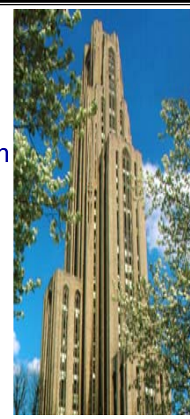
Stochastic Process and Markov Chains



David Tipper
Associate Professor
Graduate Telecommunications and Networking Program
University of Pittsburgh
dtipper@mail.sis.pitt.edu
http://www.sis.pitt.edu/~dtipper/tipper.html

Stochastic Processes

- A stochastic process is a mathematical model for describing an empirical process that changes in time according to some probabilistic forces.
- A stochastic process is a family of random variables $\{X(t), t \in T\}$ defined on a given probability space S, indexed by the parameter t, where t is in an index set T.
- For each $t \in T$, $X(t)$ *is a random variable with* $F(x,t) = P\{X(t) \leq x\}$
- A realization of X(t) is called a sample path
- Characterization of a stochastic process.
  1. State Space S,
  2. Index set T
  3. Stationarity

Characteristics of Stochastic Processes

- *State Space*
  - The values assumed by a random variable *X(t)* are called “states” and the collection of all possible values forms the “state space *S*” of the process.
  - If *X(t)=i*, then we say the process is in state *i*.
  - Discrete-state process
    - The state space is finite or countable for example the non-negative integers {0, 1, 2,…}.
  - Continuous-state process
    - The state space contains finite or infinite intervals of the real number line.

Characteristics of Stochastic Processes

- *Index parameter*
  - The index T is usually taken to be the time parameter.
  - Discrete-time process
    - A process changes state (or makes a “transition”) at discrete or finite countable time instants.
  - Continuous-time process
    - A process may change state at any instant on the time axis.
- The probability that stochastic process *X* takes on a value *i* ($i \in S$) at time = *t* is P[*X(t)=i*]
- *Stationarity*
  - A stochastic process X(t) is strict sense stationary if the statistical properties are invariant to time shifts

    f(x,t) = f(x) for all *t*.

Characteristics

- A stochastic process $X(t)$ is wide sense stationary if
  1. Mean is constant $E\{X(t)\} = K$ for all t
  2. The autocorrelation R is only a function of the time difference $R(t_1, t_2) = R(t_2 - t_1) = R(\tau)$

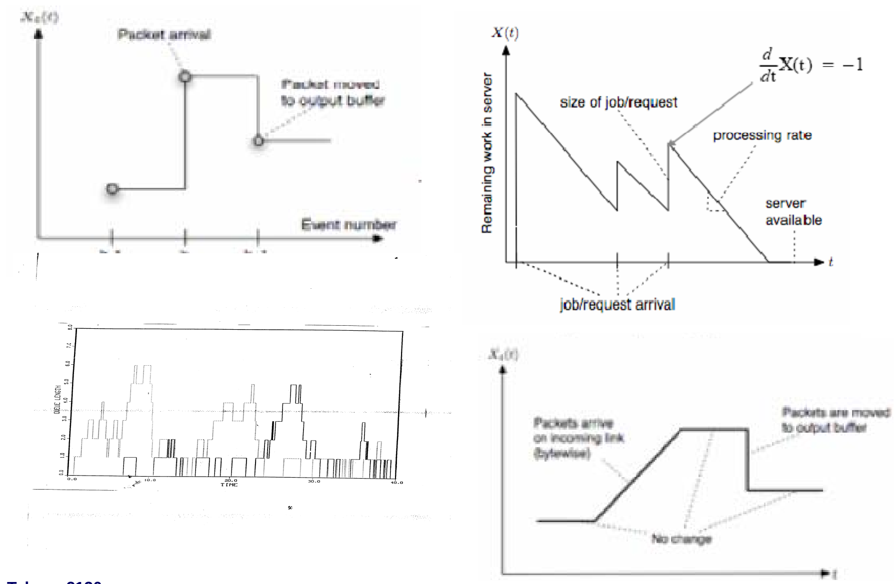
- Ergoditcity
  - A stochastic process $X(t)$ is ergodic if it's ensemble averages equal time averages
  - => Any statistic of $X(t)$ can be found from a sample path

$$E\{X(t)\} = \int_{-\infty}^{\infty} xf(x,t)dx \quad = \quad \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)dt$$

