Proceedings of the 10th WSEAS International Conference on COMMUNICATIONS, Vouliagmeni, Athens, Greece, July 10-12, 2006 (pp590-594)

Apply Markov Decision Process to Class-based Packet Buffer Management

CHING-LUNG CHANG AND JIA-KAI LIN
Department of Computer Science and Information Engineering
National Yunlin University of Science & Technology
123 Sec. 3 University Rd. Douliu, Yunlin
TAIWAN 640, R.O.C

Abstract: - With the development of Internet and rapid growths of the network bandwidth, the quality of network transmission has a significant improvement. However, in the pluralistic network applications, the dramatic network bandwidth may be occupied by some of computers or users, such as ill-computer or improper peer-to-peer users. It affects the communication quality of service (QoS) of general network users or applications. How to deal with the unfair network bandwidth usage problem becomes a critical issue on network management. In general, the network bandwidth management can be realized by packet buffer management or output packet scheduling. In this paper, we formulate the packet admission control of buffer management as the Markov Decision Process (MDP) optimization problem. All of the received packets are classified into three types: real-time class, good-user class, and bad-user class. According to the QoS of packet class and traffic condition, the allowing policy derived by the MDP decides whether accept or discard the received packet for getting an optimal packet buffer management. With computer simulation, we demonstrate that the MDP-based buffer management can confine the buffer occupancy of bad-user class packet, and decrease the discarding rate of real-time class packet and good-user class packet significantly.

Key-Words: - Quality of Service (QoS), Packet Buffer Management, Packet admission control, Markov Decision Process (MDP)

1 Introduction

With the development of Internet and rapid growths of the network bandwidth, the quality of network transmission has a significant improvement and the traffic on the Internet has caused a gradual shift from data to multimedia communication. Under the IP-based network, although the bandwidth has dramatically increased to the 10 Gbps, the real-time application stills hard to obtain the required quality of services (QoS) [1]. Large amount of network bandwidth is wasted by the ill-computer defected by the computer virus or occupied by the abnormal user who may use peer-to-peer (p2p) tools to get the large amount of music/movie files all day long. How to deal with the unfair network bandwidth usage problem becomes a critical issue in network management.

In the network bandwidth management, there are many mechanisms proposed for router design so far. Drop Tail (DT) is one of the familiar schemes which sets a buffer threshold and accepts any incoming packets until the buffer occupancy reaching the threshold. Once the buffer occupancy has over the limited threshold, the incoming packet is discarded until the buffer occupancy reduces to below the threshold value. However, there are some drawbacks of DT scheme. For real-time streaming data, such as voice or video streaming, its packet may incur large network jitter or packet loss, if the status of packet buffer is usually in full or near to full. It affects the network QoS.

In contrast to the DT scheme, Fair Queueing (FQ) [2] is even share the buffer resource to all of service classes. The advantage of FQ is that it confines some of users to occupy excessive amount of buffer resource. Thus it avoids unfair buffer resource usage. The drawback of FQ is that the unused buffer space allocated to other service class can not be shared by needed class. It results in lower buffer utilization.

In addition to the above mechanisms, Random Early Detection (RED) [3] considers the impacts of network loading on the efficiency of network. The RED scheme defines two buffer thresholds, maximum threshold and minimum threshold, for the buffer occupancy. When the buffer occupancy is less than a minimum threshold, all of incoming packets are allowed into the buffer. If the buffer occupancy greats than the maximum threshold, all of arriving packets are discarded. While the buffer occupancy is between the minimum and maximum thresholds, all the incoming packets are discarded with a probability. Based on the buffer occupancy, the discarding probability may proportional increase with exponential or linear function.

Unlike previous works, we formulate the packet admission control of buffer management as the Markov Decision Process (MDP) optimization problem. The MDP is a decision scheme used in the stochastic process [4]. In the wireless cellular networks, there were many related papers [5, 6, 7, 8] that applied the MDP technique to call admission control (CAC) and achieved better QoS. In this paper, all of incoming packets are classified into three categories. These are real-time class, good-user class, and bad-user class. And based on the QoS metrics of packet class and traffic condition, the MDP technique derives the packet allowing policy which decides whether to accept or to discard the incoming packet. The allowing policy is transfer to a decision table. Once a packet arrives, we look up the decision table with the current buffer status and the incoming packet type. The lookup results are used to decide whether to accept the incoming packet or not.

