A Decision Theoretic Approach to Measurement-based Admission Control

Jungwon Ko, Thomas M. Chen
Department of Electrical Engineering
Southern Methodist University
Dallas, TX P.O. Box 750338
jwk@engr.smu.edu, tchen@engr.smu.edu

***Abstract*— This paper presents a new approach to measurement-based admission control (MBAC) using real-time traffic classification. The network traffic is classified based on Bayesian decision theory. We derive the decision rules and cost functions considering utilization, packet loss rate, and rejection rate. The approach is compared by simulations against other measurement-based admission control approaches. Experiments show that the new method offers significant advantages particularly when the traffic is highly dynamic or complex.**

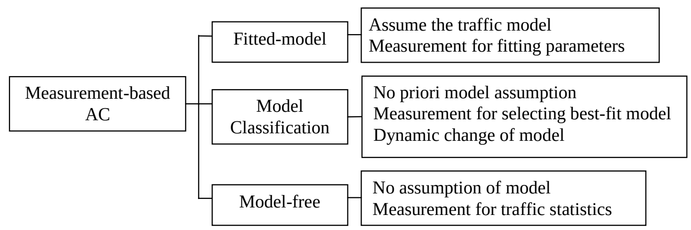

Fig. 1: Categories of MBAC approaches

I. Introduction

Regulation of new traffic flows is one of the primary means of preventing network congestion and ensuring quality of service (QoS). The basic challenge in admission control is efficient resource allocation, balancing the goal to admit more traffic for efficient bandwidth utilization against the need to reject new traffic to protect QoS guarantees. Although deterministic bounds can ensure QoS, resource allocation is conservative leading to low utilization [1]. For more efficient resource utilization, many stochastic traffic models have been developed based on historical data [2]–[6]. Queueing analysis for specific traffic models can provide accurate resource allocation if the traffic model is valid. However, it is generally difficult to know the stochastic properties of a new traffic flow *a priori* when an admission decision must be made.

Instead of assuming traffic models, measurement-based admission control (MBAC) approaches use observations of the current traffic to make decisions for admitting or rejecting new traffic. MBAC approaches can be classified into three categories as shown in Fig. 1: fitted-model, model-free, and model classification.

The fitted-model approach starts by assuming a traffic model and then finds the best-fit parameters from the observed real traffic [7]–[10]. The fitted-model approach makes use of the concept of equivalent bandwidth (EB) which is an estimate of the resources required to meet QoS requirements for statistically multiplexed traffic flows. The EB depends on the statistical characteristics of traffic sources. Guerin et al. derived EB from a two-state fluid flow model [7]. Some EB calculations assume Markovian traffic sources without long-range dependence [8], [9]. Other EB calculations are appropriate for self-similar traffic, e.g., Norros assumed fractional Brownian motion (fBm) [10], [11]. The fBm traffic was considered as input to a queue to estimate the packet loss caused by a new connection. This approach can be a good choice if the traffic structure is known and unchanging, but could lead to erroneous decisions if the assumed traffic model is not valid.

In contrast, the model-free approach does not assume any *a priori* traffic model but uses general traffic statistics [12]–[14]. Jamin et al. used traffic measurements to refine token bucket filter parameters originally declared by a traffic source [12]. Grossglauser and Tse estimated the traffic statistics (e.g. mean and variance) from traffic measurements and estimate the loss probability using a normal approximation [13], [14]. Qiu used measurements of the maximal traffic envelopes of the aggregate traffic [15]. Duffield et al. calculated packet loss in a buffer based on large deviation theory [16]–[18]. Gibbens et al. made admission decisions by comparing the current measured traffic load with a precomputed threshold [19]. The threshold was chosen to maximize the reward of increased utilization against the penalty of packet loss. The model-free approach is appealing because it completely avoids the problem of traffic modeling. However, it is doubtful whether general statistics can represent traffic characteristics as accurately as a traffic model.

