ELSEVIER

Computer Communications 25 (2002) 522–533

computer communications

www.elsevier.com/locate/comcom

Single channel occupied voice and multi-channel occupied HSCSD over GSM

Jahangir H. Sarker*, Seppo J. Halme

Communications Laboratory, Helsinki University of Technology, PO Box 2300, FIN-02015 HUT, Finland

Received 13 July 2001; accepted 13 July 2001

Abstract

Random access and traffic channel parts are utilized, respectively, for call establishment and information transmission in the uplink direction, comprising mobile to base station of the Global System for Mobile Communication (GSM) networks. Voice (V) and High Speed Circuit Switched Data (HSCSD) calls use two parallel random access parts and a common traffic channel part where a V call occupies a single traffic channel and a HSCSD call occupies a multiple number of traffic channels. A call is either rejected or blocked if it is unsuccessful either in the random access part or in the traffic channel part. The optimum number of random access slots is directly proportional to the average access arrival rate. According to the specification, four types of retransmission cut-off policies can be executed in the random access part wherein only one type can provide almost zero call rejection probability, pertaining to a limited amount of traffic arrival rate per random access slot. A complete analysis for the traffic channel utilization and call blocking probability for V and HSCSD calls with an exact number of random access slots in collaboration with four different types of retransmission cut-off schemes is introduced. The overall call success probability is devised resorting to call rejection and call blocking probabilities. Results demonstrate that the overall call success probability can be independent of random access part up to a certain call arrival rate if a special type of retransmission cut-off is utilized. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Call blocked; Call rejection; GSM; HSCSD; Random access; Traffic channel

1. Introduction

The core network of the Universal Mobile Communication Systems (UMTS) is perhaps based on Global System for Mobile (GSM) [1]. The data communications, especially Internet Protocol (IP) based data transmissions increased a lot in the last decade. Expectations show increasing demand for a wide range of services from voice, involving low to high bit rate data and advanced data services for the cellular communication system. At the beginning of this century, mobile subscribers enjoy various kinds of data services together with regular voice service. Worldwide activities preserve the continuance in implementing the third generation mobile radio system International Mobile Telecommunications by the year 2000 (IMT-2000), in the International Telecommunication Union (ITU), or Universal Mobile Telecommunication System (UMTS), in Europe. The implemented third generation system will provide service to anyone, anytime, anywhere in all radio environments. The WCDMA based third generation system will provide the services: voice and low rate data (144 kb/s) in vehicular, medium rate data (384 kb/s) for outdoor and indoor, and up to 2 Mb/s for indoor environment [2].

Billions of subscribers use the three most successful second generation TDMA based cellular systems: GSM; TDMA (136); and PDC, where speech is the main teleservice. Mobile communication network operators envisage the demand for new mobile subscribers, and they expand the networks continuously. Some operators may not be able to get license to third generation spectrum to support the new services, such as high-speed data. Some operators do not wish to consider a jump to the third generation system due to its large investment in the current second generation technology. Therefore, they look for possible solutions to achieve services comparable to the third generation in the existing second generation systems [3].

The European Telecommunications Standards Institute (ETSI) has been defined two new technologies called General Packet Radio Services (GPRS) and High Speed Circuit-Switched Data (HSCSD), and currently are operated for data services over the existing GSM system. In the existing GSM system, each voice (V) call can occupy a single

* Corresponding author. Tel.: +358-9-451-2347, 2367; fax: +358-9-431-2345.

E-mail address: jahangir.sarker@hut.fi (J.H. Sarker).

0140-3664/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0140-3664(01)00374-7

traffic channel for the whole conversation period. A V call is a circuit-switched based, whereas a GPRS call is a packet-switched based transmission system.

Traffic channel utilization is an important measure for the operators. The users are charged depending on the traffic channel utilization, implicating the other parameters. One of the most important criteria for the operators is to keep the existing V calls' blocking probability within a certain limit. The GPRS is an effective solution, which can transmit data using multiple numbers of traffic channels dynamically without increasing V call blocking probability [4–7]. Unfortunately, a GPRS call may suffer from high delay during congested periods of the network. Some services such as video stream need to transmit data using multiple traffic channels with an almost constant delay. A HSCSD call occupies multiple number of traffic channels at a time up to completing its whole data transmission time [8]. Thus, the HSCSD provides the solution [9].

The blocking probability of the existing V calls is changed after the implementation of HSCSD calls in the networks. The blocking probabilities with HSCSD calls are discussed extensively in Ref. [10] using a multi-dimensional Markov chain. The blocking probabilities and their properties with a multi-dimensional Markov model are also discussed in Refs. [11–13]. Stasiak carried out excellent work [14], where he developed the determined distributions using a combinatorial method. The determined distribution enables the calculation of blocking probabilities in multi-stage switching networks carrying different multi-channel traffic streams.

Multiple channels occupied HSCSD calls enhance the existing V call blocking probabilities. Fortunately, for the same V call blocking probabilities, the overall channel utilization with continuous multi-channel occupied HSCSD calls is higher than that without HSCSD calls. The overall channel utilization increases with increasing number of HSCSD calls pertaining to the same V call blocking probabilities [15]. Such a result encourages the operators to provide different kinds of services using HSCSD with the existing V call services and, thus, increases the different types of traffic arrival from mobile terminals. The effective traffic arrival to the traffic channel part depends on the output of the random access part.

2. Random access part

The random access part of the GSM network is a slotted ALOHA based, which is an efficient algorithm for distributed bursty traffic. Since the access arrival of a GSM base station is bursty in nature, the choice is appropriate. The slotted ALOHA is used mainly for two kinds of purpose: (1) mobile data transmission; and (2) request for a dedicated mobile channel. In the first kind of purpose, each packet should be received successfully. So, the retransmissions take place up to the successful reception of each data packet. On the other hand, for the request of a dedicated mobile channel, the successful reception of each data packet is not needed. Therefore, the retransmission cut-off occurs after a certain number of transmissions involved.

Kim [16] studied the frequency-hopped spread-spectrum random access (slotted ALOHA) considering stability while using the number of (re)transmissions, new packet generation rate and the code rate as parameters. A slotted ALOHA system is mono-stable if the number of retransmissions is limited to eight irrespective of the offered traffic load [17].

Excellent work regarding the stable operation of slotted ALOHA by usage of the retransmission cut-off has been done in Ref. [18]. It has been shown using computer simulation that without any capture, a maximum of seven retransmissions guarantee the best grade of service for a new packet generation rate of between 0 and a critical point corresponding to approximately 0.35 packets per time slot. Under the same assumptions, it is shown by analysis [19] that the retransmission cut-off is not needed when the new packet generation rate is below e^{-1} (≈ 0.368) packet per time slot. If the new packet generation rate exceeds this limit, the number of transmissions has to decrease abruptly with increasing new packet generation rate for a guaranteed grade of service. In [20], it is shown that the value of the critical point adhering to the new packet generation rate, e^{-1} is the minimum value of a more general case, comprising a finite number of users and capture effect.

Most earlier papers analyzed either the traffic channel part [10–15] or the random access part [16–20] of a TDMA based cellular system. This paper combines the two parts together. Traditionally, excess random access slots are mapped in physical channels of a base station, assuming the base station is always informed of an arrival of a new call. Therefore, the probability that a call is rejected in the random access part has a negligible impact on the traffic channel part parameters, for instance, channel utilization and call blocking probability. In particular, the excess random access slot mapped over the physical channel is a waste of the radio resource. In addition, the introduction of new services increases the overall traffic arrival and the overestimated (excess) random access slot might not be a sufficient solution to cope with the new system. Nowadays, a multiple number of services with different Quality of Service (QoS) are desired from the same mobile network. This work might be a guideline to estimate the parameters in random access and traffic channel parts with different QoS of different services in the same network.