The rest of the paper is structured as follows. The packet admission control model is described in Section 2. How to apply MDP to derive the decision table in packet admission control is presented in Section 3. In Section 4, we have a computer simulation and some of performance comparison with DT scheme and FQ scheme. Finally, we make a brief conclusion.

2 Packet Admission Control Model

Network bandwidth management can be designed with the packet buffer management or packet output scheduling. In the packet buffer management, all of the arriving packets are decided to accept or to reject into the packet buffer, as Fig. 1 shows. At the output side, we adopt the First-in First-out (FIFO) as the output packet scheduling in this paper. As Fig.1 shows, the packet buffer management can be modeled as a single-queue single-server system. Thus we can apply M/D/1 queueing model to the packet admission control which have a Poisson arrival process and a server with constant service rate. In this model, the received packet is classified into three of service types based on the required network QoS. In the current status, the packet is classified into real-time class, good-user class, and bad-user class. Here we define the bad-user as that someone whose computer may be infected by computer virus or who may improperly use network resources, such as using peer-to-peer (p2p) tools to get the large amount of music/movie files all day long. Such user packets may often result in excessive occupying network resource. The behavior of bad-user will affect the QoS of good-user class packet or real-time class packet transmission. Therefore we call these kinds of packet as the bad-user packet. In contrast to bad-user class packet, the packet generation rate of good-user is usually lower than bad-user obviously.

In this paper, according to the buffer usage we expect the MDP technique to derive an optimal decision policy for packet admission control. Based on the decision table resulted by MDP to provide optimal packet buffer management, we can limit the affect of bad-user packet and provide lower discarding rate of real-time packet and good-user packet.

In the following, we assume the traffic distribution of good-user class and bad-user class are Poisson process and the real-time class packet is generated by constant bit rate.

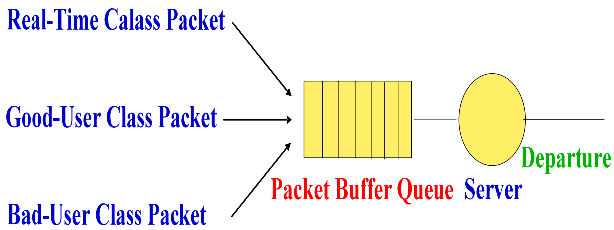

Fig.1 The model of packet buffer management

3 Markov Decision Process in Buffer Management

In general, the design issues of MDP are described as follows:

- It should consider the characteristics of system operation, observe the system in each discrete time, and present the status of the system into state x (x=0, 1, …, M), where x is the state number.
- Based on the current state x, the MDP chooses an action a when an input event arrives. The action a is one of probable action in {1, 2, …, A}.
- The selected action will affect the system cost and system reward.
- The action a chosen in the state x will decide the transition probability from state x to state y (y=0, 1, …, M), denoted as $P_{xy}(a)$.
- The goal of MDP is to derive the proper action for each state.

- Based on the system reward or cost function, an optimal action policy which meets the stationary transition probability is derived by the MDP. Thus we have minimal expected cost or maximal average reward.

According to the Fig. 1, we convert the packet admission control problem into the MDP optimization problem and apply the linear programming technique [9] to derive the optimal buffer management policy. The MDP design process in buffer management is explained as follows:

- ***Decision Epochs:*** The decision is made only at the instant of a packet arrival. At this moment, we will determine whether the packet can enter the buffer or not. In this phase, we assume there are three types of packet arrival event: real-time packet event, good-user packet event, and bad-user packet event.

- ***State Space:*** The state space X represents all possible combination of the number of packet of each packet class in the buffer, described as

$$X = \{x \mid x = (x_g, x_b, x_r), x_g \geq 0, x_b \geq 0, x_r \geq 0, \quad x_g + x_b + x_r \leq B\} \tag{1}$$

, where the x_g, x_b, and x_r stand for the packet counts of good-user packet, bad-user packet, and real-time packet in the buffer, respectively. B is the buffer size which unit is packet count.

