but less accurate, representation in terms of deterministic vector flows, where this description is exact for the first moment, but only approximate for the others.

3.4 Moment-Matching Techniques for Overflow

Representing traffic by a small set of moments, although simpler than the full stochastic description, raises another problem. If we know only the first two moments of the traffic offered to a link, how can we compute the carried and overflow moments? One method, dating back to the introduction of alternate routing in telephone networks in the mid-1950s, is called the *moment-matching technique*. Let us now describe this class of methods and the particular problems raised by its use in network analysis.

A long-standing, much-studied problem in circuit switching is calculating traffic parameters subject to multiple overflows. This problem arises when a traffic stream, generally Poisson, is offered to some trunk group. The blocked calls are then offered to a second group, the calls blocked on this second choice are offered to a third one, and so on. The object is to compute the probability of not being able to make a connection on any of these groups, and to estimate the traffic carried on each one.

In principle, this problem is readily solved by repeated application of the techniques of Section 3.3, since the overflow process is also of the renewal type. In practice, this is not done for two reasons. Historically, the problem was studied, and partly solved, long before modern renewal theory was commonly used in teletraffic theory. Thus adequate methods already existed, making recourse to the exact renewal method somewhat redundant. Also, the renewal method is generally quite difficult to apply in practical cases because of numerical stability problems, long computation times, or both. For this reason, a discussion of the classical moment-matching systems is in order since these systems are the basis of all the techniques currently used in network analysis or synthesis.

Moment-matching techniques work as follows. The arrival process is represented by a small number of parameters, generally two or at most three. A process, called the *equivalent process*, is then selected to represent the actual arrival process. The parameters of the equivalent process are chosen in such a way that the moments of the traffic it generates are equal to the moments of the real offered traffic — hence the generic name moment-matching techniques. This equivalent process is then used to compute all the quantities of interest pertaining to the group: time and call congestion, moments of overflow and carried traffics, and so forth (see Fig. 3.2). The usefulness of this technique depends strongly on the possibility of choosing an equivalent process that yields accurate parameters for the overflow and carried traffics with reasonable computation times. Also, if the method is to be usable in network calculations, it must be reasonably easy to select the parameters of the equivalent process.

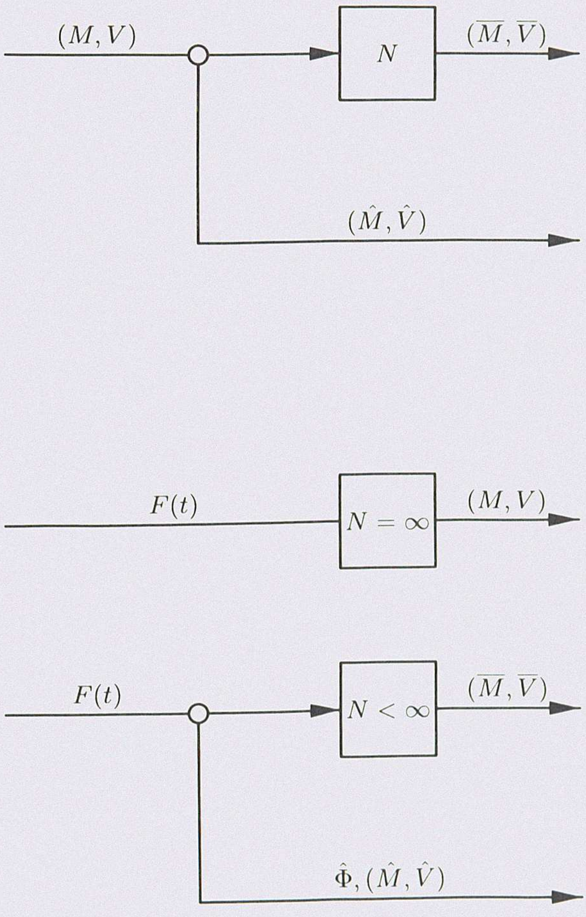

Figure 3.2 Moment-Matching Technique

Although not the only ones possible, renewal processes form an important class of equivalent processes. The general theory of renewal processes is well known, encompassing most of the equivalent methods currently used in network algorithms. For this reason, we review some of the more important processes of this type, only briefly indicating other types at the end of the discussion.