The rest of the paper is organized as follows. Section 3 describes the system model. The efficiency of random access channel is described in Section 4. A fixed number of slotted ALOHA based random access slots are mapped onto a normal GSM system for the access of V and HSCSD calls which is not optimized. The average number of channels occupied by a single channel occupied V and multiple channel occupied HSCSD calls together with the optimization of a number of random access slots can be obtained

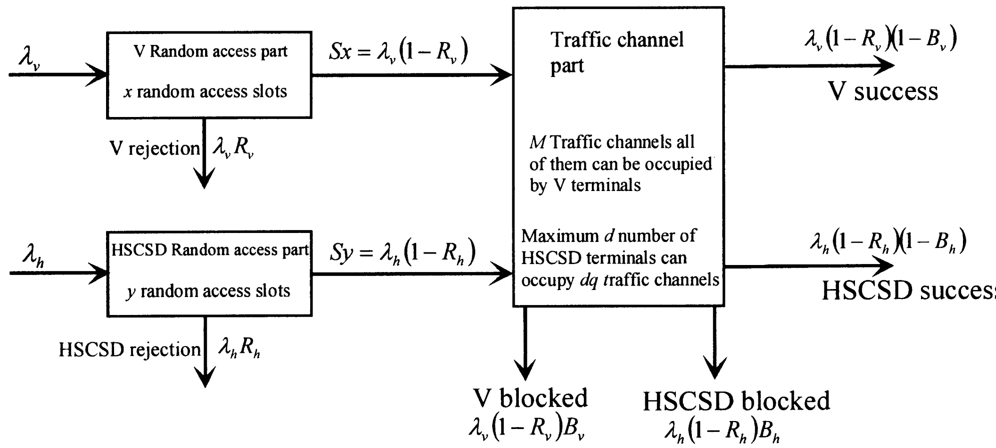

Fig. 1. System model.

from Section 5. The overall call success probability depending on two parts—the random access part and the traffic channel part—are considered in Section 6. Finally, Section 7 deals with conclusions.

3. System model

In the case of a normal GSM transmission system, one mobile terminal initiates the call using Random Access Channel (RACH) [21]. The RACH is based on slotted ALOHA. After a successful reception of an access packet through the random access part, the mobile network realizes the nature of the terminals, comprising a maximum of one channel occupied V call or a multiple (usually four) channels occupied HSCSD call. A single base station is considered, having M traffic channels, consisting of maximum M number of V calls and h number of HSCSD calls allowed to transmit information at a time.

Consider that the arrival of each kind of call has Poisson distribution with a mean λ (λ_v for V calls and λ_d for HSCSD calls). Two separate random access parts are used for initial access. The number of random access slots is x for V and y for HSCSD calls, respectively. The selection of a random access slot is random in nature. Therefore, the Poisson arrival traffic is distributed uniformly over the random access slots that are also Poisson distributed. The output process of different contention packet broadcasting systems is studied in Ref. [22]. Assuming that the output of *each* random access slot is Poisson with an average value amounting to the slotted ALOHA throughput and the addition of Poisson arrivals is also Poisson, the overall output of the random access part (also the input of the traffic channel part) is Poisson where the mean is the addition of multiple output means corresponding to each random access slot. The system model is shown in Fig. 1.

After successful reception of a V access packet through the V random access part, the base station allocates a Traffic CHannel (TCH) for that call to transmit the voice bursts if the network has that TCH available. If all M traffic channels are full, a successful V call through the random access part is blocked. On the other hand, a HSCSD call transmits the bursts occupying q ($=4$) channels if the number of the base station has this amount of free channels available to provide them and the number of existing HSCSD calls is less than h. If either of these two possibilities does not exist during that time, a successful HSCSD call through the random access part is blocked.

4. Efficiency of random access channel

A general approach can be considered for two random access parts and thus the performance of the V random access part is shown below.

4.1. The number of random access slots

Let the traffic arrival from all GSM calls be exponentially distributed with a mean rate λ. In a GSM network, a call generates access packets and transmits those to the random access slots (RACH). The number access (slotted ALOHA based) slot is x. There are five different structures of RACH [23] with approximately 400,000 and 780,000n RACH slots per hour ($n = 1, 2, 3, 4$). Let us assume that the arrival packet is uniformly distributed over the x random access slots. Thus, the new access arrival rate per slot is λ/x packet.

We assume that the access packet arrival rate from any call is equal. If the first attempt is unsuccessful, the call retransmits the access packet after a random delay. Therefore, the aggregate traffic (first attempt and the retransmitted traffic) generation rate from any call is greater than or equal to the access arrival rate. Since the selection of a random access slot is random, the aggregate traffic generation rate from any call transmitted to any slot can also be considered as equivalent. The system is assumed to be *memoryless*, where the probability of transmitting a packet in a given slot is independent of the state of the previous slot(s). In the case of an infinite number of calls, the distribution of the aggregate packets is Poisson. A packet is successfully

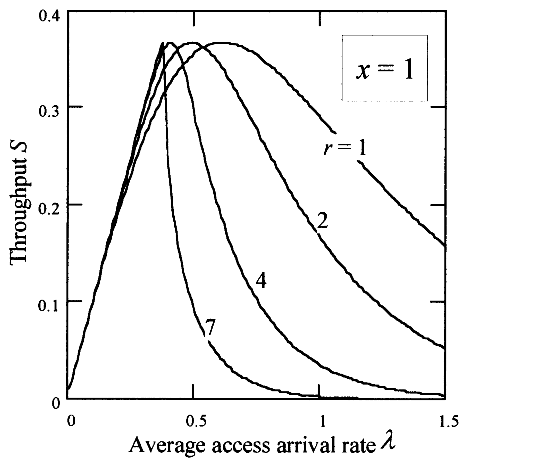

Fig. 2. Random access throughput per slot S versus average access arrival rate λ.

transmitted if there is no other interfering packet in the same time slot and the probability of that is $\exp(-G)$. Herein, G is the average aggregate traffic (first time access arrival and a certain number of retransmission arrivals) generation rate from all active calls per access slot.

According to specifications, maximum r retransmissions are allowed for each mobile terminal during the access period [21,24]. The parameter r can be set for four different possible values 1, 2, 4 or 7. Including the first transmission, the maximum allowed transmission number is $(r + 1)$. The retransmission cut-off scheme is studied in Ref. [20] and accordingly, the throughput per slot is redefined as

$$S = \frac{\lambda}{x}\left[1 - \{1 - \exp(-G)\}^{r+1}\right] \tag{1}$$

where the new packet arrival rate per random access slot is λ/x and the aggregate arrival rate in *each* access slot G is related by

$$\lambda/x = \frac{G\exp(-G)}{1 - \{1 - \exp(-G)\}^{r+1}} \tag{2}$$

The throughput or the successful access rate per random access slot is S packets per time slot. The random access throughput per slot S for different values ascribed to the number of retransmissions r, is depicted in Fig. 2.