- ***Action Space:*** The action space A is a set of vectors representing the probable action when packet arrives. The action can be accepting or rejecting the arrival packet enters the buffer. It can be represented by binary value.

$$A = \{a \mid a = (a_g, a_b, a_r); a_g, a_b, a_r \in \{0(\text{reject}), 1(\text{accept})\}\} \tag{2}$$

, where a_g, a_b, and a_r, are the action of good-user packet event, bad-user packet event, and real-time packet event, respectively. The value 0 represents to reject the incoming packet and to accept the incoming packet if the value is equal to 1.The action space for state space x can be expressed as

$$A_x = \begin{cases} (1, 1, 1), & \text{if } x = (0, 0, 0) \\ (0, 0, 0), & \text{if } x_g + x_b + x_r = B \\ \{(a_g, a_b, a_r) \mid a_g, a_b, a_r \in \{0, 1\}, \\ \text{otherwise} \end{cases} \tag{3}$$

When the packet buffer is empty, all of incoming packets allow entering buffer. On contrast, if the packet buffer is full, all of packets are reject. In other state, the admission control is decided by different policy.

- ***Reward Function:*** According to the action, the reward function of state x is expressed as

$$r(x, a) = w_g(x_g + e_g \cdot a_g) + w_b(x_b + e_b \cdot a_b) + w_r(x_r + e_r \cdot a_r) \tag{4}$$

, where w_g, w_b, w_r are the weighting value for each packet type. The packet arrival event can be represented by a vector ***e*** which consists of e_g, e_b, and e_r.

$$e = (e_g, e_b, e_r) = \begin{cases} (1, 0, 0), \text{ if good-user class packet arrival} \\ (0, 1, 0), \text{ if bad-user class packet arrival} \\ (0, 0, 1), \text{ if real-time class user packet arrival} \end{cases} \tag{5}$$

In order to meet the maximal reward and minimal system cost, we define the accepting cost is

$$c(x, a) = c_g \cdot e_g \cdot a_g + c_b \cdot e_b \cdot a_b + c_r \cdot e_r \cdot a_r \quad \text{where } c_b > c_g > c_r \tag{6}$$

, where c_g, c_b, c_r are the accepting cost for each packet type and $c_b > c_g > c_r$. Due to the bad-user class packet will affect the QoS of other user packet, therefore we set the c_b to the largest.

- ***Transition Probabilities:*** The transition probabilities are the probability from current state change to next state. The sojourn time τ(x, a) is the time stayed in next state when action $a \in A_x$ is chosen in current state *x*. The goal of sojourn time is to convert the continuous time Markov chain to the discrete time Markov chain. The definition is represented as:

$$\tau(x, a) = \{\lambda_g a_g + \lambda_b a_b + \lambda_r a_r + \mu\}^{-1} \tag{7}$$

, where λ_g, λ_b, λ_r are the arrival rate of good-user packet, bad-user packet and real-time packet, respectively. The value μ is the packet departure rate. In this paper, we adopt the FIFO output scheme. Thus μ is the output bandwidth. The transition probability from state x to state y is defined as

$$P(y \mid x, a) = \begin{cases} a_g \cdot \lambda_g \cdot \tau(x, a), \text{ if } y = x + (1,0,0), \\ a_b \cdot \lambda_b \cdot \tau(x, a), \text{ if } y = x + (0,1,0), \\ a_r \cdot \lambda_r \cdot \tau(x, a), \text{ if } y = x + (0,0,1), \\ \mu \cdot \tau(x, a), \text{ if } y = x - (1,0,0), \\ \mu \cdot \tau(x, a), \text{ if } y = x - (0,1,0), \\ \mu \cdot \tau(x, a), \text{ if } y = x - (0,0,1), \\ 0, \text{ if } y = x \end{cases} \tag{8}$$

