Network	Physic Channel per Cell, N	20
	GPRS Users, M	20
	BSC Buffer Size, K	1000 IP packets
	Channel Coding, $\mu_{servizio}$	CS-2 13.4 kbps
GSM/GPRS Traffic	GSM/GPRS arrival, $\lambda = \lambda_{GSM} + \lambda_{GPRS}$	0.1+2 /sec.
	GSM Users (%)	90 %
	GPRS Users (%)	10 %
	GSM Call average time, $1/\mu_{GSM}$	120 sec.
	GPRS Session average time, $1/\mu_{GPRS}$	300 sec.
	Average Packet size in the session	15000 byte
	Average Packet size in the singol packet	2500 byte
	Average reading time in the session	20 sec
	Classe 1 - Data Traffic (%)	20%
	Classe 2 - Data Traffic (%)	20%
	Classe 3 - Data Traffic (%)	20%
	Classe 4 - Data Traffic (%)	40 %.

Table 1 : Simulation parameters related to traffic and network

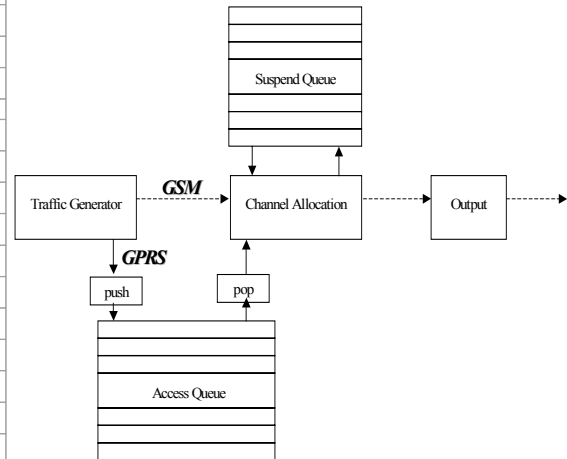


Figure 8 : Simulation Software Architecture

5. SIMULATION RESULTS

Before to cover the main topics of this work (scheduling algorithms analysis and implementation) we study the Results showed here are related to SJF (Short Job First) and FIFO (First In First Out) disciplines.

differences between dynamic and static channel allocation in order to choice the best one. On the best allocation system we have experimented scheduling algorithms. All next figures have GPRS and GSM arrive on the x axis.



Figure 9 : PDCH utilization for the static allocation

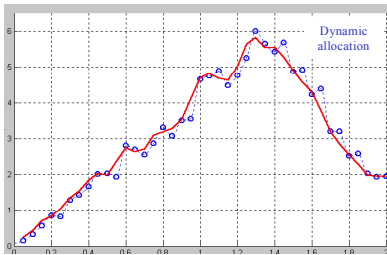


Figure 10 : PDCH utilization for the dynamic allocation

Figures 9 and 10 show channel utilization in the static case (with 2 and 4 PDCHs) and in the dynamic case. The former shows that in the case of 4 PDCHs the system is already overloaded with an arrival rate of 1 user per second; in case of 2 PDCHs the saturation is reached with a lower rate. The latter, instead, shows how the system is able to assign, in a certain interval of arrivals, more than 4 PDCHs thus allowing to serve an higher number of GPRS users. However, as soon as the load increases, the available PDCHs decrease due to the concurrence with the GSM calls, since they have a priority higher than the GSM ones. Figures 10 and 11 show the packet-loss probability in the case of BSC buffer overflow. In the former, the static cases with 2 and 4 PDCHs are compared: with 4 PDCHs the curve is lower since there are more radio resources available. The latter shows the packet-loss probability in the dynamic case. When the arrival rate is up to 1.3 users per second the probability is almost zero. With an higher rate it increases faster and faster up to overtake the static allocation curve. That is why, in the static case, the system guarantees a minimum number of PDCHs, whereas in the dynamic case there is no such guarantees and the GPRS users suffer the GSM traffic priority.

The comparison between the two allocation schemes highlights that with the dynamic approach there is a better allocation of the radio resources, whereas the advantage of the static approach relies in the implementation simplicity.

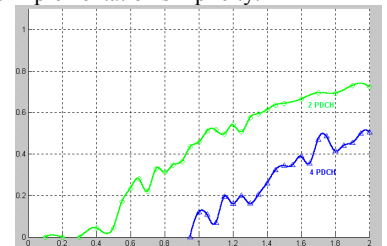


Figure 11 : Packet-loss probability in the static allocation

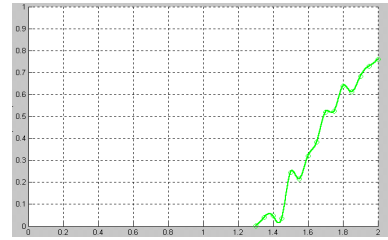


Figure 12 : Packet-loss probability for the dynamic allocation

As far as the packet-loss probability the dynamic allocation is better when arrival rate is lower than 1.6 users per second, whereas the static one is better in the case of an higher rate.

The second phase of the simulation concerns an analysis about a comparison of scheduling for general packet radio service classes. Thanks to the results achieved through the first phase, only the dynamic allocation scheme will be taken in consideration.

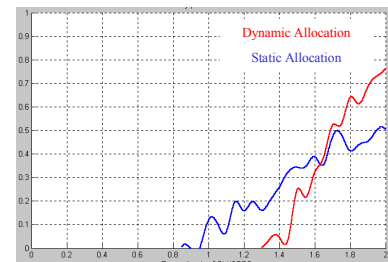


Figure 13 : Packet-loss probability comparison for the static and the dynamic allocations

6. SERVICE SCHEDULING FOR GENERAL PACKET RADIO SERVICE CLASSES

While the QoS profiles for a number of GPRS classes has been specified by ETSI, how QoS management is provided by means of traffic scheduling, traffic shaping, and connection admission control, in a GPRS network is an implementation issue that is attracting significant current research interest. In this section we presents an evaluation of several traffic scheduling methods, including FIFO, Priority FIFO, Static Priority Scheduling (SPS), Shortest Job First (SJF), Earliest Deadline First (EDF), Weighted Round Robin (W.R.R.) and Token Bank Leaky Bucket (T.B.L.B.), with the objective to compare the obtained results. The seven algorithms subject of our study are applied on Access Queue and each determines the packets allocation over PDCHs [20]. In this section we don't provide analytical details and theory regarding implemented scheduling algorithms but we depicted our modifications in order to exploit in a simulation environment of real scheduling algorithms. After this phase we show a comparative study. The main difference between theory and implementation issues is the finite length of the queue: this real situation leads packets loss towards full Access Queue.

First In First Out (FIFO) : FIFO or FCFS (First Come First Served) is the simplest scheduling and queuing method. First arrived packet is served first. We have implemented two version of this algorithms: Single buffer (for all classes) and Two Buffers (priority classes and best effort). In the first case the queue length is 1000 pkts and traffic is treated in the same manner. When two buffers are used, one is for predictive services (class 1, 2 and 3) while the other is for best-effort service (class 4). We always use a separate buffer for best-effort (250 pkts length) services in all of scheduling methods. The best-effort service is activated only if buffers (750 pkts length) for the predictive classes are empty.

Priority First In First Out (FIFO): pFIFO or pFCFS (Priority First Come First Served) is the simplest scheduling and queuing method with priority. Packets belonging to class i are running only if in the queue there aren't packets of class j , with $j < i$. Queue length is 1000 pkts and it is necessary to analyze all the queue for the sake of find different classes packets. When the queue is full packets are lost without priority indication.

Static Priority Scheduling (SPS): with SPS [21], each service class has its own buffer and it has assigned a fixed (static) service priority: highest for class 1 and lowest for class 4 (best-effort). All the queues are 250 pkts length. When the next downlink time slot is available, a class i buffer will receive service only if all class j ($j < i$) buffers are empty. Each queue is scheduled with FIFO discipline and the time for the read of all queue is function of packet numbers in the queue.

Shortest Job First (SJF): the scheduling is based on packets length. The short packets are served first of long packets. We have implemented three version of SJF: SJF standard, SJF with virtual length and SJF with virtual length and separate queue for each class of traffic. The first case is simple. We have only one queue with 1000 pkts and the length is the measure in bytes. In the second case we have only one queue but we introduce a virtual length that includes the priority class of the single packet: $virtual\ length = f(l, c) = l * c$, where l is the real length in byte of the packet and c is the class of service. In the third case there are four different queue of 250 pkts and each queue is served with SPS discipline.

Earliest Deadline First (EDF): with Earliest Deadline First (EDF) or Earliest Due Date (EDD) method, each arrived packet has its own deadline (or due-date) [22]. Packets are served according to their deadlines. Assume the arrival time of a packet is a , and the length of the packet is l . Its priority class is c ($1 \leq c \leq 3$), the time-slot capability of its destination is s , and $rate$ denote the data rate of one time-slot. The deadline is:

$$dead_line = a + r(c, l) - \frac{l}{s \times rate}$$

where, the function $r(c, l)$ represents the delay requirement of the packet with delay class c and length l calculated by.

$$r(c, l) = \begin{cases} \frac{0.5 \times l}{128} & c = 1; l \leq 128 \\ 0.5 + \frac{1.5 \times l}{1024 - 128} & c = 1; l > 128 \\ \frac{5 \times l}{128} & c = 2; l \leq 128 \\ 5 + \frac{10 \times l}{1024 - 128} & c = 2; l > 128 \\ \frac{50 \times l}{128} & c = 3; l \leq 128 \\ 50 + \frac{25 \times l}{1024 - 128} & c = 3; l > 128 \end{cases}$$

The EDF mechanism needs to sort the packet queue using at least $O(\log N)$ insertion operation for each arrived packet. This may affect its feasibility due to implementation difficulty.

In this case there are two queues: the first one for the priority traffic is 750 pkts length and the second one for the best effort traffic (class 4) of 250 pkts. In our simulation $rate$ is equal to 1675 and s equal to 2000. These values are dependent by CS-2 coding (13,4 Kbit/sec \rightarrow $data\ rate = 16\ kbps$ over the Abis circuit between BSC and BTS).

When a new packet arrives and the queue is full, it is necessary to calculate packet's deadline and the deadline of all packets present in the queue. The new packet can be inserted in the queue if its deadline is minor of the deadline of one or more packets present in the queue and the slots/place in queue for the old packets are larger than the place for new packet. If in the queue there isn't necessary slots, a sufficient number of packet

must be discarded (according to deadline discipline). The scheduling discipline process the packet following crescent deadline. In this case it is needed to analyze all the queue.

Weighted Round Robin (WRR): The implementation of this algorithm is compliant with theory details. There are four different queues each one of 250 pkts. The queues are cyclically scheduled and a different number of time slot for the transmission is assigned according to their priority. The first queue can transmit on four time slots, the second on three time slots, the third on two time slots and the four (best effort) on one time slot. In this algorithm (as well as in SPS) it is necessary to check the presence of the packets in the different queues: when there aren't packet in a queue the free time slots must be let to the early low class service [23].

