

Performance Analysis and Evaluation of Communication Systems

M.S. Okundamiya and C.E. Ojieabu
Department of Electrical and Electronic Engineering,
Ambrose Alli University, P.M.B. 14, Ekpoma, Nigeria

Abstract: A mathematical approach to obtaining closed formula for generating functions of a system's content from which most important performance measures are derived. The derivation of such generating functions was based on Queuing Theory (QT), as a means of analyzing the performance of communication systems for effective data transmission, while asymptotic approximation a standard result in the Theory of Large Deviation (TLD) in the analysis of broadband networks and linearization of their boundaries was adopted due to high transmission rates and stringent Quality of Service (QoS) guarantees of modern systems. The numerical results of the analysis is satisfactory, a confirmation of the validity of derived performance measures.

Key words: Asymptotic, overflow probability, performance models, distribution tails, traffic dynamics, buffers, packet delay

INTRODUCTION

Communication systems are broadly classified as analogue and digital systems and for comparative reasons the digital systems are mostly used. The world is demanding for wireless communication technologies than ever before as more people around the world are subscribing to wireless technologies (Nathan, 2008) hence, the need for the analysis of such systems if it must guarantee the required QoS parameters. Figure 1 shows the incremental progress of wireless technologies.

Performance analysis uncovers several perspectives on a problem or opportunity, determining any and all drivers towards (or barriers to) successful performance and proposing a solution system based on what is discovered. A performance analysis is generally called for in improving a part (or all) of an organization network, ensuring optimum system efficiency and to identify and correct any issues that may negatively impact that efficiency. It also helps the engineer to adjust components in a manner that helps the system makes the best use of available resources.

To obtain the performance measures of interest two techniques are available: simulate the system or solve the model mathematically by analytically computing the performance measures. In this study, we present analytical methods. The research begins with modeling communication systems using generating functions, which considers discrete-time queuing models as a representation of the communication systems with the aim of obtaining closed formula for generating function of a

system's content from which most important performance measures are derived. The main characteristic of this approach is that it is almost entirely analytical.

In computer science and applied mathematics particularly the analysis of algorithms, real analysis and engineering, asymptotic analysis is a method of describing limiting behaviour. Similar limiting behaviour is sometimes expressed in the language of equivalence relations.

An asymptotic function gradually approaches a constant value (the asymptote) after some period of time the function is arbitrarily close to the constant. Due to analytical tractability and potentials for conceptual clarity, tools vital for highlighting the effects of fundamental phenomena in explicit terms asymptotic approximations is also presented. They are used to link the load and resources capacity to the probability of loss due to buffer overflow.

The asymptotic approximations methods have become extremely relevant due to high transmission rates and stringent QoS guarantees of modern systems, which makes buffer overflow significant.

With the realization of multi-service broadband networks, current implementation already serve as high speed backbone infrastructures and more extensive usage accompanied by a further exploitation of these networks advanced capabilities is expected when the need to provide complex information services with increase quality guarantees increases. Hence, performance measures are based on distribution tails instead of the first moments (or mean values). In this study, both

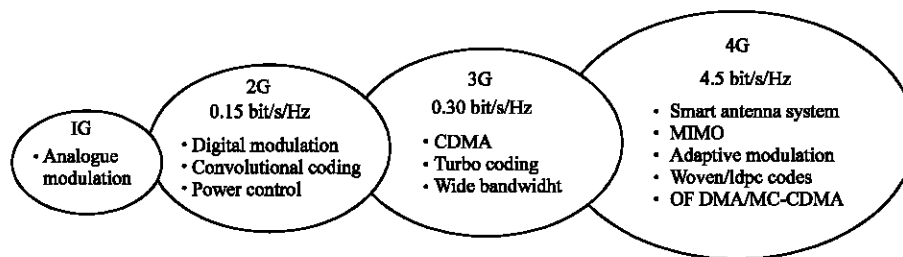


Fig. 1: Incremental progress of wireless technologies

asymptotics for multiplexers with small and large buffers: enabling the establishment of mathematical relation of load and resources capacity to the probability of loss due to buffer overflow is presented.

MATERIALS AND METHODS

In this study, analytical method is presented. We relate the performance of communication systems to the behaviour of buffers as they are temporary storage device facilities for digital information units. The buffers are mathematically analyzed, taking their contents and delay into consideration in the discrete-time and steady-state domains to obtain the required queuing models.