The overall successful arrival into the traffic channel part from all active calls is Sx. Note that the value of S should always be less than one packet per time slot, but the value of

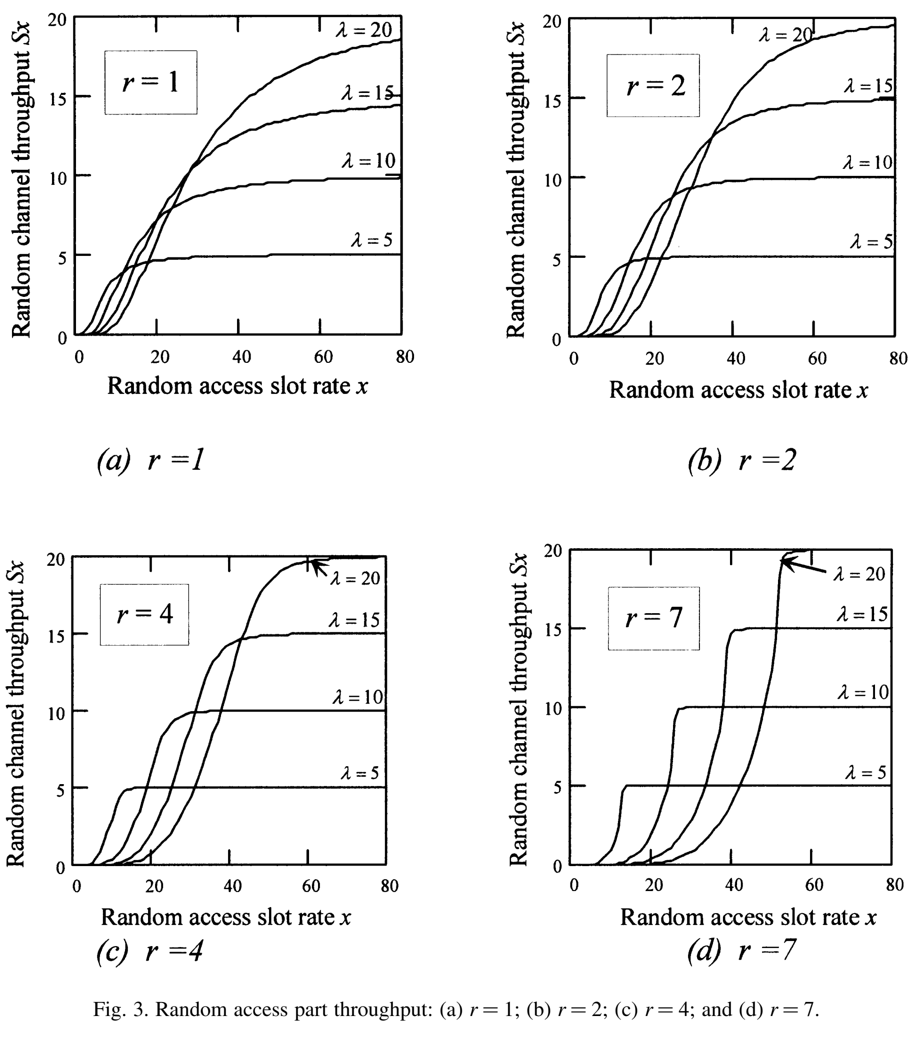

Fig. 3. Random access part throughput: (a) $r = 1$; (b) $r = 2$; (c) $r = 4$; and (d) $r = 7$.

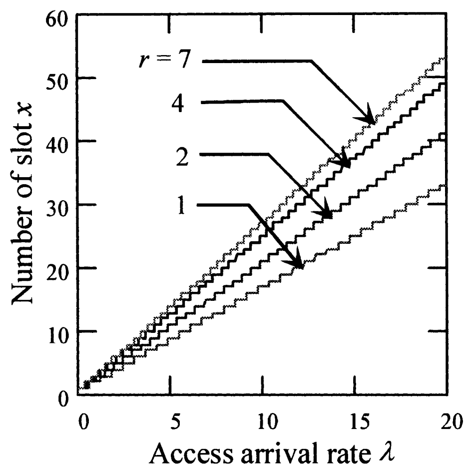

Fig. 4. Required number of random access slots x.

Sx might be more than one. However, its value should be less or equal to λ. Fig. 3 depicts the RACH throughput Sx with the variation of the number of random access slots x for different values of access arrival rate λ. The figure shows: (a) the random channel throughput increases almost linearly with increasing number of random access slots x if the retransmission number r is low ($r = 1$ or 2). If the number of random access slots x is very high, the output traffic of RACH Sx is almost similar to input traffic λ; (b) in the case of higher retransmission number r ($r = 4$ or 7), the output traffic of RACH Sx increases abruptly if the number of random access slots x crosses a certain limit. To obtain the same output Sx as the input traffic λ, the retransmission number 7 is most preferable in terms of number of slots x.

Operators are interested in this access state to give full availability for all kinds of GSM calls. In the second stage, if all traffic channels are full and the network is unable to provide any service, the network informs the mobile terminal of its *blocked* call. In contrast, if the call is *rejected* in the access stage, it cannot receive any message from the base station. Physically, it is unknown to the network that one terminal tries to access to transmit the information. In this case, the mobile terminal may receive an automatic 'access rejection' signal, which may request the mobile user to try again. The mobile user may think that the network is full and it supplicates a call later. This is true for the case of call blocking, but not for call rejection. So, *call blocked* is a better situation than that of *call rejection*. Thus, the operators have to stimulate more random access slots with augmenting retransmission number if the average access arrival rate of the GSM users is more than this limit.

It is well known that the slotted ALOHA shows the maximum throughput when the aggregate traffic generation rate from all active GSM calls per random access slot is $G = 1$. If the aggregate arrival rate is more than this value, the throughput decreases sharply and the system becomes unstable [20,25]. Considering this case in the random access part, the optimum relation between the access arrival rate and the number of slots becomes

$$x = \left\lfloor \lambda \mathrm{e}\left\{1 - (1 - \mathrm{e}^{-1})^{r+1}\right\}\right\rfloor = \lfloor w \rfloor \tag{3}$$

where $\lfloor w \rfloor$ is the smallest integer $\geq w$. The numerical representation of Eq. (3) is depicted in Fig. 4. Evidently, the required number of slots x is higher if the number of retransmissions is high. Remarkably, this higher number slot implementation stimulates the lower rejection probability and, thus, enhances the output traffic of the random access part only with the higher number of retransmission number r. In the case of lower number of retransmission number r, the rejection probability is high because of the protocol implementation and the excess random slot implementation is a waste of radio resource.

4.2. Average number of retransmissions

The collapse of the RACH is controlled in a GSM system by two different parameters: the retransmission number control; and the average time between them. The average number of retransmissions defines the average number of retransmission slots needed for a successful access packet transmission. A GSM call tries to access into the base station immediately if it has any information for transmitting. If the first access transmission is unsuccessful, it retransmits the access packet using a special algorithm. The time between the two transmissions, measured in units of RACH slots, is an integer random variable, uniformly distributed between 1 and a parameter *Tx-integer*, where *Tx-integer* can be set to some values ranging from 3 to 50 slots. In GSM, maximum r ($r = 1$, 2, 4 or 7) time retransmissions is allowed. The probability that an access packet is successfully transmitted after the k-th retransmission is

$$P_k = \{1 - \exp(-G)\}^k \exp(-G), 0 \leq k \leq r \tag{4}$$

Unambiguously, Eq. (4) is not the absolute geometric distribution for the access attempt k ($0 \leq k \leq r$). Therefore, the modified distribution is

$$P'_k = \sum_{k=0}^{\infty} P_k \frac{P_k}{\sum_{k=0}^{r} P_k}$$

which yields a modified geometric distribution as

$$P'_k = \frac{\{1 - \exp(-G)\}^k \exp(-G)}{1 - \{1 - \exp(-G)\}^{r+1}} \tag{5}$$