Token Bank Leaky Bucket (TBLB): T.B.L.B. [24], [25], [26] was the first scheduling algorithm used in wireless networks over the downlink channel. It merges both policing and scheduling functions. In the T.B.L.B. algorithm each data flow goes into LB (Leaky Bucket). This LB has a specific token rate (r). The LB holds P tokens, which are enough for one packet (assuming fixed size packets). Each arriving packet (with an arrival rate λ) is buffered at the LB input (queue D) until it can acquire enough tokens to allow its departure to the output buffer of the link, which is emptied at a constant rate μ . Unused tokens overflow the LB to the token bank of size B . Each flow that has run out tokens in its LB may borrow not more than m tokens from the token bank at a time, where m is the 'Burst Credit'. A token counter (E) is associated with each flow and it counts the number of tokens both borrowed from or deposited into the token bank. A flow is not allowed to borrow any more tokens when E falls below the 'Debt Limit'. Borrowing may resume when E exceeds the 'Creditable Threshold'. Above parameters are selected with respect to each flow in order to regulate the burstiness of the flow over the output link. The E token counter and its (E/r) rate calculate the priority that is used to borrow tokens from token bank. The TBLB scheduler is able to serve packets by distributing unused bandwidth from other connections. Otherwise packets could be discarded/arked by the per-flow LB policer. TBLB exploits the statistical multiplexing of group connections and thereby enhances the utilization of the output link bandwidth.

TBLB parameters	Packet	Session
Token Generation Rate Class 1	100 bytes/s	1000 bytes/s
Token Generation Rate Class 2	90 bytes/s	900 bytes/s
Token Generation Rate Class 3	80 bytes/s	800 bytes/s
Token Generation Rate Class 4	70 bytes/s	700 bytes/s
Token Pool Size Class 1	100 bytes	1000 bytes
Token Pool Size Class 2	90 bytes	900 bytes
Token Pool Size Class 3	80 bytes	800 bytes
Token Pool Size Class 4	70 bytes	700 bytes
Date Pool Size	10000 bytes	200000 bytes
Debt Limit	-10000 bytes	-100000 bytes
Burst Credit	10000 bytes	100000 bytes

Table 2: T.B.L.B.implementation parameters

In our implementation T.B.L.B. has eight LB: there are LBs for each priority class (the former for the sessions and the latter for the packets).

In our scenario packets have variable length. When a new packet arrives it is needed a check of the number of token and of the value of the E ($E > Debt\ Limit$) in order to borrow the right number of tokens. Hence it is also needed a check in the bank to discover if the required number of tokens are available. Both D and Access Queue are scheduled according to a FIFO discipline. The token arrival rate in several LBs is determined by packet priority class and traffic type. In the table 2 values of all parameters used in our simulation are reported. As far as the Access Queue overflow, token arrival rates have been chosen with a max bound of the Access Queue max length (1000 pkts). The policing guarantees the absence of Access Queue overflow with respect to each LB of the flows.

7. EXPERIMENTAL RESULTS

The simulator has been deployed as discrete-event: the discrete time evolves with period of 1 second. As far as the previous sessions (description of simulation parameters) the simulator generates GSM/GPRS traffic and puts this traffic over TCH or over PDCH. A single test keeps 5000 seconds of traffic. The arrival rate (λ) is variable between 0.1 and 2 calls per second. According to these parameters, measures cover packet loss percentage (or each priority class in case of BSC buffer overflow; average number of assigned PDCH is calculated by adding the PDCH assigned in measure time (5000 seconds) and dividing it by 5000 seconds; average time in the scheduler is calculated by adding the wait time in both BSC Access and in Suspend Queue and dividing the results by the total number of transmitted packets. These measures are related to following situations: 20% Type 1 Traffic and 80% Type 2 Traffic; 80% Type 1 Traffic and 20% Type 2 Traffic; 50% Type 1 Traffic and 50% Type 2 Traffic. For these three situations the simulator shows comparisons among all implemented scheduling algorithms with respect to:

- packet loss percentage for each priority class;
- average number of PDCH occupied;
- average time in the scheduler for each priority class;

Results with the best granularity of our simulator are not reported in this paper. Rather than showing all output graphics (some of these are depicted in Appendix section) in this section we focus on a comparative analysis. TBLB algorithm shows worst performances: this behaviour is imputed to the dynamic channel allocation. In this case the admission control is not appropriate because GPRS has not guaranteed resources. Output results show that EDF is the best algorithm for the first three classes, although the most part of the traffic in current GPRS networks is a "best effort" traffic. As far as this last class of service, best results are obviously experimented by F.I.F.O or S.J.F. Next tables show a comparative schema with the best algorithm with respect to each class of service and for each traffic situation.

Packet loss	Class 1	Class 2	Class 3	Class 4
20% packet 80% session	E.D.F.	E.D.F.	F.I.F.O. with two queues	S.J.F.
80% packet 20% session	S.P.S., E.D.F.	E.D.F.	-	S.J.F.
50% packet 50% session	E.D.F.	E.D.F.	priority F.I.F.O.	-

Table 3 : Packet loss

As you can read from these tables, in case of average number of PDCH occupied the performances are equal for all simulated

algorithms, with the exception of the only TBLB that shows bad performances.

Average time	Class 1	Class 2	Class 3	Class 4
20% packet 80% session	Priority F.I.F.O.	priority F.I.F.O.	priority F.I.F.O.	S.J.F.
80% packet 20% session	Priority F.I.F.O. - S.J.F. with virtual length	-	-	S.J.F.
50% packet 50% session	Priority F.I.F.O. - S.J.F. with virtual length	Priority F.I.F.O. - S.J.F. with virtual length	Priority F.I.F.O. - S.J.F. with virtual length	S.J.F.

Table 4 : Average time in the scheduler

8. CONCLUSIONS AND FUTURE WORK

This paper presents an IP level discrete-event simulator for GPRS systems developed in Matlab environment. With this simulator we provide a comprehensive performance study of radio resources sharing between circuit switched GSM connections and packet switched GPRS sessions under both a static and a dynamic channel allocation scheme. A comparison between these channel allocation schemes has been presented with the result of better performances for the dynamic scheme. The only advantage of the static scheme is in its easy implementation. Hence we studied seven scheduling algorithms in order to obtain results regarding different performances. Such results can give valuable hints for network designers on how many packet data channels should be allocated for GPRS and how many GPRS session should be allowed for a given amount of traffic in order to guarantee appropriate quality of service according to the selected scheduling algorithm. Results of implemented simulation are not valid for every situation but they can be analyzed and compared in the depicted scenario. Here a comparative analysis of seven scheduling algorithms has been presented. Results and related graphics show packet loss percentage, average time in the scheduling phase and finally average number of allocated channel for data traffic (PDCH). Future enhancements concerns simulation period (using longer simulation time) and arrivals process. Finally simulated architecture can be improved including re-transmission mechanisms caused by noisy channel. In order to exploit a generic study, it is possible to consider another workload schema with more details and greater granularity [27], [28]. We wish to consider a real example designed to tailor four QoS traffic classes defined by ETSI. Traffic type includes conversational, streaming, interactive and background traffics [8]: Conversational Traffics (i.e. voice traffic modelled by means of exponential distribution), Streaming Traffics, Interactive Traffics, Background Traffics (i.e. e-mail downloading where the packet size follow the Cauchy distribution).

9. ACKNOWLEDGEMENT

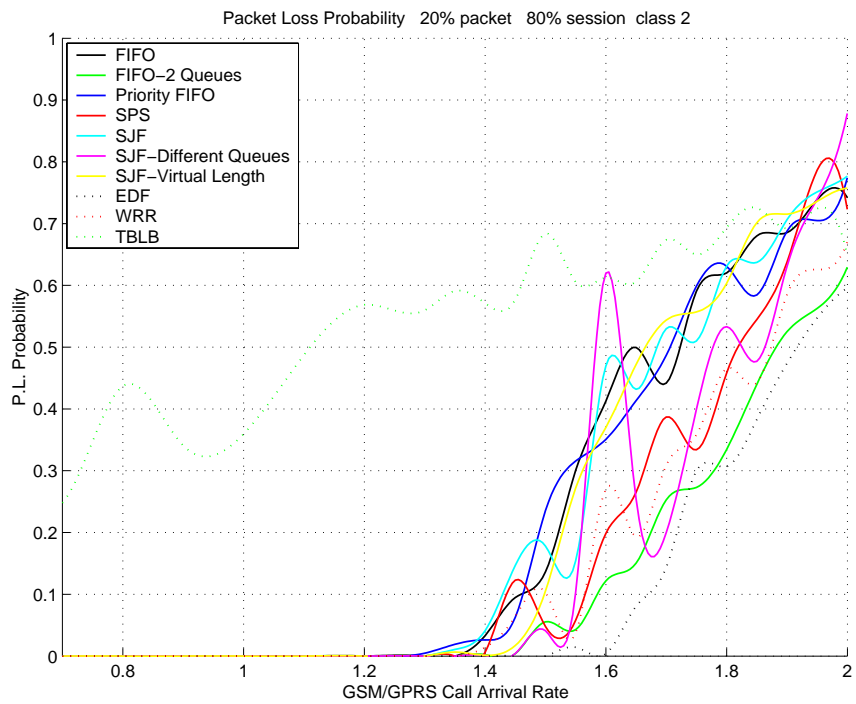
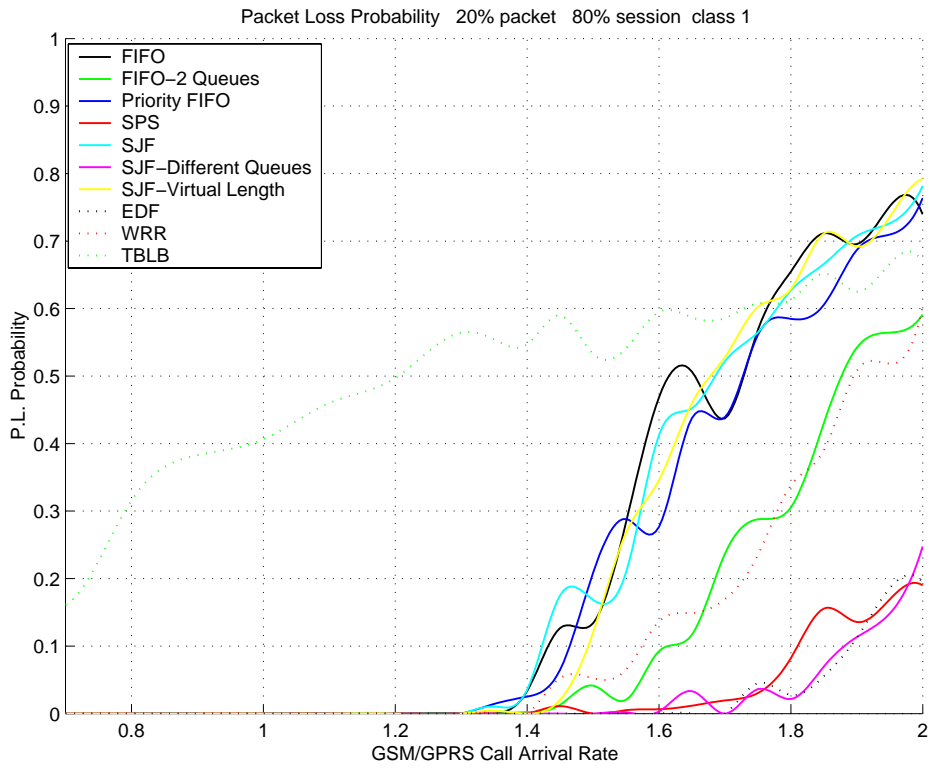
This work has been carried out partially under the financial support of the Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) in the framework of the FIRB Project "Middleware for advanced services over large-scale, wired-wireless distributed systems (WEB-MINDS)"