The analysis begins with generating functions in modeling communication systems represented by the discrete-time queuing models in order to obtain closed form formulas. We then advance to asymptotic approximation in evaluating and modeling the performance of such systems due to current and future high transmission rates and stringent QoS guarantees as these approximations are based on distribution tails and the studies asymptotic for both multiplexers with small and large buffers.

Performance modeling of communication systems using generating functions: The performance of communication systems may be closely related to the behaviour of buffers, which can be analyzed using the discrete-time and steady-state models. Buffers are used for temporary storage of digital information units that can not be transmitted to their destination immediately (Okundamiya, 2009; Prasad, 2003).

Analysis of discrete-time queuing models: Discrete-time queuing models are very natural tools for the statistical analysis of the behaviour of buffers in digital communication networks. In these models, the arrival stream of digital information into the buffer is commonly characterized by specifying the numbers of arriving packets or messages (where each message may be composed of a variable number of fixed-length packets) in the consecutive slots.

Queuing theory (Allen, 1990) plays an important role in the modeling and evaluation of communication systems (or networks). Such models appropriately describes traffic and congestion phenomenon since they reflect in a natural way the synchronous nature of modern transmission systems where time is segmented into slots (or interval) of fixed length and information packets are transmitted at a discrete sequence of time points (Okundamiya, 2009). The arrival of stream of digital information into the buffer is commonly characterized by specific numbers of arriving packets during the consecutive slots which are assumed in basic models as Independent and Identically Distributed (IID) discrete random variables and the corresponding arrival process is called independent or uncorrelated arrival process.

The storage capacity of the buffers is usually modeled as unlimited such that the loss probabilities are very small, this assumption acceptable in most communication systems facilitates with analytical analysis techniques. The transmission of information units from the buffer is characterized by the (Prasad, 2003; Hoang and Motani, 2004)

- Number of output channels of the buffers
- Availability of output channels
- Distribution of the transmission times of the information units
- Order of the transmission

In some applications however, such descriptions are not sufficient to characterize the possibly very complicated arrival streams that may occur (for instance, in integrated-services digital networks which carry a great variety of information types and services). Therefore, more advanced models will allow the numbers of arrivals in consecutive slots to be non independent. Such models are referred to as correlated arrival processes.

Teady-state queuing analysis and modeling using generating functions: An overview of the fundamental techniques for the analysis in the steady state of a wide

range of discrete-time queuing models is presented. The main features of this technique are that an extensive use of Probability Generating Functions (PGF) is being made and they are almost entirely analytical (except for a few minor calculations).

Thus, steady-state only exists if the mean number of packet arrivals per slot is strictly less than the mean number of packets per slot that can be transmitted. The behaviour of the queuing system is commonly analyzed in terms of the probability distribution of the buffer content and packet delay.

Buffer content: If S_k is defined as the buffer content of the beginning of slot k then:

$$s_{k+1} = s_k + e_k - t_k \quad (1)$$

Where:

- e_k = The total number of packet arrivals during slot k and the random variable
- t_k = Represents the packets that leaves the buffer system at the end of slot k
- e_k = Depends on the specific nature of the arrival process
- t_k = Depends on the characteristics of the output process and cannot be larger than s_k

Only those packets present in the buffer at the beginning of a slot are eligible for transmission during the slot. In a scenario where permanently available output channels, uncorrelated arrival from slot to slot and constant transmission times of one slot each are assumed, the system Eq. 1 reduces to:

$$s_{k+1} = e_k + (s_k - c)^+ \quad (2)$$

- c = The number of output channels
- e_k and s_k = Statistically independent of each other
- $(...)^+$ = Max. (0, ...)

System Eq. 2 by means of standard z transform techniques (Bruneel, 1993) can be transformed to the z domain to obtain the PGF's for S_{k+1} and $s_k(z)$. If the PGF for s_k is:

$$S_k(z) = E[Z^{s_k}] \quad (3)$$

$$S_{k+1}(z) = E(z)z^{-c} \left\{ \sum_{j=0}^{c-1} (z^c - z^j) \text{Prob}[s_k = j] + S_k(z) \right\} \quad (4)$$

- $E(z)$ = The PGF of the number of packet arrivals in the slot

$S_{k+1}(z)$ and $S_k(z)$ = Will converge in the steady state to a common limiting function $S(z)$

The PGF of the buffer content s at the beginning of an arbitrary slot in the steady state $k \rightarrow \infty$;

$$S_{k+1}(z) = \frac{E(z) \sum_{j=0}^{c-1} (z^c - z^j) \text{Prob}[s = j]}{z^c - E(z)} \quad (5)$$

$\text{Prob}[s = j]$, $0 \leq j \leq c-1$, in Eq. 5 can be determined by invoking the analyticity of the PGF $S(z)$ inside the unit disk $\{z: |z| \leq 1\}$ of the complex z -plane, which implies that any zero of the denominator of Eq. 5 in this area must necessary also be a zero of the numerator together with the normalization condition $S(1) = 1$ of the buffer-contents distribution. This results in a set of c linear equations in c unknown probabilities and allows one to obtain $S(z)$ explicitly.

