RC25149 (C1104-020) April 18, 2011
Computer Science

IBM Research Report

A Generic Admission Control Methodology for Packet Networks

Chi Harold Liu
IBM Research Division
China Research Laboratory
Building 19, Zhouguancun Software Park
8 Dongbeiwang West Road, Haidian District
Beijing, 100193
P.R.China

Athanasios Gkelias, Kin K. Leung
Imperial College
London, SW7 2BT
UK

A Generic Admission Control Methodology for Packet Networks *

Chi Harold Liu
IBM Research - China
Beijing, 100193, China
chiliu@cn.ibm.com

Athanasios Gkelias
Imperial College
London, SW7 2BT, UK
a.gkelias@imperial.ac.uk

Kin K. Leung
Imperial College
London, SW7 2BT, UK
kin.leung@imperial.ac.uk

ABSTRACT

In packet networks with multiples concurrent data connections, admission control (AC) is required to guarantee quality of service (QoS). The admission of a new connection requires the knowledge of the network capacity as *a priori*. However, the estimation of network capacity and more important its parametrization by a single or a set or variables is not feasible in most networks, due to different operational characteristics of the underlying communication protocols, channel and mobility statistics (in wireless networks), data traffic dynamics, and multiple QoS requirements. In this work, to tackle these challenges, the network between any source and destination node pairs is modeled as a "black box", and a generic mathematical function is used to map the input variables to a single output parameter that captures the whole set of QoS requirements. A generic AC (GAC) methodology is proposed, that by using the Taylor expansion, predicts the impact of the potential admission of a new connection on the QoS experience of existing connections.

1. INTRODUCTION

Modern internet applications and services must simultaneously satisfy and support multi-dimensional QoS metrics, such as, end-to-end (ETE) throughput, packet delay, packet-error-rate (PER), etc; finding the optimal resource allocation to jointly satisfy multiple QoS dimensions is NP-hard [3]. On the other hand, the limited network resources cannot allow arbitrarily large number of connections with strict QoS requirements admitted into the network, since they can easily jeopardize the QoS of ongoing connections and make the overall network inefficient and fragile. Moreover, the estimation of network capacity is known to be a difficult and tedious task that requires various assumptions and abstractions even for simple networks which makes the use of any mathematical capacity formulation impractical. This becomes even more challenging in the case of ETE heterogeneous networks where several networking standards and protocols, highly dynamic traffic with multiple QoS metrics, wireless channel statistics and mobility patterns need to be considered. All these challenges are the main motivation of this work.

The main contribution of this work is the development of a novel admission control framework for any generic packet network with multi-dimensional QoS metrics. First, a novel concept of *the QoS performance index* is introduced to aggregate the multiple QoS metrics into a single parameter. Second, a generic mathematical process, based on Taylor approximation, is used to model any unknown network as a "black box" where no knowledge of the inherent network parameters is required. Multiple variables can be used as inputs to the black box model, while an estimated value of the QoS performance index produced as a single output. Third, the admission of a new connection is then based on the impact that it will have on the output of the black box model, i.e, on the estimated value of the QoS performance index.

Related Work: Much work has been done for AC algorithm in a variety of packet networks [1, 2]. Recently, [6] studied the design of optimal joint AC and routing to maximize the overall "revenue" but the QoS guarantees are not addressed. In [5], an adaptive admission control (AAC) protocol estimates the resource availability in contention-based WLAN MAC layer to control QoS. [7] estimates the achievable capacity with only packet loss constraint, assuming traffics arrive according to Gaussian distribution. In a summary, existing AC algorithms do not provide the degree of transparency to the underlying communication protocols in use and AC schemes are developed for specific network settings. Furthermore, none of these have accurately estimate the admission impact in terms of QoS experience; and these become the central of our paper and the research path in general.

The rest of this paper is organized as follows. The system model is presented in Section 2, including the the QoS performance index and the mathematical presentation of our black box model. The admission control procedure is demonstrated in Section 3 and numerical results and detailed analysis are given in Section 4. Conclusions are drawn in Section 5.

2. SYSTEM MODEL

Consider an ETE packet network that comprises a set of (possibly heterogeneous) subnetworks, as shown in Fig-

*The major part of this work has been done during Chi Harold Liu's evolvement at Imperial College, UK. This work is submitted to IFIP Performance 2011 for review.