Finally, the expected number of retransmissions needed for a successful transmission of an access packet is

$$RTx = \sum_{k=0}^{r} k P'_k$$

$$= \frac{\exp(-G)}{1 - \{1 - \exp(-G)\}^{r+1}} \sum_{k=0}^{r} k\{1 - \exp(-G)\}^k$$

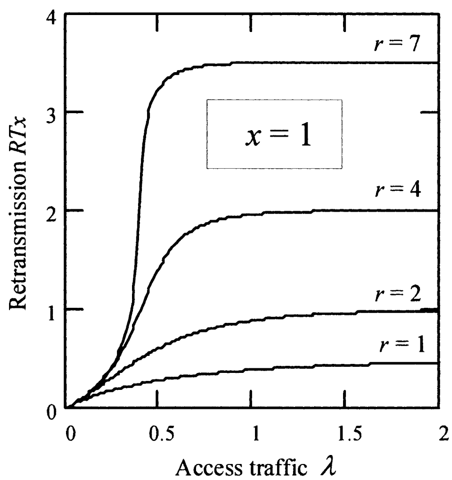

Fig. 5. Average number of retransmissions.

After simplification

$$RTx = \frac{\{1-\exp(-G)\}[1-(r+1)\{1-\exp(-G)\}^r + r\{1-\exp(-G)\}^{r+1}]}{\exp(-G)[1-\{1-\exp(-G)\}^{r+1}]} \quad (6)$$

where the relationship between the aggregate traffic arrival rate G packet per time slot and the average access packet arrival rate λ/x packet per time slot can be obtained from Eq. (2). The average number of retransmissions of a random access call (packet) is depicted in Fig. 5.

4.3. Call rejection probability

The interesting parameter in the RACH part is the call rejection probability. It defines the probability that an access packet is rejected. If an access packet is rejected, a mobile terminal cannot inform the network of the wanted service. For this reason, the network or

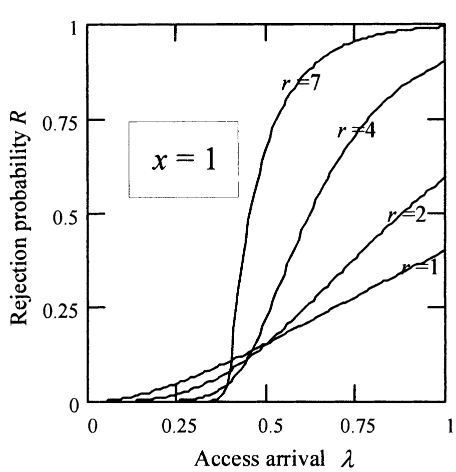

Fig. 6. Call rejection probability.

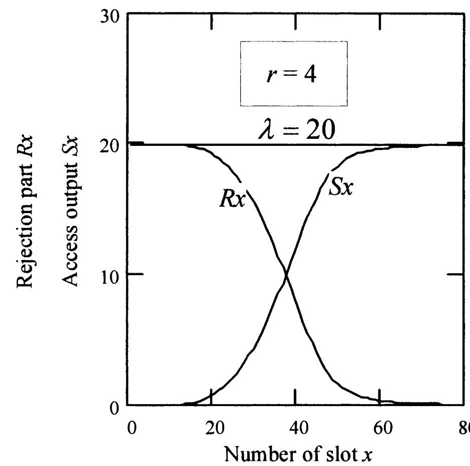

Fig. 7. Call rejection traffic and access part output traffic.

base station cannot inform the mobile terminal of the access failure. This parameter can be calculated in the following way. The probability that an access packet is successfully transmitted after the first transmission is $\exp(-G)$. The probability of failure after the first transmission is $\{1-\exp(-G)\}$. Assuming that the failure probability in each transmission time is independent, and the transmission takes place $(r+1)$ times (maximum limit), the call rejection probability is

$$R = \{1-\exp(-G)\}^{r+1} \quad (7)$$

The numerical representation of call rejection probability is shown in Fig. 6. Clearly, the access arrival traffic λ is the sum of two different parts: rejection traffic Rx; and the access part output traffic Sx. The numerical representation of these two different traffics with the variation of number of random access slot is depicted in Fig. 7. A careful inspection shows that most of the arrival traffics are rejected if the number of random access slots is less than 20 and most of the arrival traffics are converted to a random part output traffic if the number of random access slots is more than 40. In other words, the arrival traffic switches linearly from rejected traffic to random access part output traffic while the number of random access slots varies from 20 to 40.

5. Single channel occupied voice and multi-channel occupied HSCSD transmission system

The output traffic of random access part directly works as the input traffic of traffic channel part. The aggregate of x branch Poisson processes is another Poisson process with a mean arrival rate Sx. The call blocking probability is defined as the successful random access, but the network has no traffic channel to serve a new call. Consider only V calls, where a V call can hold only one traffic channel at a time. The access

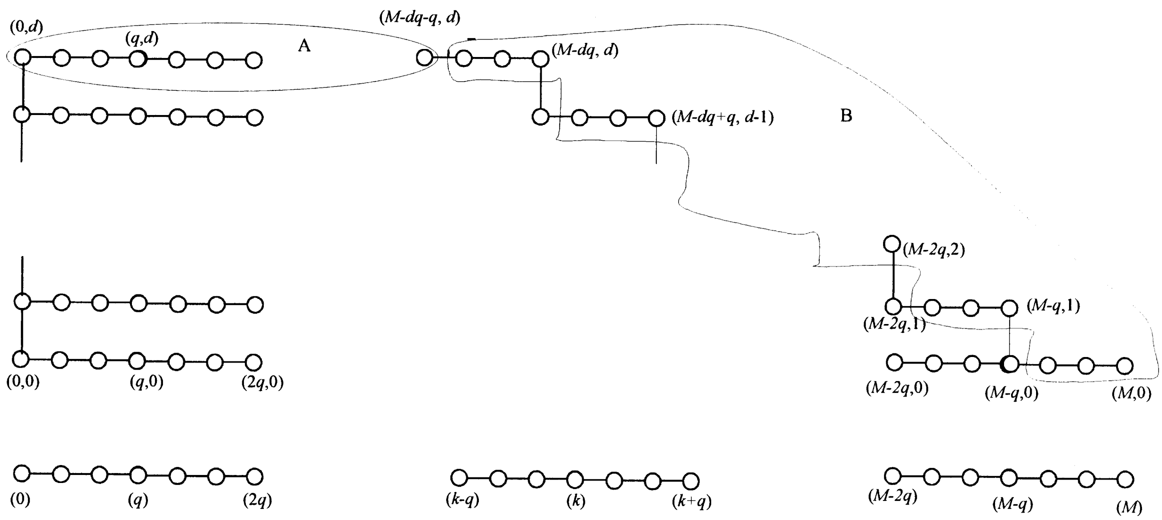

Fig. 8. States from two-dimensional to one-dimensional models.

arrival rate of V calls is Poisson with a mean λ_v. Let x random access slots be mapped onto a time period for all V calls in a base station. If the output of each random access slot is S_v, then the average arrival rate of the traffic channel part is $S_v x$. If a V call is successful (neither rejected nor blocked), it holds a traffic channel up to the entire conversation period of the user. Let us consider that the channel holding time of each V call is independent and exponentially distributed with a mean $1/\mu_v$. The V call traffic intensity of the traffic channel part is $S_v x/\mu_v$. The number of traffic channels is M in the considered base station. The maximum M number of V calls is allowed to transmit voice bursts at a time.