The model classification MBAC approach in this paper differs from the other two MBAC approaches. We use traffic modeling but do not assume an *a priori* model. Instead, we select the best-fit traffic model from a “library”of candidate models based on the observed traffic. After a traffic model is selected, it is applied to a queueing system to decide on the acceptance or rejection of new traffic flows. While traffic model classification has been proposed before [20], [21], we present a new method here following a decision theoretic

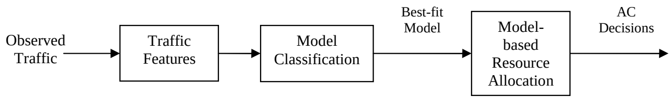

Fig. 2: MBAC using model classification

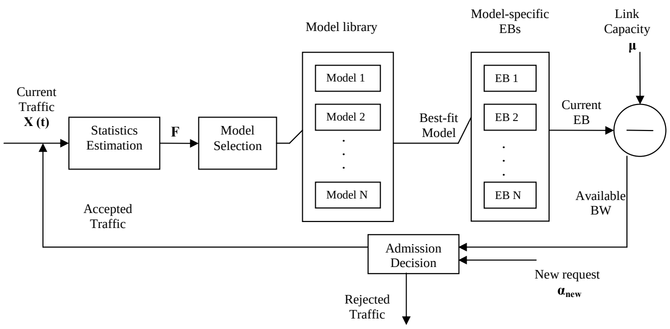

Fig. 3: Model library for MBAC

approach [22], [23]. The best-fit model is selected based on optimizing a Bayesian cost function. As traffic is continually monitored, the selected traffic model may change to another candidate in the model library. At any time, the selected traffic model represents our best knowledge of the traffic.

The remainder of this paper is organized as follows. Section II describes our general approach and an overview of its operation. We derive a set of Bayes decision rules and cost functions to optimize the admission decisions. Section III presents the experimental results comparing our MBAC approach with two other MBAC approaches (fitted-model and model-free) using different types of simulated traffic.

II. Measurement-based AC

An overview of our decision theoretic approach to MBAC system is shown in Fig. 2. The first stage of the system extracts features of the observed traffic as the basis to decide which traffic model among the candidate models is the best-fit model. Each candidate traffic model is associated with a model-specific EB estimate as shown in Fig. 3. Selection of the best-fit model is chosen according to the decision theoretic formulation described below. As the real traffic is monitored continuously, the best-model selection is recalculated periodically. The model-specific EB is used to estimate the available bandwidth, which is compared to the bandwidth requirements of new traffic flows to make acceptance and rejection decisions.

A. Overview of System

Our MBAC system is shown in Fig. 3. New traffic flows can not be observed before their acceptance, so their bandwidth requirements are based on assumptions. After new flows are accepted, then they can be observed and their EB can be determined more accurately. As an initial estimate, we assume that new traffic flows are regulated by a leaky bucket with parameters $(\lambda_p, \lambda_s, b_s)$ where λ_p is the peak rate, λ_s is the mean rate and b_s is the maximum burst size. A regulated traffic flow is approximated by an on-off process with exponential on periods with mean length $1/b$ and exponential off periods with mean length $1/a$, where

$$a = \frac{\lambda_p \, \lambda_s}{b_s \left(\lambda_p - \lambda_s\right)} \text{ and } b = \frac{\lambda_p}{b_s} \quad (1)$$

The effective bandwidth for a new traffic flow is

$$\alpha_{new} = \frac{\lambda_p \theta - (a+b) + \sqrt{\left(\lambda_p \theta - a - b\right)^2 + 4a\lambda_p \theta}}{2\theta} \quad (2)$$

where $\theta = \frac{-\log(\epsilon)}{B}$, ϵ is the target QoS, and B is buffer size [7].

The number of admissible new traffic flows depends on the available bandwidth. The current consumed bandwidth is determined by choosing a model from a library of N candidate traffic models. The traffic models can be represented as the stochastic processes $X_j(t)$ for $j = 1, 2, ..., N$. Each traffic model is associated with a model-specific EB. If the cumulative amount of traffic from a source up to time t, $X(t)$, is a random process with stationary increments with $\hat{X}_m$ representing the amount of traffic that has arrived in the time period m ($X(t) = \sum_{m=1}^{t} \hat{X}_m$), the EB is

$$\alpha_X(s,t) = \frac{1}{st} \log E\left[e^{sX(t)}\right] \quad (3)$$

where s and t are the parameters for space scale and time scale, respectively [24]. The space and time scales of interest depend not only on the source characteristics, but also on the link capacity μ, buffer size B, and target QoS (buffer overflow probability) ϵ. If a constraint for buffer overflow probability is

$$P\left(Q(\infty) > B\right) \approx e^{-I} \leq \epsilon, \quad (4)$$

where

$$I = \inf_{t \geq 0} \left(\sup_{0 \leq s} \left(s(B + \mu t) - st\alpha_X(s,t) \right) \right) \quad (5)$$

the space and time scales of interest [25] are the solutions to

$$\sup_{t \geq 0} \left(\inf_{0 \leq s} \left(st\alpha_X(s,t) - s(B + \mu t) \right) \right) \approx \gamma = \log(\epsilon) \quad (6)$$

The solution of (6) denoted as (s^*, t^*) is referred to as the critical point of the system. At the critical point, the traffic source behaves as if it were a constant traffic stream of rate $\alpha_X(s^*, t^*)$. Each traffic model $X_j(t)$ in the library is associated with an EB $\alpha_j(s^*, t^*)$.

