1. The Kleinrock Independence Approximation

We now formulate a framework for approximation of average delay per packet in telecommunications networks. Consider a network of communication links as shown in Figure 1. Assume that there are several packet streams, each following a unique path that consists of a sequence of links through the network. Let x_s, in packets/sec, be the arrival rate of the packet stream *s*. Then the total arrival rate at link (*i*, *j*) is

$$l_{ij} = \sum_{\substack{\text{all packet streams } s \\ \text{crossing link } (i,j)}} x_s$$

The preceding network model is well suited for virtual circuit networks, with each packet stream modelling a separate virtual circuit. For datagram networks, it is sometimes necessary to use more general model that allows bifurcation of the traffic of a packet stream. Here are again several packet streams, each having a unique origin and destination. However, there may be several paths followed by the packets of a stream (see Figure 2). Assume that no packets travel in a loop, let x_s denote the arrival rate of packet stream *s*, and let $f_{ij}(s)$ denote the fraction of the packets of stream s that go through link (*i*, *j*). Then the total arrival rate at link (*i*, *j*) is

$$l_{ij} = \sum_{\substack{\text{all packet streams } s \\ \text{crossing link } (i,j)}} f_{ij}(s) x_s$$

We know from the special case of two tandem queues that even if the packet streams are Poisson with independent packet lengths at their point of entry into the network, this property is lost after the first transmission line. To resolve the dilemma, it was suggested by Kleinrock that merging several packet streams on a transmission line has an effect akin to restoring the independence of interarrival time and packet lengths.

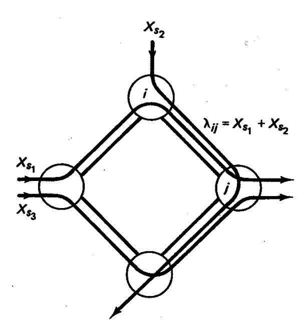

Figure 1 Model suitable for virtual circuit networks

It was concluded that it is often appropriate to adopt an *M/M/1* queueing model for each communication link regardless of the interaction of traffic on this link with traffic on other links. This is known as Kleinrock independence approximation and seems to be a reasonably good approximation for systems involving Poisson stream arrivals at the entry points, packet lengths that are nearly exponentially distributed, a densely connected network and moderate-to-heavy traffic load. Based on this M/M/1 model, the average number of packets in queue or service at (*i, j*) is

$$N_{ij} = \frac{l_{ij}}{m_{ij} - l_{ij}} \tag{1.1}$$

where $1/m_{ij}$ is the average packet transmission time on link (*i, j*). The average number of packets summed over all queues is

$$N = \sum_{(i,j)} \frac{l_{ij}}{m_{ij} - l_{ij}} \tag{1.2}$$

so by Little's Theorem, the average delay per packet (neglecting processing and propagation delays) is

$$T = \frac{1}{g} \sum_{(i,j)} \frac{l_{ij}}{m_{ij} - l_{ij}} \tag{1.3}$$

where $g = \sum_s x_s$ is the total arrival rate in the system. If the average processing and propagation delay d_{ij} at link (*i, j*) is not negligible, this formula should be adjusted to

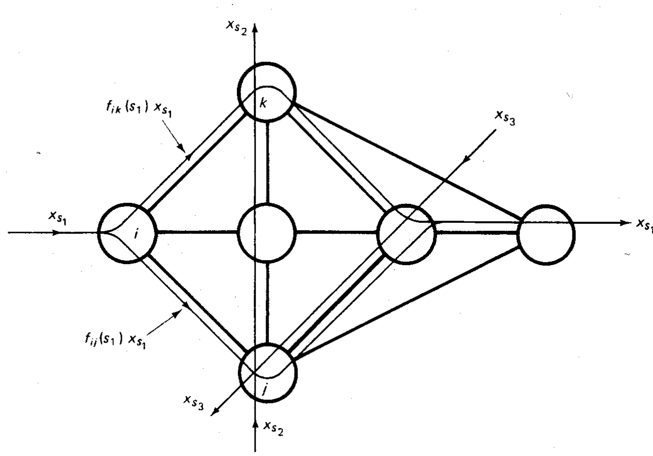

Figure 2 Model suitable for datagram networks

$$T = \frac{1}{g}\sum_{(i,j)} \frac{l_{ij}}{m_{ij} - l_{ij}} + l_{ij} d_{ij} \qquad (1.4)$$

Finally, the average delay per packet of a traffic stream traversing a path *p* is given by

$$T_p = \frac{1}{g} \sum_{\substack{all\,(i,j) \\ on \quad path \quad p}} \frac{l_{ij}}{m_{ij}(m_{ij} - l_{ij})} + \frac{1}{m_{ij}} + d_{ij} \qquad (1.5)$$

where the three terms in the sum above represent average waiting time in queue, average transmission time and processing and propagation delay, receptively.

In many networks, the assumption of exponentially distributed packet lengths is not appropriate. Given a deferent type of probability distribution of the packet lengths, one may keep the approximation of independence between queues but use the P-K formula for average number in the system in place of the *M/M/*1 formula (1.1). Equation (1.2) to (1.5) for average delay would then be modified in an obvious way.