Categories of Stochastic Processes

Time Parameters	State Space	
	Discrete State	Continuous State
Discrete Time	Discrete time stochastic chain	Discrete time stochastic process
Continuous Time	Continuous time stochastic chain	Continuous time stochastic process

Stochastic Processes



Stochastic Processes

- Important Stochastic Processes for Queueing System Analysis
  - Markov Chains
  - Markov Process
  - Counting Process - Poisson Process
  - Birth Death Process

- In 1907 A.A. Markov defined and investigated a particular class of stochastic processes – now know as Markov processes/chains

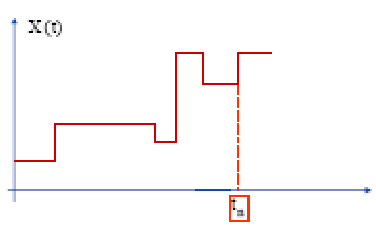
Markov Process

- For a ***Markov process*** $\{X(t), t \in T, S\}$, *with state space S,* its future probabilistic development is dependent only on the current state, how the process arrives at the current state is irrelevant.
- Mathematically
  - The conditional probability of any future state given an arbitrary sequence of past states and the present state depends only on the current state
- *Memoryless property* - The process starts afresh at the time of observation and has no memory of the past.

Discrete Time Markov Chains

- The Discrete time and Discrete state stochastic process $\{X(t_k), k \in T\}$ is a Markov Chain if the following conditional probability holds for all *i*, *j* and *k*. (note X_i means $X(t_i)$)

  $P[X_{k+1} = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_{k-1} = i_{k-1}, X_k = i]$
  $= P[X_{k+1} = j \mid X_k = i]$
  $= p_{ij}(k)$ ← state transition probability at k^{th} time step
- The future probability development of the chain depends only on its current state (k^{th} instant) and not on how the chain has arrived at the current state (i.e., memoryless)



Discrete Time Markov Chains (2)

- $p_{ij}(k)$ is (one-step) transitional probability, which is the probability of the chain going from state i to state j at time step t_k
- $p_{ij}(k)$ is a function of time t_k. If it does not vary with time (independent of k), then the chain is said to have stationary transition probabilities and is *time homogeneous* $p_{ij}(k) = p_{ij}$ *for all k*

  p_{ij} *is the one step transition probability of going from state i to state j*
- The state transition matrix $P = [p_{ij}]$ characterizes the Markov chain.

Discrete Time Markov Chains (2)

- The one step state transition matrix $P = [p_{ij}]$ is a stochastic matrix
  1. $0 \le p_{ij} \le 1$ All elements between zero and one
  2. $\sum_{j \in S} p_{ij} = 1$ Each row sums to one and is a density function
  3. $\lambda_1 = 1$ is an eigenvalue of P and $|\lambda_j| < 1$ j = 2, 3, …

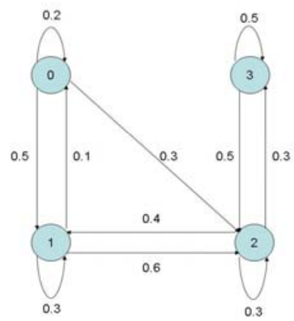
$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} & \cdots & \cdots & p_{0K-1} & p_{0K} \\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14} & \cdots & \cdots & p_{1K-1} & p_{1K} \\ \vdots & p_{21} & p_{22} & p_{23} & p_{24} & \cdots & \cdots & p_{2K-1} & p_{2K} \\ \vdots & \vdots & p_{32} & p_{33} & p_{34} & \cdots & \cdots & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots & \vdots & \vdots \\ p_{K,0} & p_{K,1} & p_{K,2} & p_{K,3} & p_{K,4} & \cdots & \cdots & p_{K,K-1} & p_{K,K} \end{bmatrix}$$

Discrete Time Markov Chains (2)

- For small state space can represent state transition matrix P = [p_{ij}] as a state transition diagram

Consider the Time Homogeneous Markov Chain with one step transition matrix for the states {0, 1, 2, 3} given below.

$$P = \begin{bmatrix} .2 & .5 & .3 & 0 \\ .1 & .3 & .6 & 0 \\ 0 & .4 & .3 & .3 \\ 0 & 0 & .5 & .5 \end{bmatrix}$$