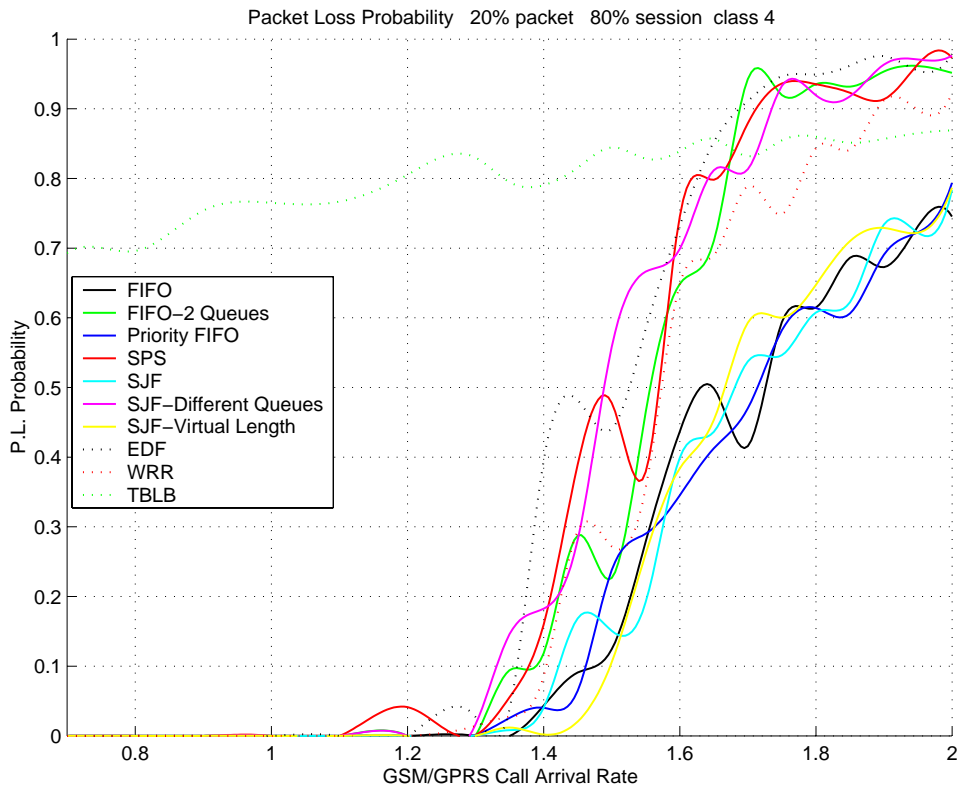
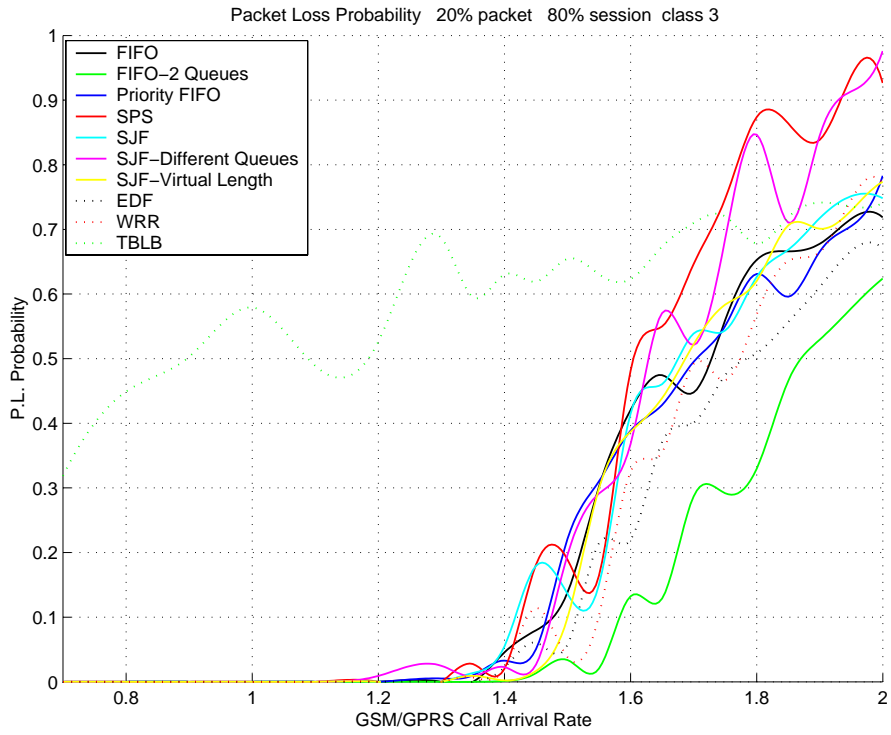
10. REFERENCES

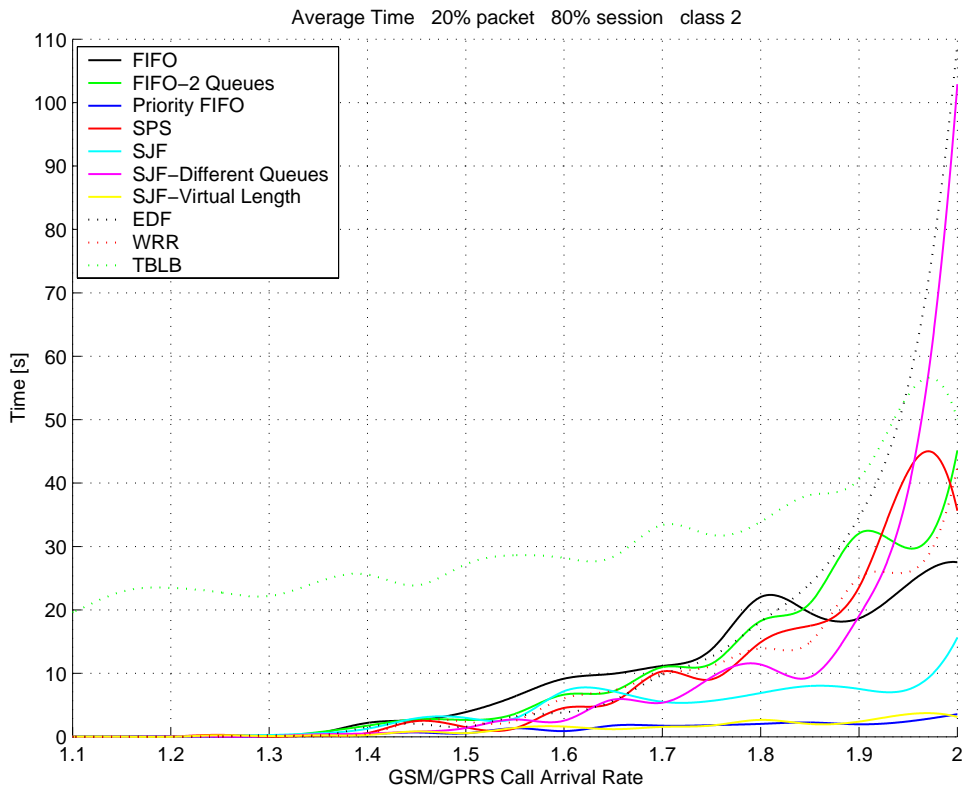
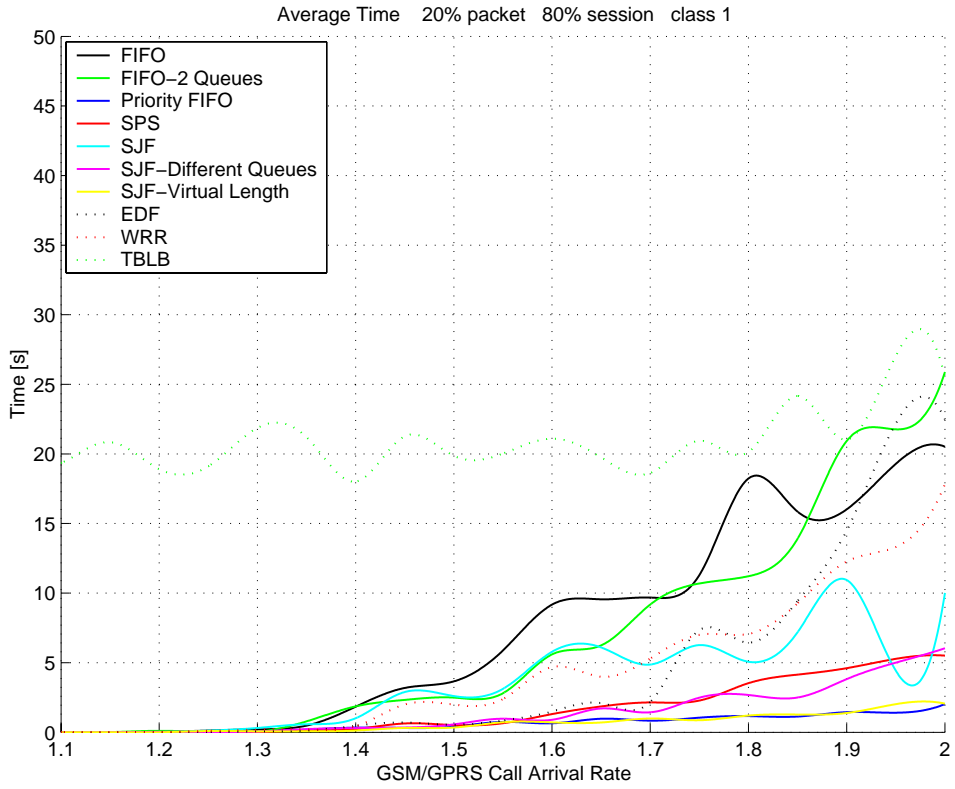
- [1] ETSI, Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Service description; Stage 2, GSM recommendation 03.60, 1999.

