ABSTRACT

In the Global System for Mobile Communications (GSM) Phase 2+ development, one major activity is to specify the general packet radio service (GPRS). The standardization is expected to be completed at the end of 1997. The purpose of GPRS, like other packet data services, is to efficiently accommodate data sources that are bursty in nature. Another important goal of the technology is to make it possible for GSM license holders to share physical resources on a dynamic, flexible basis between packet data services and other GSM services. In this article the authors present a detailed description of the GPRS system, emphasizing the radio link protocol, and demonstrate its performance. The performance measures have been obtained from the WINLAB GPRS simulator.

General Packet Radio Service in GSM

Jian Cai and David J. Goodman, Rutgers University

GSM (the Global System for Mobile Communications) is the pan-European digital cellular standard published by ETSI (the European Telecommunications Standards Institute). It is the world's most popular second-generation cellular system. In 1997, there are more than 40 million GSM subscribers in more than 100 countries. In addition to cellular operation in the 900 MHz band, GSM technology is used in the 1800 MHz and 1900 MHz personal communications frequency bands. In these bands the technology is referred to as Digital Cellular System at 1800 and 1900 MHz, or DCS1800 and DCS1900, respectively. Technology creation for GSM has proceeded in three phases [1]. The technology standardized in phase 1 was sufficient for the introduction of commercial GSM services, including telephony and short message services, in 1992. In 1996, phase 2 completed the original GSM design task and established a framework for ongoing technology enhancement. GSM standardization is now in phase 2+, which consists of a large number of projects including improved voice coding and advanced data transmission services. Two data services are high-speed circuit-switched data service (HSCSD) and the subject of this article, the general packet radio service (GPRS).

Initial work on GPRS began in 1994 with the published standard expected at the end of 1997. ETSI has produced a sequence of drafts of the GPRS specification. This article is based on the GPRS air interface protocol proposed by the Special Mobile Group (SMG) in March 1996 [2]. Although some further modifications to GPRS have been made [3] and others are expected, the final product will be substantially similar to the protocol described in this article. The purpose of GPRS, like other packet data services, is to efficiently accommodate data sources that are bursty in nature. Another important goal of the technology is to make it possible for GSM license holders to share physical resources on a dynamic, flexible basis between packet data services and other GSM services. Consequently, GPRS shares GSM frequency bands with telephone and circuit-switched data traffic, and makes use of many properties of the physical layer of the original GSM system, most importantly the time-division multiple access (TDMA) frame structure, modulation technique, and structure of GSM time slots.

To evaluate GPRS and study the effects of various operational parameters, WINLAB has developed and validated a GPRS simulator. The purpose of this article is to describe the GPRS protocol and demonstrate its performance, as measured by throughput, delay, and blocking, under various conditions. The performance measures have been obtained from the WINLAB GPRS simulator.

GENERAL GSM CONCEPT

The GSM system architecture includes three standard interfaces: the air interface (Um), the Abis Interface, and the A interface (Fig. 1). The GSM functionalities are divided between the mobile stations (MS), base station subsystem (BSS), and mobile switching center (MSC). The BSS includes two types of elements: the base transceiver station (BTS), which handles the radio interfaces toward the MS, and the base station controller (BSC), which manages the radio resources and controls handovers. A BSC can manage several BTSs. Through the MSC, the GSM system communicates with

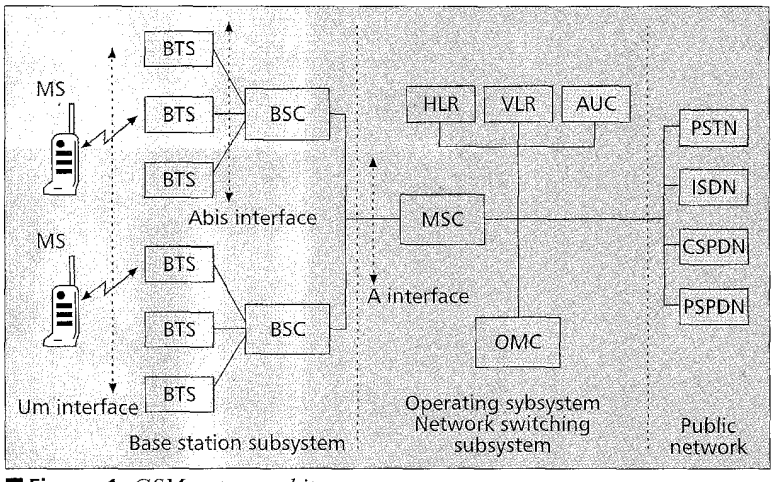


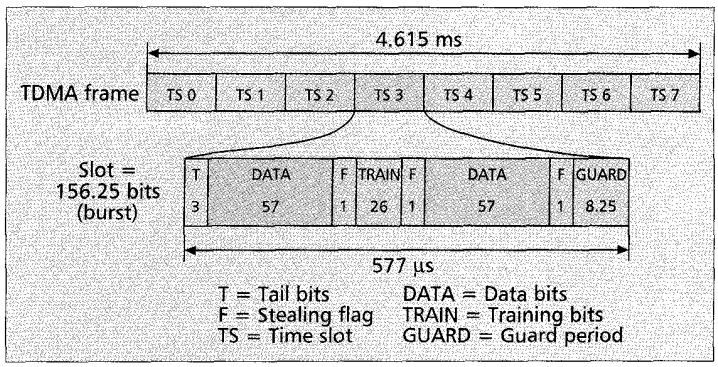
Figure 1. *GSM system architecture.*

0163-6804/97/$10.00 © 1997 IEEE

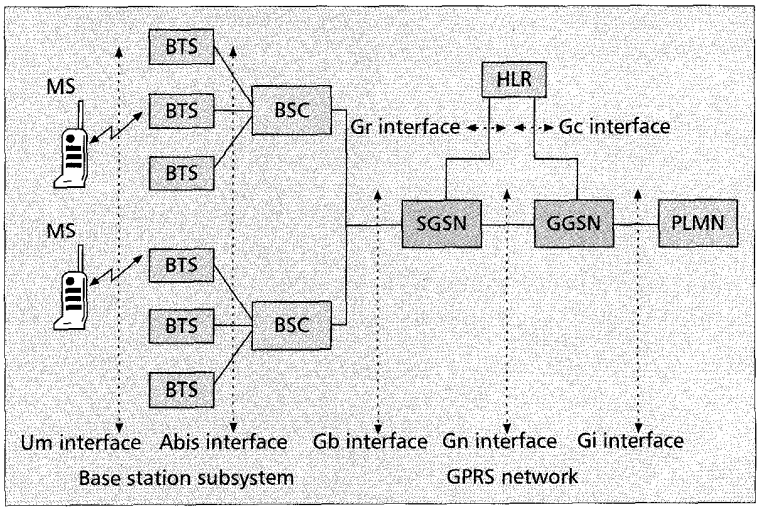
other networks such as the public switched telephone network (PSTN), integrated services digital network (ISDN), circuit-switched public data network (CSPDN), and packet-switched public data network (PSPDN). In addition, GSM specifies three databases, the home location register (HLR), visitor location register (VLR), and authentication center (AUC).