Markov Chain Analysis Summary

- Markov chain $\{\tilde{x}(t_k) : k \in N\}$
  - one step transition matrix $P = [p_{ij}]$
  - state probabilities $\pi_j^{(n)} = p\{\tilde{x}(t_n) = j\}$
- General/transient behavior

1) $\pi^{(n)} = \pi^{(0)} p^{(n)} = \pi^{(0)} (p)^n$ with computation $O(nN^2)$

2) $\pi^{(n)} = \pi^{(n-1)} p$ with computation $O((\log_2 n) N^2)$

3) $\pi^{(n)} = n^{(0)} p^{(n)} = \pi^{(0)} T^{-1} [\text{diag}(\lambda_i^n)] T$ with computation $O(N^3)$

where $T = m^{-1} \leftarrow [\text{modal matrix}]^{-1}$

Also $P^{(n)}$ the n step transition matrix is $P^{(n)} = (P)^n$

- Steady state behavior $\pi = \lim_{n \to \infty} \pi^{(n)}$

4) $\pi = \pi p$ and $\sum_{i \in S} \pi_i = 1$

Markov Chain Analysis Summary

- 5. First Passage Time

  The first passage time T_{ij} is the number of transitions required to go from state *i* to state *j* for the first time (recurrence time if *i* = *j*).

  Let $f_{ij}^{(n)} = P\{T_{ij} = n\}$. That is probability the first passage time is n steps

$$f_{ij}^{(n)} = P\{x(t_{k+n}) = j, x(t_{k+r}) \neq j, r = 1, 2, , , n-1 \mid x(t_k) = i\}$$

$$f_{ij}^{(1)} = p_{ij}$$

$$f_{ij}^{(2)} = p_{ij}^{(2)} - f_{ij}^{(1)} p_{jj}$$

$$\vdots$$

$$f_{ij}^{(n)} = p_{ij}^{(n)} - \sum_{k=1}^{n-1} f_{ij}^{(k)} p_{jj}^{(n-k)} \qquad n = 2, 3, \ldots.$$

Markov Chain Analysis Summary

- 6. Mean First Passage Time

  The Mean First Passage Time $E\{T_{ij}\}$ is the average number of transitions required for the Markov Chain to go from state *i* to state *j* for the first time (recurrence time if *i* = *j*).

$$E\{T_{ij}\} = \sum_{n=1}^{\infty} n f_{ij}^{(n)}$$

  Using probability generating function approach get

$$E[T_{ij}] = \pi_j^0 (I - R_j)^{-1} e$$

Where p_j^0 = [0, …0, 1, 0, …0] is one only in the ith element and

$R_j = [P_{ik}]$ $i \neq j, \ k \neq j$ ← one step transition matrix P *without* row j and column j

Markov Chain Analysis Summary

- 7 Transform approach to $\pi^{(n)}$ and $P^{(n)}$

$$\pi^{(n)} = \pi^{(n-1)} P$$

$$\pi(z) = Z\{\pi^{(n)}\} = Z\{\pi^{(n-1)} P\}$$

$$\pi(Z) - \pi^{(0)} = z\pi(Z)P$$

$$\Rightarrow \pi(Z) = \pi^{(0)}[I - zP]^{-1}$$

$$\Rightarrow P^{(n)} = [I - zP]^{-1}$$

Markov Chain Analysis Summary

- 8. State Holding Times

Let H_i be the random variable that represents the number of time slots the Markov Chain spends in state i

$$P\{H_i = n\} = P\{x(t_{k+n}) = j, x(t_{k+r}) = i, r = 1, 2,, n-1 \mid x(t_k) = i\}$$

$$= P\{x(t_{k+n}) = j \mid x(t_{k+n-1}) = i\} P\{x(t_{k+n-1}) = i \mid x(t_{k+n-2}) = i\} .. P\{x(t_{k+1}) = i \mid x(t_k) = i\}$$

$$= (1 - p_{ii}) p_{ii}^{n-1}$$

Note the holding time has a geometric distribution which is the only memoryless discrete distribution.

Markov Chain Example

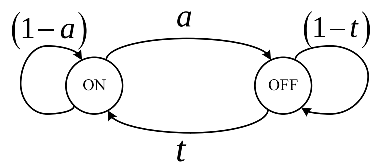
One model of a discrete time bursty ATM traffic source is a two state markov chain with one state representing ON and the other state representing OFF. When the source is in the ON state a cell is generated in every slot, when the source is in the OFF state no cell is generated.