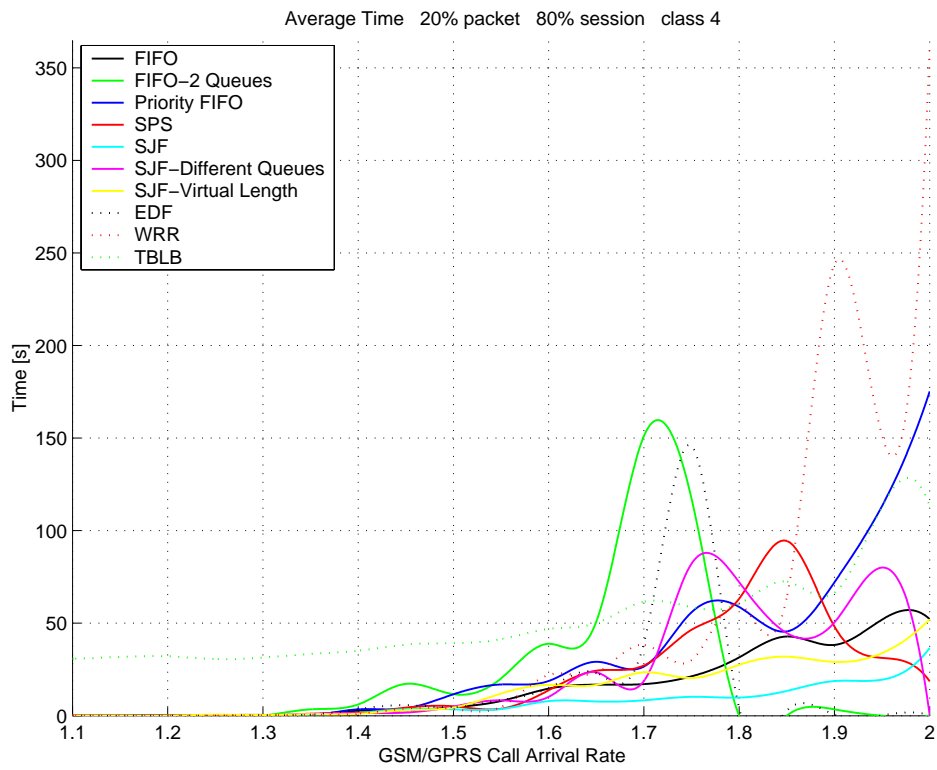
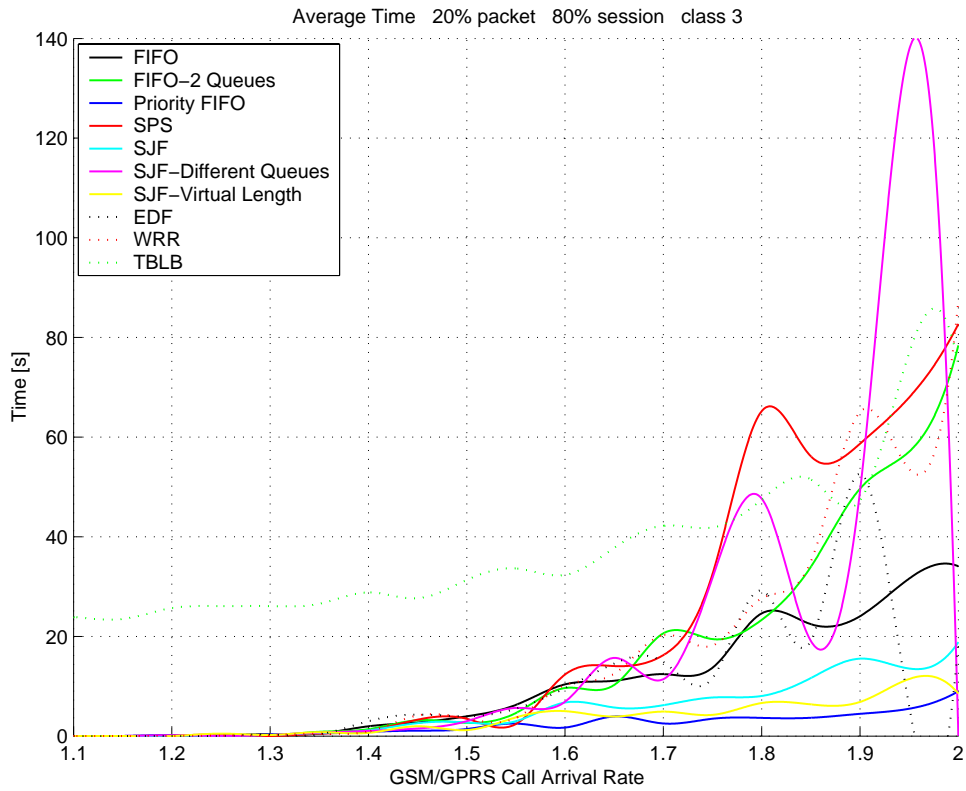
- [2] G. Bianchi, A. Capone, L. Fratta, L. Musumeci, Dynamic Channel Allocation Procedures for Packet Data Services over GSM Networks, Proc. ISS'95, Berlin, Germany, 246-250, 1995.
- [3] C. Lindemann and A. Thümmler, Evaluating the GPRS Radio Interface for Different Quality of Service Profiles. In Proc. 12th GI/ITG Fachtagung Kommunikation in Verteilten Systemen (KiVS), Hamburg, Germany, pp. 291-301, Feb. 2001
- [4] ETSI Doc. Draft EN, GSM 04.60: "GPRS, MS-BSS interface, RLC/MAC Protocol", ver. 7.0.0, 07/1999.
- [5] ETSI Doc. Draft EN, GSM 02.60: "GPRS, Service Description Stage1", ver. 6.2.0, 04/1999.
- [6] R. Kalden, I. Meirick, M. Meyer, Wireless Internet Access Based on GPRS, IEEE Personal Comm. 7, 8-18, Apr. 2000.
- [7] ETSI Doc. TS, GSM 03.64: "Overall description of the GPRS radio interface", ver. 6.2.0, 05/1999.
- [8] Wen-Tsuen Chen Jaw-Liang Lo Hung-Chang Hsiao, Multiple Traffic Scheduling for Enhanced General Packet Radio Service
- [9] C.Li, R.Bettati, and W.Zhao, "Static priority scheduling for ATM networks", Proc. 18th IEEE Real-time Systems Symposium, pp.264-273, 1997.
- [10] T.Hou and C.Chen, "Jitter-EDD implementation for transporting MPEG-2 video streams on ATM network", Proc.12th International Conf. on Information Networks, pp.155-160, Jan. 1998.
- [11] H.Zhang, "Service disciplines for guaranteed performance service in packet-switching networks", Proceedings of the IEEE, vol.83, no.10, pp.1374-96, Oct.1995.
- [12] P.Goyal, H.M.Vin and H.Cheng, "Start-time fair queuing: a scheduling algorithm for integrated services packet switching networks", Proc. of SIGCOMM'96, 1996.
- [13] N.R.Figueira and J.Pasquale, "Leave-in-time: a new service discipline for real-time communications in a packet-switching network", Proc. of SIGCOMM'95, pp.207-18, 1995.
- [14] D.D.Clark, S.Shenker and L.Zhang, "Supporting real-time applications in an integrated services packet network: architecture and mechanism", Proc. ACM SIGCOMM'92, pp.14-26, 1992.
- [15] L.Zhang, "Virtual clock: a new traffic control algorithm for packet switching networks", Proc. ACM SIGCOMM'90, pp.19-29, Sep. 1990.
- [16] J.C.R.Bennert and H.Zhang, "WF 2 Q: worst-case fair weighted fair queueing", INFOCOM'96, 1996.
- [17] S.Golestani, "A self-clocked fair queueing scheme for broadband applications", Proc. IEEE INFOCOM'94, pp.636-646, Jun.1994.
- [18] H.Zhang and D.Ferrari, "Rate-controlled static priority queueing", Proc. IEEE INFOCOM'93, pp.227-236, Apr.1993.
- [19] C.Liu and J.Layland, "Scheduling algorithms for multiprogramming in a hard real-time environment", J.ACM, vol.20, pp.46-61, Jan. 1973.
- [20] E. Nowicki, J. Murphy, Resource allocation for interactive traffic class over GPRS, Dublin City University
- [21] J. Sau, C. Scholefield, Scheduling and Quality of Service in the General Packet Radio Service, Motorola Cellular Infrastructure Group, 1998
- [22] Q. Pang, A. Bigloo, V. C. M. Leung, C. Scholefield, Service Scheduling for General Packet Radio Service Class, 1999
- [23] Chuck Semeria, Supporting Differentiated Service Classes: Queue Scheduling Disciplines, Juniper Network, 2001
- [24] William K. Wong, Victor C. M. Leung, TBLB Algorithm for Servicing Real-time Multimedia Traffic Streams, University of British Columbia Vancouver, B.C., Canada
- [25] William K. Wong, Victor C. M. Leung, Scheduling for Integrated Services in Next Generation Packet Broadcast Networks, University of British Columbia Vancouver Canada, 1998
- [26] William K. Wong, Yi Qian, Victor C. M. Leung, Scheduling for Heterogeneous Traffic in Next Generation Wireless Networks, University of British Columbia, Vancouver Canada and Xybridge Technologies, Inc., Richardson, Texas, 2000
- [27] D. Staehle, K. Leibnitz, P. Tran-Gia, Source traffic modeling of wireless applications; Institut für Informatik, Universität Würzburg, June 2000
- [28] Norbert Vicari, Stefan Köhler, Joachim Charzinski, The Dependence of Internet User Traffic Characteristics on Access Speed, University of Würzburg, 2000
- [29] J. Sau, C. Scholefield, Scheduling and Quality of Service in the General Packet Radio Service, Motorola Cellular Infrastructure Group, 1998
- [30] Q. Pang, A. Bigloo, V. C. M. Leung, C. Scholefield, Service Scheduling for General Packet Radio Service Class, 1999
- [31] V. Sivaraman, F. Chiussi, M. Gerla, End-to-End Statistical Delay Service under GPS and EDF Scheduling: A Comparison Study. In Proceedings of INFOCOM '01, Anchorage (Alaska). April 22 - 26, 2001
- [32] A. I. Elwalid and D. Mitra. Design of generalized processor sharing schedulers which statistically multiplex heterogeneous qos classes. In Proceedings of INFOCOM '99, New York, NY, March 1999.
- [33] K. Kumaran, G. E. Margrave, D. Mitra, and K. R. Stanley. Novel techniques for the design and control of generalized processor sharing schedulers for multiple qos classes. In Proceedings of INFOCOM 2000, Tel Aviv, Israel, March 2000.
- [34] M. A. Marsan, M. Franceschinis, E. Leonardi, F. Neri, A. Tarello, Instability Phenomena in Underloaded Packet Networks with QoS Schedulers. In Proceedings of INFOCOM 2003, San Francisco California, USA, March 30- April 3, 2003.
- [35] M. A. Marsan, P. Laface, M. Meo, Packet Delay Analysis in GPRS Systems. In Proceedings of INFOCOM 2003, San Francisco California, USA, March 30- April 3, 2003.
- [36] G. Brasche and B. Walke, Concepts, Services and Protocols of the New GSM Phase 2+ General Packet Radio Service", IEEE Communications Magazine, August 1997.
- [37] B. Walke, Mobile Radio Networks, Networking and Protocols (2. Ed.), John Wiley & Sons, Chichester 2001
- [38] C. Bettsetter, H. Vogel, J. Eberspacher, GSM Phase 2+ General Packet Radio Service GPRS: Architecture, Protocols, and Air Interface", IEEE Communication surveys Third Quarter 1999, Vol.2 No.3.
- [39] M. Meyer, TCP Performance over GPRS", In Proceedings of IEEE WCNC, pages 1248-1252, 1999
- [40] P. Stuckmann, Quality of Service Management in GPRS-Based Radio Access Networks", Telecommunication Systems, Kluwer Academic Publishers, 19:3,4, pages 515-546, 2002
- [41] J-H. Park, Wireless Internet Access for Mobile Subscribers Based on the GPRS/UMTS Network", IEEE Communications Magazine, April 2002.
- [42] R. Ludwig et al., Multi-Layer Tracing of TCP over a Reliable Wireless Link", In Proceedings of ACM SIGMETRICS, 1999.
- [43] An Introduction to the Vodafone GPRS Environment and Supported Services, Issue 1.1/1200, December 2000, Vodafone Ltd., 2000.