As a hybrid frequency-division/time-division system, GSM organizes radio transmissions by assigning carriers and time slots to logical channels. The frame duration is 4.615 ms, and each frame is divided into eight time slots, as illustrated in Fig. 2.

There are two categories of logical channels in GSM, traffic channels and signaling channels. The signaling channels can be further separated into three groups: *broadcast channels*, *common control channels*, and *dedicated control channels* (Table 1). A base station uses broadcast channels to transmit the same information to all terminals in a cell. Three types of broadcast channels are described in the standard: the broadcast control channel (BCCH), the frequency correction channel (FCCH), and the synchronization channel (SCH). The common control channels carry information to and from specific terminals. There are three different common control channels: the paging channel (PCH) and the access grant channel (AGCH), which are forward link channels, and the random access channel (RACH), which is a reverse link channel. The dedicated control channels use physical channels that are assigned to specific terminals. There are three types of dedicated control channels in GSM. Standalone dedicated control channels (SDCCH) are used for transferring signaling messages between MSs and BSSs when a call is not in progress. The slow- and fast-associated control channels (SACCH and FACCH) are used for transmissions of control messages between the mobile station and the base station during a call.



■ **Figure 2.** *GSM burst structure.*



■ **Figure 3.** *GPRS system architecture.*

General Packet Radio Service

A System Point of View

Figure 3 illustrates the GPRS system architecture [4–6]. Compared to GSM, two new elements (shadowed objects) are introduced in order to create an end-to-end packet transfer mode. In addition, the HLR is enhanced with GPRS subscriber data and routing information. Two services are provided:

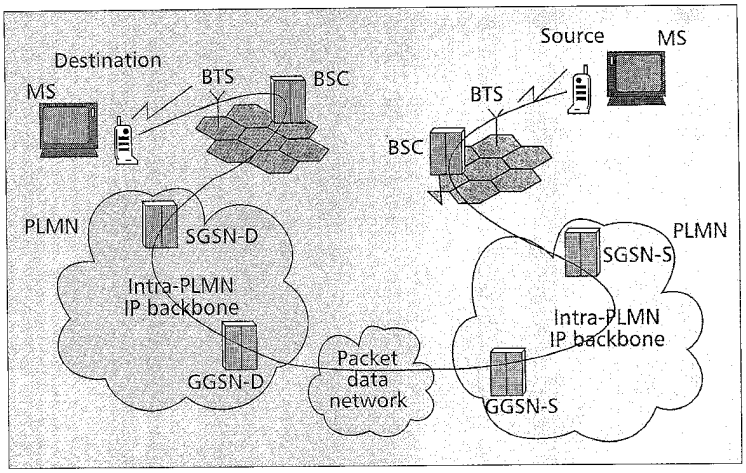
- Point-to-point (PTP)
- Point-to-multipoint (PTM)

Independent packet routing and transfer within the public land mobile network (PLMN) is supported by a new logical network node called the *GPRS support node* (GSN). The gateway GPRS support node (GGSN) acts as a logical interface to external packet data networks. The serving GPRS support node (SGSN) is responsible for the delivery of packets to the MSs within its service area. Within the GPRS network, protocol data units (PDUs) are encapsulated at the originating GSN and decapsulated at the destination GSN. In between the GSNs, the Internet Protocol (IP) is used as the backbone to transfer PDUs. This whole process is defined as *tunneling* in GPRS. The GGSN also maintains routing information used to tunnel the PDUs to the SGSN currently serving the MS. All GPRS user-related data needed by the SGSN to perform the routing and data transfer functionality is stored within the HLR.

Figure 4 shows a simple example of routing in a mobile originated transmission. The serving SGSN of the source mobile (SGSN-S) encapsulates the packets transmitted by the MS and routes them to the appropriate GGSN (GGSN-S). Based on the examination of the destination address, packets are then routed to the GGSN-D through the packet data network. The GGSN-D

Group	Name	Function
Broadcast channels	BCCH	Broadcast control
	FCCH	Frequency correction
	SCH	Synchronization
Common control channels	PCH	Paging
	AGCH	Access grant
	RACH	Random access
Dedicated control channels	SDCCH	Standalone dedicated control
	SACCH	Slow-associated control
	FACCH	Fast-associated control

■ **Table 1.** *GSM signaling channels.*



■ **Figure 4.** *A simple example of routing.*

checks the routing context associated with the destination address and determines the serving SGSN (SGSN-D) and relevant tunneling information. Each packet is then encapsulated and forwarded to the SGSN-D, which delivers it to the destination mobile.

The Protocol Architecture

Figure 5 shows the proposed transmission plane up to the network layer according to the International Organization for Standardization/open systems interconnection (ISO/OSI) reference model. Above the network layer, widespread standardized protocols may be used. The selection of these protocols is outside of the scope of GPRS specification.

Between two GSNs, the GPRS Tunnel Protocol (GTP) tunnels the PDUs through the GPRS backbone network by adding routing information. Below the GTP, the Transmission Control Protocol/User Datagram Protocol (TCP/UDP) and the Internet Protocol (IP) are used as the GPRS backbone network-layer protocols. Ethernet, ISDN, or asynchronous transfer mode (ATM)-based protocols may be used below IP depending on the operator's network architecture.

Between the SGSN and MS, the Subnetwork Dependent Convergence Protocol (SNDCP) maps network-level protocol characteristics onto the underlying logical link control and provides functionalities like multiplexing of network-layer messages onto a single virtual logical connection, encryption, segmentation, and compression.

Radio communication between an MS and the GPRS network, indicated by the shaded area in Fig. 5, covers physical and data link layer functionality.

Between the MS and BSS, the data link layer has been separated into two distinct sublayers: the logical link control (LLC) and radio link control/medium access control (RLC/MAC) sublayers. The LLC layer is the higher sublayer of the data link layer. It operates above the RLC/MAC layer and provides a logical link between the MS and SGSN. Protocol functionality is based on link access procedure-D (LAPD) (see GSM 04.22) used within the GSM signaling plane with support for PTM transmission.

The RLC/MAC layer provides services for information transfer over the physical layer of the GPRS radio interface. It defines the procedures that enable multiple MSs to share a common transmission medium which may consist of several physical channels.

The RLC layer is responsible for the transmission of data blocks across the air interface and the backward error correction (BEC) procedures consisting of selective retransmission of uncorrectable blocks (ARQ, automatic repeat request).

