A Predictive Flow Control Mechanism to Provide QoS and Efficient Network Utilization

Dongyu Qiu and Ness B. Shroff

Network Engineering and Wireless Systems (NEWS) group
School of Electrical and Computer Engineering
Purdue University, West Lafayette, IN 47907-1285
http://www.ece.purdue.edu/~shroff/NEWS
E-mail: shroff@purdue.edu

June 21, 2000

Introduction

- High-speed networks will carry both delay sensitive (DS) traffic and delay insensitive (DI) traffic.
- DS traffic: voice, video, multimedia applications.
- DI traffic: e-mail, ftp, web-browsing.
- Transient oscillations in network traffic can lead to network congestion $\Rightarrow$ excessive delays $\Rightarrow$ poor QoS performance, especially for DS traffic.
- We must provide low delay service to DS traffic, while at the same time efficiently utilizing the network.

Possible Approach

- Approach: Give the DS traffic priority over the DI traffic.
- Advantage:
  - Simple
  - Can provide QoS for DS traffic since it is not delayed by DI traffic.
- Drawback: By itself this is not enough.
- Reason: The DI traffic is not controlled $\Rightarrow$ Unnecessary retransmissions may take place that would reduce the overall throughput.
- A number of control schemes have been developed in the literature to control the DI traffic

Existing Works

- Binary feedback schemes
  - Ramakrishnan88, Bonomi95
- Two-bit feedback schemes
  - Lapsley96
- Explicit-rate control schemes
  - Benmohamed93, Jain94, Zhao97, Altman98

Difficulties

- Queueing system is non-linear.
- Control system is non-linear since the controlled DI traffic rate can never be less than zero.
- In some explicit-rate control schemes, a linear system is used as an approximation (e.g. [Zhao97], [Altman98]) but this could result in poor network performance.
- In our work, the non-linear properties of both the queueing system and the control system are considered in developing and analyzing the performance and stability of our flow control system.
- **Goal:** To develop a predictive flow control algorithm that minimizes congestion under a throughput constraint.

Model

- Focus on a single bottleneck link within the network
- Infinite buffer, Discrete-time, fluid queue
- At time $n = 0$, $q(0) = 0$
- $V(n)$ is the aggregate amount of DS traffic that arrives at the bottleneck queue at time n.
- $V_{max} := \sup_{n \geq 0} \{V(n)\}$ is finite
- $V(n)$ is stationary in the mean and $\bar{V} := \mathrm{E}\{\mathrm{V(n)}\}$
- $N :=$ the number of DI traffic sources
- $\mu :=$ the service rate of the link.

Model

- $n_i, i = 1, \ldots, N$ is the round trip delay between the ith DI source and the destination
- $a_i(n)$ is the ith DI traffic rate computed at time n based on the prediction of DS traffic rate. This explicit rate information will be sent back to the ith DI traffic source
- The DI traffic sources always have enough data to transmit
- $a(n) = \sum_{i=1}^{N} a_i(n - n_i)$ is the aggregate DI traffic arrival to the queue at time n
- $\hat{V}_i(n)$ is the predicted value of $V(n)$ based on the history of V upto time $n - n_i$.