11. APPENDIX

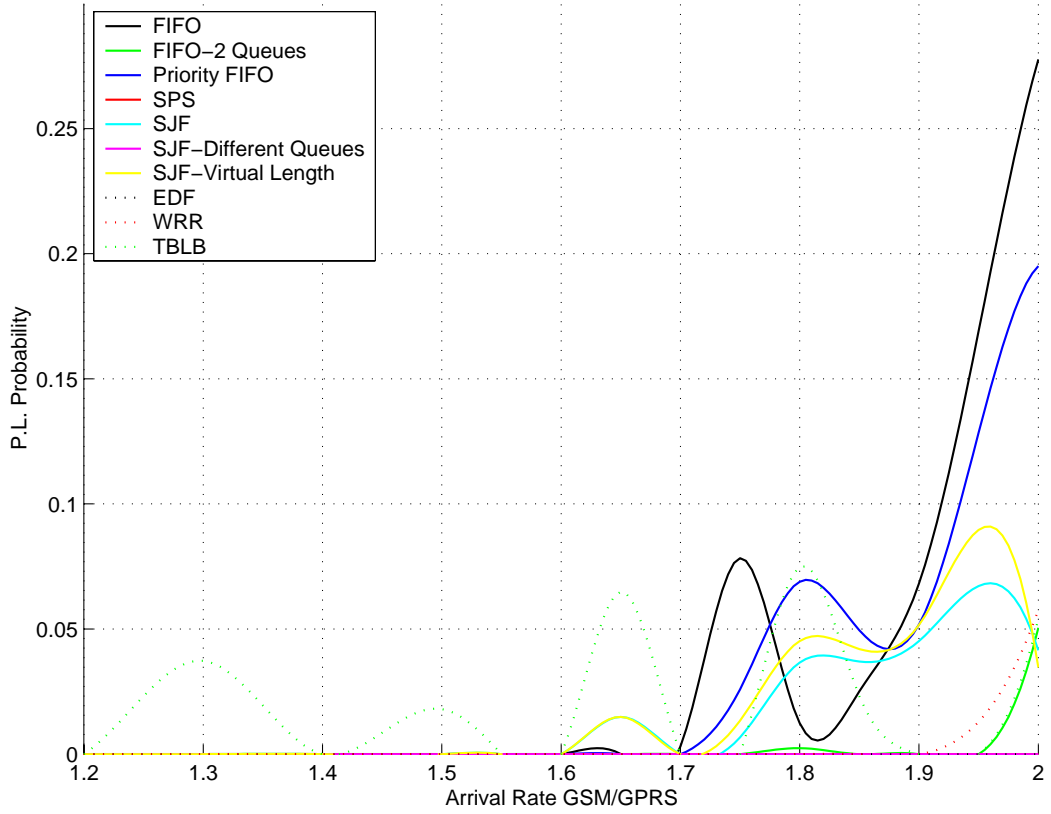




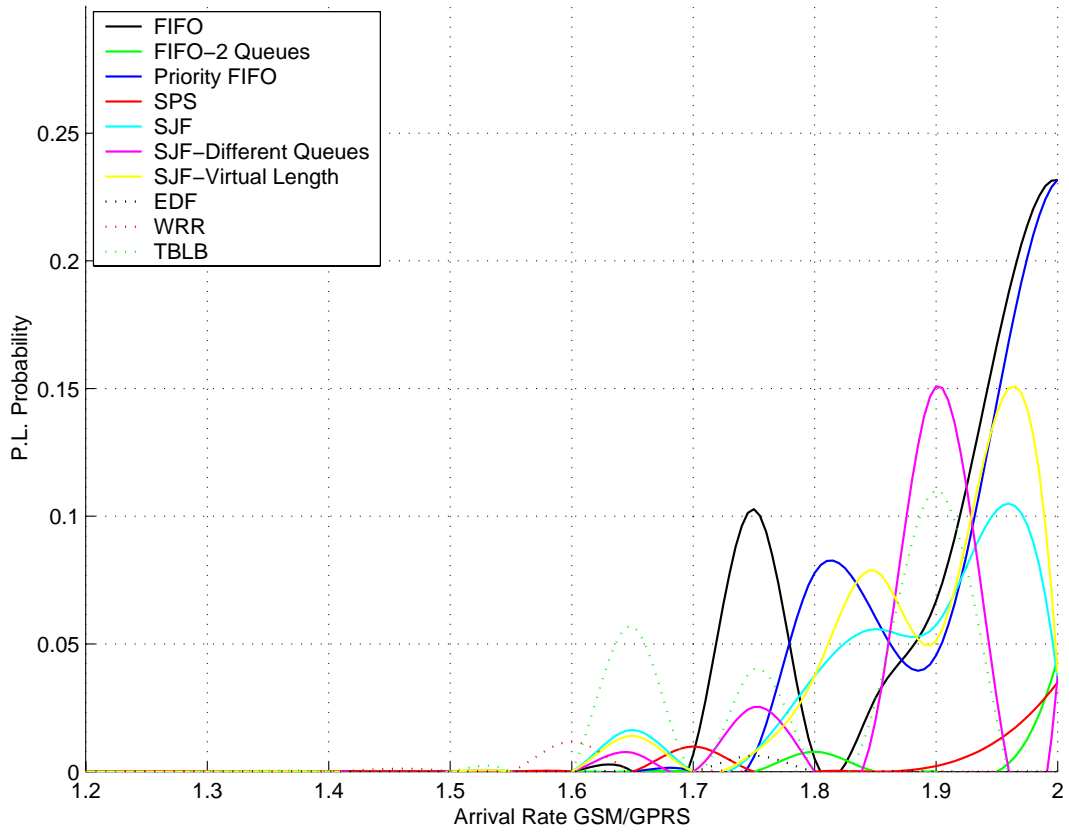




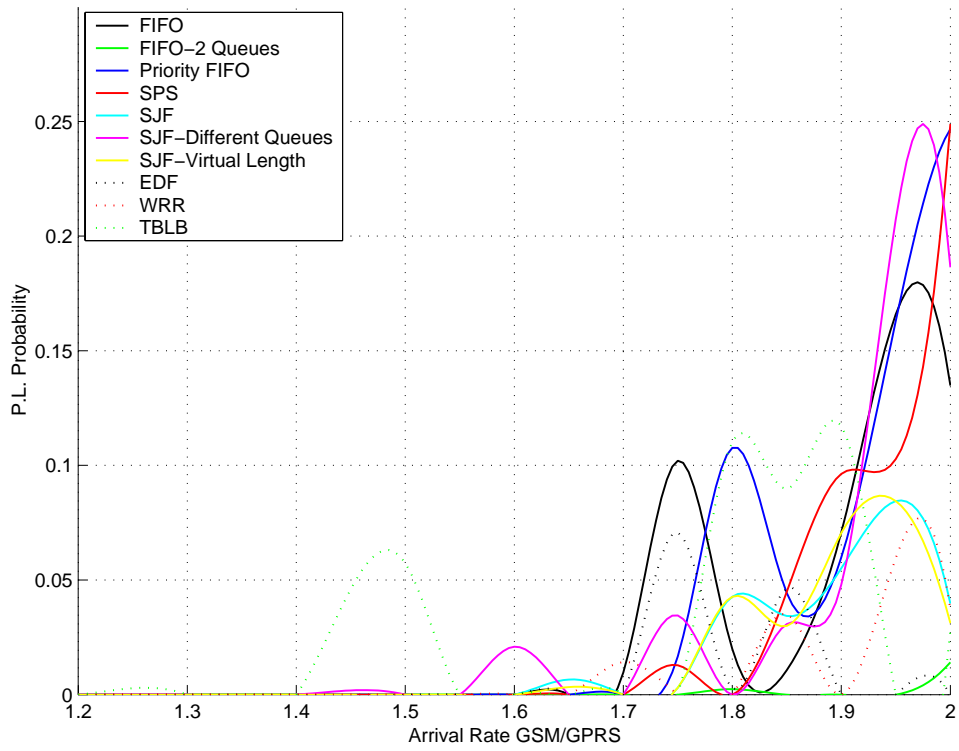
Packet Loss Probability 80% packet 20% session class 1



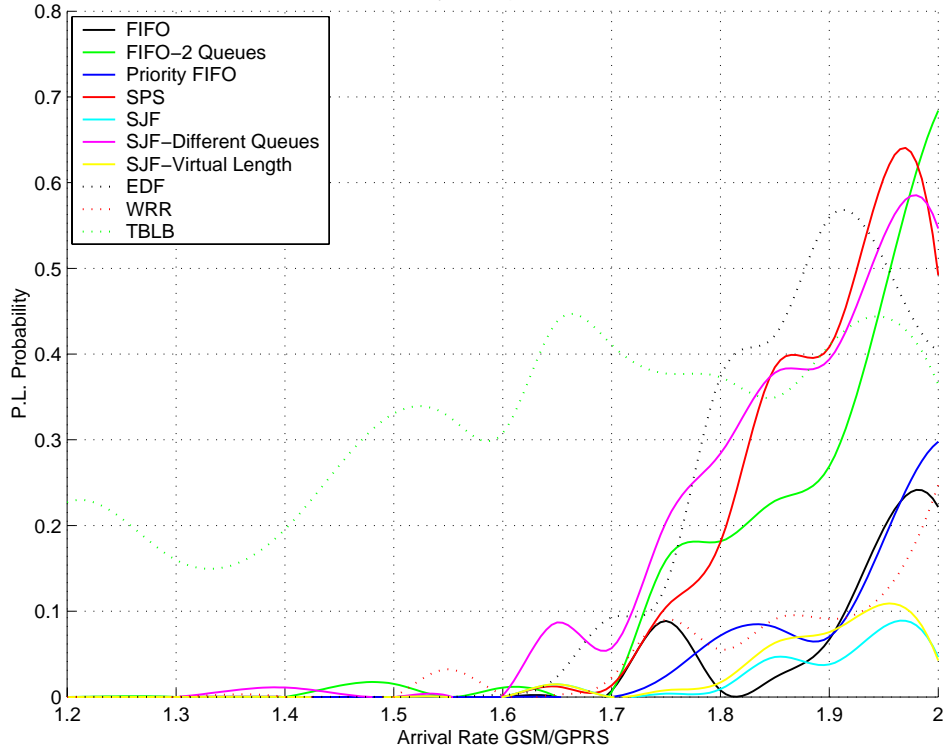
Packet Loss Probability 80% packet 20% session class 2