B. Model Selection

The best-fit model is chosen at the decision times $\{T_i : i = 0, 1, 2, ...\}$ where $T_{i+1} = T_i + \Delta$. The time window Δ determines the rate for updating the traffic model and model-specific EB. At each decision time, we have w samples of the observed traffic represented by the data vector $\mathbf{x} = (x_1, x_2, ..., x_w)$ where $x_i := \sum_{m=(i-1)\tau+1}^{i\tau} \hat{X}_m$ for $i = 1, 2, ..., w$.

The observed data $\mathbf{x}$ is vector random variable with given conditional probability distribution $p(\mathbf{x}|X_j)$, for $j =$

$1, 2, ..., N$. If we denote by $P(X_j)$ the *a priori* probability that the true model is X_j, then the *a posteriori* probability of model X_j is given by Bayes formula

$$P(X_j|\mathbf{x}) = \frac{P(X_j)p(\mathbf{x}|X_j)}{p(\mathbf{x})} \tag{7}$$

where

$$p(\mathbf{x}) = \sum_{j=1}^{N} P(X_j)p(\mathbf{x}|X_j) \tag{8}$$

is the probability distribution function of $\mathbf{x}$.

For a given observed data $\mathbf{x}$ at time T_i ,we can take one action from the set of possible actions $\mathbf{A} = \{a_1, a_2, ..., a_N\}$ where action $a_j \triangleq$ Choose X_j. We define $\mathcal{C}_{jk}$ as the *cost* for choosing action a_k when the real traffic model is X_j (and the correct action is a_j). We define the *conditional risk* associated with the particular action a_k as

$$r_k(\mathbf{x}) = r(a_k|\mathbf{x}) = \sum_{j=1}^{N} \mathcal{C}_{jk} P(X_j|\mathbf{x}) \tag{9}$$

We formulate a decision rule $a(\mathbf{x})$ which assigns one of actions from $\mathbf{A}$ given the data observed vector $\mathbf{x}$. We are interested in a decision rule to minimize the *total risk*

$$R = \int_{\mathbf{R}^w} r(a(\mathbf{x})|\mathbf{x})p(\mathbf{x})\, d\mathbf{x}. \tag{10}$$

This is achieved by applying the general *Bayes decision rule* which can be stated as: Given a sample vector $\mathbf{x}$, define

$$a(\mathbf{x}) = \min_{1 \le k \le N} \{r(a_k|\mathbf{x})\} \tag{11}$$

that is, $a(\mathbf{x})$ is the action which minimizes the total risk (10). The minimum risk associated with Bayes decision rule is called the *Bayes risk*.

Theorem 1 (Bayes decision rules for minimum risk): Given a vector $\mathbf{x}$, the Bayes decision rules can be stated as follows: $\mathbf{x}$ is model X_i if

$$r_i(\mathbf{x}) < r_k(\mathbf{x}) \text{ for all } k \text{ except } i \tag{12}$$

Proof: [22], [23] ■

Bayes decision rule assigns to each $\mathbf{x}$ the action a_i (choose the traffic model X_i) from $\mathbf{A}$ with minimum conditional risk r_i. The observed data $\mathbf{x}$ is classified as process X_i from the library of model candidates.

C. Cost Functions

The key element of the decision rule is the cost functions. We consider that a reward ($\mathcal{R}_1$) is gained for accepting traffic satisfying the QoS. At the same time, penalties are imposed for rejecting traffic (penalty $\mathcal{P}_1$) or traffic violating the QoS (penalty $\mathcal{P}_2$). One expects the the penalty in the case of admitting too many new requests which will induce QoS to be greater than the penalty associated with rejecting requests that could have been admitted will under-utilize network resource, i.e. $\mathcal{P}_2 \ge \mathcal{P}_1$.

Suppose new requests arrive according to a Poisson process with rate λ. If the best-fit model to the observed traffic is X_j, the current bandwidth can be estimated to be α_j, and the probability of QoS violation can be calculated to be ϵ_j. The number of admissible new flows during the interval (T_i, T_{i+1}), given chosen model X_j, will be

$$n_j = \left\lfloor \frac{\mu - \alpha_j}{\alpha_{new}} \right\rfloor \tag{13}$$

where $\lfloor y \rfloor$ is the largest integer number less than y.