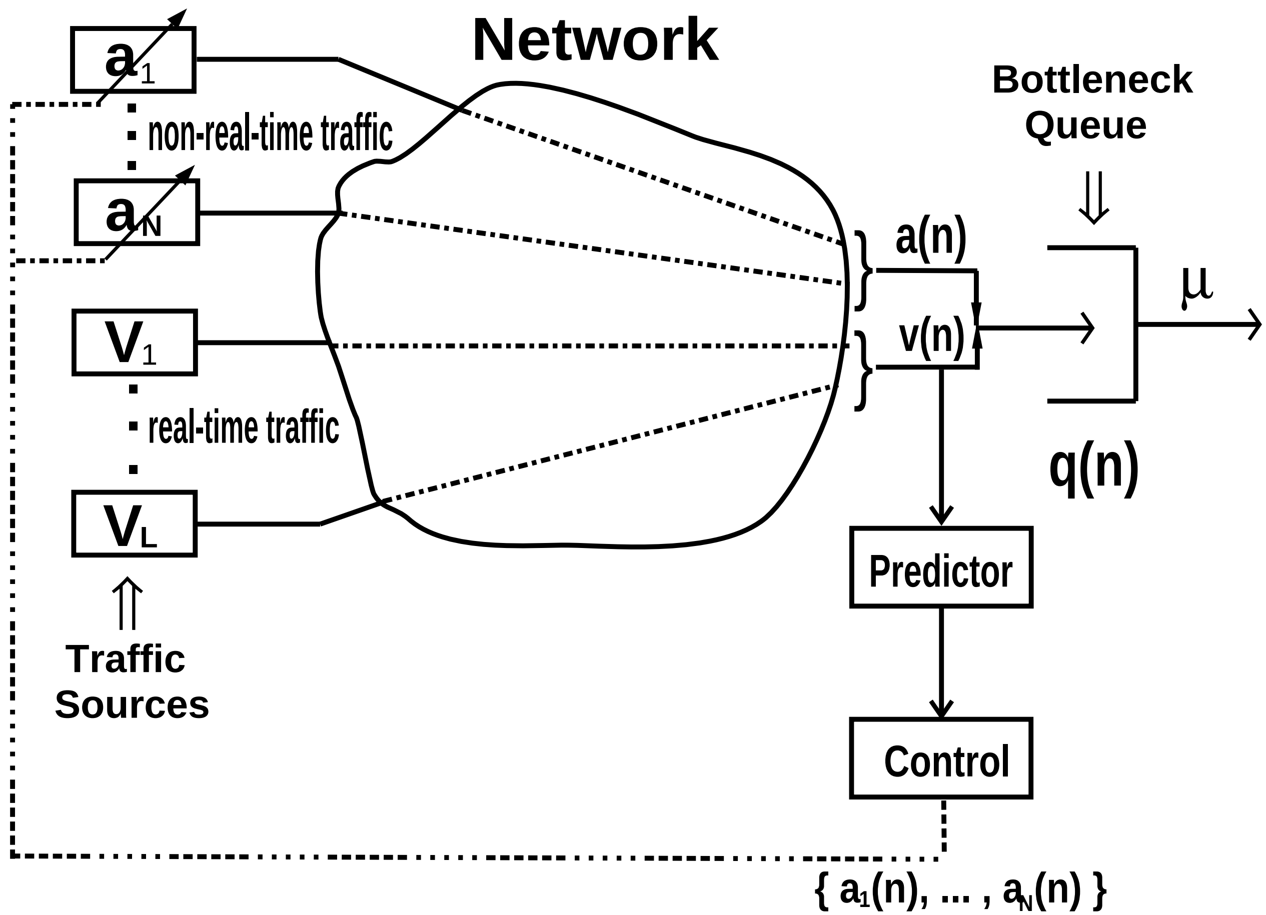

Figure 1: System diagram of predictive flow control

Model

- Let $V(z)$ be the Z-transform of $V(n) - \bar{V}$ and let $\hat{V}_i(z)$ be the Z-transform of $\hat{V}_i(n) - \bar{V}$.
- The above predictor is linear (e.g. AR, ARMA, etc.).

$$\hat{V}_i(z) = z^{-n_i} H_i(z) V(z),$$

- $H_i(z)$ is a causal, stable linear time invariant system.
- The predictor above is unbiased, i.e., $\mathrm{E}\{\hat{V}_i(n)\} = \bar{V}$.

Single DI traffic Source Case

- Ideally, we would like to achieve $a(n) + V(n) = \mu$ at all time n.
- This is not possible for two reasons:
  - Prediction Error. Then what about

$$a_1(n) + \hat{V}_1(n + n_1) = p\mu?$$

    Here p is the percentage of output link capacity that we would like to utilize.

  - Still not possible because $V(n) > \mu$ and $\hat{V}_1(n + n_1) > \mu$ for some values of $n \Rightarrow a(n) < 0$ for those values of n, which is clearly infeasible.

Single DI traffic Source Case

- So, considering the possibility of prediction error and the possibility that $V(n) > \mu$, a simple control algorithm could be

$$a_1(n) = [p\mu - \hat{V}_1(n + n_1)]^+,$$

where $x^+ = x$ if $x > 0$, or $x = 0$, otherwise.

- Drawback: does not take into account the queueing behavior. When there is a substantial backlog in the queue, even if $p\mu > \hat{V}_1(n + n_1)$, $a_1(n)$ should be zero.

Single DI traffic Source Case

- **Proposed Control Algorithm ($N = 1$ case)**

  1. Define an auxiliary process $S_1(n)$ and set $S_1(0) = 0$.
  2. $S_1(n) = [S_1(n-1) + \hat{V}_1(n) - p\mu]^+$. For $n \leq 0$, we let $V(n) = 0$. p is the percentage of output link capacity that we would like to utilize.
     ($S_1(n)$ corresponds to the queue length of a system with $\hat{V}$ as input and $p\mu$ as link rate).
  3. $a_1(n) = [p\mu - \hat{V}_1(n + n_1) - S_1(n + n_1 - 1)]^+$. For $n \leq 0$, we let $a_1(n) = 0$.

- $a_1(n)$ is non-zero only when there is no predicted backlog in the queueing system.