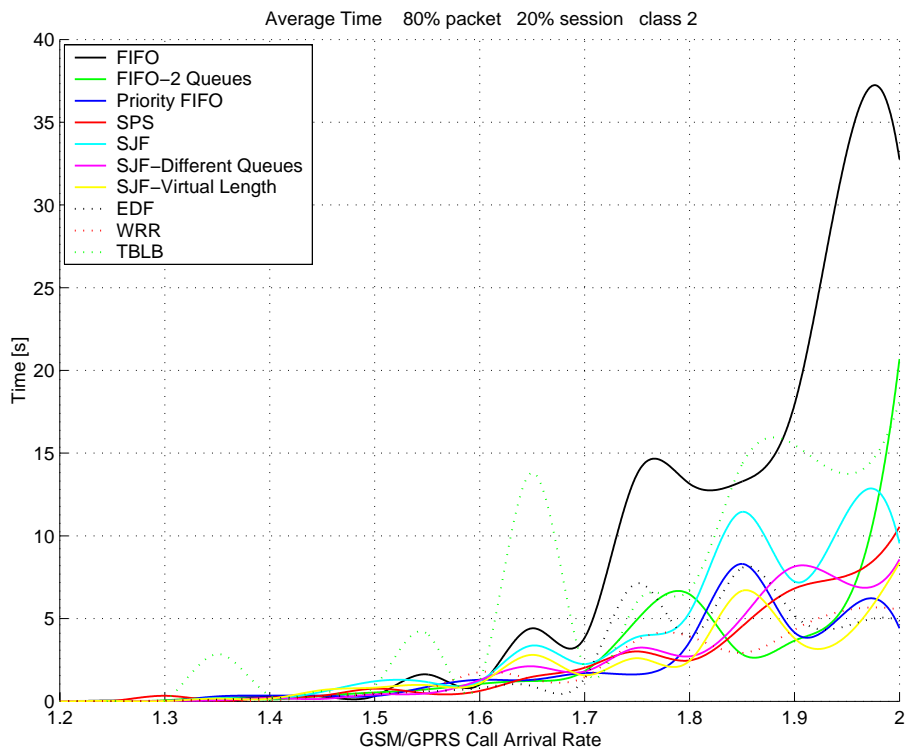
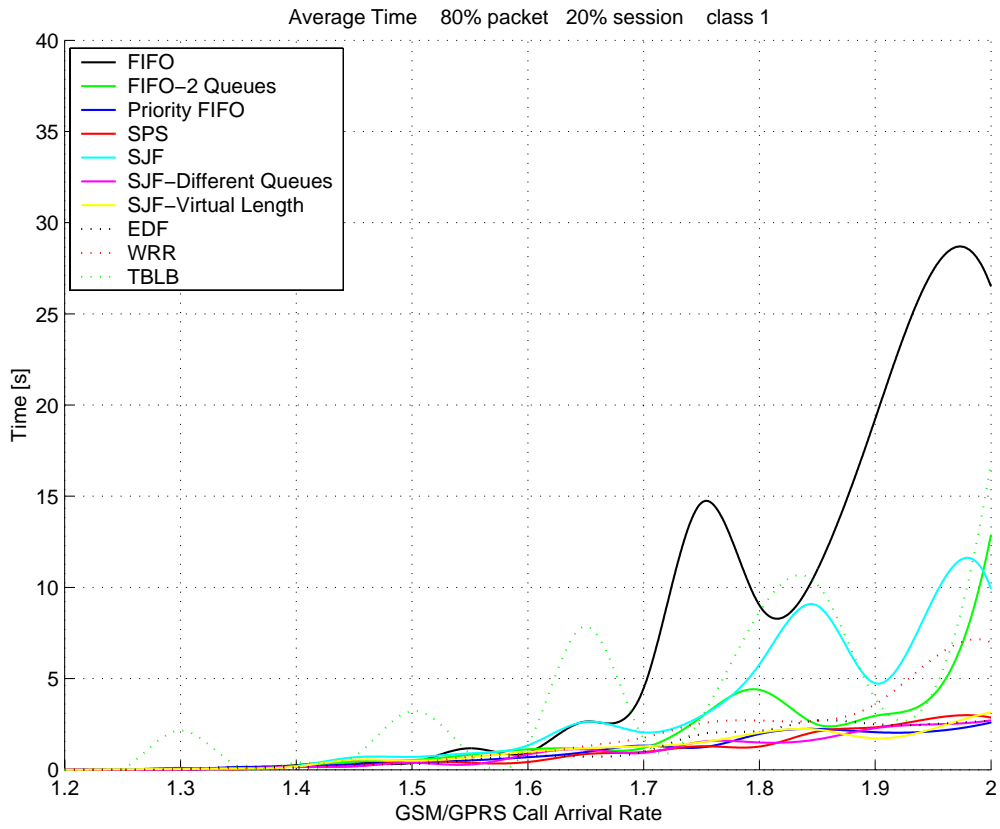


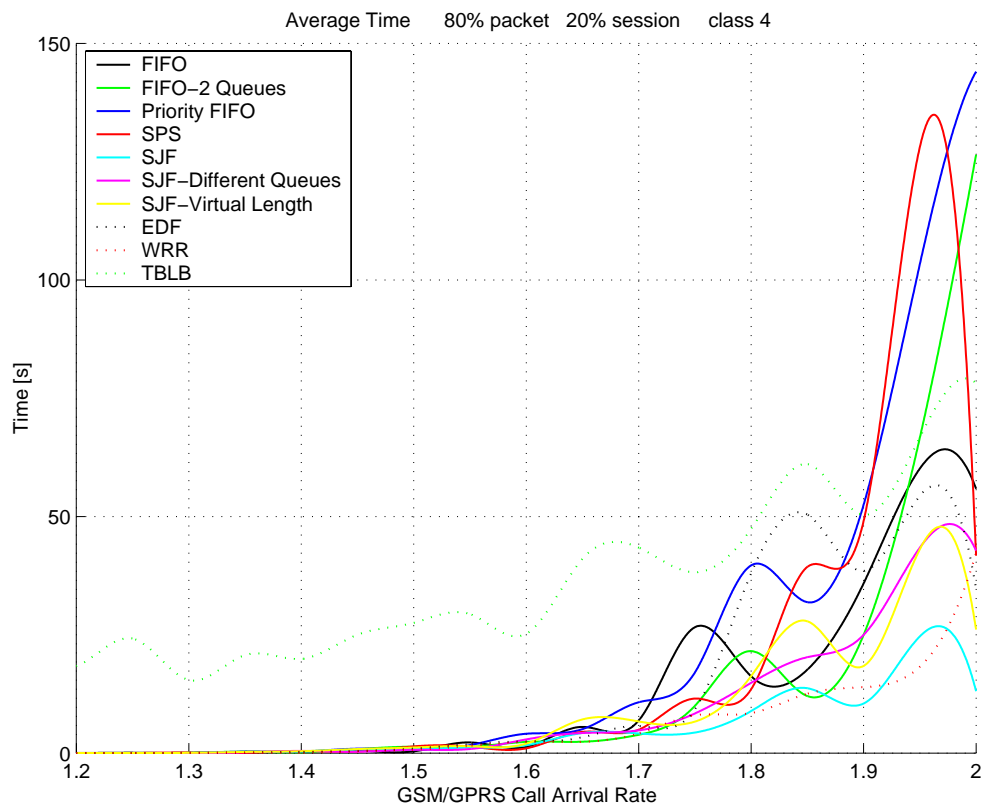
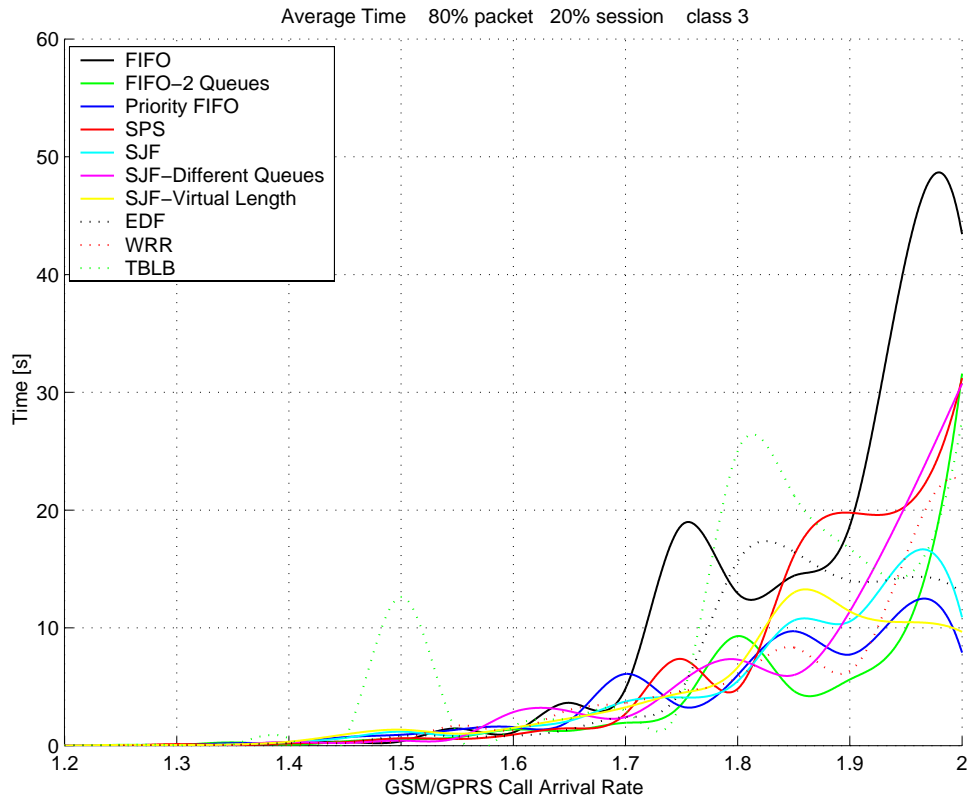
Packet Loss Probability 80% packet 20% session class 3

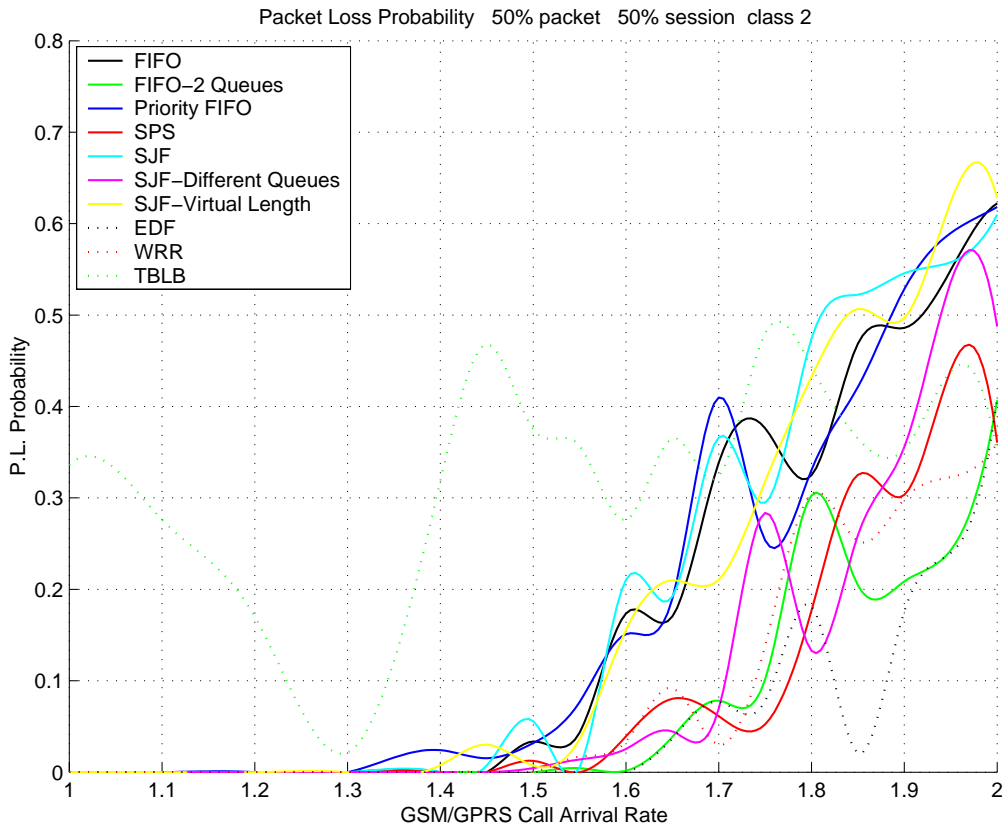
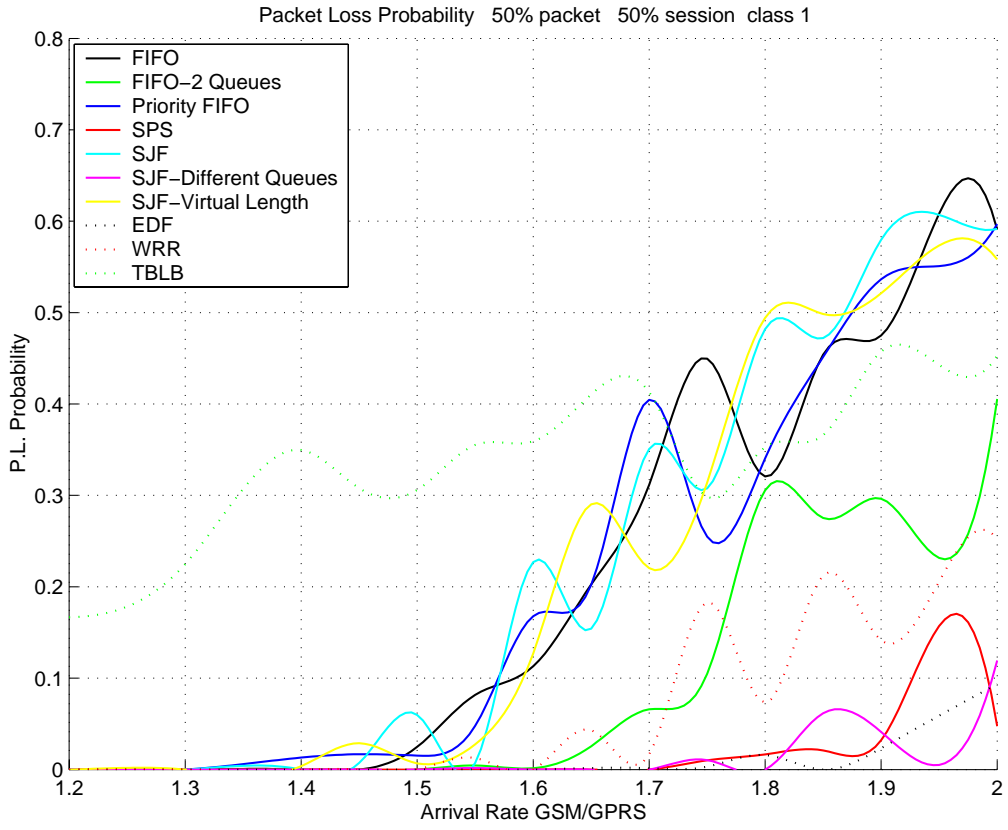


