Call-level Multi-Rate Teletraffic Loss Models

Half-Day Tutorial by Michael D. Logothetis and Ioannis D. Moscholios
WCL, Dept. of Electrical & Computer Engineering, University of Patras, 265 04 Patras, Greece.
Tel. +30 2610 996 433. Fax: +30 2610 991855, +30 2610 997336. E-mail:{m-logo,moscholios}@wcl.ee.upatras.gr
http://www.wcl.ee.upatras.gr/m-logo/, or http://www.wcl.ee.upatras.gr/m-logo/Publications.htm

Synopsis

Call-level multi-rate teletraffic loss models aim at assessing the call-level QoS of IP based networks with resource reservation capabilities but also for the emerging and future all-optical core networks, such as MPλS/GMPLS, or 3G wireless networks (e.g. UMTS) networks. This assessment is important for the bandwidth allocation among service-classes (guaranteeing QoS), the avoidance of too costly over-dimensioning of the network and the prevention, through traffic engineering mechanisms, of excessive throughput degradation. Despite of its importance, the call-level performance modelling and QoS assessment is a challenge in the highly heterogeneous environment of modern telecom networks, due to the presence of elastic traffic, or complicated call arrival process. The key call-level performance index is the Call Blocking Probabilities (CBP), while the efficient CBP calculation is a sine qua non of a teletraffic model, in order to cope with the high bandwidth capacities of network links. We aim at presenting efficient call-level teletraffic loss models leading to recurrent CBP calculations in multirate systems by taking into account that calls can be distinguished as follows.

- According to the call arrival process, into:
  - (i) Random calls – Random traffic (infinite number of traffic sources).
  - (ii) Quasi-random calls – Quasi-random traffic (finite number of traffic sources)
  - (iii) Batch Poisson arrivals (infinite number of traffic sources). Calls from different service-classes arriving in batches, while batches arriving randomly.
- According to bandwidth requirements upon arrival, into:
  - (i) Calls with fixed bandwidth requirements.
  - (ii) Calls with several, alternative, contingency ("elastic") bandwidth requirements.
- According to their behavior while in service, into:
  - (i) Calls with constant use of their assigned bandwidth (constant-bit-rate) during their lifetime.
  - (ii) Calls that, although constantly use the assigned bandwidth, can tolerate bandwidth compression/expansion while in service.
  - (iii) Calls that alternate between transmission periods (ON) and idle periods (OFF).

Combining the aforementioned differentiations the following six categories of multi-service loss systems have been analyzed:

- A: Random arriving calls with either fixed (certain) or elastic bandwidth requirements upon arrival, and constant use of the assigned bandwidth (constant-bit-rate/stream traffic) while in service.
- B: Random arriving calls with either fixed or elastic bandwidth requirements upon arrival, and elastic bandwidth while in service.
- C: Quasi-random arriving calls with either fixed or elastic bandwidth requirements upon arrival, and constant use of the assigned bandwidth (constant-bit-rate/stream traffic) while in service.
- D: Quasi-random arriving calls with either fixed or elastic bandwidth requirements upon arrival, and elastic bandwidth while in service.
- E: Batched Poisson arriving calls with fixed bandwidth requirements and continuous use of the assigned bandwidth (constant-bit-rate/stream traffic) while in service.
- F: Random or Quasi-random or Batched Poisson arriving calls that, while in service, alternate between transmission periods (ON) and idle periods (OFF).