Single DI traffic Source Case

- **Lemma 1** *Let* $q_1(n)$ *be the workload at time* n *in a queueing system with input* $\hat{V}_1(n)$ *and service rate* μ *and let* $q_2(n)$ *be the workload at time* n *in a queueing system with input* $a(n) + \hat{V}_1(n)$ *and service rate* μ*. Then* $q_2(n) = q_1(n)$ *for any* $n \geq 0$*.*

- Lemma 1 shows that if we have perfect prediction, then the queue length using our scheme will be the same as the queue length of a queueing system with only $\hat{V}_1(n)$ as input.

Single DI traffic Source Case

- **Proposition 1** *For our predictive flow control algorithm defined in steps 1–3 earlier, under the condition $\bar{V} < p\mu$, we have*

$$\lim_{n\to\infty} \frac{\sum_{j=1}^{n} a(j)}{n} + \bar{V} = p\mu$$

- Proposition 1 tells us that, using our scheme, the average aggregate input rate to the queue is given by $p\mu$.

- When we fix p in our control algorithm, the output link utilization is also fixed and does not depend on other predictor parameters.

- Note: We need the condition $\bar{V} < p\mu$, otherwise our utilization would have to be larger than p.

Single DI traffic Source Case

- **Theorem 1** *Let $q_0(n)$ be the workload at time n in a queueing system with input $V(n)$ and service rate μ and let $q(n)$ be the workload at time n in a queueing system with input $a(n)+V(n)$ and service rate μ. Then under the predictive flow control algorithm defined above, if $\frac{p\mu-\bar{V}}{\mu-\bar{V}} \leq H_1(1) \leq 1$, there exists a constant C_1 such that $q(n) \leq q_0(n) + 2C_1$ for any $n \geq 0$.*

- $q_0(n)$ corresponds to the queue length without DI traffic, and is expected to be a loose lower bound on the queue length generated by any control scheme.

- Using our algorithm, $q(n)$ is only an additive constant (not dependent on the time n) larger than $q_0(n)$.

Multiple DI traffic Source Case

- **Proposed Control Algorithm**

  1. Set $S_i(0) = 0, 1 \leq i \leq N$.
  2. $S_i(n) = [S_i(n-1) + \hat{V}_i(n) - p\mu]^+$. For $n \leq 0$, we let $V(n) = 0$.
  3. $a_i(n) = \frac{1}{N}[p\mu - \hat{V}_i(n+n_i) - S_i(n+n_i-1)]^+$. For $n \leq 0$, we let $a_i(n) = 0$.

- For fairness, we assign the same arrival rate to each DI traffic. This is by no means a necessary condition, and can be relaxed by modifying step 3 above in favor of a more unfair system if the situation so required it.

Multiple DI traffic Source Case

- **Theorem 2** *Let* $q_0(n)$ *be the workload at time* n *in a queueing system with input* $V(n)$ *and service rate* μ *and let* $q(n)$ *be the workload at time* n *in a queueing system with input* $a(n) + V(n)$ *and service rate* μ*. Then under the predictive flow control algorithm defined above, If* $\frac{p\mu - \bar{V}}{\mu - \bar{V}} \leq H_i(1) \leq 1$*, for all* $1 \leq i \leq N$*, we will have* $q(n) \leq q_0(n) + 2C$*, where* $C = \max_{1 \leq i \leq N} C_i$.

- The only requirement for the DS traffic is that $V_{max} < \infty$. Hence, our the result can be applied to both short range dependent and long range dependent DS traffic.

Asymptotic Properties of the Overflow Probability

- By correctly choosing $H_i(1)$ in our scheme, we have

$$\mathrm{P}\{\mathrm{q(n)} > \mathrm{x}\} \leq \mathrm{P}\{\mathrm{q_0(n)} > \mathrm{x} - 2\mathrm{C}\}.$$

  - Let Q and Q_0 be stationary versions of $q(n)$ and $q_0(n)$, resp.
  - When $V(n)$ is short range dependent, for a large class of short range dependent traffic, we have $\mathrm{P}\{\mathrm{Q_0} > \mathrm{x}\} \sim \mathrm{Ae^{-Bx}}$, when x is large (A, B are constants). Hence,

$$\lim_{x\to\infty} \frac{\mathrm{P}\{\mathrm{Q} > \mathrm{x}\}}{\mathrm{P}\{\mathrm{Q_0} > \mathrm{x}\}} \leq \lim_{x\to\infty} \frac{Ae^{-B(x-2C)}}{Ae^{-Bx}} = e^{2BC}$$

- Asymptotically, our scheme produces an overflow probability that does not diverge from the overflow probability of the system without DI traffic.