Consider a base station where the HSCSD calls arrive at a base station with a Poisson distribution having a mean λ_h. The base station allocates y random access slots for HSCSD calls in the random access part as depicted in Fig. 1. A HSCSD call holds multiple q channels at a time in a time frame. Assuming that the multiple q channels holding by each HSCSD call is independent and exponentially distributed with a mean $1/\mu_d$. The data traffic intensity of the traffic channel part for HSCSD calls is $S_h y/\mu_h$. The maximum d number of HSCSD calls is allowed to transmit information at a time.

The state (i, j) defines the probability that i number of V calls and j number of HSCSD calls are in the system. Yielding the probability that i number of V calls and j number of HSCSD calls are in the system leads to [26]

$$p(i,j) = \frac{1}{i!}(S_v x/\mu_v)^i \frac{1}{j!}(S_h y/\mu_h)^j p(0,0), 0 \le i \le d, \quad 0 \le j \le M, 0 \le iq + j \le M \tag{8}$$

We consider the two-dimensional system state model as a one-dimensional system state probability model as depicted in Fig. 8. Since there are M number of traffic channels in the system, the system states will be from 0 to M.

Assuming $Q(k)$ as the probability that k channels are occupied by either HSCSD or V calls in Fig. 8, the probability that the system is idle is

$$Q(0) = p(0,0) \tag{9}$$

One HSCSD call needs at least q channels. Thus, the probability that 1 to q channels are occupied is

$$Q(1) = p(0,1) = Q(0)(S_v x/\mu_v) \tag{10a}$$

$$Q(2) = p(0,2) = Q(0)\frac{(S_v x/\mu_v)^2}{2!} \tag{10b}$$

$$\vdots$$

$$Q(q) = p(0,q) + p(1,0) = Q(0)\left\{\frac{(S_v x/\mu_v)^q}{q!} + \frac{S_h y/\mu_v}{1!}\right\} = Q(0)\sum_{j=0}^{1}\frac{(S_v x/\mu_v)^{q-qj}}{(q-qj)!}\frac{(S_h y/\mu_h)^j}{j!} \tag{10c}$$

$$Q(q+1) = p(0,q+1) + p(1,1) = Q(0)\left\{\frac{(S_v x/\mu_v)^{q+1}}{(q+1)!} + \frac{(S_v x/\mu_v)^1}{1!}\frac{(S_h y/\mu_h)}{1!}\right\} = Q(0)\sum_{j=0}^{1}\frac{(S_v x/\mu_v)^{q+1-qj}}{(q+1-qj)!}\frac{(S_v x/\mu_v)^j}{j!} \tag{10d}$$

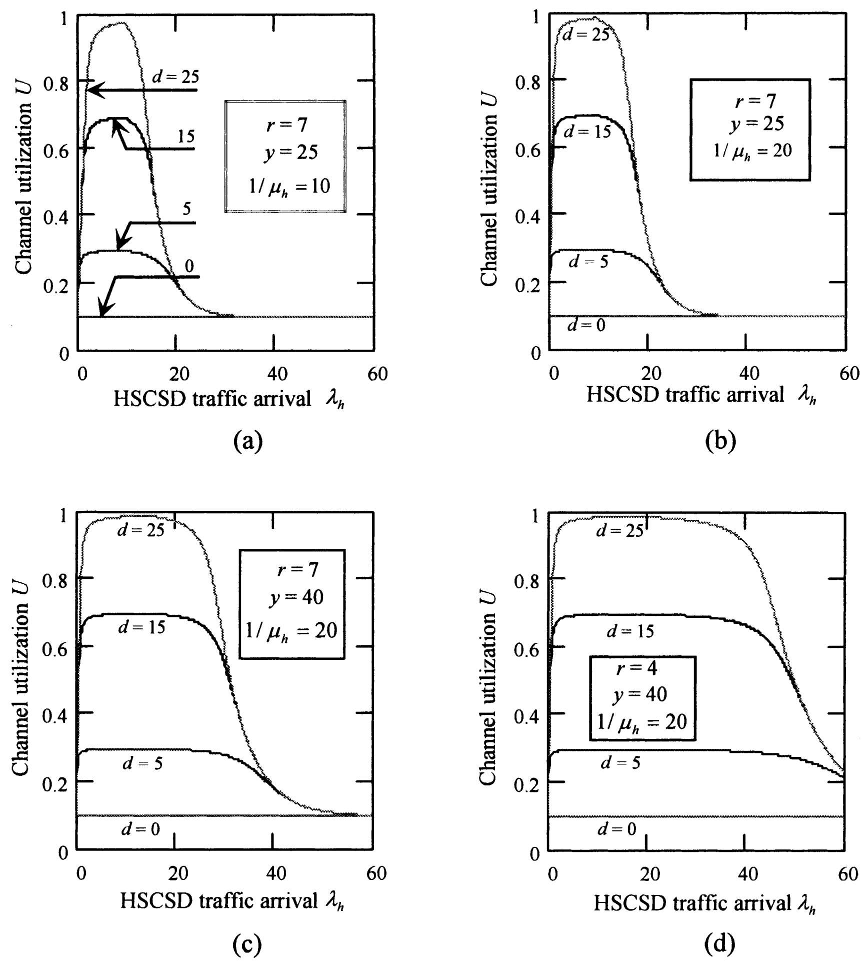

Fig. 9. Channel utilization with the variation of HSCSD traffic arrival rate. $M = 100$, $q = 4$, $\lambda_v x = 10$, and $1/\mu_v = 2$ considered for all diagrams.

$$Q(2q) = p(0, 2q) + p(1, q) + p(2, 0)$$

$$= Q(0)\left\{ \frac{(S_v x/\mu_v)^{2q}}{(2q)!} + \frac{(S_v x/\mu_v)^{q}}{q!}\frac{(S_h y/\mu_h)}{1!} + \frac{(S_h y/\mu_h)^2}{2!} \right\}$$

$$= Q(0) \sum_{j=0}^{2} \frac{(S_v x/\mu_v)^{2q-qj}}{(2q-qj)!}\frac{(S_h y/\mu_h)^j}{j!} \tag{10e}$$

Therefore, for the states ranging from 0 to dq, the probability that k channels are occupied by HSCSD and V calls is

$$Q(k) = Q(0) \sum_{i=0}^{\lceil \frac{k}{q} \rceil} \frac{(S_v x/\mu_v)^{k-qj}}{(k-qj)!}\frac{(S_h y/\mu_h)^j}{j!}, 0 \le k \le dq \tag{11}$$

where $\lceil u \rceil$ is the largest integer $\le u$.