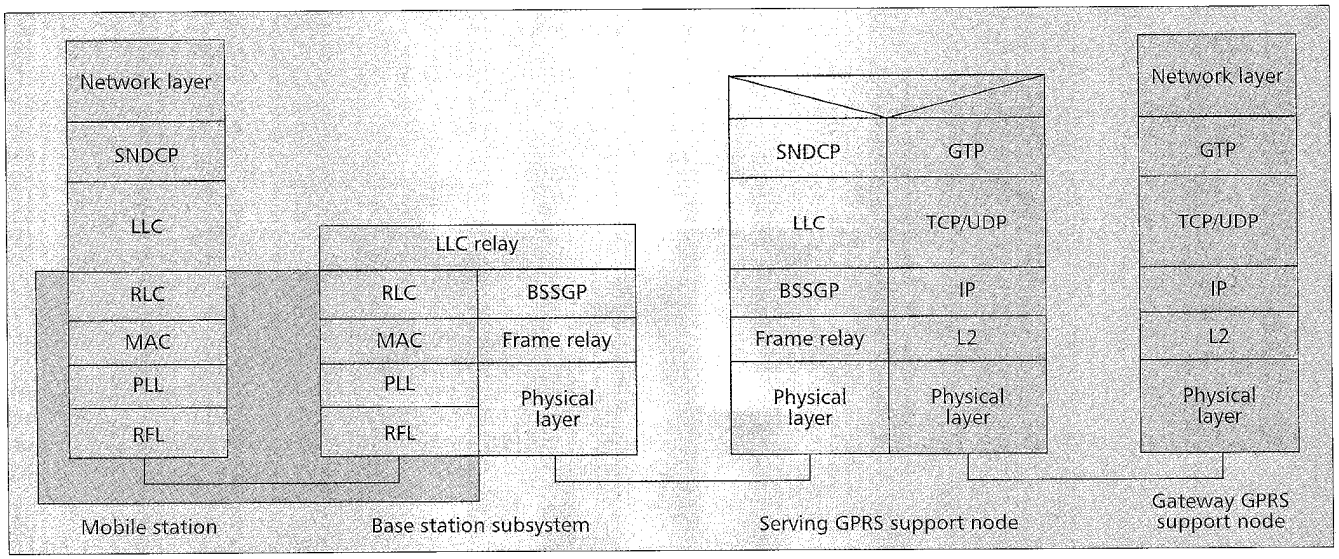
The MAC layer itself is derived from a slotted ALOHA protocol and operates between the MS and BTS. It is responsible for access signaling procedures for the radio channel governing the attempts to access the channel by the MSs, and the control of that access by the network side. It performs contention resolution between channel access attempts, arbitration between multiple service requests from different MSs, and medium allocation to individual users in response to service requests.

The physical layer is split up into a physical link sublayer (PLL) and a physical RF sublayer (RFL).

The PLL provides services for information transfer over a physical channel between the MS and the network. These functions include data unit framing, data coding, and the detection and correction of physical medium transmission errors. The PLL uses the services of the physical RFL.

The PLL is responsible for:

- Forward error correction (FEC) coding allowing detec-



■ **Figure 5.** *GPRS transmission plane.*

tion and correction of errors in transmitted codewords and the signaling of uncorrectable codewords
- Rectangular interleaving of one radio block over four bursts in consecutive TDMA frames (as defined in GSM 05.03)
- Procedures for detecting physical link congestion

The RFL is part of a complete GSM system that delivers a range of services including GPRS. The RFL performs the modulation and demodulation of the physical waveforms and conforms to the GSM 05 series of recommendations which specify, among other things:
- Carrier frequencies and GSM radio channel structures (GSM 05.02)
- Modulation of transmitted waveforms and the raw data rates of GSM channels (GSM 05.04)
- Transmitter and receiver characteristics and performance requirements (GSM 05.05)

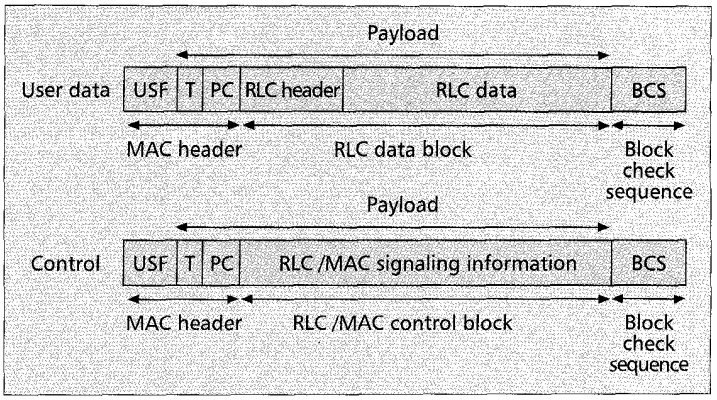
In the network, the LLC is split between the BSS and SGSN. The BSS functionality is called *LLC relay*. Between the BSS and SGSN, the BSS GPRS Protocol (BSSGP) conveys routing and quality of service (QoS)-related information, and operates above frame relay.

Figure 6 shows the radio block structures for user data and control messages. Each radio block consists of a MAC header, an RLC data block or RLC/MAC control block, and a block check sequence (BCS). It is always carried by four normal bursts.

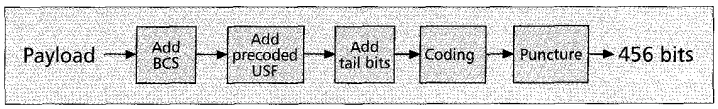
The MAC header consists of uplink state flag (USF), block type indicator (T), and power control (PC) fields. The RLC data block consists of the RLC header and RLC data. The RLC/MAC control block contains the RLC/MAC signaling information elements.

Channel coding is currently specified in GSM 05.03. Four different coding schemes, CS-1 to CS-4, are defined for the radio blocks carrying RLC data blocks. Figure 7 shows the coding procedure, where the payload is defined in Fig. 6. The details of the codes are shown in Table 2.

Between the MS, BSS, and SGSN, the same protocols are used for data transmission up to the SNDCP protocol. At the network layer, a specific mobility management protocol is required within the MS and SGSN.



■ **Figure 6.** *GPRS radio block structure.*



■ **Figure 7.** *GPRS coding procedure.*

GPRS Air Interface Protocol

The GPRS air interface protocol [2, 4, 7] is concerned with communications between the MS and BSS at the physical, MAC, and RLC protocol layers, as indicated by the shadowed area in Fig. 5. The RLC/MAC sublayers allow efficient multiuser multiplexing on the shared packet data channel(s) (PDCH) and utilize a selective ARQ protocol for reliable transmissions across the air interface.

The physical channel dedicated to packet data traffic is called a packet data channel (PDCH). A cell that supports GPRS may allocate one or more shared PDCHs which are taken from the common pool of physical channels available to the cell and otherwise used as traffic channels (TCHs). The allocation of TCHs and PDCHs is done dynamically according to the "capacity-on-demand" principles described below.

Master–Slave Concept — At least one PDCH (mapped on one physical time slot), acting as a master, accommodates packet common control channels (PCCCHs) which carry all necessary control signaling for initiating packet transfer as well as user data and dedicated signaling. The others, acting as slaves, are only used for user data transfer.