When the arrival process is correlated, the random variables of Eq. 2 are no longer statistically independent as the set $\{S_k\}$ no longer follow the Markov chain and a more-dimensional state description of the system containing extra information about the arrival process has to be used, hence the above analysis techniques need modification. For a discrete-time queuing model with one output channel ($c = 1$) permanently available using a simple correlated arrival process if the successive on-periods and off-periods of a source are assumed to be independent and geometrically distributed with α and β parameters, respectively, then e_k can be derived from e_{k-1} (Bruneel, 1988) in Eq. 2 as:

$$e_k = \sum_{t=1}^{e_{k-1}} c_t + \sum_{t=1}^{N-e_{k-1}} d_t \quad (6)$$

c_t and d_t are two independent sets of IID and their Bernoulli random variables with PGF becomes:

$$c(z) = 1 - \alpha + \alpha z \quad (7)$$

$$d(z) = (1 - \beta)z + \beta \quad (8)$$

If $P_k(x, z)$ is the joint PGF of the (two-dimensional) state vector (e_{k-1}, S_k) :

$$P_k(x, z) = E[x^{e_{k-1}} z^{s_k}] \quad (9)$$

It was established (Kostovasilis *et al.*, 2002) by means of state equation that:

$$P_k(x, z) = [d(xz)]^N E \left[\left(\frac{c(xz)}{d(xz)} \right)^{s_k-1} z^{(s_k-1)^+} \right] \quad (10)$$

$$P_k(x, z) = \frac{[d(xz)]^N}{z} \left\{ P_k \left(\frac{c(xz)}{d(xz)}, z \right) + (z-1) \text{Prob}[s_k = 0] \right\} \quad (11)$$

A functional equation for the limiting function $P(x, z)$ is obtained as $k \rightarrow \infty$, which typically contains the P-function on both sides with different arguments. A similar analysis is possible for a variety of correlated arrival processes such as the general on/off sources (Wittevrongel and Bruneel, 1997) train arrivals (Wittevrongel and Bruneel, 1998) and the correlated train arrival (De Vuyst and Bruneel, 2001) to mention but a few and an approximation technique (Wittevrongel and Bruneel, 1997) can be used based on the observation that a buffer contents equal to n at the beginning of the slot: implying that not more than n packets have entered the buffer during previous slot.

For more complicated models such as when the general transmission times are considered, additional information about the amount of services already received by the packet(s) in the transmission, if any is needed in the state description.

Packet delay: This refers to the number of slots between the end of the slot of arrival of the packet and the end of the slot when the packet leaves the buffer.

The analysis of packet delay in the case of First-Come-First-Serve (FCFS) queuing, typically involves the derivation of a relationship between a tagged packet delay and the total number of packets present in the buffer just after the arrival slots of the tagged packet and to be transmitted before the tagged packet.

However, for a discrete-time queuing systems with one permanently available output channel, constant transmission times of one slot, a possibly correlated arrival process and a FCFS queuing protocol, the PGF $D(z)$ (Kostovasilis *et al.*, 2002) of the packet delay becomes:

$$D(z) = \frac{S(z) - S(0)}{1 - S(0)} \quad (12)$$

Eq. 12 makes a full delay analysis superfluous once the buffer contents have been analyzed. Comparable relationships such as non-deterministic service times (Vinck and Bruneel, 1996a) and multiple servers (Vinck and Bruneel, 1996b) also exist.