Asymptotic Properties of the Overflow Probability

- When $V(n)$ is long range dependent, for a large class of traffic models, $\mathrm{P}\{\mathrm{Q}_0 > \mathrm{x}\} \approx \mathrm{Ae}^{-\mathrm{Bx}^{\mathrm{d}}}$, when x is large, where A, B, d are constants and $0 < d < 1$.

$$\lim_{x\to\infty} \frac{\mathrm{P}\{\mathrm{Q} > \mathrm{x}\}}{\mathrm{P}\{\mathrm{Q}_0 > \mathrm{x}\}} \leq \lim_{x\to\infty} \frac{Ae^{-B(x-2C)^d}}{Ae^{-Bx^d}} = 1$$

- Asymptotically, the overflow probability of our system converges to that of the system without DI traffic.
- If $H_i(1)$ does not satisfy the above condition, $\lim_{x\to\infty} \frac{\mathrm{P}\{\mathrm{Q}>\mathrm{x}\}}{\mathrm{P}\{\mathrm{Q}_0>\mathrm{x}\}}$ may go to ∞.

Stability

- The system is stable if the queueing system is stable.
- If $H_i(z)$ are stable and $H_i(1) = 1$ for any i, $q(n)$ will not exceed $q_0(n)$ plus a constant.
- If $q_0(n)$ is stable, the queueing system $q(n)$ will also be stable.
- The stability does not depend on the predictor parameters once $H_i(1) = 1$ is set.
- This stability is robust to the round trip delay errors as long as $H_i(1) = 1$ is maintained.

Simulations

- Assume that $V(n)$ is wide sense stationary.
- $\Delta(j) \triangleq V(j) - \hat{V}_i(j)$.
- $X_{n,l} \triangleq \sum_{j=n-l+1}^{n} \Delta(j)$.
- $\mathrm{E}\{\mathrm{X_{n,l}}\} = 0$.
- $V_l = \mathrm{Var X_{n,l}}$ does not depend on n.
- Choose predictor parameters such that we minimize $\lim_{l \to \infty} V_l$ under the condition $H_i(1) = 1$.
- Our scheme is abbreviated as MPQ.

Simulations

- The "Minimum Mean Squared Error (MMSE)" predictor is designed as follows:
  - First, a low pass filter $H_{LPF}(z)$ is applied to the DS traffic.
  - Next, a standard minimizing mean square error linear predictor $H_{MMSE}(z)$ is calculated based on the low frequency part of the DS traffic.
  - The final "MMSE" predictor is $H_{LPF}(z)H_{MMSE}(z)$.

Simulations

- H_2 represents the scheme used in [Zhao97].
- In the H_2 scheme, the low frequency part of the DS traffic is considered as a deterministic step function rather than a stochastic process. Thus,

$$H_i(z) = z^{-n_i} H_{LPF}(z)$$

- H_2 scheme satisfies the condition $H_i(1) = 1$. So, in many cases, the H_2 scheme is actually pretty good.
- However, because the stochastic property of the DS traffic is not considered, in some cases, we can see a substantial difference between the H_2 Scheme and MPQ.

Simulations

- "No DI" represents the Overflow probability of q_0, i.e., the queueing system with only DS traffic.
- In all simulations, five DI traffic sources are simulated.
- The round trip delays of the DI traffic loops are from 5 to 9.