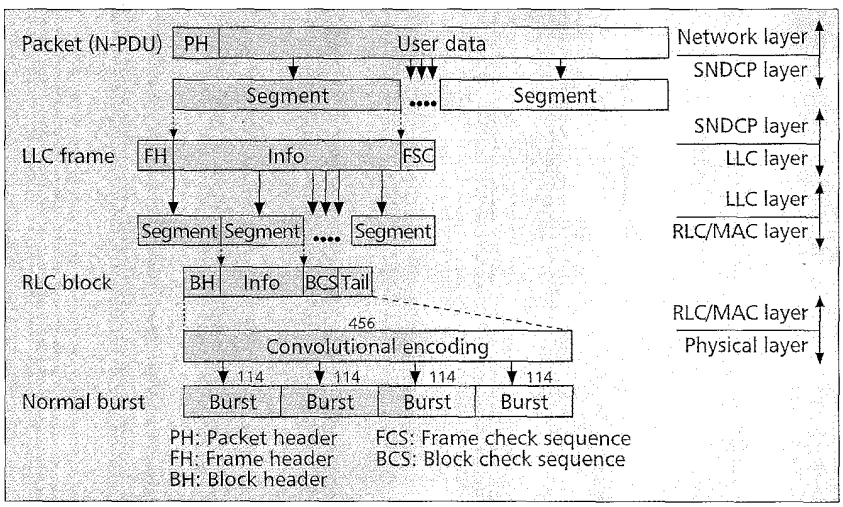
Capacity on Demand — In order to allow GPRS service in cells where there are few (or no) GPRS users without the need for any permanently allocated resources, the concept of capacity on demand has been introduced. The operator can decide whether to dedicate some PDCHs for GPRS traffic. Load supervision is done in the MAC layer to monitor the load on the PDCH(s), and the number of allocated PDCHs in a cell can be increased or decreased according to demand. Unused channels can be allocated as PDCHs to increase the overall QoS for GPRS. If other services with higher priority request resources, deallocation of PDCHs can take place.

However, the existence of PDCH(s) does not imply the existence of PCCCH. When no PCCCH is allocated in a cell, all GPRS-attached MSs automatically camp on the existing GSM CCCH as they do in the idle state. When a PCCCH is allocated in a cell, all GPRS-attached MSs camp on it. The PCCCH can be allocated either as the result of the increased demand for packet data transfer or whenever there are enough available physical channels in a cell. If the network releases the PCCCH, the MSs return to the CCCH.

Multiframe Structure for PDCH — A multiframe structure is needed for the PDCH in order to accommodate paging groups and possibly blocks for broadcasting

Scheme	Code rate	Payload	BCS	Pre-coded USF	Tail bits	Coded bits	Punct. bits	Data rate (kb/s)
CS-1	1/2	181	40	3	4	456	0	9.05
CS-2	≈ 2/3	268	16	6	4	588	132	13.4
CS-3	≈ 3/4	312	16	6	4	676	220	15.6
CS-4	1	428	16	12	0	456	0	21.4

■ **Table 2.** *GPRS coding schemes.*



■ **Figure 8.** *Packet transformation data flow.*

GPRS system information. The multiframe structure of both 51 TDMA frames and 52 TDMA frames are specified in GSM 05.01.

Data Flow

The network-layer protocol data units (N-PDUs or packets) received from the network layer are transmitted across the air interface between the MS and the SGSN using the LLC protocol. First, the SNDCP transforms packets into LLC frames. The process includes optional header/data compression, segmentation, and encryption. The maximum amount of user data in an LLC frame is 2 kbytes[1] in our study [2].

An LLC frame is then segmented into RLC data blocks which are formatted into the physical layer. Each block comprises four normal bursts in consecutive TDMA frames. Figure 8 summarizes the data flow in GPRS.

GPRS Logical Channels

Table 3 lists the GPRS logical channels and their functions. A detailed description for each channel is presented below.

Packet Broadcast Control Channel — The PBCCH transmits system information to all GPRS terminals in a cell.

Packet Common Control Channel — The *packet random access channel* is used by MSs to initiate packet transfers or respond to paging messages. On this channel MSs transmit access bursts with long guard times. On receiving access bursts, the BSS assigns a timing advance to each terminal.

The *packet paging channel* is used to page an MS prior to downlink packet transfer.

The *packet access grant channel* is used in the packet transfer establishment phase to send resource assignment to an MS prior to the packet transfer.

The *packet notification channel* is used to send a PTM-multicast (PTM-M) notification to a group of MSs prior to a PTM-M packet transfer. The notification has the form of a resource assignment for the packet transfer.

Packet Traffic Channels — The *packet data transfer channel* is a channel allocated for data transfer. One MS may use more than one PDTCH in parallel (multislot operation) for individual packet transfers.

The *packet associated control channel* is used to convey signaling information related to a given MS such as acknowledgments (ACK) and power control (PC) information. It also carries resource assignment and reassignment messages, either for allocation of a PDTCH or further occurrences of a PACCH. One PACCH is associated with one or several PDTCHs concurrently assigned to one MS.

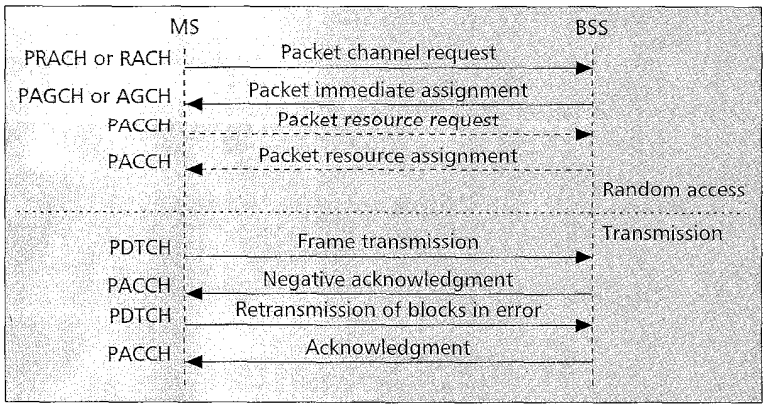
Model of Operation

The RLC/MAC layer protocol proposed is a bitmap-based selective ARQ type protocol with a slotted ALOHA random-access-based packet reservation mechanism for uplink transmission. High flexibility is ensured by implementing multislot MAC-layer operation. More than one PDCH (each corresponding to a time slot in a TDMA frame) can be used by one MS for packet data transfer.

The GPRS radio interface consists of asymmetric and independent uplink and downlink channels. In a certain TDMA time slot, a PDCH uplink may carry data from one MS and the downlink data to another MS.

A packet that is to be transmitted over the air interface is formatted into one or more LLC frames, which are then segmented into RLC data blocks. A selective ARQ protocol between the MS and BSS provides retransmission of blocks in error. When a complete frame is successfully transferred across the RLC layer, it is forwarded to the LLC layer.

The method of implementing selective ARQ on the RLC level requires frame numbering. When LLC frames are segmented, a temporary frame identity (TFI) is assigned to each data frame transmitted to/from an MS. The assigned TFI is unique among concurrent frame transfer sequences in a cell and basically replaces the combination of a complete MS identity (e.g., temporary logical link identity, TLLI) and a frame sequence number. The TFI is included in every block belonging to a particular frame, including retransmitted blocks determined by the ARQ protocol. Furthermore, blocks belonging to frames to/from different MSs can be multiplexed on the downlink/uplink based on the TFI.