Performance measures: The results of the above analysis is useful in deriving simple and accurate (approximate or exact) formulas as such mean and variance of buffer occupancies and delays and packet loss probabilities are of a wide range of practical importance. The mean system contents and mean packet delay in the steady state can be found by evaluating $S(z)$ and $D(z)$ at $z = 1$. The tail distribution of the buffer contents is often approximated by a geometric form based on the dominant pole z_0 of the PGF of the buffer contents due to computational complexity. For large values of n , tail distribution of the buffer contents is approximated by (Bruneel *et al.*, 1994) as:

$$\text{Prob}[s = n] \approx -\frac{\theta}{z_0 z_0^n} \quad (13)$$

Where θ is the residue of $S(z)$ for $z = z_0$. An approximation for the packet loss ratio can be obtained from Eq. 13 for a buffer with finite waiting space S and the same arrival statistics (Steyaert and Bruneel, 1995) if the probability of the buffer contents (in the infinite buffer) exceeds a given threshold S .

It is not always possible to calculate the PGF $S(z)$ of the buffer contents explicitly as in Eq. 11, a technique has been developed (Wittevrongel and Bruneel, 1997, 2000) to derive results concerning the tail distribution and moments of the buffer contents and associated functional equation, which involves the consideration of those values for which the first argument (s) of $P(z)$ functions in both sides of the function equation becomes equal.

Asymptotic approximation for evaluating the performance of large broadband networks: Due to high speed and the need of proving individualized and stringent QoS guarantees very rare events becomes significant thus most relevant performance metrics must be based on distribution tails rather than mean values. Also, most bandwidth-demanding traffic types appearing on broadband network are volatile, such typical queuing effects is shown in Fig. 2 (Kostovasilis *et al.*, 2002).

Figure 2 depicts the overflow probability at a network multiplexer or switch loaded by superposition of volatile traffic streams as a function of the buffer size. Two distinct regions are clearly identified: In the first region (area with small buffer sizes), the rate correlations do not become apparent as the traffic is primarily characterized by properties of the individual inter-arrival times between successive packets and the overflow probability decays rapidly with increasing buffer size (at an exponentially fast rate since the graph uses a log-linear scale). In the second region (area with larger buffer sizes), the correlation details becomes noticeable resulting

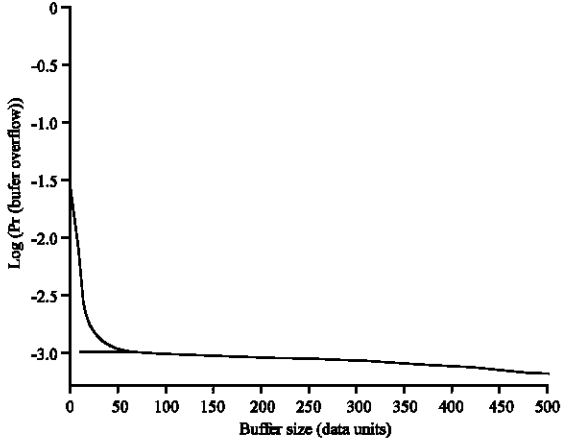


Fig. 2: Graph of overflow probability against buffer size (as a log-linear scale)

in a quite smaller rate of decay of the overflow probability. Accurate prediction of tail probabilities requires the use of sophisticated models being able of providing sufficiently precise characterization of traffic at both packet and burst levels. Such detailed and associated analysis models suffer from the state space explosion problem; precisely the state spaces of models for all (excluding the simplest) traffic patterns have to be rather large in order to captured both packet-level and burst-level behaviour. The scenario is worse when it is realized that in virtually all congestion phenomena of interest, the aggregate traffic load consists of superposition of a large number of individual streams and that the state space of the model for the aggregate traffic depends on the already large spaces of the constituents.

The fluid-flow models of traffic proposed by (Kostovasilis *et al.*, 2002), attempts to partially alleviate this difficulty.

Such an approach claimed to have been quite successfully employed towards the accurate representation of burst-level traffic dynamics with a reduced set of model parameters as illustrated by the dash of Fig. 2, representing overflow probability curve corresponding to the fluid-flow counterpart of the original traffic and which matches quite satisfactorily with the exact result over burst level region. Although, the fluid-flow concept works for reducing the complexity of individual traffic streams model, it can not alleviate the state space explosion due to superposition hence many performance-related mechanisms especially those that must operate within a short time-frame or over a combinatorial large domain cannot rely on classical (or fluid-flow) techniques.

The Theory of Large Deviations (TLD), (Dembo and Zeitouni, 1998) may be used to compute the rate of

exponential decay in the probabilities of interest and to determine the way in which these rare events occur as in congestion-related phenomena such as multiplexing gain and buffering gain. In the absence of significant buffering, multiplexing gain is the only mechanism through which QoS may be attained while using less bandwidth than peak-rate. The two regimes relate to either no buffer or a large buffer so that either multiplexing gain or the buffering gain dominates, respectively.