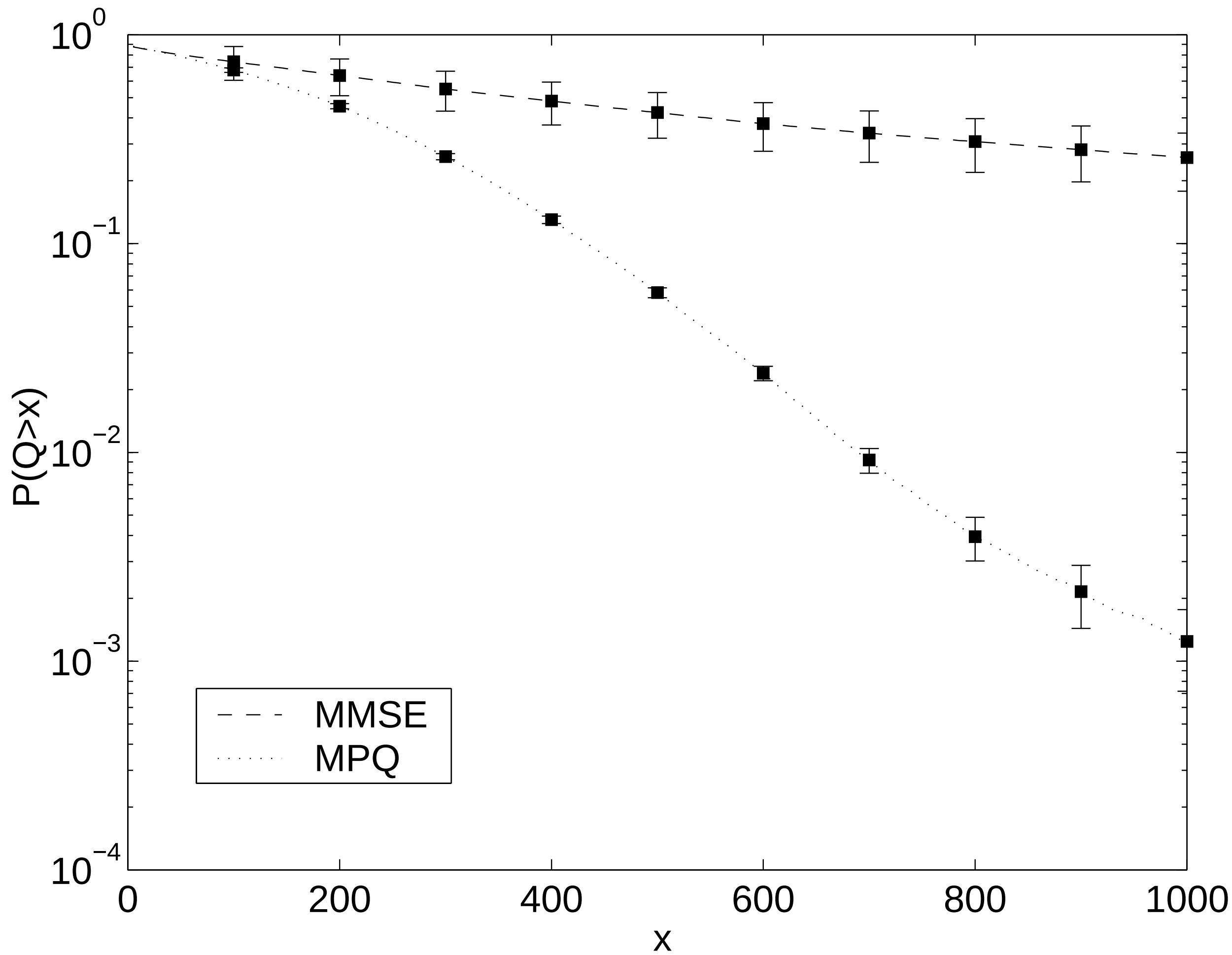

Figure 2: Overflow probability with $V(n)$ Gaussian process, $C_v(k) = 1600 \times (|k|+1)^{-0.75}$, $\bar{V} = 150$, link capacity = 200, utilization = 98%.