Let us consider the case when $k > dq$. The probability that $dq + 1$ channels are occupied is

$$Q(dq+1) = p(0, dq+1) + p(1, dq+1-q) + p(2, dq+1-2q) + \cdots + p(d-1, q+1) + p(d, 1)$$

$$= Q(0) \sum_{i=0}^{d} \frac{(S_v x/\mu_v)^{dq+1-qj}}{(dq+1-qj)!}\frac{(S_h y/\mu_h)^j}{j!} \tag{12a}$$

Similarly, the probability that $dq + 2$ channels are occupied is

$$Q(dq+2) = p(0, dq+2) + p(1, dq+2-q) + p(2, dq+2-2q) + \cdots + p(d-1, q+2) + p(d, 2)$$

$$Q(dq+2) = Q(0) \sum_{j=0}^{d} \frac{(S_v x/\mu_v)^{dq2-qj}}{(dq+2-qj)!}\frac{(S_h y/\mu_h)^j}{j!} \tag{12b}$$

In general, for values of k between dq and M, the following

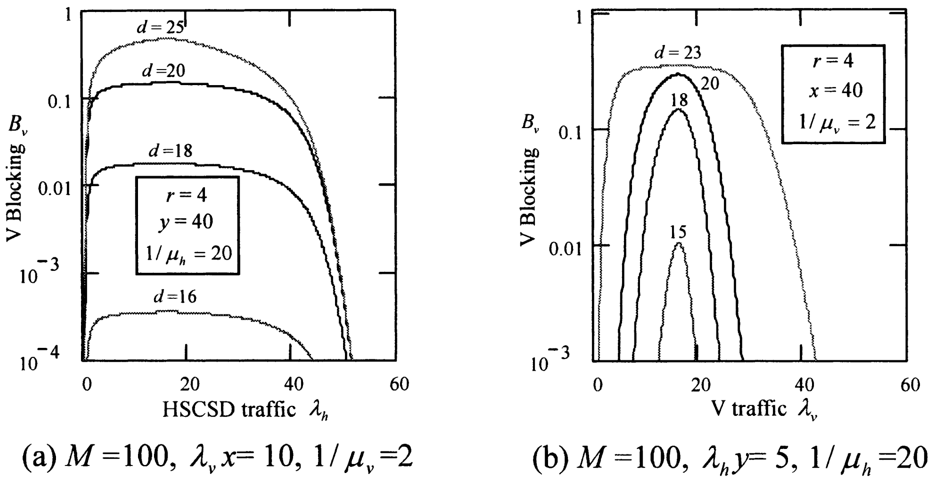

Fig. 10. V call blocking probability. (a) $M = 100$, $\lambda_v x = 10$, and $1/\mu_v = 2$. (b) $M = 100$, $\lambda_h y = 5$, and $1/\mu_h = 20$.

relation is developed

$$Q(k) = Q(0)\sum_{j=0}^{d} \frac{(S_v x/\mu_v)^{k-qj}}{(k-qj)!}\frac{(S_h y/\mu_h)^j}{j!}, dq \le k \le M \quad (13)$$

Combining Eqs. (11) and (13), the probability that k channels out of M channels are occupied is

$$Q(k) = \begin{cases} Q(0)\displaystyle\sum_{j=0}^{\lceil k/q \rceil} \frac{(S_v x/\mu_v)^{k-qj}}{(k-qj)!}\frac{(S_h y/\mu_h)^j}{j!}, & 0 \le k \le dq \\ Q(0)\displaystyle\sum_{j=0}^{d} \frac{(S_v x/\mu_v)^{k-qj}}{(k-qj)!}\frac{(S_h y/\mu_h)^j}{j!}, & dq \le k \le M \end{cases} \quad (14)$$

and from Eq. (9) $Q(0) = p(0,0)$ where

$$Q(0) = p(0,0) = \left\{ \sum_{i=0}^{d} \frac{(S_v x/\mu_v)^i}{i!} \sum_{j=0}^{M-iq} \frac{(S_h y/\mu_h)^j}{j!} \right\}^{-1} \quad (15)$$

Traffic channel utilization or system efficiency of the traffic channel part can be defined as the ratio of the average number of traffic channels to the total number of traffic channels. Hence, the channel utilization can be written as

$$U = \frac{1}{M}\sum_{k=0}^{M} kQ(k) \quad (16)$$

Fig. 9 shows the numerical results of traffic channel utilization with the variation of HSCSD traffic arrival rate λ_h, where $M = 100, q = 4, \lambda_v x = 5$ and $1/\mu_v = 2$ considered for all four diagrams. The four diagrams demonstrate four characteristics. (a) The channel utilization can be increased by increasing the number of allowed HSCSD calls amounting to d. After a certain limit of traffic arrival rate λ_h, the channel utilization U decreases abruptly. The reason is that the number of random access slots y ($= 25$) is not enough and most of the traffics are rejected in the random access part. (b) A given HSCSD call occupies q ($= 4$) traffic channels at a time. Therefore, as the HSCSD call channel holding time $1/\mu_h$ increases from 10 to 20, the multiple channels occupy longer time without changing the arrival or random access part. Higher traffic channel utilization can be obtained with a higher number of allowed HSCSD calls equivalent to d. (c) As mentioned earlier, the abruptly decreased throughput results from the sudden increasing rejection probability, due to enhanced HSCSD traffic arrival rate. The introduction of more number random access slots x (here y) can solve this problem. (d) If the arrival traffic is sufficiently high and the number of random access slots x is not sufficient for the higher traffic channel or the enhanced random access output, the reduced number of retransmissions r may provide the solution.

Setting the maximum HSCSD call d to zero indicates that the base station operates only the V calls and the traffic channel utilization is only for voice traffic. In Fig. 9, the voice traffic is constant and its value is 10. A similar characteristic can be obtained if the traffic channel utilization is justified with the variation of V traffic arrival rates.

If either HSCSD calls and/or V calls occupy all M channels, a new V call is lost or blocked. The probability for that is termed as *blocking probability for V calls* identified by B_v, and can be derived from Eqs. (14) and (15) as

$$B_v = Q(M) = \frac{\displaystyle\sum_{j=0}^{d} \frac{(S_v x/\mu_v)^{M-qj}}{(M-qj)!}\frac{(S_h y/\mu_h)^j}{j!}}{\displaystyle\sum_{i=0}^{d} \frac{(S_v x/\mu_v)^i}{i!} \sum_{j=0}^{M-iq} \frac{(S_h y/\mu_h)^j}{j!}} \quad (17)$$

The V call blocking probability is depicted in Fig. 10.

According to the above mentioned system model, a maximum of dq channels are available for HSCSD calls. The HSCSD calls are blocked for two different reasons: firstly, if HSCSD calls already occupy dq channels; secondly, if more than $M - q + 1$ channels are already occupied, i.e. q channels are not available for one HSCSD call. In the

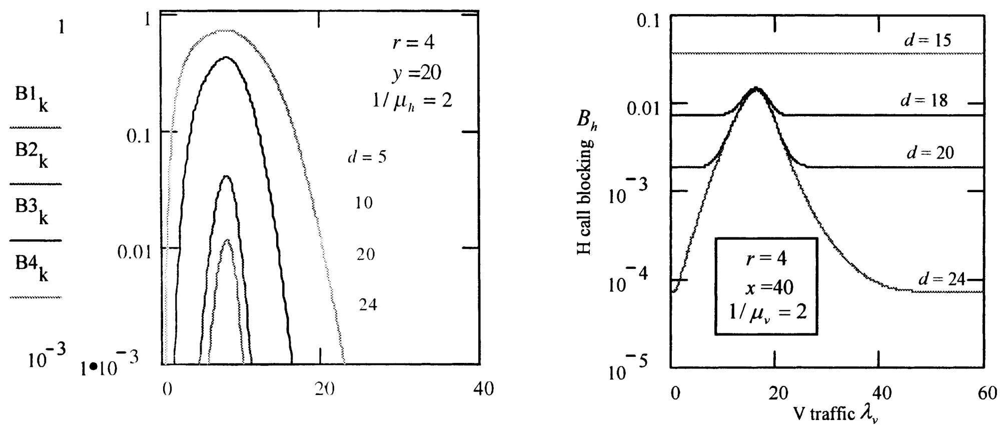

Fig. 11. HSCSD call blocking probability. (a) $M = 100$, $\lambda_v x = 3$, and $1/\mu_v = 2$. (b) $M = 100$, $\lambda_h y = 2$, and $1/\mu_h = 5$.

system state probability diagram, the HSCSD call blocking probability states can be shown in areas A and B, as evoked in Fig. 8, and can be defined as

$$B_h = \sum_{j=0}^{M-dq-q} p(d,j) + \sum_{k=M-q+1}^{M} Q(k) = \frac{\dfrac{(S_h y/\mu_h)^d}{d!} \displaystyle\sum_{j=0}^{M-dq-q} \dfrac{(S_v x/\mu_v)^j}{j!} + \displaystyle\sum_{k=M-q+1}^{M} \sum_{j=0}^{d} \dfrac{(S_v x/\mu_v)^{k-qj}}{(k-qj)!} \dfrac{(S_h y/\mu_h)^j}{j!}}{\displaystyle\sum_{i=0}^{d} \dfrac{(S_v x/\mu_v)^i}{i!} \sum_{j=0}^{M-iq} \dfrac{(S_h y/\mu_h)^j}{j!}} \tag{18}$$

Fig. 11 exhibits the HSCSD call blocking probabilities. A detailed discussion of the call blocking probabilities together with the call rejection probabilities is provided in the next section.