The springboard of call-level modelling is the *Erlang Multirate Loss Model* (EMLM), in category A (random – stream traffic), where calls of different service-classes arrive to a single link of certain bandwidth capacity according to a Poisson process, and compete for the available link bandwidth under the complete sharing (CS) policy (a call is admitted if link bandwidth is available at the call-arrival time). We call it EMLM, as it provides the same results with the famous Erlang B-Formula, when only one service-class is accommodated to the link. The EMLM has a product form solution (PFS) that leads to an accurate recurrent calculation of the link occupancy distribution and CBP, known as Kaufman-Roberts recursion. An extension of the EMLM allows a blocked call to retry once or several times, requesting for less resources each time (*retry models*). In a further extension, calls arrive to the link with several possible resource requirements, while their request is made according to thresholds (*threshold models*), which indicate the total number of occupied resources. Retry and threshold models do not have a PFS. The *Connection-Dependent Threshold Model* (CDTM) generalizes the *retry* and *threshold models* (as well as the EMLM) by individualizing the thresholds among different service-classes.

In the *Extended-EMLM* (E-EMLM), the first model of category B (random – elastic traffic), although calls request for certain bandwidth upon arrival, bandwidth modification is applied to the bandwidth allocated to all in-service calls. Thus, when a shortage of link bandwidth occurs, a new call can be accepted in the link by compressing the bandwidth allocated to all in-service calls. Similarly, when a call departs, the remaining in-service calls expand their bandwidth. In the *Extended CDTM* (E-CDTM), the call arrival process and the call admission control is the same with the CDTM; however, in service calls can tolerate bandwidth compression/expansion. These models do not have a PFS. The E-CDTM covers in a great extent the notion of elastic traffic.

In category C (quasi-random – stream traffic), the models are based on another extension of the EMLM, the *Engset Multirate Loss Model* (EnMLM), where each service-class has finite population of traffic sources, whereas all other assumptions coincide with those of the EMLM. We characterize it “Engset”, since it provides the same CBP results with the well-known Engset formula for the time congestion probability (CBP that an outside observer measures), when only a single service-class is considered. It is proved that the EnMLM has a PFS. This model is extended to the *Finite-CDTM* (f-CDTM), that is, the call admission control is the same with the CDTM, while the call arrival process is quasi-random.

According to the above models description, category D (quasi-random – elastic traffic) comprises the so called *Extended EnMLM* (E-EnMLM) and its generalization, the *Extended Finite CDTM* (Ef-CDTM).

In category E (Batched Poisson arrivals – stream traffic), calls of the same service-class arrive in batches, while batches arrive randomly. The batch size (number of calls per batch) is arbitrarily distributed. Concerning batch call acceptance, if the available bandwidth does not suffice for all calls comprising the batch to be accepted, the system accepts as many calls as are necessary to occupy the available bandwidth (partial batch blocking discipline). This model, named *Batched Poisson Multirate Loss Model* (BPMLM) has a PFS and therefore the call loss and time congestion probabilities are accurately calculated using a recurrent formula. The BPMLM has been extended to include: 1) non-random batch call arrival dependant on the system’s state and 2) the complete batch blocking discipline, according to which on the occasion that even one call cannot be accepted by the system then all calls comprising the batch are rejected and the batch arrival is unsuccessful.

In category F (ON-OFF calls) the first model is an extension of the EMLM where calls after having accepted for service can alternate between transmission periods (ON) and non-transmission periods (OFF). While entering the system are in state ON; at the end of an ON period the call switches to an OFF period, allowing other calls to take over the previously occupied bandwidth. When the time comes for a call to re-enter the ON state, it once again requests its initial bandwidth. If the requested bandwidth is available, then the call proceeds otherwise it remains at the OFF state (the so-called burst blocking occurs). This *ON-OFF model* has a PFS. The second model of category F allows for blocked calls to retry one or more times in order to enter the system, with reduced bandwidth and increased mean service time requirements (*ON-OFF retry model*). The third model is the ON-OFF model with quasi-random input and is called *Finite ON-OFF model* (f-ON-OFF model). The last model is the *Batched Poisson ON-OFF model* (BP-ON-OFF model), that is, the ON-OFF model with Batched Poisson input process.

To guarantee a certain QoS for each service-class, the *Bandwidth/Trunk Reservation* (BR) policy rather than the CS policy must be applied to a multi-service loss system. We shall present most of the above models under the BR policy.