Therefore, for deriving the optimal packet admission control policy, we formulate the packet buffer management problem into linear programming as follows:

$$\text{Maximize} \sum_{x \in X} \sum_{a \in A_x} [r(x, a)\text{-}c(x, a)]\, \tau(x, a)\pi(x, a) \tag{9}$$

$$\text{subject to} \sum_{a \in A_y} \pi(y, a) - \sum_{x \in X} \sum_{a \in A_x} P(y|x, a)\pi(x, a) = 0, \ \forall y \in X \tag{10}$$

$$\sum_{x \in X} \sum_{a \in A_x} \tau(x, a)\pi(x, a) = 1, \tag{11}$$

$$\pi(x, a) \geq 0, \ x \in X, \ a \in A_x \tag{12}$$

The term r(x, a)-c(x, a) in the objective function is used to meet the maximal system utilization when the accepting cost is the minimal. The term $\pi(x, a)$ and $\pi(y, a)$ are the decision variable which represent the steady state probability of state *x* or state *y* and state *y* is the next state of state *x* when chooses action *a* in state *x*.

For solving the above linear programming problem, an optimal simulation tool-Lingo is adopted in this paper. The computation results of decision space A_x for each state *x* is applied to the packet admission control for obtaining the optimal buffer management policy.

4 Simulation

In this section, we use computer simulation to reveal the performance of proposed MDP-based buffer management and to compare with DT and FQ methods in the metric of packet discarding rate. The policy of MDP-based buffer management is based on the results of linear programming tool - Lingo described in Section 3. The associated MDP parameters are showed in Table 1 which affects the QoS of each service class. Based on the Fig. 1 network architecture, a simulation tool NS2 (Network Simulator Version 2) is used in this paper. The traffic pattern of good-user packet and bad-user packet is Poisson process in this simulation. And we use constant bit rate to generate real-time packet. A FIFO technique is applied in the packet output which service rate is μ.

The related simulation parameters and values are shown in Table 1. The unit of packet buffer size ***B*** is depicted in the number of packet and we use different buffer size {12, 120, and 12000} in the simulation. The simulation time ***T*** is set to 60 seconds. All of the packet size ***S*** generated in the simulation is same as 1000 bytes.

Table 1. SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
λ_g	6M B/sec	c_r	5
λ_b	10M B/sec	w_g	10
λ_r	5M B/sec	w_b	6
μ	11M B/sec	w_r	12
c_g	6	T	60 second
c_b	10	S	1000
B	(12, 120, 12000)		

Figure 2 to Fig. 4 illustrates the simulation results with different packet buffer capacity. As Fig. 2 shows, when the buffer size is small to store merely 12 packets, DT and FQ result in the higher discarding rate for all types of service class. For the real-time class packet, the scheme of FQ also has 26% discarding rate. On contrary, the discarding rate of MDP-based buffer management has a significant decrease in good-user packet and real-time packet, although it sacrifices the service of bad-user packet.

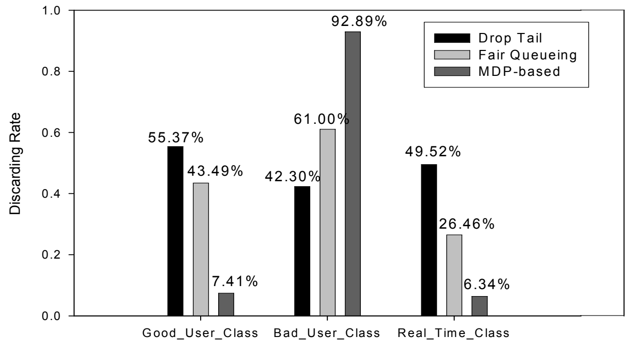

Fig. 2 The capacity of packet buffer is 12 packets

When the buffer size increased, the packet discarding rate of MDP-based for good-user packet and real-time packet will near reduce to zero as Fig. 3 shows. Meanwhile, the performance enhancement with DT and FQ is limited when the packet buffer increases. As Fig. 4 depicts, when the capacity of packet buffer is up to 12000 packets, there are no any packet loss of good-user class packet and real-time class packet with MDP-based buffer management. However, the performance of good-user class packet and real-time class packet with DT and FQ is not improved. Thus MDP-based can guarantee the transmission quality of good-user class packet and real-time class packet

.