Let a be the probability of transition from ON to OFF

Let t be the probability of transition from OFF to ON

The probability of making a transition from a state back to itself are $1-a$ and $1-t$ respectively for ON and OFF

The state transition diagram and state transition matrix P are



$$p = \begin{array}{c} \\ \text{ON} \\ \text{OFF} \end{array} \begin{array}{c} \begin{array}{cc} \text{ON} & \text{OFF} \end{array} \\ \begin{bmatrix} (1\text{-a}) & a \\ t & (1\text{-t}) \end{bmatrix} \end{array}$$

Markov Chain Example

Determining the steady state probabilities $\pi = [\pi_{ON} \;,\; \pi_{OFF}]$

$$\pi = \pi p \quad \Rightarrow \quad \begin{array}{l} \pi_{ON} = \pi_{ON}(1-a) + \pi_{OFF}\, t \\ \pi_{OFF} = \pi_{ON}\, a + \pi_{OFF}(1-t) \end{array} \quad \Rightarrow \quad \pi_{OFF} = \frac{a}{t}\pi_{ON}$$

$$\pi e = 1 \quad \Rightarrow \quad \pi_{ON} + \pi_{OFF} = 1 \qquad \pi_{ON}\left(1 + \frac{a}{t}\right) = 1$$

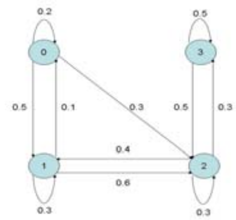
substituting from above

$$\pi_{ON} = \frac{t}{a+t} \quad \Rightarrow \quad \pi_{OFF} = \frac{a}{a+t}$$

Example

Consider the Time Homogeneous Markov Chain with one step transition matrix for the states {0, 1, 2, 3} given below.

$$P = \begin{bmatrix} .2 & .5 & .3 & 0 \\ .1 & .3 & .6 & 0 \\ 0 & .4 & .3 & .3 \\ 0 & 0 & .5 & .5 \end{bmatrix}$$



Find $P^{(2)}$, $P^{(4)}$, and $P^{(16)}$, what is noticeable about $P^{(16)}$?

From C-K equation $P^{(n)} = (P)^n$

$P^{(2)} = P*P =$

|0.09	**0.37**	**0.45**	**0.09|**
|0.05	**0.38**	**0.39**	**0.18|**
|0.04	**0.24**	**0.48**	**0.24|**
|0	**0.2**	**0.4**	**0.4 |**

Example

Find $P^{(2)}$, $P^{(4)}$, and $P^{(16)}$, what is noticeable about $P^{(16)}$?

From C-K equation $P^{(n)} = (P)^n$

$P^{(4)} = P^{(2)}*P^{(2)} =$

|0.0446	0.2999	0.4368	0.2187|
|0.0391	0.2925	0.4299	0.2385|
|0.0348	0.2692	0.438	0.258 |
|0.026	0.252	0.43	0.292 |

$P^{(16)} = (P)^{16} =$

.034015	0.272114	0.433674	0.2601960
.034015	0.272112	0.433674	0.2602000
.034014	0.272109	0.433674	0.2602040
.034012	0.272104	0.433674	0.260210

Notice all rows become about the same and approach steady state probability π

Example

Determine $\pi^{(n)}$ for n = 1, 2,… 10 given the initial condition $\pi^{(0)}$ = [0,0,0,1]

From $\pi^{(n)} = \pi^{(n-1)} P$

$\pi^{(1)} = \pi^{(0)}P$=[0, 0, 0.5, 0.5]

$\pi^{(2)} = \pi^{(1)}P$=[0, 0.2, 0.4, 0.4]

$\pi^{(3)} = \pi^{(2)}P$=[0.02, 0.22, 0.44, 0.32]

$\pi^{(4)} = \pi^{(3)}P$=[0.026, 0.252, 0.43, 0.292]

$\pi^{(5)} = \pi^{(4)}P$=[0.0304, 0.2606, 0.434, 0.275]

$\pi^{(6)} = \pi^{(5)}P$=[0.03214, 0.26698, 0.43318, 0.2677]