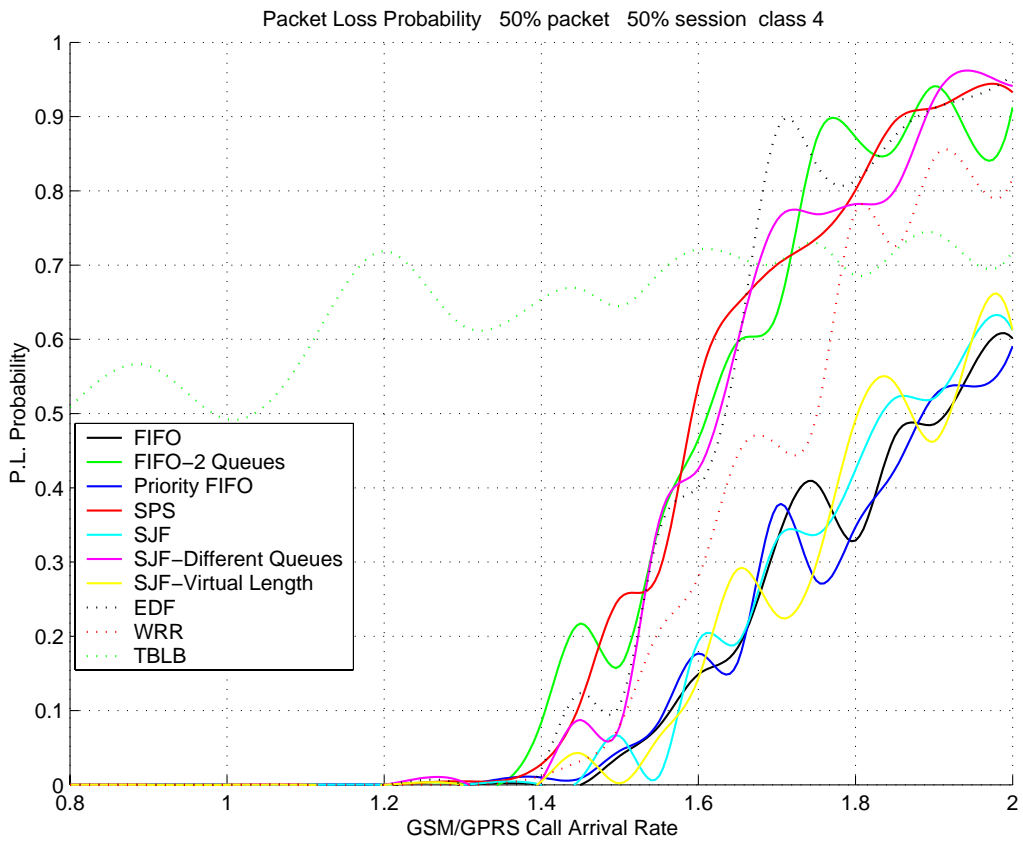
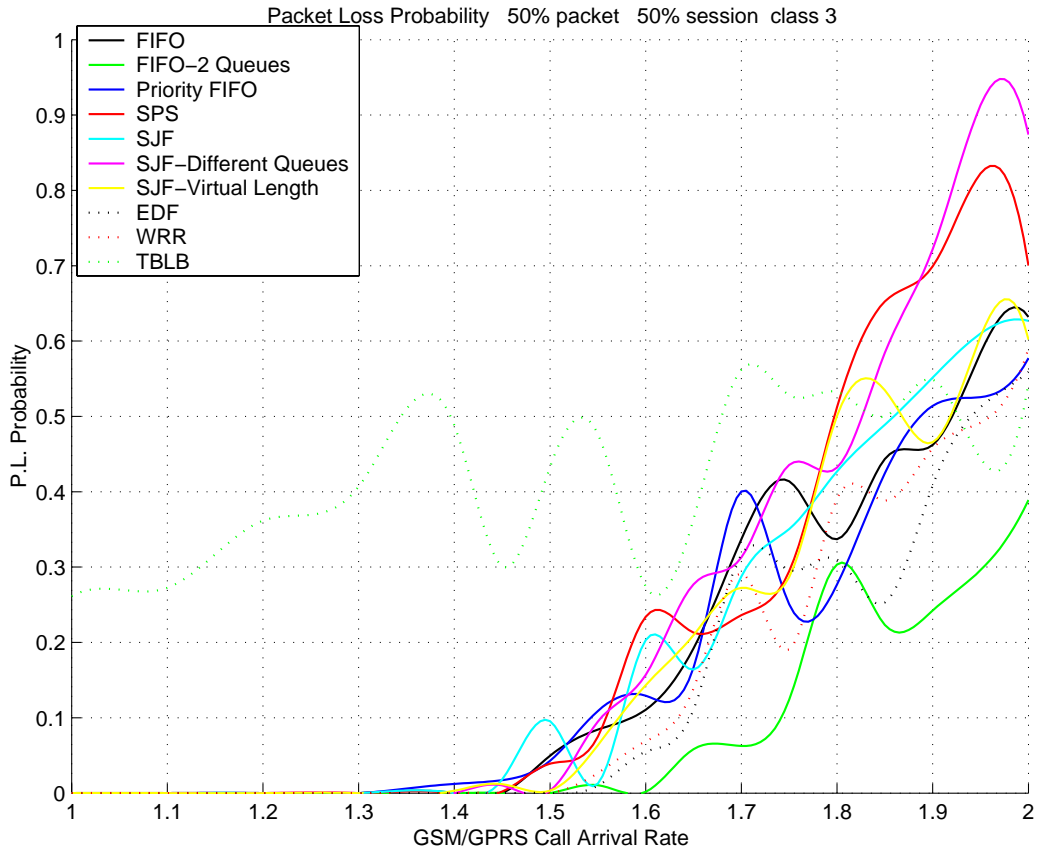
Packet Loss Probability 80% packet 20% session class 4