■ **Figure 9.** *MAC layer: random access and transmission for uplink data transfer.*

[1] *In the most recent version of GPRS standard [3], this limit is changed to 1600 bytes (discussed later).*

MOBILE-ORIGINATED PACKET TRANSFER

Multiple Access — An MS initiates a packet transfer by making a packet channel request on the PRACH or RACH. The network responds on PAGCH or AGCH, respectively. It is possible to use a one- or two-phase packet access method. Figure 9 illustrates the uplink data transfer procedure. The dashed arrows are the additional messages needed for the optional procedure in the two-phase packet access method.

In one-phase access, the network responds to the packet channel request with the packet immediate assignment, reserving the resources on PDCHs for uplink transfer of a number of radio blocks.

In two-phase access, the network responds to the packet channel request with the packet immediate assignment, which reserves the uplink resources for transmitting the packet resource request. The packet resource request message carries the complete description of the requested resources for the uplink transfer. Thereafter, the network responds with the packet resource assignment, reserving resources for the uplink transfer.

If there is no response to the packet channel request within a predefined time period, the MS retries after a random backoff time.

Packet data traffic is bursty in nature. Momentarily, the BSS will receive more packet channel requests than it can serve within a certain time limit. However, on the mobile side, if the timer expires before a packet resource assignment is received, the MS may contend again even though its last packet channel request was already correctly received. This could lead to a new wave of packet channel requests. The proposed solution to avoid this is to notify the sender of the packet channel request that its message is correctly received and that it will receive a packet resource assignment later. In this way, the system builds a queue of MSs which wait for their turn to receive a packet resource assignment to send a frame.

Uplink Data Transfer — Efficient and flexible utilization of the available spectrum for packet data traffic (one or more PDCHs in a cell) can be obtained using a multislot channel reservation scheme. Blocks from one MS can be sent on different PDCHs simultaneously, thus reducing the packet delay for transmission across the air interface. The bandwidth may be varied by allocating one to eight time slots in each TDMA frame depending on the number of available PDCHs, the multislot capabilities of the MS, and the current system load.

As mentioned earlier, the master-slave channel concept requires mechanisms for efficient utilization of PDCH uplink(s). Therefore, the uplink state flag (USF) is used on PDCHs. The 3-bit USF at the beginning of each radio block sent on the downlink points to the next uplink radio block. It enables the coding of eight different USF states which are used to multiplex the uplink traffic. The channel reservation command includes the list of allocated PDCHs and the corresponding USF state per channel. To an MS, the USF marks the part of the channel it can use for transmission. An MS monitors the USF and, according to the USF value, identifies PDCHs assigned to it and starts transmission. This allows efficient multiplexing of blocks from a number of MSs on a single PDCH. Additionally, the channel reservation command can be sent to the MS even before the total number of requested PDCHs is free. Thus, the status flags not only result in a highly dynamic reservation but also allow interruption of transmission in favor of pending or high-priority messages. One USF value is used to denote PRACH (`USF=FREE`). The other USF values, `USF=R1/R2/...R7`, are used to reserve the uplink for different MSs.

After the blocks have been transmitted in the reserved time slots, an acknowledgment should follow from the BSS and sent on the PACCH.

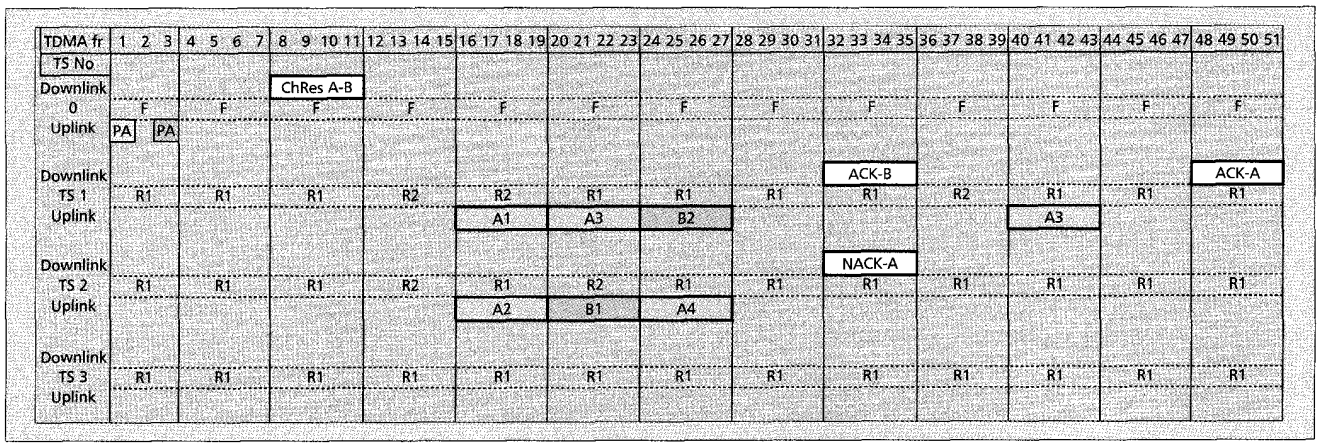
In the case of a negative acknowledgment, which includes a bitmap of erroneous blocks, a packet resource assignment for retransmission, timing advance and power control, only those blocks listed as erroneous are retransmitted.

If the MS does not receive an ACK within a certain time, a frame transfer recovery is started by sending a new packet channel request for one block. If a packet resource assignment is granted, the first block of the current frame will be sent to the BTS. Thus, the BTS knows that the last ACK belonging to this frame transfer was not correctly received and should be retransmitted.

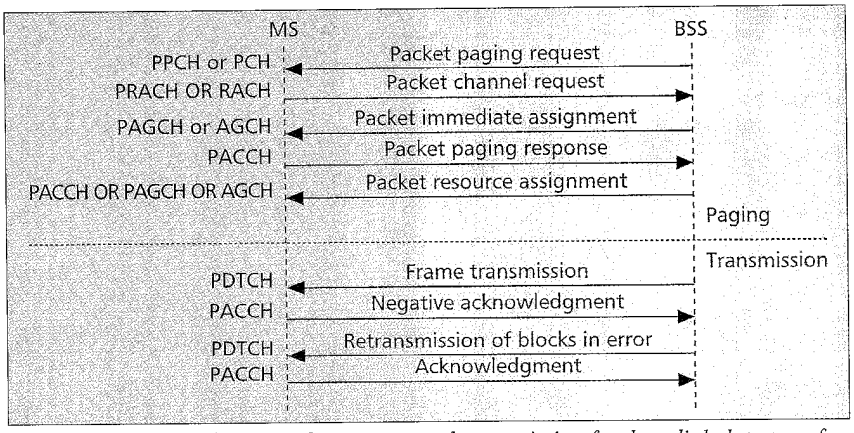
Figure 10 gives an example of the multislot and multiplexing schemes of the GPRS system. Four PDCHs are