The cost functions $\mathcal{C}_{jk}$ for choosing model X_k instead of X_j at T_i can be defined as follows:

1) If $j = k$, there is no cost for right decision.

$$\mathcal{C}_{jk} = 0 \tag{14}$$

2) If $j \neq k$, The cost functions $\mathcal{C}_{jk}$ become

$$\mathcal{C}_{jk} = \mathcal{P}_1 A_k\, \epsilon_j + \mathcal{P}_2 R_k - \mathcal{R}_1 A_k (1 - \epsilon_j) \tag{15}$$

where the accepted traffic $A_k = \alpha_j + n_k \cdot \alpha_{new}$ and the rejected traffic $R_k = (\lambda \cdot \Delta - n_k) \cdot \alpha_{new}$.

III. Experimental Results with Two-Model System

In this section, we compare our MBAC method with two other MBAC methods in terms of realizing high network utilization while maintaining the QoS. The three compared methods represent the three different categories of MBAC approaches in Fig. 1:

1) Model-free measurement-based admission control (MFMB): this approach takes advantage of measured traffic load without assuming a specific traffic model. The buffer overflow probability is calculated from large deviation theory, for example, as calculated by Duffield [16]–[18].
2) Fitted-model measurement-based admission control (FMMB): this approach assumes a traffic model and fits the model parameters to the observed traffic. In the experiments, the assumed model is fractional Brownian motion as used by Norros [10], [11].
3) Our MBAC system with a two-model library (MBMC): in the experiments, the system described in Section II includes two models. The first model is a Poisson process $X_1(t)$, and the second model is fractional Brownian motion (fBm) $X_2(t)$ with Hurst parameter H 0.9. However, we could have chosen any other traffic models for the system.

A. 2-Model MBAC System

Before the comparison with two other MBAC approaches, we describe our two-model MBAC system. First, consider the decision rule with two models. For Poisson and fBm models, the two conditional probability density functions $p(\mathbf{x}|X_1)$ and

$p(\mathbf{x}|X_2)$ are given by

$$p(\mathbf{x}|X_1) = \prod_{i=1}^{w} \frac{(m\tau)^{x_i} e^{-m\tau}}{x_i!} = \frac{(m\tau)^{\sum_{i=1}^{w} x_i} \; e^{-m\Delta}}{x_1! x_2! \cdots x_w!} \tag{16}$$

$$p(\mathbf{x}|X_2) = \prod_{i=1}^{w} \frac{e^{-\frac{(x_i - m\tau)^2}{2\sigma^2\tau^{2H}}}}{\sqrt{2\pi\sigma^2\tau^{2H}}} = \left(\frac{1}{2\pi\sigma^2\tau^{2H}}\right)^{\frac{w}{2}} \; e^{\frac{-\sum_{i=1}^{w}(x_i - m\tau)^2}{2\sigma^2\tau^{2H}}} \tag{17}$$

where m is the mean rate and σ is the standard variation of the current traffic and H is the Hurst parameter.

By applying (9),(7) and (12) we get the decision rule

$$(\mathcal{C}_{12} - \mathcal{C}_{11})\, p(\mathbf{x}|X_1) P(X_1) \underset{X_2}{\overset{X_1}{\gtrless}} (\mathcal{C}_{21} - \mathcal{C}_{22})\, p(\mathbf{x}|X_2) P(X_2) \tag{18}$$

Substituting the cost functions from (14) and (15), the decision rule can be re-expressed as

$$\frac{p(\mathbf{x}|X_1)}{p(\mathbf{x}|X_2)} \underset{X_2}{\overset{X_1}{\gtrless}} \frac{\mathcal{P}_1 A_2 \; \epsilon_1 + \mathcal{P}_2 R_2 - \mathcal{R}_1 A_2 (1-\epsilon_1) P(X_2)}{\mathcal{P}_1 A_1 \; \epsilon_2 + \mathcal{P}_2 R_1 - \mathcal{R}_1 A_1 (1-\epsilon_2) P(X_1)} \tag{19}$$

In this two-model system, the priori probabilities are assumed to be equal (i.e. $P(X_1) = P(X_2) = 1/2$).