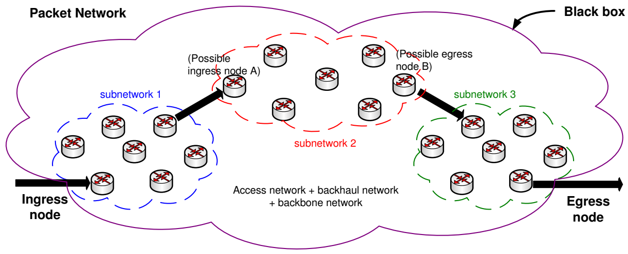

Figure 1: Black box model for packet networks.

ure 1. For instance, each of them can be an access network (wired or wireless), backhaul network, backbone network, etc. Consider an ingress node that attempts to initiate a data connection with an egress node, routing through the set of different subnetworks, and let $q(D_q^r, T_q^r, E_q^r)$ represent the new data connection with certain ETE packet delay D_q^r, throughput T_q^r and PER E_q^r QoS requirements.

To overcome the difficulties of dealing with a variety of operational communication protocols in different subnetworks, we treat the ETE packet network between a pair of ingress and egress nodes as a black box as shown in Figure 1, where the detailed operational contexts for protocols beneath the application layer are not transparent to its I/O. We opt to adopt a runtime analysis such that connections are *probed* (or monitored) by the ingress and egress nodes for the satisfactory QoS experience. In other words, the benefit a packet network can provide to each served connection across several heterogenous subnetworks is constantly monitored at the egress node and informs back to the ingress node through the feedback loop, so that the availability of the current ETE packet network to support certain QoS is known.

The input parameters to the black box model may include, but not limited to, the number of ongoing data connections, multiple QoS metrics, etc., while a single parameter is used at the output of the black box which reflects the degree of QoS satisfaction at the egress node. This output parameter is called *QoS performance index* and will be described in the following section. We emphasize here that the applicability of the proposed black box model is not limited to the considered generic ETE network but can be applied to any single subnetwork (or subset of subnetworks), as shown in Figure 1.

2.1 QoS Performance Index

In order to jointly consider multiple QoS metrics, we are using the novel utility function based on the *QoS outage ratio* R parameter:

$$R_q^D = \beta_D \frac{D_q^a}{D_q^r}, R_q^T = \beta_T \frac{T_q^r}{T_q^a}, R_q^E = \beta_E \frac{E_q^a}{E_q^r}, \quad \forall q \in \mathcal{Q}, \quad (1)$$

where $\mathcal{Q}$ is the set of connections currently serviced in the network. The subscripts a and r denote the attained and required values, respectively, and $\{\beta_D, \beta_T, \beta_E,\} \geq 1$ are a set of error margins to provide a safe guard for imperfect resource estimations and system fluctuations.

Let us now introduce the *QoS performance index* $\mathrm{I}(q)$, a simple, but representative and quantitative, scalar to uniquely identify the level of overall QoS performance of connection q:

$$\mathrm{I}(q) = g\Big(R_q^D, R_q^T, R_q^E\Big), \quad \forall q \in \mathcal{Q}. \quad (2)$$

For instance, $g(\cdot) \triangleq \max(\cdot)$ can be one of many realizations of $g(\cdot)$, and used later.

LEMMA 2.1. *For any connection $q \in \mathcal{Q}$, its multi-dimensional QoS requirements are simultaneously satisfied within a packet network if and only if* $\mathrm{I}(q) \in [0, 1]$.

For every connection q in the network, the index $\mathrm{I}(q)$ is constantly monitored by the egress node of ongoing connection q and passed to the corresponding ingress node through a feedback loop. The ingress node will keep the most recent values of $\mathrm{I}(q)$ stored in I_t:

$$\mathrm{I}_t = \gamma \mathrm{I}(q) + (1-\gamma)\mathrm{I}_{t-1}, \forall q \in \mathcal{Q}, \quad (3)$$

where $\gamma \in (0, 1)$ is the weight factor and $\mathrm{I}_0 = 0$. In our proposed model, every egress node constantly monitors the QoS performance index of each received data connection, given by (2), to obtain an updated value $\mathrm{I}(q)$, and passes the value to the ingress node through a feedback loop. The ingress node calculates the exponentially smoothed value according to (3) and keeps a record of the most recent values of I_t to be used for the AC process. It is worth noting that the main essence of (2) is to hide the heterogenous network protocols in use, but only use attained and required QoS values to calculate a scaler. Therefore, index $\mathrm{I}(q)$ represents and reflects the degree of resource availability for maintaining certain QoS performance in real-time, in a simple mathematical way.