$\pi^{(7)} = \pi^{(6)}P$=[0.033126, 0.269436, 0.433634, 0.263804]

$\pi^{(8)} = \pi^{(7)}P$=[0.033569, 0.270847, 0.433592, 0.261992]

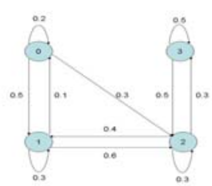
$\pi^{(9)} = \pi^{(8)}P$=[0.033799, 0.271475, 0.433653, 0.261074]

$\pi^{(10)} = \pi^{(9)}P$=[0.033907, 0.271803, 0.433658, 0.260633]

Example

Consider the Time Homogeneous Markov Chain with one step transition matrix for the states {0, 1, 2, 3} given below.

$$P = \begin{bmatrix} .2 & .5 & .3 & 0 \\ .1 & .3 & .6 & 0 \\ 0 & .4 & .3 & .3 \\ 0 & 0 & .5 & .5 \end{bmatrix}$$



Determining the steady state probability vector we get $\pi = \lim_{n\to\infty} \pi^{(n)}$

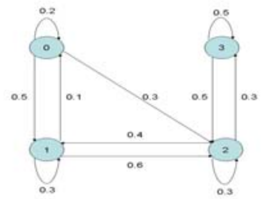
Solving $\pi = \pi p$ and $\sum_{i \in S} \pi_i = 1$

Results in π = [0.034013, 0.27210, 0.43367, 0.260204]

Example

Find the probability of the first passage time from state 3 to state 1 in 3 steps $f_{31}^{(3)}$

$$P = \begin{bmatrix} .2 & .5 & .3 & 0 \\ .1 & .3 & .6 & 0 \\ 0 & .4 & .3 & .3 \\ 0 & 0 & .5 & .5 \end{bmatrix}$$



Determining the first passage times we use

$$f_{ij}^{(1)} = p_{ij}$$

$$f_{ij}^{(2)} = p_{ij}^{(2)} - f_{ij}^{(1)} p_{jj}$$

$$\vdots$$

$$f_{ij}^{(n)} = p_{ij}^{(n)} - \sum_{k=1}^{n-1} f_{ij}^{(k)} p_{jj}^{(n-k)} \qquad n = 2, 3, \ldots$$

$f_{31}^{(1)}=0$

$f_{31}^{(2)} = P_{31}^{(2)} - f_{31}^{(1)} P_{11}^{(1)} = 0.2$

$f_{31}^{(3)} = P_{31}^{(3)} - (f_{31}^{(1)} P_{11}^{(2)} + f_{31}^{(2)} P_{11}^{(1)}) = 0.16$

Exampe

- Determine the mean first passage time from 3 to 1

$E[T_{31}] = \pi_1^{(0)}(I-R_1)^{-1}e$

$\pi_1^{(0)} = [0, 0, 1]$

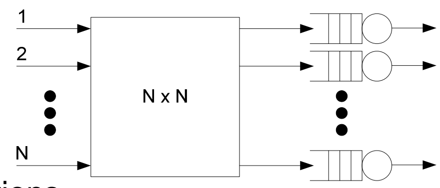
$$P = \begin{bmatrix} .2 & .5 & .3 & 0 \\ .1 & .3 & .6 & 0 \\ 0 & .4 & .3 & .3 \\ 0 & 0 & .5 & .5 \end{bmatrix}$$

$$R_1 = \begin{vmatrix} 0.2 & 0.3 & 0 \\ 0 & 0.3 & 0.3 \\ 0 & 0.5 & 0.5 \end{vmatrix}$$

$E[T_{31}] = \pi_1^{(0)}(I-R_1)^{-1}e = 6$

Markov Chain Example

- Analyze N x N non-blocking output buffered switch



- Assumptions
  - Arrival streams are independent
  - Bernoulli arrival process
  - Service time deterministic – D
  - Buffer size fixed – SS
  - Uniform distribution of traffic