Secondly, the model-specific EBs are calculated. The EB for the Poisson model is

$$\alpha_1(s_1^*) = m + \frac{m\, s_1^*}{2} \tag{20}$$

where $s_1^* = \frac{\gamma}{B}$ with $\gamma = \log(\epsilon)$. The EB for the fBM model is

$$\alpha_2(s_2^*, t_2^*) = m + \frac{\sigma^2\, s}{2}\, t^{2H-1} \tag{21}$$

where $s_2^* = 2(1-H)\frac{\gamma}{B}$ and $t_2^* = \left(\frac{H}{1-H}\right)\frac{B}{\mu - m}$.

Now we evaluate the traffic classification by measuring the costs paid for wrong decisions and right decisions. During the time window Δ, the number of new flows arriving is Poisson with mean $\lambda{\cdot}\Delta$. The maximum number of acceptable requests is

$$\mathcal{N} = \left\lfloor \frac{\mu - \alpha_{current}}{\alpha_{new}} \right\rfloor \tag{22}$$

The possible number n_i of accepted requests under the model i follows the truncated Poisson distribution with expected value $\lambda\,\Delta$ in $[0, \mathcal{N}]$

$$p(n_i) = \begin{cases} \frac{\frac{(\lambda\Delta)^{n_i} \; e^{-\lambda\Delta}}{n_i!}}{1 - \sum_{k=0}^{\mathcal{N}} \frac{(\lambda\Delta)^k \; e^{-\lambda\Delta}}{k!}} & n_i = 0, 1, ..., \mathcal{N} \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

Four costs $\widehat{\mathcal{C}_{ij}}$ for $i, j = 1, 2$ are measured during the time window Δ. The current EB (α_{CEB}), true EB (α_{TEB}) and the number of accepted requests (n_{AC}) and new EB (α_{NEB}

Measured Costs	True Model	Classified Model	α_{TEB}	α_{CEB}	n_{AC}	α_{NEB}
$\widehat{\mathcal{C}_{11}}$	X_1	X_1	α_1	α_1	n_1	$\alpha_1 + \alpha_{new} n_1$
$\widehat{\mathcal{C}_{12}}$	X_1	X_2	α_1	α_2	n_2	$\alpha_1 + \alpha_{new} n_2$
$\widehat{\mathcal{C}_{21}}$	X_2	X_1	α_2	α_1	n_1	$\alpha_2 + \alpha_{new} n_1$
$\widehat{\mathcal{C}_{22}}$	X_2	X_2	α_2	α_2	n_2	$\alpha_2 + \alpha_{new} n_2$