2.2 Mathematical Model

In the following we introduce a generic mathematical model that will be used later for the network capacity estimation and AC. Let us consider a generic function $f : \mathbb{R}^M \to \mathbb{R}$, which maps a M-dimensional input vector $\underline{x}_t = (x_t^1, x_t^2, ..., x_t^M)$ $\in \mathbb{R}^M$ to a scalar output $y_t \in \mathbb{R}$, i.e.,

$$y_t = f(\underline{x}_t). \quad (4)$$

We characterize any change in the input variables $\underline{x}_t$, by $\Delta\underline{x}_t = (\Delta x_t^1, \Delta x_t^2, ..., \Delta x_t^M)$, which will result to a change of the output to

$$\tilde{y}_t = f(\underline{x}_t + \Delta\underline{x}_t). \quad (5)$$

Since f is an unknown generic function, the exact mathematical mapping can not be expressed in a closed-form expression. However, an approximation of y_t can be obtained by considering a Taylor expansion of $f(\underline{x}_t)$. Only the first order (the long-term average) and second order (the fast change) derivatives are used for the approximation:

$$\tilde{y}_t \approx f(\underline{x}_0) + (\Delta\underline{x}_t^i)^T \mathrm{D}f(\underline{x}_0) + \frac{1}{2}(\Delta\underline{x}_t^i)^T \mathrm{D}^2 f(\underline{x}_0)\Delta\underline{x}_t^i, \quad (6)$$

where $\mathrm{D}f(\underline{x}_0)$ and $\mathrm{D}^2 f(\underline{x}_0)$ are the gradient and Hessian matrix of f evaluated at $\underline{x}_0$, respectively. These two values can be approximated through the limit definition, i.e., first order derivatives can be derived by using three adjacent measurements on the shape of curve produced by the mapping f, while second order statistics can be derived by four adjacent measurements.

The generic function f will be used in the following to model any generic ETE network as a black box, and will

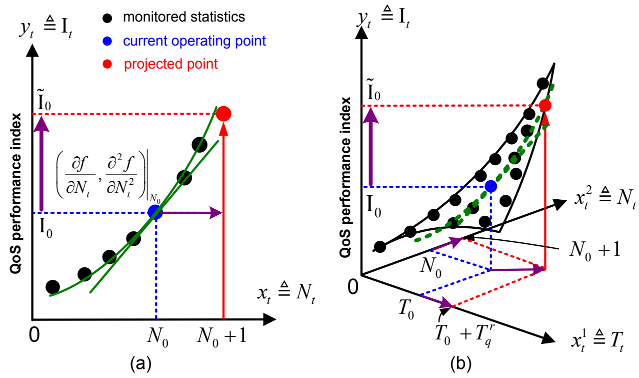

Figure 2: An illustrative example for admission estimation on mapping f, where (a) scaler input N_t, i.e., $M = 1$, and (b) two-dimensional inputs, i.e., $M = 2$.

enable us to map multiple network input parameters to a single scalar output related to network capacity that can be used for admission control.

3. GENERIC ADMISSION CONTROL

In this section, the mathematical model of Section 2.2 is used to estimate the impact of different network parameters on the QoS performance index $\mathrm{I}(q)$, and allow us to perform admission control in the network.

Single input parameter: For illustration purposes, we start by using a single scaler input as an example, i.e., the number of ongoing connections $\underline{x}_t \triangleq N_t \in \mathbb{R}$ being serviced in the network (see Figure 2(a)), and the mapping f becomes $\mathrm{I}_t = f(N_t)$. Let us now consider N_0 connections simultaneously being serviced in an ETE network, and a new connection $q(D_q^r, T_q^r, E_q^r)$, arrives at an ingress node, awaiting for admission. The possible admission of the new connection will result in an input change of $\Delta x_0 = \Delta N_0 = 1$ which will impose a change of the output to $\widetilde{\mathrm{I}}_0 = f(N_0 + 1)$. An estimation of the new network output can be obtained by using (6), and is given by

$$\widetilde{\mathrm{I}}_0 \approx \mathrm{I}_0 + \frac{\partial f}{\partial N_t} + \frac{1}{2}\frac{\partial^2 f}{\partial N_t^2}, \tag{7}$$

where the partial derivatives are taken at the current operating point N_0. This is the expected new *QoS performance index* for the ingress node caused by the admission of the new data connection with the given QoS metrics. Admission control will be performed based on Lemma 2.1, to verify if $\widetilde{\mathrm{I}}_0 \leq 1$, which ensures that there are adequate network resources to incorporate the admission of the new connection. If $\widetilde{\mathrm{I}}_0 \leq 1$ the new connection will be admitted to the network, otherwise the connection will be rejected.