For virtual circuit networks (Figure 1), the main approximation involved in the *M/M/1* formula (1.2) is due to the correlation of the packet lengths and the packet interarrival times at the various queues in the network. If somehow this correlation was not present (*e.g.*, if a packet upon departure from a transmission line was assigned a new length drawn from an exponential distribution), then the average number of packets in the system would be given indeed by the formula

$$N = \sum_{(i,j)} \frac{l_{ij}}{m_{ij} - l_{ij}}$$

In datagram networks that involve multiple path routing from some origin-destination pairs (Figure 2), the accuracy of the *M/M/1* approximation deteriorates for another reason.

2. Burke's Theorem

Consider an M/M/1, M/M/m or M/M/∞ system with arrival rate λ. Suppose that the system starts in steady-state. Then the following hold true:

1) The departure process is Poisson with rate λ.
2) At each time *t*, the number of customers in the system is independent of the sequence of departure times prior to *t*.

3. Networks of Queues – Jackson's Theorem

The main difficulty with analysis of networks of transmission lines is that the packet interval times after traversing the first queue are correlated with their lengths. It turns out that if somehow this correlation were eliminated (which is the premise of the Kleinrock independence approximation) and randomization is used to divide traffic among different routes, then the average number of packets in the system can be derived as if each queue in the network were *M/M/1*. This is an important result known as Jackson's Theorem.

Consider a network of K first-come first-serve, single-server queues in which customers arrive from outside the network at each queue i in accordance with independent Poisson processes at rate r_i. We allow the possibility that $r_i = 0$, in which case there are no external arrivals at queue i, but we require that $r_i > 0$ for at least one i. Once a customer is served at queue i, it proceeds to join each queue j with probability P_{ij} or to exit the network with probability $1-\sum_{j=1}^{K} P_{ij}$.

The routing probabilities P_{ij} together with the external input rates r_j can be used to determine the total arrival of customers l_j at each queue j, that is, the sum of r_j and the arrival rate of customers coming from other queues. Calculating l_j is fairly easy when the network is of the acyclic type as follow;

$$l_j = r_j + \sum_{i=1}^{K} l_i P_{ij} \qquad j = 1, \dots, K \quad (1.6)$$

These equations represent a linear system in which the rates l_j, $j = 1,\dots,K$, constitute a set of K unknowns. To guarantee that they can be solved uniquely to yield l_j, $j = 1,\dots,K$ in terms of r_j, P_{ij}, $i,j = 1,\dots,K$, we make a fairly natural assumption that essentially asserts that each customer will eventually exit the system with probability 1.

The service times of customers at the jth queue are assumed exponentially distributed with mean $1/m_i$ and are assumed mutually independent and independent of the arrival process at the queue. The utilization factor of each queue is denoted

$$r_j = \frac{l_j}{m_j} \qquad j = 1, \dots, K$$

and we assume that $r_j < 1$ for all j.

Jackson's Theorem*:* Assuming that $r_j < 1$, $j = 1, \dots, K$, we have for all $n_1,\dots,n_K \geq 0$,

$$P(n) = P_1(n_1)P_2(n_2)\dots P_K(n_K)$$

Where $n = (n_1,\dots,n_K)$ and

$$P_j(n_j) = r_j^{n_j}(1 - r_j), \qquad n_j \geq 0$$

Extension of Jackson's Theorem – State-dependent service rates: The model for Jackson's Theorem assumed so far requires that all queues have a single server. An extension to the multiserver case can be obtained by allowing the service rate at each queue to depend on the number of customers at the queue (homework2, question2!). Thus the model is same as before but the service time at the *j*th queue is exponentially distributed with mean $1/m_j(m)$, where *m* is the number of customers in the *j*th queue just before the customer's departure (*m* includes the customer). The single-queue version of this model includes as special cases the M/M/m and M/M/ queues, and can be analyzed by means of a Markov chain. The corresponding network of queues model can also be analyzed by means of a Markov chain, and is characterized by a product from structure for the stationary distribution.

Let us define

$$r_j(m) = \frac{l_j}{m_j(m)} \qquad j = 1,\ldots,K, \quad m = 1,2,\ldots$$

where lj is the total arrival rate at the *j*th queue determined by Eq. (1.6). Let us also define

$$\hat{P}_j(n_j) = \begin{cases} 1 & \text{if} \quad n_j = 0 \\ r_j(1)\, r_j(2) \ldots r_j(n_j) & \text{if} \quad n_j > 0 \end{cases}$$

We have:

Jackson's Theorem for State-Dependent Service Rates: We have for all states $n = (n_1,\ldots,n_K)$

$$P(n) = \frac{\hat{P}_1(n_1)\ldots\hat{P}_K(n_K)}{G}$$

assuming $0 < G < \infty$, where the normalizing constant G is given by

$$G = \sum_{n_1=0}^{\infty}\ldots\sum_{n_K=0}^{\infty}\hat{P}_1(n_1)\ldots\hat{P}_K(n_K)$$