TABLE I: EB based on the two-model classifier

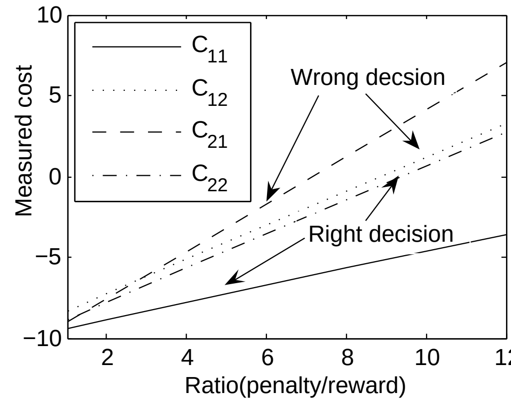

Fig. 4: Comparison of costs of four decisions

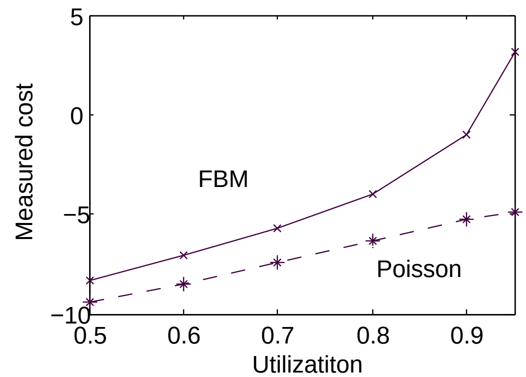

Fig. 5: Measured cost for Poisson traffic

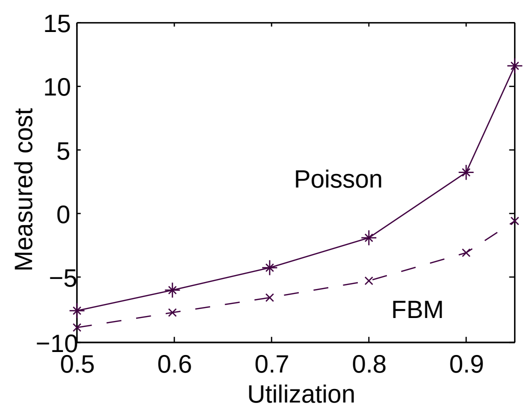

Fig. 6: Measured cost for fBm traffic

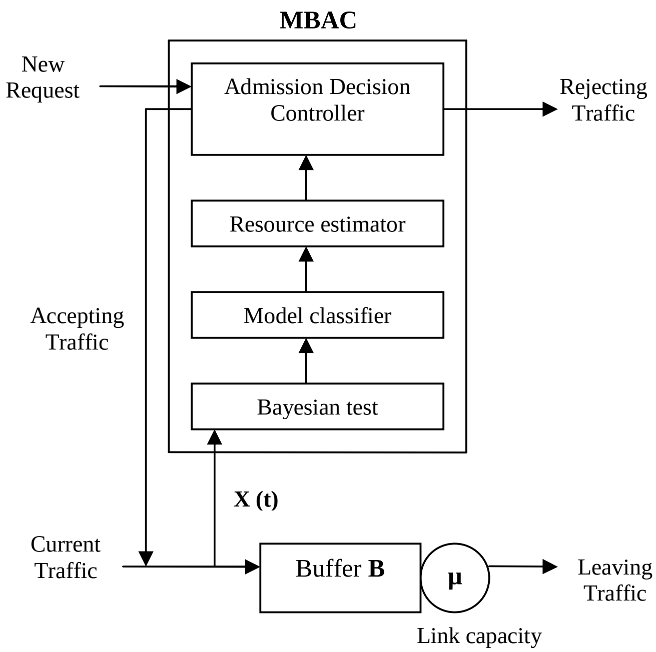

Fig. 7: Queue model with proposed MBAC

which means $\alpha_{TEB} + \alpha_{new} \cdot n_{AC}$), as listed in Table I are obtained over Δ. Now, the measured costs are derived as

$$\text{Measured cost} = (\alpha_{NEB} - \mu)\mathcal{P}_1 I_{\{}\alpha_{NEB} > C\} + (\lambda \cdot \Delta - n_{AC})\alpha_{new}\mathcal{P}_2 - \alpha_{NEB}\mathcal{R}_1 \quad (24)$$

The comparison of measured costs show in Fig 4 with $B = 50, \mu = 100, \epsilon = 10^{-3}$. We give the same value to the reward $\mathcal{R}_1$ for the accepted traffic and penalty $\mathcal{P}_1$ for the rejected traffic ($\mathcal{R}_1 = \mathcal{P}_1$) but the penalty $\mathcal{P}_2$ for the traffic violating the QoS is higher that that $\mathcal{P}_2$ for the rejected traffic ($\mathcal{P}_2 = k * \mathcal{P}_1 (k \geq 1)$). The new flows arrive according to a Poisson process with rate $\lambda = 0.5$ and the window size $\Delta = 100$. The measured cost of two right decisions are lower than that of the wrong decision. Notice that the cost $\widehat{\mathcal{C}_{21}}$ is a little higher than that $\widehat{\mathcal{C}_{12}}$. The wrong decision of choosing the fBm instead of Poisson costs more than the wrong decision of choosing Poisson because the EB for fBm is greater than that for Poisson and leads to buffer overflow. Fig 5 presents the measured costs in terms of utilization when the true model is Poisson. The measured cost of the Poisson process is lower than that of the fBm process, which means classifier choose the true model when it choose the lowest cost. When true model is fBm process, Fig 6 shows the same result. Classifier chooses the fBm as the best-fit model of observed traffic and costs lower than the wrong choice.

B. Experimental Results

We used discrete event simulation to evaluate the performance for each approach. Simulations were carried out using Matlab. We simulated the single-server queueing model in Fig. 7. New flows arrive with mean rate 1 per second and flow lifetimes are exponentially distributed with mean of 200 seconds. Each new flow is admitted or rejected according to the specific MBAC algorithms. In all experiments, the link capacity is 120 packets and there is buffering for 50 packets at the link. The packet loss probability is constrained to 10^{-4}. Simulations were run for 5000 simulation seconds. Each simulation was repeated 5 times with different seeds for the random number generator. Our experiments use three different types of simulated traffic

- Poisson traffic (as an example of short-range dependent traffic)
- fBm traffic with Hurst parameter H=0.9 (representing the long-range dependent traffic)
- Alternating traffic with alternating intervals of Poisson and fBm