Group	Name	Direction	Function
PBCCH	PBCCH	Downlink	Broadcast
PCCCH	PRACH	Uplink	Random access
	PPCH	Downlink	Paging
	PAGCH	Downlink	Access grant
	PNCH	Downlink	Multicast
PTCH	PDTCH	Downlink and uplink	Data
	PACCH	Downlink and uplink	Associated control

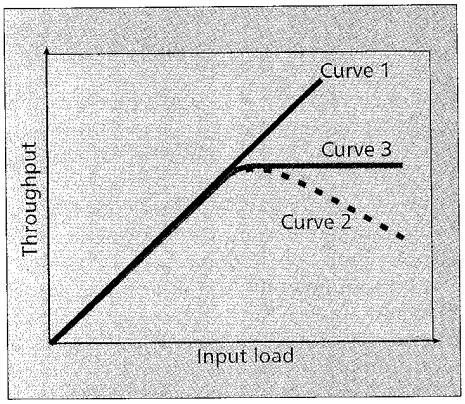
■ **Table 3.** *GPRS logical channel.*



■ **Figure 10.** *Message sequence across the air interface for mobile-originated traffic.*



■ **Figure 11.** *MAC layer: random access and transmission for downlink data transfer.*



■ **Figure 12.** *Throughput vs. input load in a multi-access system.*

available in this cell, where TS 0 is used as PRACH and TSs 1–3 are used as PDTCHs. An R1/R2 USF is used on the PDTCHs.

Mobile Terminated Packet Transfer

A BSS initiates a packet transfer by sending a packet paging request on the PPCH or PCH downlink. The MS responds to the page by initiating a procedure for page response very similar to the packet access procedure described earlier. The paging procedure is followed by the packet resource assignment for downlink frame transfer containing the list of PDCHs to be used. Figure 11 illustrates the paging and packet transfer procedures.

Since an identifier (e.g., TFI) is included in each radio block, it is possible to multiplex radio blocks destined for different MSs on the same PDCH downlink. It is also possible to interrupt a data transmission to one MS if a higher-priority data or pending control message is to be sent to some other MS. Furthermore, if more than one PDCH is available for the downlink traffic, and provided the MS is capable of monitoring multiple PDCHs, blocks belonging to the same frame can be transferred on different PDCHs in parallel.

The network obtains acknowledgments for downlink transmission by polling the MS. The MS sends the ACK/NACK message in the reserved radio block which is allocated in the polling process. In the case of a negative acknowledgment, only those blocks listed as erroneous are retransmitted.

Numerical Results

Using OPNET, we have created a computer simulation of GPRS in order to examine the effects of protocol performance of various protocol parameters and operating conditions.

Our simulation study of GPRS focuses on three well-known performance measures, defined as follows.

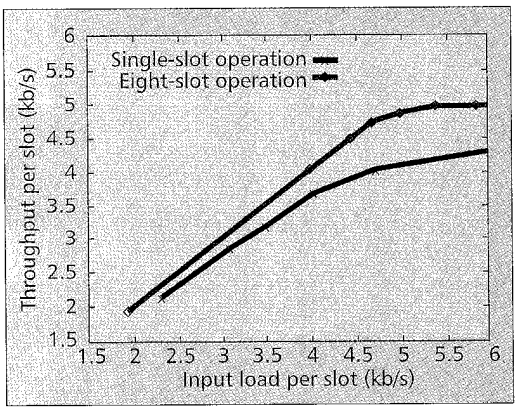
- Throughput — the amount of error-free user data that reaches the destination (kb/s)
- Delay — the time in seconds from the arrival of a frame at the source until the whole frame is correctly received at the destination
- Blocking — the fraction of random access attempts that fail because either the number of contentions or the random access time exceeds a limit

We examine each measure as a function of input load (kb/s). Figure 12 shows that in a multi-access system, the graph of throughput as a function of input load has three possible shapes.

When a multi-access system has unlimited capacity, the throughput always equals the input load (curve 1, Fig. 12). In a slotted ALOHA model

Parameter	Default Value
Traffic Model	Railway
FEC Coding Rate	1/2
Number of Mobiles	20
Number of PDCH	8
Channel SIR producing transmission errors	Log-normal with mean 16.1 dB standard deviation 7 dB
Max. Time Waiting for a Res_Asgn	0.1 sec
Max. Time for Random Access Attempts per frame	5 sec
Max. Number of Random Access Attempts per frame	7
Max. Time Waiting for an Acknowledgment	0.1 sec
Max. Number of Retransmissions per frame	10
The Response Time at the BS for a Channel Request	18.5 ms[1]
The Response Time at the BS for an Acknowledgment	18.5 ms[1]
The Response Time at the MS for a Resource Assignment	18.5 ms[1]
[1] 4 TDMA frames	

■ **Table 4.** *System parameters used in simulation.*



■ **Figure 13.** *Multislot vs. single-slot operation: throughput comparison.*

where a packet is a short burst, the throughput vs. input load is indicated by curve 2 in Fig. 12. When the system is stable (before it reaches the maximum throughput), the throughput equals the input load when no dropping occurs. When the system is overloaded, most of the packets collide during the random access process and remain backlogged. As the load increases, more packets collide, and fewer packets go through. Thus, the system throughput decreases as the load increases.

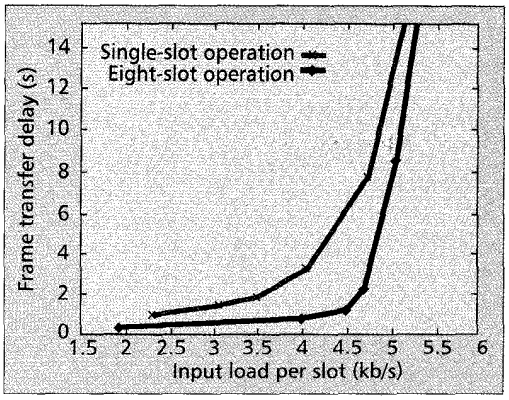


Figure 14. *Multislot vs. single-slot operation: delay comparison.*

The GPRS system has the characteristics of a slotted ALOHA system but with several distinguishing features:

- With the queued random access requests, and packet capture, the random access process is not a bottleneck in GPRS.
- At high input loads, the channel reservation mechanism and long packets create enough traffic to keep the channel fully utilized.

Therefore, the throughput of the GPRS system is expected to behave like curve 3 in Fig. 12. When the system is stable, it performs as a slotted ALOHA system. When the system is overloaded, the throughput saturates at a maximum value.

Table 4 defines a nominal GPRS configuration as defined by ETSI [8, 9]. The Railway traffic model includes packet sizes that conform to a truncated exponential distribution with mean value of 170 bytes. ETSI also defines a FUNET (Finnish University Network) traffic model with packet sizes conforming to a truncated Cauchy distribution (mean value 800 bytes), and a Mobitex model with a uniform packet size distribution (mean value 30 bytes). The section on traffic model effects shows that the efficiency of GPRS depends strongly on the nature of the data traffic model. The simulation results presented in the following sections use the parameters in Table 4 unless stated otherwise. In many cases, we vary one parameter of the system and hold the others fixed at the values shown in Table 4.

