1997年度日本オペレーションズ・リサーチ学会
春季研究発表会

Determining the optimum packet length on the network layer

01205220 NTT Multimedia Networks Laboratories Masaaki Shinohara

1. Introduction

Two models for packet length optimization problem for IP traffic are presented. They are M/G/1 model and $M^{[X]}/G/1$ model. The objective is the minimization of queuing-time related measure, and retransmission of packets is not considered. The packet length optimization problem is formulated and the optimality condition is derived. For certain cases, explicit expressions of the optimum packet length are given.

2. Packet transmission model

Fig.1 shows the packet transmission model treated in this paper.

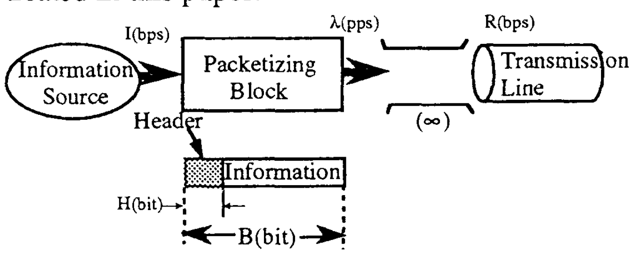

Fig.1 Packet transmission model

Information stream is generated from the source with rate I(bps, bits per second) and is packetized into packets of constant length B(bit). Those packets are offered to a single transmission line with rate R(bps). The mean arrival rate of the packets is λ (pps, packets per second). A packet of length B (bit) consists of a header part of length H(bit) and an information part. The queuing time related measure T given by Eq.(1) is considered.

$$T=W_q+kh \qquad (1)$$

$$h=B/R \qquad (2)$$

Here, W_q is the queuing time, h is the time length of packet and k is a real number. The capacity of buffer for waiting packets is assumed to be infinite. Since header length H(bit) is constant, transmission efficiency decreases as the packet length B(bit) is decreased. On the other hand, queuing time related measures such as queuing + servicing time (k=1), and packetizing + queuing + servicing time (k=2) tend to decrease as B is decreased. As a tradeoff of these two effects, there must be the optimum packet length B(bit) which minimizes the queuing time related measure T for a given information generation rate I(bps).

3. M/G/1 model

The packet transmission model of Fig.1 is remodeled as M/G/1 model of Fig.2 by inserting the randomizing mechanism between the packetizing block and the transmission line. Since the packet output from the randomizing mechanism is Poissonian, the transmission line part can be modeled as M/G/1.

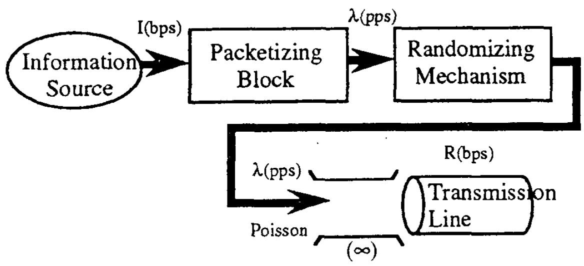

Fig.2 M/G/1 model

(Example 3.1) Case of M/D/1 with H=0 and k=0. The queuing time $W_q(x)$ is given by Eq.(3).

$$T(x)=W_q(x)=\frac{x}{2}\frac{a}{1-a} \qquad (3)$$

x=B/R is a decision variable and a=λx=I/B is a given variable (0≤a<1). In this case, $T=W_q(x)$ decreases as the time length of packet is decreased. That is, $T=W_q(x)$ approaches zero as the packet length approaches zero. Note that this property holds also with a case of M/G/1 and k>0.

(Example 3.2) Case of M/D/1 with H≠0 and k=0. The queuing time $W_q(x)$ is given by Eq.(4).

$$T(x) = W_q(x) = \frac{x+h_0}{2} \frac{\frac{a}{x}(x+h_0)}{1-\frac{a}{x}(x+h_0)} \quad (4)$$

Here, h_0=H/R(sec) is the time length of header.

By solving $\frac{d}{dx}T(x)=0$, the optimum packet time length x_{opt} is given by Eq.(5).

$$x_{opt} = h_0 \frac{1+a}{1-a} \quad (5)$$

Corresponding packet queue delay $W_q(x_{opt})$ and transmission utilization U_{opt} are given as below.

$$W_q(x_{opt}) = h_0 \frac{2a}{(1-a)^2} \quad (6)$$

$$U_{opt} = a\frac{x_{opt}+h_0}{x_{opt}} = \frac{2a}{1+a} \quad (7)$$

(Example 3.3) Case of M/G/1 with H≠0 and k≠0.

$$x_{opt} = \frac{ah_0}{1-a} + \sqrt{\left(\frac{ah_0}{1-a}\right)^2 - \frac{ah_0^2\{a(2Q-1)-1\}}{(1-a)\{a+2Q(1-a)\}}} \quad (8)$$

$$Q = \frac{k}{1+C_h^2} \quad (9)$$

C_h : coefficient of variation for holding time $x+h_o$

4. $M^{[X]}$/G/1 model

The whole model of Fig.1, from the information source to the transmission line, is modeled with a batch arrival queue $M^{[X]}$/G/1 model. Information generation follows a Poissonian batch arrival with mean rate λ and mean batch volume gh, where g is mean batch size (number of packets in a batch) and h is mean packet time length. The queueing time $W_q(x)$ is given by Eq.(10).

$$W_q = \frac{1}{2(1-\lambda gh)}\left\{\lambda gh^{(2)} + h\left(\frac{g^{(2)}}{g}-1\right)\right\} \quad (10)$$

(Example 4.1) Case of $M^{[X]}$/D/1 with Binomial distributed batch size X, H=0 and k=0.
Putting $h^{(2)}=h^2$, $g^{(2)}=2g^2$ and λgh=a(constant) into Eq.(10), Eq.(11) is obtained.

$$T(x) = W_q(x) = -\frac{1}{2}h + \frac{a}{\lambda(1-a)} \quad (11)$$

In this case, $W_q(x)$ increases as h is decreased while keeping a and λ constant, contrary to the case in Example 3.1 of M/D/1 model.

(Example 4.2) Case of $M^{[x]}$/D/1 with H≠0 and k≠0 (λgx=a and λ are constant). The queuing time related measure T(x) is given by Eq.(12).

$$T(x) = W_q(x) + k(x+h_0) \quad (12)$$

$$W_q(x) = -\frac{1}{2}(x+h_0) + \frac{(1+C_g^2)\left(\frac{a}{\lambda}+\frac{ah_0}{\lambda x}\right)}{2\left(1-a-\frac{ah_0}{x}\right)} \quad (13)$$

By solving $\frac{d}{dx}T(x)=0$, x_{opt} is obtained.

$$x_{opt} = \frac{2a^2h_0^2\left(k-\frac{1}{2}\right) - (1+C_g^2)\frac{a}{\lambda}h_0}{2a(1-a)(k-\frac{1}{2})h_0 - \sqrt{\frac{2a(1-a)^2}{\lambda}(k-\frac{1}{2})h_0(1+C_g^2)}}$$

$$k \geq \frac{1}{2} \quad (14)$$

5. Conclusion

The optimum packet length is decided by trading off overhead effect and queuing effect. The effect of packet retransmission is not considered in this paper, which should be considered, if necessary, in the performance evaluation on the transport layer or TCP layer. Since the propagation delay through transmission line is constant, the result in this paper is not influenced by its existence.