Figs. 8-10 show the utilization realized by the different MBAC approaches (MFMB, FMMB, MBMC) as a function of normalized offered load, given a buffer overflow probability constraint of 10^{-4} and different types of input traffic. For short-range dependent traffic, the different MBAC approaches appear to perform similarly without significant differences. For long-range dependent traffic, the MBAC approaches using traffic models achieve higher utilization than approaches which only use general traffic load. As might be expected, our MBMC approach has a considerable advantage in the case when the input traffic consists of alternating SRD and LRD intervals, because the alternating intervals can be well tracked

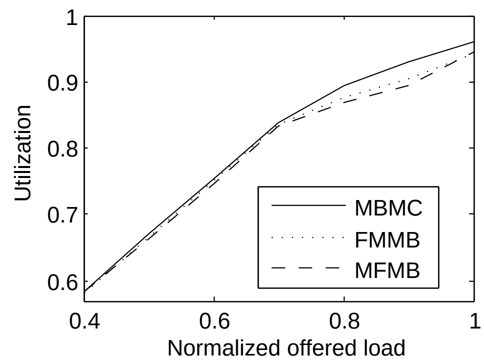

Fig. 8: Utilization for SRD traffic

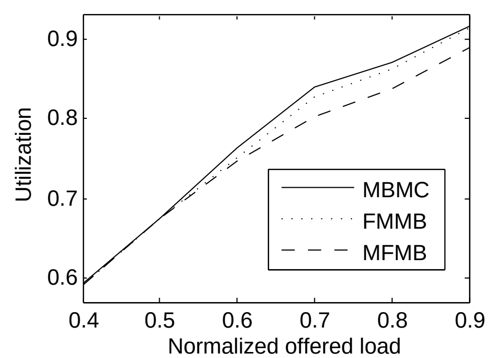

Fig. 9: Utilization for LRD traffic

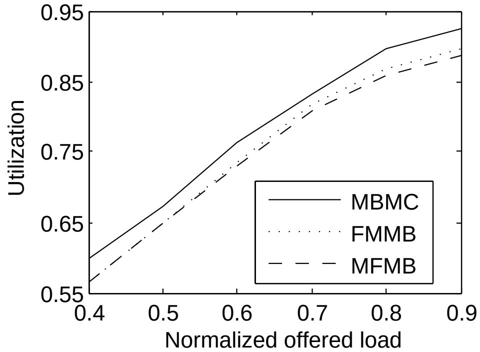

Fig. 10: Utilization for alternating traffic

Normalized Offered load	SRD	LRD	Alternating
0.4	0.95	0	0.57
0.5	0.96	0.1	0.54
0.6	0.97	0.1	0.52
0.7	0.98	0.1	0.55
0.8	0.93	0.1	0.54
0.9	0.96	0.1	0.49
1	0.95	0.1	0.62

TABLE II: The fraction of time the traffic is classified as SRD

by the model classifier. This type of dynamic traffic points out the weaknesses in the other approaches which are limited to a single pre-assumed traffic model or do not use a traffic model.

Table II shows the performance of classification. Each columns shows that the each simulated traffic was classified as the exact traffic model. Table III shows that all three MBAC approaches were able to meet the target QoS 10^{-4}.

IV. Conclusions

This paper has presented a new MBAC approach with traffic classification formulated in a decision theoretic approach. The traffic classification works by decision rules to optimize the network utilization while satisfying the target QoS. We derived a decision rule and defined cost functions taking into account utilization, packet loss, and rejection rate.

Our MBAC approach was compared by simulations with other MBAC approaches. It appears from experimental results that our MBAC approach has significant advantages particularly when the traffic is highly dynamic. We will continue more simulations to evaluate the system with more types of traffic.

Another issue for future research is practicality of implementation. We would also like to evaluate the feasibility of real-time operation.

	Target QoS:10^{-4}		
Algorithms	SRD	LRD	Alternating
MFMB	0.1529×10^{-4}	0.3792×10^{-4}	0.1969×10^{-4}
FMMB	0.3817×10^{-4}	0.0302×10^{-4}	0.3729×10^{-4}
MBMC	0.4681×10^{-4}	0.0743×10^{-4}	0.1160×10^{-4}

TABLE III: Buffer overflow probability

References

[1] B. T. Doshi, "Deterministic rule based traffic descriptors for broadband ISDN: Worst case behavior and connection acceptance control," in *Proc. IEEE GLOBECOM'93*, Dec. 1993, pp. 1759–1764.

[2] K. Siram and W. Whitt, "Characterizing superposition arrival processes packet multiplexer for voice and data," *IEEE J. Select. Areas Commun.*, vol. 4, no. 6, pp. 833–846, Sept. 1986.