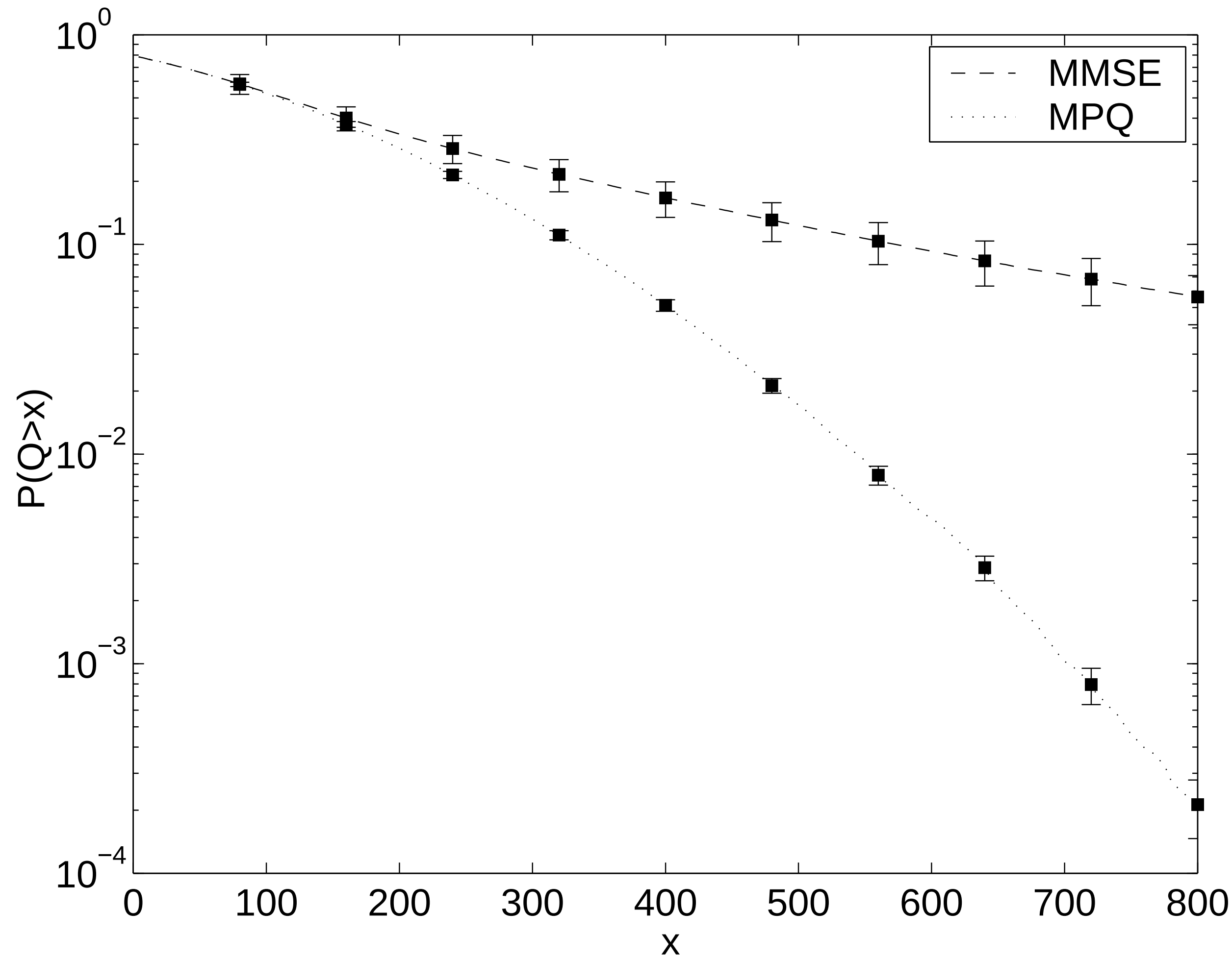

Figure 3: Overflow probability with actual mpeg video as real-time traffic, $\bar{V} = 11.6 Mbps$, link capacity $= 15 Mbps$, utilization $= 98\%$.