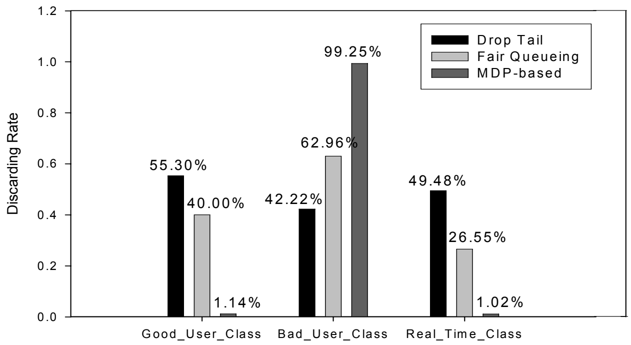

Fig. 3 The capacity of packet buffer is 120 packets

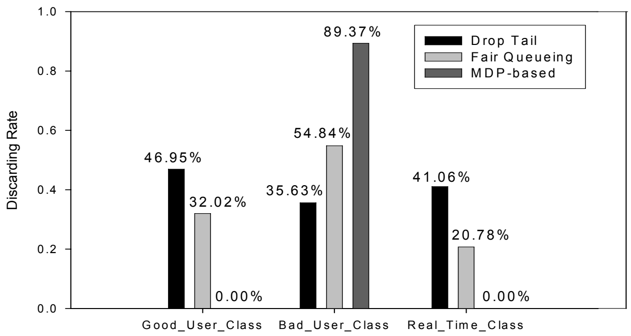

Fig. 4 The capacity of packet buffer is 12000 packets

5 Conclusion

In this paper, we engage class-based buffer management for network bandwidth management and propose a MDP-based buffer management scheme to derive optimal packet admission control policy. The packet buffer management is modeled to M/D/1 model in this paper and the packet admission control is formulated as linear programming problem based on the MDP technique. Using the optimization tool Lingo, we solve the formulated linear programming and derive the action space for each buffer state. Applying the action space to the packet admission control and using the NS2 to do computer simulation, the results reveal that the MDP-based buffer management can confine bad-user class packet occupied network bandwidth efficiently, and decrease the discarding rate of real-time class packet and good-user class packet significantly.

References:

[1] F.M. Chiussi, et al. "Providing QoS guarantees in packetswitches", *IEEE Global Telecommunications Conference-Globecom'99*, May 1999, pp. 1582-1590.

[2] A. Demers, S. Keshav, and S. Shenker, "Analysis and Simulation of a Fair Queueing Algorithm", *In Proceedings of the ACM SIGCOMM Austin, Texas*, September 1989.

[3] S. Floyd, and V. Jacobson, "Random Early Detection gateways for Congestion Avoidance", *IEEE/ACM Transactions on Networking*, August 1993.

[4] H. Tijms, *Stochastic Modelling and Analysis: A Computational Approach*. New York: Wiley, 1986.

[5] Huan Chen, and Chih-Chuan Cheng, "Priority Access for Hybrid Handoff Wireless Systems", *IEEE Vehicle Technology Conference*, Vol. 2, 26-29 Sept. 2004, pp. 1028 – 1032.

[6] Huan Chen, Chih-Chuan Cheng, Mei-Hsiu Chi, and Hsi-Hsun Yeh, "A sensitivity analysis approach to the optimal policy range of the resource access control in wireless systems", *IEEE Wireless Telecommunications Symposium*, 28-30 April 2005 pp. 100 – 105.

[7] N. Nasser, and H. Hassanein, "An optimal and fair call admission control policy for seamless handoff in multimedia wireless networks with QoS guarantees", *IEEE Global Telecommunications Conference-GLOBECOM '04*, Vol. 6, 29 Nov.-3 Dec. 2004, pp. 3926 – 3930.

[8] Jihyuk Choi, Taekyoung Kwon, Yanghee Choi, and M. Naghshineh, **"**Call admission control for multimedia services in mobile cellular networks: a Markov decision approach**"**, *Computers and Communications, 2000. Proceedings. ISCC 2000. Fifth IEEE Symposium* on 3-6 July 2000, pp. 594 – 599.

[9] Shu-Cherng Fang, Sarat Puthenpura, *Linear optimization and extensions: theory and algorithms*. Englewood Cliffs: Prentice Hall, 1993.

[10] The network simulator- *NS2 Notes and documentation and source code*. [Online] http://www.isi.edu/nsnam/ns/, retrieved 2006/5/6