6. Call success probability

Here, we discuss the subsequent development advances with a recourse to call rejection and blocked probabilities concerning the random access and the traffic channel parts. Considering a general case for V and HSCSD calls, a call is rejected in the random access part with a probability R. Consequently, the input probability of the traffic channel part is $(1 - R)$. A call is blocked in the traffic channel part with a probability B. Therefore, the probability that a call is successfully transferred (neither rejected nor blocked) is given by

$$T = (1 - R)(1 - B) \tag{19}$$

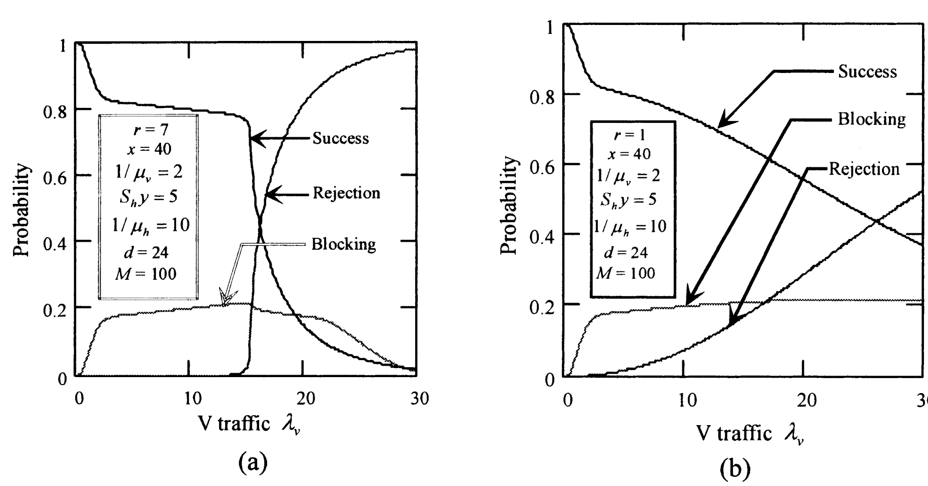

Fig. 12. Call rejection, blocking and success probabilities for V calls. Two regions are distinguished for $r = 7$. If $\lambda < xe^{-1}\{1 - (1 - e^{-1})^8\}^{-1}$, then the first region and if $\lambda \geq xe^{-1}\{1 - (1 - e^{-1})^8\}^{-1}$, then the next region.

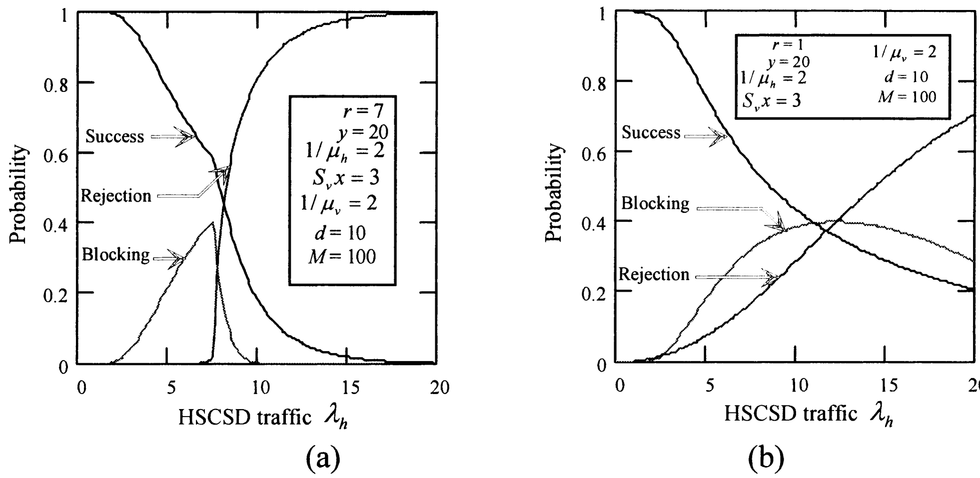

Fig. 13. Call rejection, blocking and success probabilities for H calls. Two regions are evident for $r = 7$. If $\lambda < x\mathrm{e}^{-1}\{1 - (1 - \mathrm{e}^{-1})^{8}\}^{-1}$, then the first region and if $\lambda \geq x\mathrm{e}^{-1}\{1 - (1 - \mathrm{e}^{-1})^{8}\}^{-1}$, then the next region.

A numerical representation of Eq. (19) is depicted in Fig. 12 for V calls and Fig. 13 for HSCSD calls, respectively, for a few sets of specific values of different parameters that constitute two separate views.

1. Call rejection probability increases almost linearly after a certain limit of arrival rate λ if the number of retransmissions r is low (1 or 2). Call blocking probability depends on other parameters. The overall call success probability follows the relation represented by Eq. (19).
2. In the case of higher number of retransmissions r (≥ 7), two separate regions are distinguished.

If $\lambda < x(\text{or } y)\mathrm{e}^{-1}\{1 - (1 - \mathrm{e}^{-1})^{r+1}\}^{-1}$: (a) the call rejection probability is almost zero; (b) the call blocking probability depends on other parameters; and (c) the overall call success probability is inversely proportional to the call blocking probability because of almost zero call rejection probability.

If $\lambda \geq x(\text{or } y)\mathrm{e}^{-1}\{1 - (1 - \mathrm{e}^{-1})^{r+1}\}^{-1}$: (a) the call rejection probability increases abruptly; (b) thus, most of the calls are rejected in the random access part; (c) the arrivals to the traffic channel part decrease and thus the call blocking probability also decreases; and (d) the overall call success probability decreases abruptly because of a sharp increase in the call rejection probability.

Finally, it must be emphasized that the overall call success probability can definitely be calculated using an multi-dimensional Erlang formula if and only if the number of random access slots (x or y) is more than $\lfloor \lambda \mathrm{e}^{-1}\{1 - (1 - \mathrm{e}^{-1})^{r+1}\}\rfloor$: where $r \geq 7$.

7. Conclusions

The estimation of random access slots in a GSM based network is analyzed. The call rejection probability of random access part depends on the number of retransmissions r. If r is low (1 or 2), the call rejection probability is almost linear after a certain limit of call arrival rate λ. If the maximum number of retransmissions r is more than 7, the call rejection probability of the random access part depends on the number of random access slots (x or y). The required number of random access slots is directly proportional to the average access arrival rate. A reduction in the retransmissions r can extenuate the requisite random access slots (x or y) at the expense of a higher rejection probability.