Asymptotics for multiplexers with small buffers: For a multiplexer with negligibly small buffer and serving traffic through an output link of capacity, C it has been shown that the log of the probability of overflow in the permissible range is:

$$\log \Pr\{r(t) > C\} \leq -\sup_{s \geq 0} \{Cs - \phi(s)\} \quad (14)$$

And that for a stable system ($C > Er(t)$), the maximum over nonnegative real of this Chernoff's bound coincides with the maximum over the entire real line, which implies that the Fenchel-Legendre transform of $\phi(\cdot)$ at C is:

$$\forall C > Er(t) \equiv \bar{r}, \quad \sup_{s \geq 0} \{Cs - \phi(s)\} = \sup_{s \in \mathbb{R}} \{Cs - \phi(s)\} \equiv I(C) \quad (15)$$

The Chernoff bound of Eq. (14) is conservative, allowing for safe performance-related decisions and the bound is asymptotically tight as the number of sources $n \rightarrow \infty$. Applying Cramer's Theorem (Dembo and Zeitouni, 1998).

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \Pr\{r_n(t) > nc\} = -I(c) \quad (16)$$

From Eq. 16 when n is large enough, the probability of overflow is e^ϵ , where the quality level, $\epsilon = nI(c) + o(n)$, which was shown (Kostovasilis *et al.*, 2002) to be:

$$\epsilon = n \left(\beta \log \frac{\beta}{p} + (1 - \beta) \log \frac{1 - \beta}{1 - p} \right) \quad (17)$$

and

$$p < \beta \equiv \frac{C}{nr} < 1 \quad (18)$$

Asymptotics for large buffers: For an effective bandwidth, the bandwidth theory proposed by Kostovasilis *et al.* (2002) is recommended which assumes that for a large (but finite) size buffer, B if the traffic mix contains n_I streams of class I for $I = 1, \dots, k$ then the aggregate Effective Bandwidth Function (EBF):

$$a(\theta) = \sum_i n_i a_i(\theta) \quad (19)$$

the bandwidth requirement becomes:

$$a(\frac{C}{B}) = \sum_{i=1}^k n_i a_i(\frac{C}{B}) \leq C \quad (20)$$

For Markovian on/off fluid models with peak rate r and exponentially distributed on and off sojourns with mean durations τ and σ , respectively Eq. 19 becomes:

$$a(\theta) = \frac{1}{2} \left\{ r - \left(\frac{1}{\tau} + \frac{1}{\sigma} \right) \frac{1}{\theta} + \sqrt{ \left(r - \left(\frac{1}{\tau} + \frac{1}{\sigma} \right) \frac{1}{\theta} \right)^2 + \frac{4r}{\sigma\theta} } \right\} \quad (21)$$

If $\phi_+(s)$ and $\phi_-(s)$ represents, respectively the log-moment generators corresponding to the on and off sojourns for any $\theta > 0$ Eq. (21) reduces to:

$$a(\theta) = \frac{u(\theta)}{\theta} \quad (22)$$

where $u(\theta)$ is the unique positive solution of ;

$$\phi_+(r\theta - u) + \phi_-(-u) = 0 \quad (23)$$

Adopting traffic model in trying to determine the EBF is not always feasible however, alternative approaches (Kostovasilis *et al.*, 2002) by passes modeling by targeting the direct measurement of EBF.

RESULTS AND DISCUSSION

We consider a statistical multiplexer with a variable number of fixed-length packets arrival at a rate of one packet per slot (train arrival), which results in a primary correlation in the packet arrival process.

The arrival process contains an additional secondary correlation of new messages with variable environment having two possible states: A and B, each with geometrically distributed sojourn times (De Vuyst and Bruneel, 2001). Figure 3 show a comparison of the correlated train arrival model $E[s]$ with their uncorrelated counterpart, i.e., uncorrelated train arrival $E[s_{pnm}]$ and uncorrelated packet arrival, $E[s_{un}]$.

In Fig. 3, all the curves for $E[s_{pnm}]$ coincides with the one representing $E[s]$, for $k=1$ uncorrelated environment. For uncorrelated packet arrivals, $E[s]$ slightly increases with higher values of K , although not in a drastic way as $E[s]$ in case of correlated train arrival.