Basic References

[1] J. S. Kaufman, “Blocking in a shared resource environment”, *IEEE Trans. Commun.* COM 29 (10) (1981) 1474-1481.

[2] J. W. Roberts, “Teletraffic models for the Telecom 1 Integrated Services Network”,Proceedings of ITC-10, 1983.

[3] L.E.N. Delbrouck, “On the steady state distribution in a service facility with different peakedness factors and capacity requirements”, IEEE Trans. Commun. COM-31 (11) (1983) 1209-1211.

[4] G. Stamatelos, J. Hayes, Admission control techniques with application to broadband networks, Comput. Commun., Vol. 17, No. 9, pp. 663-673, September 1994.

[5] J.S. Kaufman and K.M. Rege, “Blocking in a shared resource environment with batched Poisson arrival processes”, *Performance Evaluation,* Vol. 24, issue 4, pp. 249-263, Feb. 1996.

[6] G. Stamatelos, V. Koukoulidis, “Reservation-based bandwidth allocation in a radio ATM network”, *IEEE/ACM Trans. Networking* 5 (3) (1997), pp. 420–428.

[7] M. Mehmet Ali, “Call-burst blocking and call admission control in a broadband network with bursty sources”, *Performance Evaluation*, vol. 38, 1999, pp. 1-19.

[8] M. Stasiak, M. Glabowski, “A simple approximation of the link model with reservation by one-dimensional Markov chain” *Performance Evaluation* 41 (2000) pp. 195-208.

[9] S. Racz, B. P. Gero, G. Fodor, “Flow Level Performance Analysis of a Multi-service System Supporting Elastic and Adaptive Services”, *Performance Evaluation* 49 (2002) pp. 451-469.

[10] I. Moscholios, M. Logothetis, G. Kokkinakis, "Connection Dependent Threshold Model: A Generalization of the Erlang Multiple Service Loss Model," *Performance Evaluation,* Vol. 48, issue 1-4, pp. 177-200, May 2002.

[11] I. Moscholios, P. Nikolaropoulos and M. Logothetis, “Call level blocking of ON-OFF traffic sources with retrials under the complete sharing policy”, *Proc. International Teletraffic Congress, ITC-18,* Berlin, 31 Aug. - Sep. 5, 2003.

[13] I. D. Moscholios and M. D. Logothetis, "An ON-OFF multirate loss model of Poisson traffic under the bandwidth reservation policy", *Proc. 3rd Polish-German Teletraffic Symposium (MMB & PGTS 04),* Dresden, Germany, 12-15 September, 2004.

[14] I. D. Moscholios, M. D. Logothetis and P. I. Nikolaropoulos, "Engset Multi-Rate State-Dependent Loss Models", *Performance Evaluation*, Vol. 59, issue 2-3, pp. 247-277, February, 2005.

[15] I. D. Moscholios, M. D. Logothetis and G. K. Kokkinakis, "Call-burst blocking of ON-OFF traffic sources with retrials under the complete sharing policy", *Performance Evaluation*, Vol. 59, issue 4, pp. 279-312, March, 2005.

[16] I. Moscholios, M. Logothetis and M. Koukias, "An ON-OFF Multi-Rate Loss Model with a Mixture of Service-Classes of Finite and Infinite Number of Sources", *Proc. ICC 2005*, Seoul, Korea, 16-20 May, 2005.

[17] I. Moscholios and M. Logothetis, “Call level blocking of ON-OFF traffic sources in a shared resource environment with batched Poisson arrival processes”, *Proc. International Teletraffic Congress, ITC-19,* Beijing, 29 Aug. - Sep. 2, 2005.

[18] I. Moscholios, M. Logothetis and M. Koukias, “A State-Dependent Multi-Rate Loss Model of Finite Sources with QoS Guarantee for Wireless Networks”, *Mediterranean Journal of Computers and Networks,* Vol. 2, No. 1, pp. 10-20, January 2006.