The anticipated number of random access slots x is obtained by adjusting the optimum aggregate traffic generation rate per time slot $G = 1$. The reasoning is as follows. Retransmission cut-off is strongly recommended for the stable operation of random access if the average access arrival rate per time slot λ/x, is more than e^{-1}, otherwise the throughput decreases abruptly [20,25]. This transition point is obtained from the optimum aggregate traffic generation rate. Thus, the required/optimum number of random access slots can be calculated from the optimum aggregate traffic generation rate that maximizes the random access throughput.

Traffic channel utilization for different numbers of random access slots is derived. The utilization deteriorates if *less* random access slots are involved, particularly with a higher number of retransmissions. The reason is that more access is rejected in the access part. 'Excess' number of random access slots can decrease the call rejection probability and, hence, increase the traffic channel utilization, entangling an acceleration to the blocking probability.

The traffic channel utilization depreciates if *less* retransmissions r are used. The reason is that a moderate amount of traffic is rejected in the access part. A higher number of retransmissions (≥ 7) accompanied by the 'excess' number of random access slots can decrease the call rejection probability, rectifying the traffic channel utilization and, hence, expanding the blocking probability.

The overall call success probability experiences detrimental effect if the number of access slots is less than $\lfloor \lambda e^{-1}\{1 - (1 - e^{-1})^{r+1}\}\rfloor$, where $\lfloor w \rfloor$ is the smallest integer $\geq x$. The number of random access slots in the random access part depends exclusively on the average access arrival rate λ and the number of retransmissions r.

References

[1] P. Chaudhury, W. Mohr, S. Onoe, The 3GPP proposal for IMT-2000, IEEE Communications Magazine 37 (12) (1999) 72–81.

[2] T. Ojanperä, R. Prasad (Eds.), Wideband CDMA for Third Generation Mobile Communications, Artech House, Boston, 1998.

[3] U. Amin, S. Chennakeshu, The evolution of TDMA to 3G, IEEE Personal Communications (1999) 6–7.

[4] S. Faccin, H. Liangchi, R. Koodli, L. Khiem, R. Purnadi, GPRS and IS-136 integration for flexible network and services evolution, IEEE Personal Communications 6 (3) (1999) 48–54.

[5] C. Jain, D.J. Goodman, General packet radio service in GSM, IEEE Communications Magazine 35 (10) (1997) 122–131.

[6] G. Brasche, B. Walke, Concepts, services and protocols of the new GSM phase 2 + general packet radio service, IEEE Communications Magazine 35 (8) (1997) 94–104.

[7] GSM 03.64, V5.1.0, Digital cellular telecommunication system (phase 2 +); general packet radio service (GPRS); overall description of the GPRS radio interface (stage 2), November 1997.

[8] GSM 02.34, V5.1.0, Digital cellular telecommunication system (phase 2 +); high speed circuit switched data (HSCSD)—stage 1, March 1997.

[9] J. Dunlop, Potential for compressed video transmission over the GSM HSCSD service, Electronics Letters 33 (2) (1997) 121–122.

[10] D. Calin, D. Zeghlache, Analysis of GSM HSCSD service with channel allocation constraints, in: K. Leung, B. Vojcic (Eds.), Multiaccess, Mobility, and Teletraffic for Wireless Communications, vol. 3, Kluwer Academic, Boston, 1999, pp. 237–252.

[11] A.A. Nilsson, M. Perry, A. Gersht, V.B. Iversen, On multi-rate Erlang-B computations, Proc. ITC 16 (1999) 1051–1060.

[12] J.S. Kaufman, Blocking in a shared resource environment, IEEE Transactions on Communications COM-29 (10) (1981) 1474–1481.

[13] J.M. Aein, A multi-class, blocked-calls-cleared, demand access model, IEEE Transactions on Communications COM-26 (3) (1978) 374–385.

[14] M. Stasiak, Combinatorial considerations for switching systems carrying multi-channel traffic streams, Proc. ITC 14 (1994) 1229–1238.

[15] J.H. Sarker, S.H. Halme, M. Rinne, Performance analysis of GSM traffic channel capacity with(out) high speed circuit switched data, IEEE VTC2000 Fall, September 24–28, Boston, USA, 2000, pp. 1603–1609.

[16] S.W. Kim, Frequency-hopped speed-spectrum random access with retransmission cutoff and code rate adjustment, IEEE Journal of Selected Areas in Communication 10 (2) (1992) 344–349.

[17] K. Sakakibara, H. Muta, Y. Yuba, The effect of limiting the number of retransmission trials on the stability of slotted ALOHA systems, IEEE Transactions on Vehicular Technology 47 (4) (2000) 1449–1453.

[18] C. Luders, R. Haferbeck, The performance of the GSM random access procedure, IEEE 44th Veh. Tech. Conference 94 (VTC'94), Stockholm, Sweden, June 8-10, 1994, pp. 1165–1169.

[19] J.H. Sarker, S.J. Halme, The prudence transmission method I (PTM 1): a retransmission cut-off method for contention based multiple-access communication systems, IEEE 47th Veh. Tech. Conference 97 (VTC'97), Phoenix, AZ, USA, May 4–7, 1997, pp. 397–401.

[20] J.H. Sarker, S.J. Halme, An optimum retransmission cut-off scheme for slotted ALOHA, Wireless Personal Communication 13(1–2) (2000) 185–202.

[21] M. Mouly, M.-B. Pautet, The GSM for Mobile Communications, Published by the authors, 1992.

[22] H. Takagi, L. Kleinrock, Output process in contention packet broadcasting systems, IEEE Transactions on Communications COM-33 (11) (1985) 1191–1199.

[23] ETSI-GSM Technical Specification, GSM 05.02, Version 4.2.0, Physical Layer on the Radio Path, Multiplexing and Multiple Access on the Radio Path, April 1993.

[24] ETSI-GSM Technical Specification, GSM 04.08, Version 4.4.0, Mobile Radio Interface—Layer 3 Specification, April 1993.

[25] B. Bing, Stabilization of the randomized slotted ALOHA protocol without the use of channel feedback information, IEEE Communications Letters (2000) 249–251.

[26] H. Akimaru, K. Kawashima, Teletraffic: Theory and Applications, Springer-Verlag, Berlin, 1993.

FX1. ***Jahangir H. Sarker*** *was born in Dhaka, Bangladesh, on November 12, 1967. He received his BSc degree in electrical and electronics engineering form Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 1991, and MSc and Licentiate degrees in electrical and communication engineering from Helsinki University of Technology, Finland, in 1996 and 2000, respectively. From July 1994 to September 1996, he was a research assistant in communication laboratory of Helsinki University of Technology. Since October 1996, he has been serving as a research scientist in communication laboratory of Helsinki University of Technology. Currently, he is involved in one of the GETA projects, SYTE project, and ACCESS project. His research interests include packet access, radio resource allocation, and queueing theory. He is a member of IEEE.*

FX2. ***Seppo J. Halme*** *was born in Ruokolahti, Finland, on May 17, 1938. He received the degree of Diploma Engineering (Honors) in 1962, and the degree of Licentiate in Engineering in 1966, both in electrical engineering, from the Helsinki University of Technology, Helsinki, Finland, and the PhD degree in communications from the Massachusetts Institute of Technology, Cambridge, in 1970. He was nominated Assistant Professor in 1970 and Professor of Communication Engineering in 1972. He has also served as Dean of Electrical Engineering Department in Helsinki University of Technology. Dr Halme has published over 500 articles, reports, and books.*