Multiple input parameter: Let us now demonstrate how our model will perform for multiple input parameters. For instance, let us consider inputs as both the number of ongoing connections N_t and the total served throughput T_t, i.e, $\underline{x}_t = (N_t, T_t) \in \mathbb{R}^2$ and,

$$\mathrm{I}_t = f(N_t, T_t), \tag{8}$$

as shown in Figure 2(b). Similar to the single input case, a potential new connection admission will result in an input change of $\Delta\underline{x}_0 = (1, T_q^r)$ and a change of the output to $\widetilde{\mathrm{I}}_0 = f(N_0 + 1, T_0 + T_q^r)$. An estimation of the new network output can be obtained by using (6), and is given by (due to the space limit, we omitted the detailed derivation of the expansion):

$$\widetilde{\mathrm{I}}_0 \approx \mathrm{I}_0 + \frac{\partial f}{\partial N_t} + T_q^r\frac{\partial f}{\partial T_t} + \frac{1}{2}\frac{\partial^2 f}{\partial N_t^2} + \frac{3}{2}(T_q^r)^2\frac{\partial^2 f}{\partial T_t^2}, \tag{9}$$

where partial derivatives at taken at state $\underline{x}_0 = (N_0, T_0)$, and if this connection is accepted, the packet network would operate at state $(\underline{x}_0 + \Delta\underline{x}_0)$. In network scenarios, where the shape of the curve produced by the mapping f is smooth enough around current operating point $\underline{x}_0 = (N_0, T_0)$, the second order derivatives will become negligible and the only first order statistics are sufficient. In this case we may simplify (9) as:

$$\widetilde{\mathrm{I}}_0 \approx \mathrm{I}_0 + \frac{\partial f}{\partial N_t} + T_q^r\frac{\partial f}{\partial T_t}. \tag{10}$$

Again, the admission control will be based on Lemma 2.1, to verify if $\widetilde{\mathrm{I}}_0 \leq 1$.

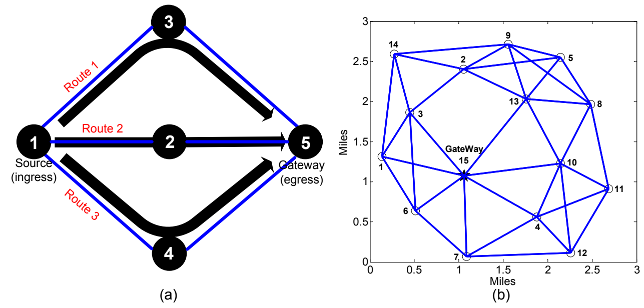

Figure 3: (a) A five-node WMN setting. (b) A complete simulation setting.

4. SIMULATION RESULTS

Without loss of generality, a wireless mesh network (WMN) is used as evaluation platform where the integrated QoS scheduling and routing protocol (IQoSR, [4]) is used in network and MAC layers to provide sub-optimal solution for QoS. Rayleigh fading channel model, adaptive modulation and coding scheme, and directional antennas are used in PHY Layer. We first assess our GAC methodology in a simple five-node WMN setting as shown in Figure 3(a), where node 1 serves as the ingress node to generate connections and node 5 serves as the egress node as the intended receiver. From Table 1, we observe that higher order statistics successfully improve the volume of maximum supported throughput and connection number by 13% and decrease the prediction error, QoS outage, and blocking probabilities, since higher order statistics aids to admit the most *appropriate* connection (in term of throughput requirement) with the knowledge of satisfactory QoS criterion in Lemma 2.1, while maintaining QoS satisfactions to all ongoing connections.