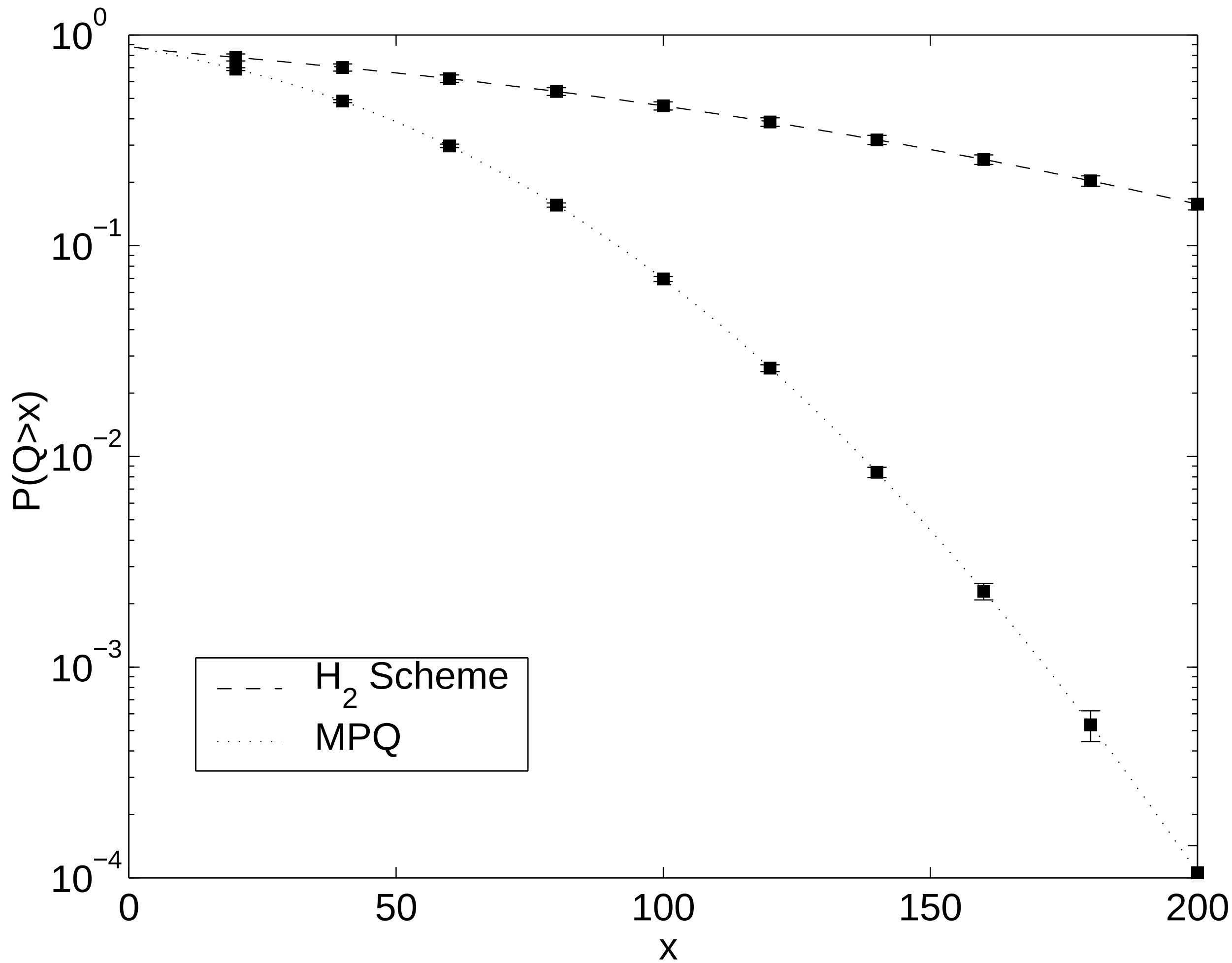

Figure 4: Overflow probability with voice sources as real-time traffic, $\bar{V} = 341.2$, link capacity $= 500$, utilization $= 99.8\%$.

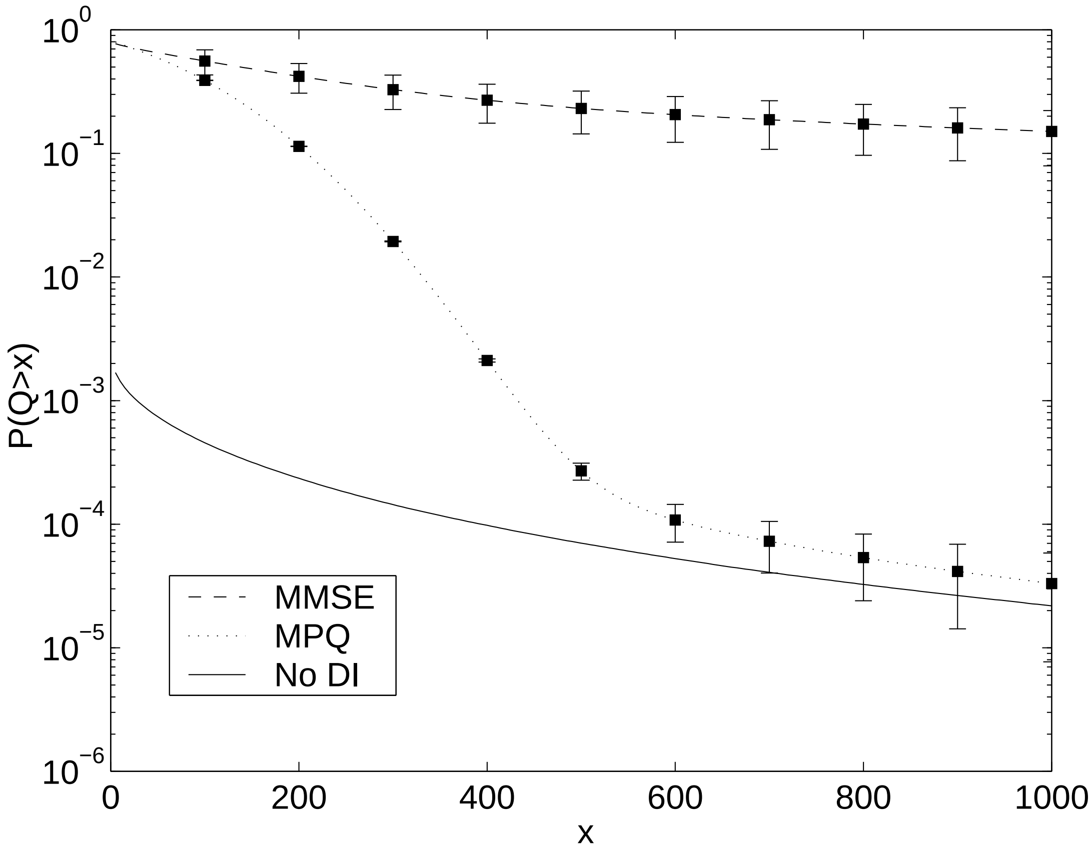

Figure 5: Overflow probability with $V(n)$ Gaussian process, $C_v(k) = 479.599 \times 0.9990005^{|k|} + 161.787 \times 0.9899995^{|k|} + 498.033 \times 0.9^{|k|}$, $\bar{V} = 100$, link capacity $= 200$, utilization $= 98\%$.