Multislot Operation Effects

GPRS is different from the original GSM specification in the sense that it allows a single MS to transmit data simultaneously on multiple time slots.

Here we compare two types of operation.

- When there is only one packet data channel (PDCH) available in the cell, it is considered to be a *single-slot operation*. This master PDCH (MPDCH) supports both data traffic and random access.
- When there are up to eight PDCHs available in a cell, and the MSs are able to operate on eight slots simultaneously, *eight-slot operation* occurs. Channel 0 is used as the MPDCH and supports both control signaling and data transmission. The other channels are used as slave PDCHs (SPDCHs) and support only data traffic.

Figure 13 shows the throughput comparison between single-slot and eight-slot operation. Figure 14 illustrates the delay comparison between the two cases. To compare the efficiency of channel utilization, the overall input load and throughput are divided by the number of slots used. The base station is capable of capture, and both uplink and downlink errors are included.

In single-slot operation, as the load exceeds 4 kb/s per slot, the throughput reaches the maximum value of 4 kb/s. The delay becomes unbounded at this point. In eight-slot operation, when the input load increases, the maximum throughput reaches 5 kb/s per slot, 40 kb/s total. The delay explodes when the input load reaches 5 kb/s per slot. In single-slot operation, the maximum throughput is lower because the channel used by GPRS has to handle both traffic and control information. With multiple slots, the additional channels only have to carry traffic.

We also observe from Fig. 14 that when the system is stable, the delay in eight-slot operation is lower than that of the single-slot operation.

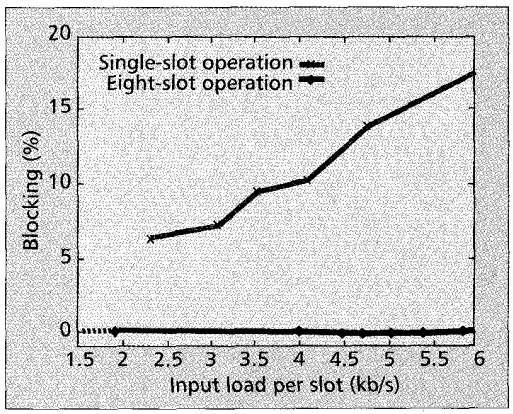


Figure 15. *Multislot vs. sisngle-slot operation: blocking comparison.*

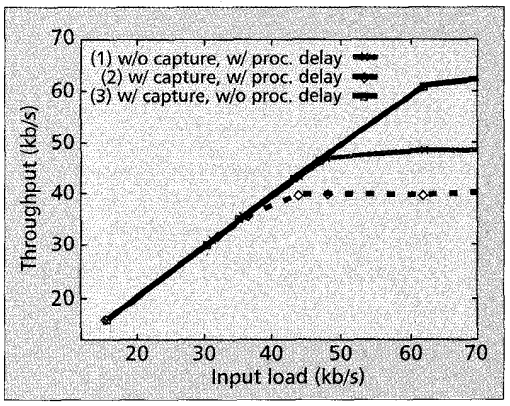


Figure 16. *Capture and processing delay effect: throughput comparison.*

Figure 15 shows the blocking comparison between the two cases. Generally, blocking increases as the load goes up. For eight-slot operation, the blocking rate is very low (less than 0.1 percent). The blocking remains the same even when the delay becomes intolerable. On the other hand, with single-slot operation, blocking becomes a serious problem. When the input load is as low as 2.5 kb/s per slot, the blocking is already more than 5 percent. As the load increases, the blocking goes higher.

To summarize the comparison of single-slot and eight-slot operation, we recall that throughput is defined in terms of the user data that is received successfully. In single-slot operation, part of the channel is used for random access, so there is less user data transmitted per unit time, and the data needs to wait longer to be served. With the service rate lower in the single-slot case, the delay is longer, there is more blocking, and the maximum throughput is lower.

Capture and Processing Delay Effects

It is well known that capture enhances the performance of ALOHA systems ([10]). Slotted ALOHA is embedded in the random access process of the GPRS system. It is interesting to know whether the general statement still holds up in GPRS. Also, in the real system, processing delay is a factor that cannot be ignored; but how much effect does it have in GPRS?

To study those effects on the GPRS system, we conducted the following simulations, assuming that there are no transmission errors:

- Model 1, without capture and with the processing delay
- Model 2, with capture and with the processing delay
- Model 3, with capture and without the processing delay

The capture model is defined in [8]. Three types of processing delay are introduced here as the response time for channel request (random access), resource assignment, and ACK messages. (See the last three rows of Table 4).

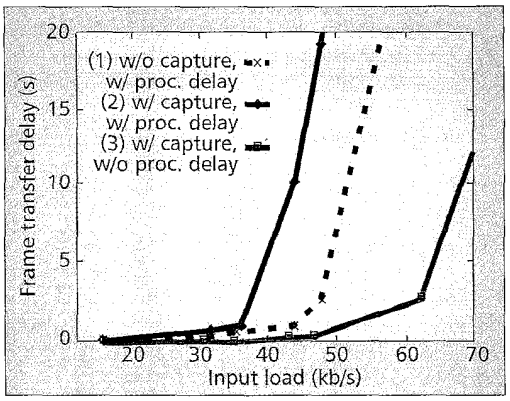
The simulation results are shown in Figs. 16–18.

Figure 16 shows the throughput as a function of input load. In all three cases, as the load increases, the throughput increases and then saturates at the maximum value. With capture, without processing delay and without transmission errors, the GPRS system can reach a maximum throughput of approximately 64 kb/s assuming the Railway traffic model (truncated exponential packet length distribution). When introducing processing delay, the maximum throughput reaches approximately 48 kb/s, which is a 25 percent drop. Keeping the processing delay and error-free operation, the throughput drops to 40 kb/s without capture.

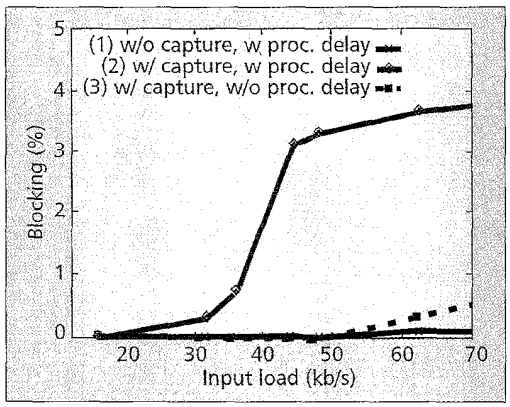
We observe corresponding effects in Fig. 17. When the load increases, the delay of model 1 explodes first. At the input load of 48 kb/s, the delay of model 2 also starts to increase without bound. With model 3, even at the load of 64 kb/s, we obtain a tolerable delay of 2.5 s.