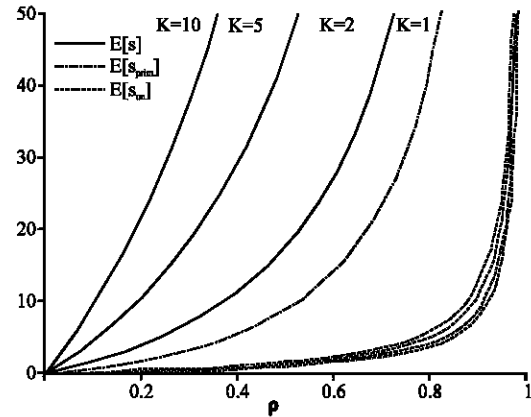


Fig. 3: The graph of mean buffer contents against total load for different values of

In A slot, the number of new messages has geometric distribution with mean 2, while no new message is generated during the B slot. This shows clearly, the severe underestimation of the buffer contents when the different levels of correlation in the arrival process are neglected.

CONCLUSION

In general, we have analytically established simple and accurate (exact or approximate) formulas for a wide variety of useful performance measures of practical interest in communication systems.

REFERENCES

- Allen, A.O., 1990. Statistic, Probability and Queuing Theory with Computer Science Applications. Part II. 2nd Edn., Academic Press Ltd., UK., ISBN: 9780120510511, pp: 245-370.
- Bruneel, H., 1988. Queuing behaviour of statistical multiplexers with correlated inputs. IEEE Trans, Commun. IEEE Trans. Commun., 36: 1339-1341.
- Bruneel, H., 1993. Performance of discrete-time queuing systems. Comput. Operations Res. Comput. Oper. Res., 20: 303-320.
- Bruneel, H., B. Steyaert, E. Desmet and G.H. Petit, 1994. Analytic derivations for tail probabilities for queue lengths and waiting times in atm multi-server queues. Eur. J. Oper. Res., 76: 563-572.
- De Vuyst, S.S.W. and H. Bruneel, 2001. Statistical multiplexing of correlated variable-length packet trains: An analytical performance study. J. Operational Res. Soc., 52: 318-327.

- Dembo, A. and O. Zeitouni, 1998. Large Deviation Techniques and Applications. Chapt. 4. 2nd Edn., Springer, New York, ISBN-13: 9780387984063, pp: 115-125.
- Hoang, A.T. and M. Motani, 2004. Buffer and channel adaptive transmission over fading channels with imperfect channel state information. Proc. IEEE, Wireless Commun. Networking Conf. IEEE WCNC-04, 3: 1891-1896.
- Kostovasilis, K., S. Wittvrongel, H. Bruneel, B. van Houdt and C. Blondia, 2002. Performance of telecommunication systems: Selected topics. Proceedings of IFIP World Congress, August 2002, Canada, pp: 61-93.
- Nathan, P.M.C., 2008. Wireless Communications. Eastern Economy Edn., Prentice-Hall of India, (PHI), New Delhi, pp: 114 - 120.
- Okundamiya, M.S., 2009. Principles of Computer Communication Networks. 1st Edn., Stemic Publishers, Benin City, Nigeria, pp: 1-6.
- Prasad, K.V., 2003. Principle of Digital Communication Systems and Computer Networks. Chapt. 1. 1st Edn., Charles River Media, Cengage, USA.
- Steyaert, B. and H. Bruneel, 1995. Accurate Approximation of the Cell Loss Ratio in Atm Buffers with Multiple Server. Performance Modeling and Evaluation of ATM Networks. Vol. 1, Chapman and Hall, London.
- Vinck, B. and H. Bruneel, 1996a. Delay analysis for server queues. IET Electronics Lett., 32: 802-803.
- Vinck, B. and H. Bruneel, 1996b. Delay analysis for multi-server atm buffers. IET Electronics Lett. Electron. Lett., 32: 1352-1353.
- Wittvrongel, S. and H. Bruneel, 1997. Deriving the tail distribution of buffer contents in a statistical multiplexer with general heterogeneous on/off sources. Proceedings of the International Conference on the Performance and Management of Complex Communication Networks, Nov. 17-21, Tsukuba Science City, Japan, pp: 37-56.
- Wittvrongel, S. and H. Bruneel, 1998. Correlation effects in atm queues due to data format conversation. Performance Eval. PE, 32: 35-56.
- Wittvrongel, S. and H. Bruneel, 2000. Discrete-Time Atm Queues with Independent and Correlated Arrival Streams. Performance Evaluation and Applications of ATM Networks. Chapt. 16, Kluwer Academic Publishers, Boston, ISBN: 978-0-7923-7851-8, pp: 387-412.