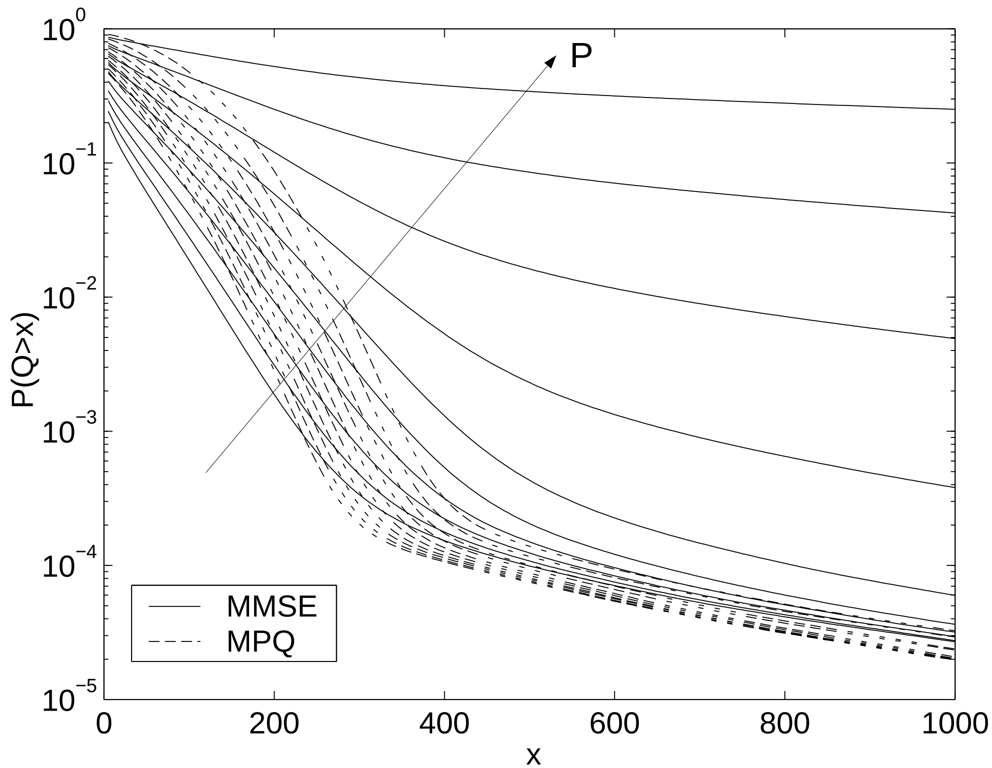

Figure 6: Overflow probability with $V(n)$ Gaussian process, $C_v(k) = 479.599 \times 0.9990005^{|k|} + 161.787 \times 0.9899995^{|k|} + 498.033 \times 0.9^{|k|}$, $\bar{V} = 100$, link capacity = 200, single DI traffic with round trip delay 5, p goes from 90% to 99% with a step of 1%.

Conclusion And Future Work

- For a given level of utilization, our algorithm results in a significantly smaller overflow probability than other schemes in the literature.
- Our control system is stable and the stability is robust to round trip delay errors.
- We show how to choose appropriate predictor parameters and show that when such parameters are chosen, the performance of our scheme is very close to the best achievable performance.
- Release the condition that V_{max} is finite.
- Find an optimal predictor.
- Extend the work to handle multiple bottlenecks in the network.

References

[1] K. K. Ramakrishnan and R. Jain, "A binary feedback scheme for congestion avoidance in computer networks with a connection-less network layer," *Proceedings of SIGCOMM'88*, ACM, August 1988.

[2] F. Bonomi, D. Mitra, and J. B. Seery, "Adaptive algorithms for feedback-based flow control in high-speed wide-area ATM networks," *IEEE Journal on Selected Areas in Communications*, 13(7):1267-1283, September, 1995.

[3] D. E. Lapsley and M. Rumsewicz, "Improved buffer efficiency via the No Increase flag in EFCI flow control," *In Proceedings of the IEEE ATM'96 Workshop*, August, 1996.

[4] L. Benmohamed and S. Meekov, “Feedback Control of Congestion in Store-and-Forward Networks: The case of Single Congestion Node”, *IEEE/ACM Trans. On Networking*, Vol. 1, No. 6, Dec. 1993, pp. 693-708.

[5] R. Jain, S. Kalyanaraman and R. Viswandathan, “The OSU Scheme for Congestion Avoidance using Explicit Rate Indication”, OSU Technical Report, Sept. 1994.

[6] Y. D. Zhao, S. Q. Li and S. Sigarto, “A linear dynamic model design of stable explicit-rate ABR control scheme,” *In Proceedings of IEEE INFOCOM*, 1997, pp. 283-292.

[7] E. Altman, T. Basar, and R. Srikant, “Robust rate control for ABR sources,” *In Proceedings of IEEE INFOCOM*, 1998, pp. 166-173.