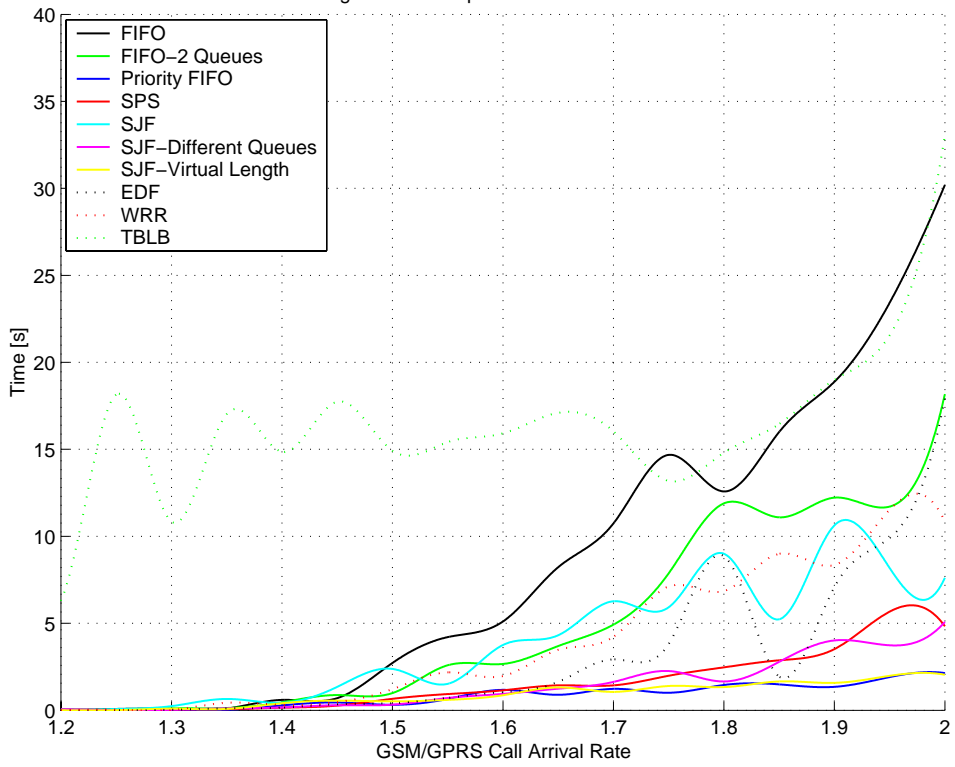




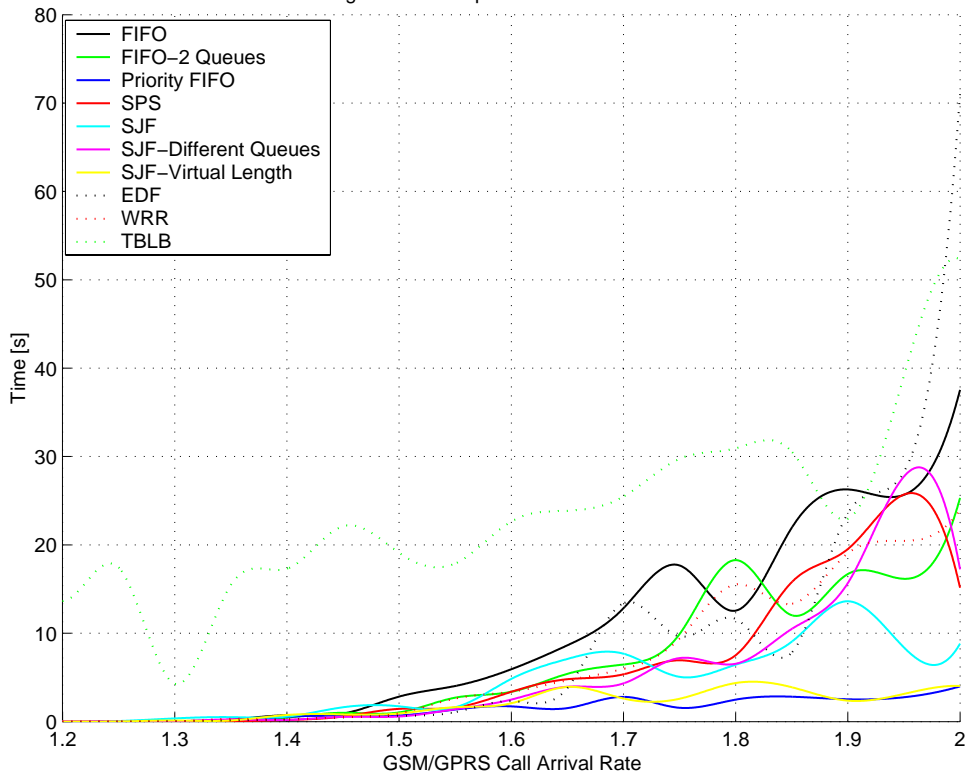


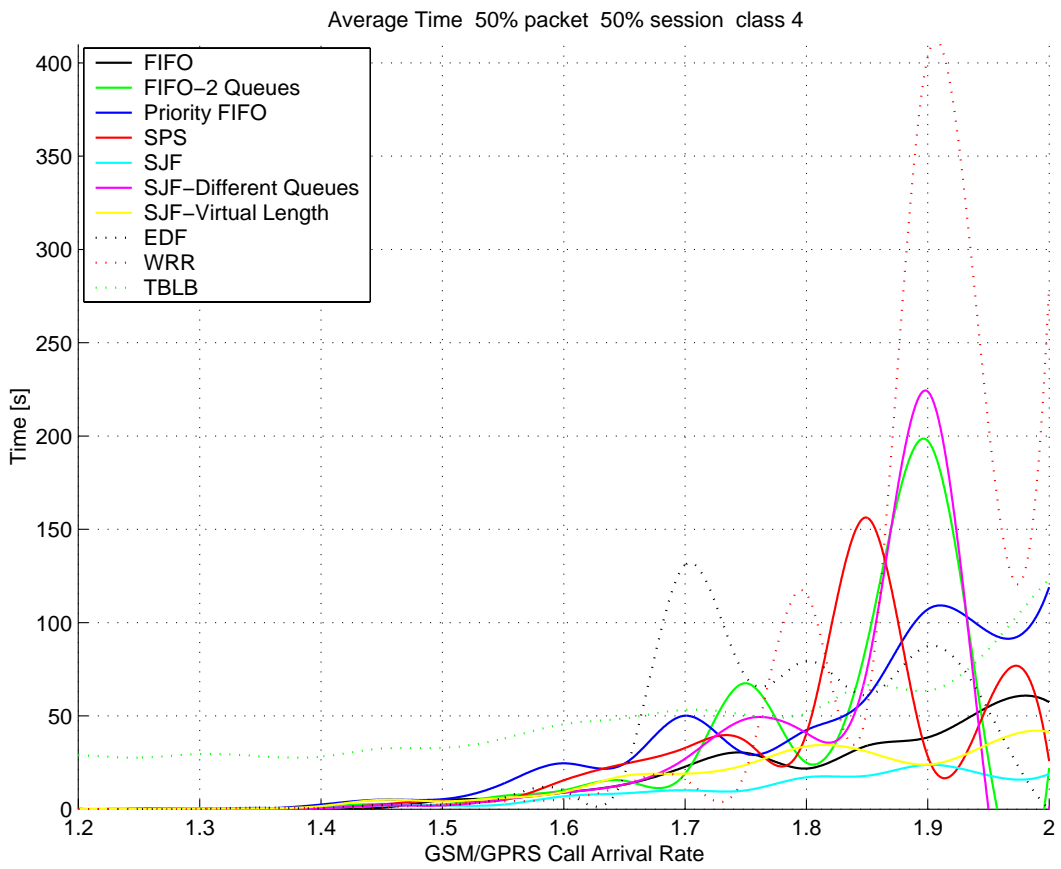
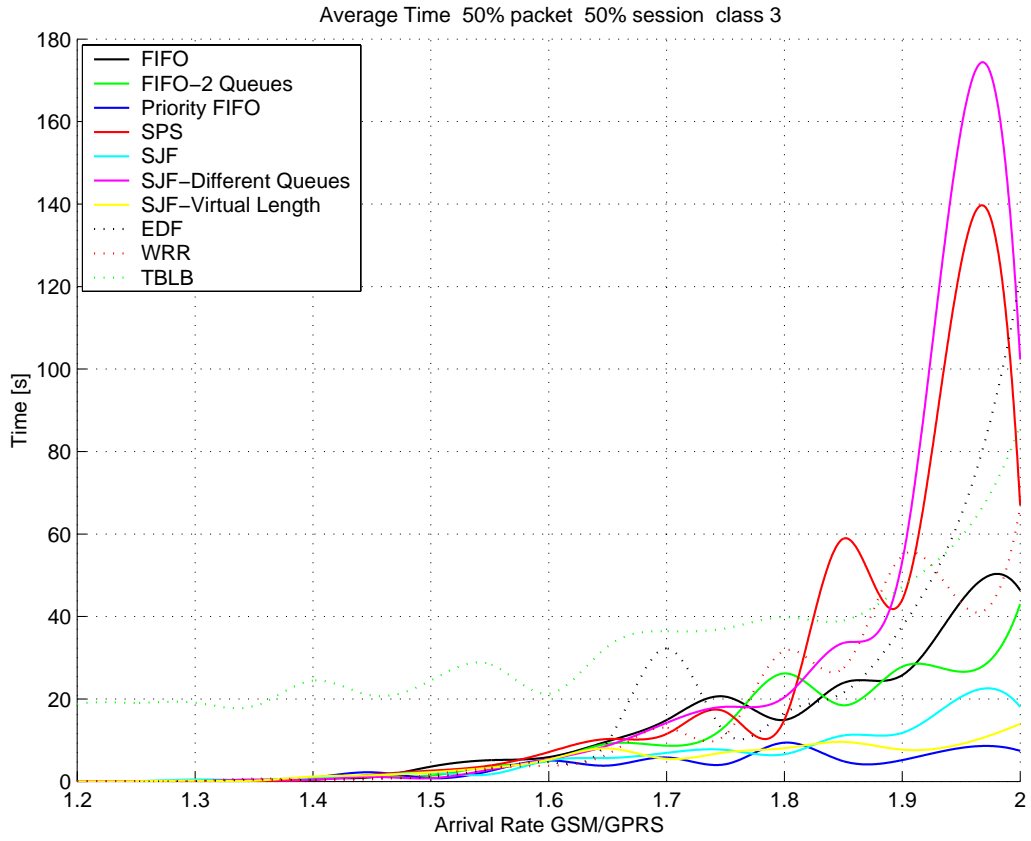


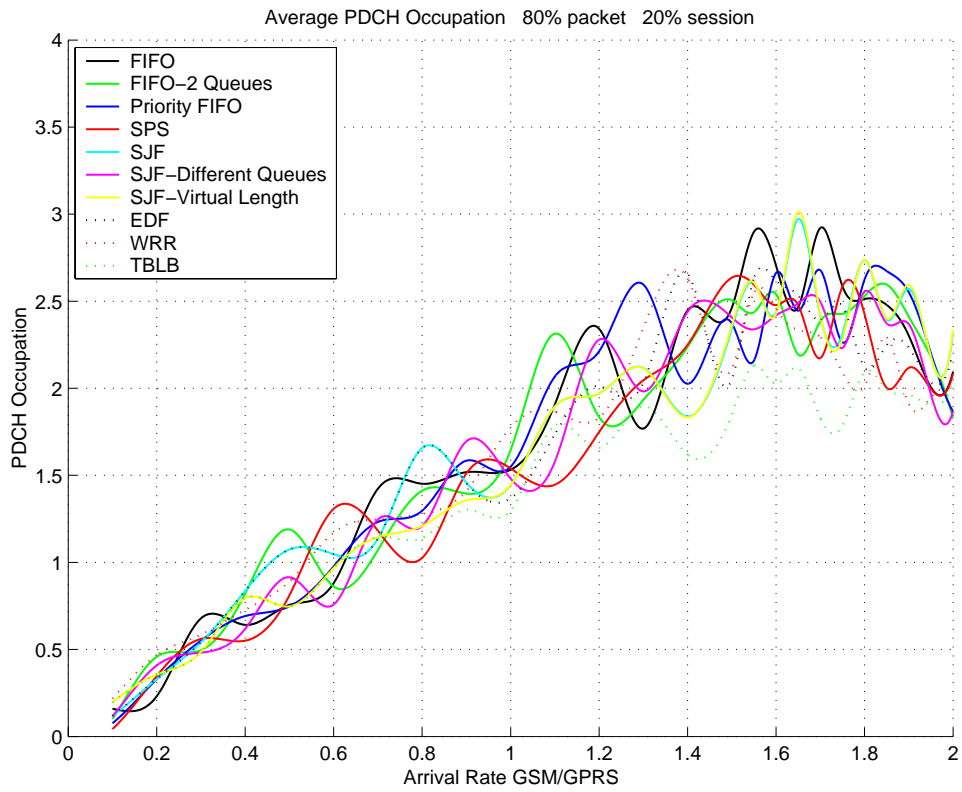
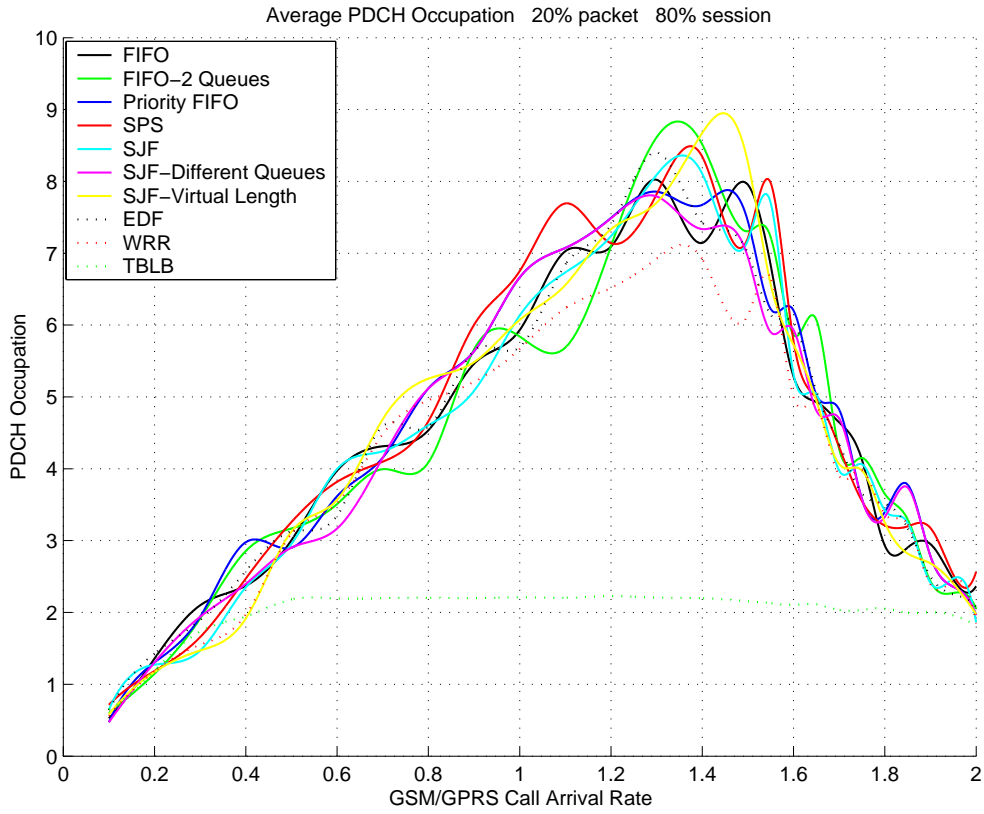
Average Time 50% packet 50% session class 1



Average Time 50% packet 50% session class 2







Average PDCH Occupation 50% packet 50% session