Figure 18 illustrates the significant improvement in blocking due to the capture effect. Without capture, as the load increases, the blocking goes as high as 3.5 percent. With capture, the blocking is always lower than 0.5 percent when the system is stable. The processing delay does not have a noticeable effect on blocking when the load is less than 50 kb/s.

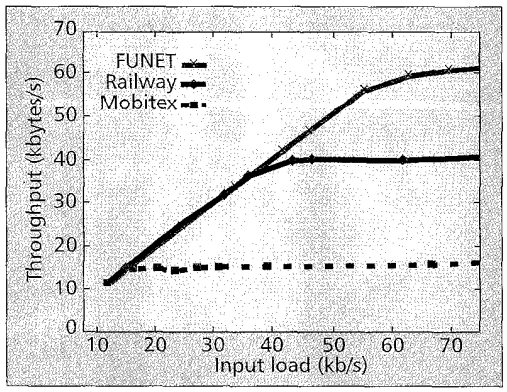
In Fig. 18, we notice that with capture at high load there is less blocking with processing delay than without delay. By definition, a user is blocked either because he tries too many times or he tries too long. Without processing delay, the failed mobile can make more random access retries within a certain time. This reduces delay but at higher load, it generates extra random access traffic, leading to congestion and increased blocking.



■ **Figure 17.** *Capture and processing delay effect: delay comparison.*



■ **Figure 18.** *Capture and processing delay effect: blocking comparison.*



■ **Figure 19.** *Traffic model effects: throughput comparison.*

Traffic Model Effects

In the GPRS protocol, the random access and transmission is frame based. The overhead (control signaling, frame header, etc.) is approximately fixed per frame. Unavoidably, the average frame length will have a strong impact on the overall performance in the GPRS system. According to [8], we studied three different traffic models: FUNET,[2] Railway,[3] and Mobitex.[4]

Figure 19 shows the throughput of different traffic models as a function of input load. Eight-slot operation and capture are assumed. Transmission errors exist in both the up- and downlink.

With a rate 1/2 convolutional code, the data rate per channel is 9.05 kb/s (Table 2). With eight-slot operation, theoretically, the maximum throughput in GPRS approaches 76 kb/s. With FUNET traffic, the maximum throughput reaches about 80 percent of the channel capacity (64 kb/s), while this is only 50 percent for the Railway model, and merely 20 percent for the Mobitex model.

The relationships among the three maximum throughput numbers are due to the average frame lengths. This is the longest in the FUNET model and shorter in the Railway and Mobitex models. With shorter frames there is more overhead per unit time, and consequently lower throughput.

Conclusions

In this article, we summarize the basic concepts of the GPRS systems, and evaluate the GPRS air interface protocol by simulation based on the draft standard. The simulation shows that multislot operation implemented in GPRS leads to efficient channel utilization over the air interface. Capture is helpful in the GPRS system. It reduces the blocking rate and improves both throughput and delay. If possible, processing delay should be reduced to achieve a better performance. However, the GPRS protocol we studied here is not suitable for small data packet

[2] *Truncated Cauchy packet length distribution.*

[3] *Truncated exponential packet length distribution.*

[4] *Uniform packet length distribution.*

transmission. A large portion of GPRS channels are wasted on signaling when the traffic is dominated by short packets. This problem has been recognized within ETSI, and countermeasures are presented in the latest standard [3].

In continuing work on GPRS, we are examining the effects of several additional parameters and operating conditions, including up- and downlink error effects, random access limits, and retransmission limits. Since the standardization of GPRS is still going on, some of the details we have studied to date are likely to be modified in the final standard. For example, the limit on LLC frame size has been changed from 2000 bytes to 1600 bytes. The selective ARQ scheme is not based on the frame, but on the flow (could be a combination of several frames if frames are short). Future work will simulate the system specified in the final standard.

Acknowledgments

The authors acknowledge the assistance and advice of D. Turina and E. Schoster of Ericsson.

References

[1] M. Mouly and M. B. Pautet, "Current Evolution of the GSM Systems," *IEEE Pers. Commun.*, vol. 2, no. 5, Oct. 1995, pp. 9–19.
[2] D. Turina *et al.*, "A Proposal for Multi-Slot MAC Layer Operation for Packet Data Channel in GSM," *Proc. Int'l. Conf. Universal Pers. Commun.*, vol. 2, Sept.–Oct. 1996.
[3] ETSI, "GSM 03.64 Overall Description of the GPRS Radio Interface, Stage 2," v. 2.1.1, May 1997.
[4] ETSI, "GSM 03.60 General Packet Radio Service: Service Description, Stage 2," v. 2.0.0, May 1997.
[5] G. Brasche, "Analysis of Multi-Slot MAC Protocols Proposed for the GSM Phase 2+ General Packet Radio Service," *Proc. Vehic. Tech. Conf.*, May 1997.
[6] G. Brasche, "Evaluation of a MAC Protocol Proposed for a General Packet Radio Service in GSM," *Proc. Vehic. Tech. Conf.*, May 1996.
[7] D. Turina, "Performance Evaluation of a Single-Slot Packet Data Channel in GSM," *Proc. Vehic. Tech. Conf.*, 1995.
[8] GPRS Ad-Hoc ETSI/STC SMG2, "Evaluation Criteria for the GPRS Radio Channel," Tech. Rep., Feb. 1996.
[9] GPRS Ad-Hoc ETSI/STC SMG2, "Simulations of Multislot MAC Layer Operation in GPRS (2)," Tech. Rep., Feb. 1996.
[10] D. J. Goodman and A. A. M. Saleh, "The Near/Far Effect in Local ALOHA Radio Communications," *IEEE Trans. Vehic. Tech.*, vol. VT-36, no. 1, Feb. 1987.

Biographies

JIAN CAI (jiancai@lucent.com) received her B.S. degree from the Electrical Engineering Department of Southeast University, China, in 1989. In 1995 she joined WINLAB, Rutgers University, to pursue the M.S. degree at the Electrical and Computer Engineering Department. Her research, supervised by Dr. Goodman, was on the performance of the general packet radio service proposed for GSM phase 2+. She is receiving her Master's degree in October 1997, and is currently working on CDMA base station software development at Lucent Technologies.

David J. Goodman [M '67-SM '86-F '88] (dgoodman@winlab.rutgers.edu) received a B.S. from Rensselaer Polytechnic Institute, Troy, New York, an M.S. from New York University, and a Ph.D. from Imperial College, University of London, United Kingdom, all in electrical engineering. Since September 1988, he has been a professor in the Department of Electrical and Computer Engineering at Rutgers, the State University of New Jersey. He is also director of WINLAB, the Rutgers Wireless Information Network Laboratory. Prior to joining Rutgers, he spent 20 years at Bell Laboratories, where he was department head of Communications Systems Research. His research has spanned many areas of digital communications, including wireless information networks, digital signal processing, digital coding of speech, and speech quality assessment. He is currently chair of the National Research Council Committee on Untethered Communications. In 1995 he was a research associate at the Program on Information Resources Policy at Harvard University.