[19] I. D. Moscholios and M. D. Logothetis, "Engset Multi-Rate State-Dependent Loss Models with QoS Guarantee", *International Journal of Communications Systems,* Vol. 19, Issue 1, pp. 67-93, February 2006.

[20] V. G. Vassilakis, I. D. Moscholios and M. D. Logothetis, “The Extended Connection Dependent Threshold Model for Elastic and Adaptive Traffic”, *Proc. of Communication Systems, Networks and Digital Signal Processing - CSNDSP' 2006*, Patras, Greece, 19-21 July 2006.

[21] V. G. Vassilakis, I. D. Moscholios, M. D. Logothetis and John S. Vardakas, “Call-level Multi-rate Loss Models for Elastic Traffic”, *45th FITCE Congress*, Athens, Greece, 30 August – 2 September 2006.

[22] G. A. Kallos, V. G. Vassilakis, I. D. Moscholios and M. D. Logothetis, “Performance Modelling of W-CDMA Networks supporting Elastic and Adaptive Trafic”, *Proc. 4th International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks (HET-NETs '06)*, Ilkley, West Yorkshire, U.K, 11-13 Sep. 2006.

[23] I. Moscholios, M. Logothetis and M. Koukias, "An ON-OFF Multi-Rate Loss Model of Finite Sources", *IEICE Trans. Commun.*, 2007 (conditionally accepted).

[24] V. G. Vassilakis, I. D. Moscholios and M. D. Logothetis, “The Extended Connection-Dependent Threshold Model for Call-level Performance Analysis of Multi-rate Loss Systems”, *Performance Evaluation* (under review).

Potential audience and prerequisite knowledge

This tutorial will enable engineers and telecom/computer network managers, even those who are not experts in teletraffic to learn and put the latest teletraffic models into practice. Researchers (postgraduate students) will learn how teletraffic models are built and they will be motivated for creating models satisfying new specifications, or for applying the existing teletraffic models to new networking technologies. No special prerequisite knowledge of audience is expected rather than basic teletraffic theory, that is, notion of traffic load and properties, call origination processes, holding (service) time distribution and elementary analysis of Markovian loss systems.

Why the proposed topic is interesting and timely

The access control of different service-classes to network resources and especially the bandwidth allocation among service-classes has been widely recognized as a necessary solution for QoS guarantee in next generation networks and the emerging and future all-optical core networks, such as MPλS/GMPLS. To this end, teletraffic models are of great assistance*;* therefore, they should appeal to a large audience.

Biographical sketch of the speaker

Dr. Michael D. Logothetis was born in Stenies, Andros, Greece, in 1959. He received his Dipl.-Eng. degree and Ph.D in Electrical Engineering, both from the University of Patras, Patras, Greece, in 1981 and 1990, respectively. From 1982 to 1990, he was a Teaching and Research Assistant at the Laboratory of Wire Communications, University of Patras, and participated in many national and EC research programmes, dealing with telecommunication networks, as well as with natural language processing. From 1991 to 1992 he was Research Associate in NTT’s Telecommunication Networks Laboratories, Tokyo, Japan. Afterwards, he was a Lecturer in the Department of Electrical & Computer Engineering of the University of Patras, and since 2003 he is an Associate Professor in the same Department. His research interests include teletraffic engineering, traffic/network control, simulation and performance optimization of telecommunications networks. He organised the 5th international conference CSNDSP 2006. He served/is serving on the Technical Program Committee of several international conferences. He is a member of the IEEE, IEICE, FITCE and the Technical Chamber of Greece (TEE).

Prior tutorial experience

“Optimal Resource Management in ATM Networks”, by M. Logothetis, *Third IFIP Workshop on Performance Modelling and Evaluation of ATM Networks*, Ilkley, West Yorkshire UK, 2-6 July, 1995.