[3] P. Sen, B. Maglaris, R. N, and D. Anastassiou, "Models for packet switching of variable-bit-rate video sources," *IEEE J. Select. Areas Commun.*, vol. 7, no. 5, pp. 834–844, June 1989.

[4] W. E. Leland and D. V. Wilson, "High-time resolution measurement and analysis of LAN traffic: Implication for LAN interconnection," in *Proc. IEEE INFOCOM'91*, Balhour, FL, Apr. 1991, pp. 1360–1366.

[5] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger, "Long-range dependence in variable- rate video traffic," *IEEE/ACM Trans. Networking*, vol. 43, pp. 1566–1579, Apr. 1995.

[6] M. Crovella and A. Bestavros, "Self-similarity in WWW traffic: evidence and possible causes," *IEEE/ACM Trans. Networking*, vol. 5, no. 6, pp. 835–946, Dec. 1997.

[7] R. Guerin, H. Ahmadi, and M. Naghshineh, "Equivalent capacity and its application to bandwidth allocation in high-speed networks," *IEEE J. Select. Areas Commun.*, vol. 1, no. 7, pp. 968–981, Sept. 1991.

[8] A. I. Elwalid and D. Mitra, "Effective bandwidth of general Markovian traffic sources and admission control of high-speed networks," *IEEE/ACM Trans. Networking*, vol. 1, no. 3, pp. 329–343, June 1993.

[9] C. S. Chang, "Stability, queue length and delay of deterministic and stochastic queueing networks," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 913–931, 1994.

[10] I. Norros, "A storage model with self-similar input," *Queueing Systems*, vol. 16, pp. 387–396, 1994.

[11] ——, "On the use of fractional Brownian motion in the theory of connectionless networks," *IEEE J. Select. Areas Commun.*, vol. 6, no. 13, pp. 953–962, Aug. 1995.

[12] S. Jamin, P. B. Danzig, S. J. Shenker, and L. Zhang, "A measurement-based admission control algorithm for integrated-service packet network," *IEEE/ACM Trans. Networking*, vol. 5, no. 1, pp. 56–70, Feb. 1997.

[13] M. Grossglauser and D. N. C. Tse, "A framework for robust measurement-based admission control," *ACM computer Commun. Review*, vol. 27, no. 4, pp. 237–248, Oct. 1997.

[14] ——, "A time-scale decomposition approach to measurement-based admission control," *IEEE/ACM Trans. Networking*, vol. 11, no. 4, pp. 550–563, Aug. 2003.

[15] J. Qiu and E. W. Knightly, "Measurement-based admission control with aggregate traffic envelopes," *IEEE/ACM Trans. Networking*, vol. 9, no. 2, pp. 199–210, 2001.

[16] D. D. Botvich and N. G. Duffield, "Large deviations, economics scale, and the shape of the loss curve in large multiplexer," *Queueing Syst. Theory Appl.*, vol. 20, no. 3-4, pp. 293–320, Aug. 1995.

[17] N. G. Duffield and N. O'Connell, "Large deviations and overflow probabilities for the general single server queue with application," *Math. Proc. the Cambridge Phil. Society*, vol. 118, no. 26, pp. 363–374, 1995.

[18] N. G. Duffield, J. T. Lewis, N. O. Connel, R. Russel, and F. Toomey, "Entropy of atm traffic streams: A tool for estimating qos parameters," *IEEE J. Select. Areas Commun.*, vol. 13, no. 6, pp. 981–990, Aug. 1995.

[19] R. J. Gibbens, F. P. Kelly, and P. B. Key, "A decision-theoretic approach to call admission control in atm network," *IEEE J. Select. Areas Commun.*, vol. 13, no. 6, pp. 1101–1114, Aug. 1995.

[20] Y. Zeng and T. M. Chen, "Automatic model classification of measured internet traffic," in *Proc. IEEE IPOM'02*, Oct. 2002, pp. 197–201.

[21] ——, "Measurement-based real-time traffic model classification," in *Proc. IEEE ICC'04*, June 2004, pp. 1857–1861.

[22] J. Berger, *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, 1985.

[23] K. Fukunaga, *Introduction to Statistical Pattern Recognition.* Academic Press, 1990.

[24] F. P. Kelly, "Notes on effective bandwidths," in *Stochastic Networks: Theory and Applications*. Oxford University Press, 1996, pp. 141–168.

[25] C. Courcoubetis and R. Weber, "Effective bandwidth for stationary sources," *Probability in Engineering and Informational Sciences*, vol. 9, no. 2, pp. 285–294, 1995.