Performance Evaluation

- Define embeded Markov Chain at slot times

  $\pi_{i,j}$ = Prob{ *i* class '1' cells, *j* class '2' cells}

  $\pi_n = [\pi_{0,n}, \pi_{1,n-1}, \ldots, \pi_{n,0}]$

  $\pi = [\pi_0, \pi_1, \pi_2, \ldots, \pi_k]$

- Solve for steady-state probabilities

  $\pi = \pi \cdot P$

  where P is state transition matrix

$$P = \begin{bmatrix} \alpha_{00} & \alpha_{01} & \alpha_{02} & 0 & 0 & \cdots & \cdots & 0 & 0 \\ \alpha_{10} & \alpha_{11} & \alpha_{12} & 0 & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \alpha_{21} & \alpha_{22} & \alpha_{23} & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \alpha_{32} & \alpha_{33} & \alpha_{34} & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & \alpha_{K,K-1} & \alpha_{K,K} \end{bmatrix}$$

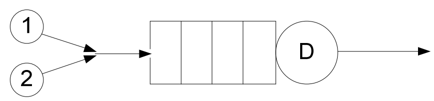
  Also use normalization condition

  $\pi \cdot \underline{e} = 1$ where $e^T = [1, 1, 1, \ldots, 1]$

- Exact form of *P* depends on space priority scheme modeled - for details see posted Infocom paper

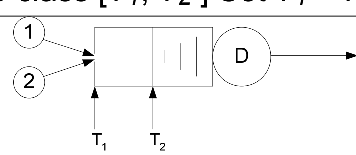
Performance Evaluation

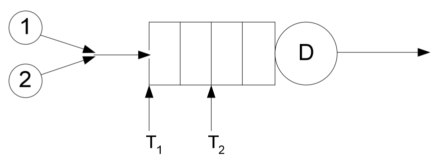
- No Priority Scheme
  - Cells accepted into the buffer in FCFS fashion.
  - When buffer is full, all packets are rejected.



Performance Evaluation

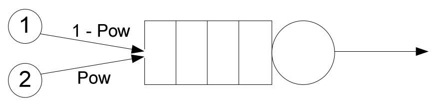
- Partial Buffering Scheme – (Nested Thresholds)
  - Define a threshold T_i for each class *i*
  - If number in the system ≥ T_i ,all new class *i* packets dropped
  - Here two class [T_1, T_2] Set T_1 = K





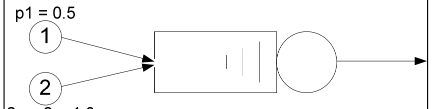
Performance Evaluation

- Pushout with overwrite probability (*Pow)*
  - Admit all packets until buffer full
  - If buffer full class ‘1’ pushout class ‘2’ with probability *1-Pow*
  - If buffer full class ‘2’ pushout class ‘1’ with probability *Pow*



Performance Evaluation

- Validate Analytical Model with Simulation
- Experiment 1



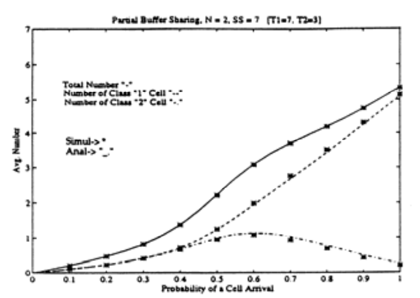


Figure 5: Average Number Curves: Nested Threshold Scheme (h=0.5)

Performance Evaluation

- Experiment 2
  - Define grade of service requirements
  - Δ_1 Acceptable Loss Probability for class '1' cells
  - Δ_2 Acceptable Loss Probability for class '2' cells
- For specific traffic mixture (% class 1, % class 2)
  - Determined maximum offered load (MOL)

Performance Evaluation

$\Delta_1 = 10^{-10}$	MOL for (% class 1 , % class 2) traffic mix								
$\Delta_2 = 10^{-6}$	10,90	20,80	30,70	40,60	50,50	60,40	70,30	80,20	90,10
No Priority	0.43052	0.35350	0.32219	0.30939	0.30829	0.31776	0.34042	0.38569	0.48840
Partial Buffer [7,6]	0.67295	0.55555	0.49824	0.46794	0.45435	0.45439	0.46928	0.50654	0.59414
Pushout Threshold	0.73945	0.62802	0.57086	0.54019	0.52667	0.52749	0.54407	0.58386	0.67303
Pushout Pow	0.73945	0.62802	0.57086	0.54019	0.52667	0.52749	0.54407	0.58386	0.67303
% improvement over Partial Buffer	9.88%	13.04%	14.57%	15.44%	15.92%	16.08%	15.94%	15.44%	13.67%