A complete simulation topology is shown in Figure 3(b), where our algorithm, "IQoSR+GAC", is compared with "IQoSR" alone, SAC ([7], "IQoSR+SAC"), and round robin scheduler and AODV routing protocol. Figure 4(a) shows that "IQoSR+GAC" outperforms all other schemes in terms of the overall gateway goodput even for small traffic inter-

Table 1: Effects of using different combinations of partial derivatives for admission estimation

	First order	Higher order
Max. Supported Thrpt	22Mbps	25Mbps
Error Bound	±2Mbps	±500Kbps
Max. No. of Connections	30	35
Error Bound	±5	±2
QoS Outage	≈ 13%	≈ 10%
Blocking Probability	≈ 12%	≈ 8%

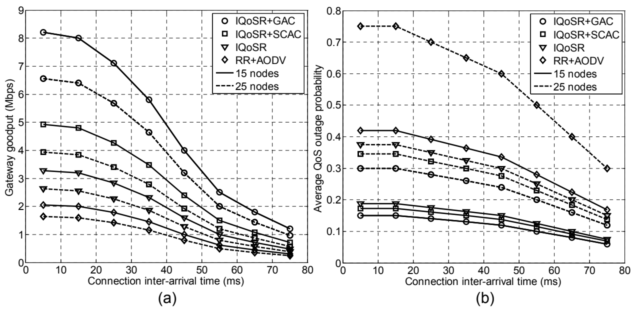

Figure 4: (a) The overall gateway goodput, and (b) the average QoS outage probability, w.r.t. different connection inter-arrival time and node density.

arrival time (heavy load conditions), i.e., 1.7 times, 2.5 times, and 4.1 times gains are achieved if compared with "IQoSR +SAC", "IQoSR", and "RR+AODV", respectively. Due to the assumption of Gaussian arrival process and only PLR considered, we find "SAC" scheme sometimes makes wrong admission decisions which turns into less goodput and higher QoS outage. Furthermore, we observe that the gateway goodput saturates when the traffic load becomes higher. Finally, when we increase the node density in a fixed network area, gateway goodput decreases by 20%. Figure 4(b) illustrates the QoS outage probability, defined as the probability of any connection's QoS requirements to fail during their lifetime. Our proposal guarantees 85% of the QoS satisfaction, compared with only 81% for "IQoSR", 82% for "IQoSR+SAC", and 58% for "RR+AODV". This is because the impact of the newly admitted connections on existing ones' QoS experience has been accurately estimated in the parameter of the updated QoS performance index $\widetilde{\mathrm{I}}_0$ which reflects the new QoS experience the ETE packet network can provide to all connections, new and old.

5. CONCLUSIONS

In this paper, a novel AC methodology is proposed that can apply to any packet network or set of heterogeneous networks. A new QoS performance index was introduced to combine multiple QoS metrics into a single variable, the abstraction of the packet network by a black box has been proposed, and a generic mathematical process is used to represent it. In this way, the heterogenous operational features of the inherent network protocols in use are hidden, so that a degree of transparency is successfully provided to end nodes. AC decision is made for the new data connections by estimating the potential impact of the new connection on the QoS performance index by Taylor expansion in limited orders. Extensive simulations showed significant performance gains compared with other schemes, i.e., we accommodate higher number of connections with satisfactory QoS.

Acknowledgement

This research was sponsored by US Army Research laboratory and the UK Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US Army Research Laboratory, the U.S. Government, the UK Ministry of Defense, or the UK Government. The US and UK Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

6. REFERENCES

[1] M. H. Ahmed. Call admission control in wireless networks: a comprehensive survey. *IEEE Comm. Surveys & Tutorials*, 7(1):49–68, Qtr. 2005.

[2] O. Baldo. A cross-layer distributed call admission control. In *IEEE WIMOB'09*, pages 441–446, 2009.

[3] C. Lee, J. Lehoczky, D. Siewiorek, R. Rajkumar, and J. Hansen. A scalable solution to the multi-resource qos problem. pages 315–326, 1999.

[4] C. H. Liu, A. Gkelias, and K. K. Leung. A cross-layer framework of QoS routing and distributed scheduling for mesh networks. In *IEEE VTC'08 Spring*, pages 2193 –2197, Singapore, May 2008.

[5] Q. Shen, X. Fang, P. Li, and Y. Fang. Admission control for providing QoS in wireless mesh networks. In *IEEE ICC'08*, pages 2910–2914, May 2008.

[6] S. Zhang, F. R. Yu, and V. C. M. Leung. Joint connection admission control and routing in ieee 802.16-based mesh networks. In *IEEE Trans. on Wireless Comm.*, volume 9, pages 1370 –1379, 2010.

[7] H. Zhu, V. O. K. Li, Z. Ma, and M. Zhao. Statistical connection admission control framework based on achievable capacity estimation. In *IEEE ICC'06*, volume 2, pages 748–753, Istanbul, June 2006.