Equivalent Random Theory

Equivalent random theory (ERT) is the first application of the moment-matching

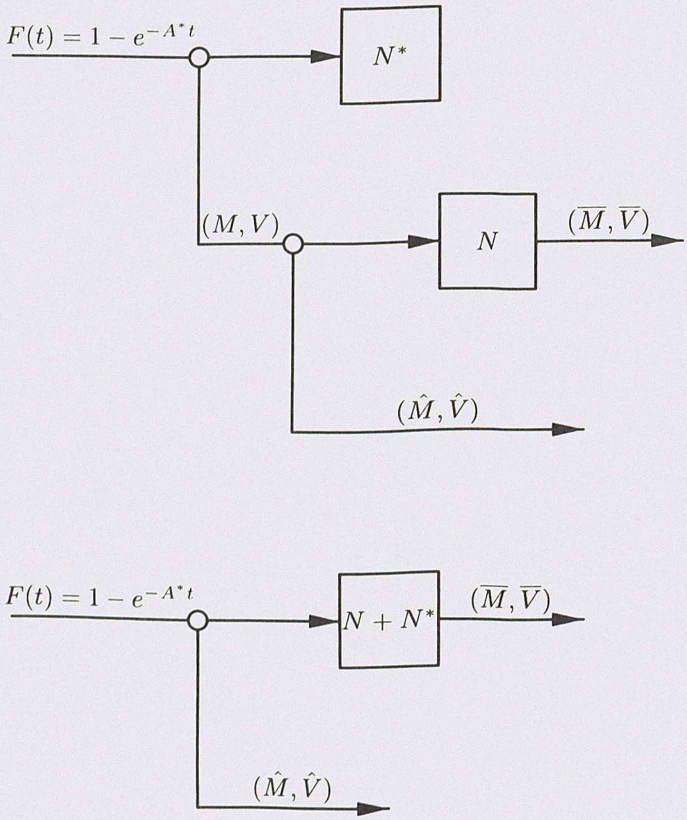

Figure 3.3 Equivalent Random Theory System

method to the analysis of overflow systems. As we have seen, the traffic overflowing a group offered Poisson traffic is always peaked. Thus it is reasonable to use the overflow model, where a primary group of N^* equivalent circuits offered some equivalent Poisson traffic A^*, as a generator of the (M, V) traffic whenever $V > M$. A^* and N^*, is chosen such that the parameters of its overflow are precisely M and V (see Fig. 3.3). From the Kosten model, we know that A^* and N^* must satisfy

$$M = A^* E(A^*, N^*) \tag{3.87}$$

$$V = M \left(1 - M + \frac{A^*}{N^* + 1 + M - A^*}\right) \tag{3.88}$$

Note that, in general, N^* is not an integer, which requires the use of a generalized Erlang B function. A fundamental question about the system defined by Eqs. (3.87) and (3.88) is whether there exist solutions $A^*, N^* \geq 0$ for an arbitrary pair M, V, with $V \geq M$. Although no formal proof seems to exist, this seems to be the case and the solution seems to be unique. Approximate values have been given by Rapp [2], which can either be used as such or as good starting values in the numerical solution:

$$A^* \approx V + 3Z(Z-1) \tag{3.89}$$

$$N^* \approx \frac{A^*(M+Z)}{M+Z-1} - M - 1. \tag{3.90}$$

Having computed the parameters of the equivalent system, we can compute the mean and variance of the overflow traffic using the Brockmeyer model; these values are given by

$$\hat{M} = A^* E(A^*, N + N^*) \tag{3.91}$$

$$\hat{V} = \hat{M}\left(1 - \hat{M} + \frac{A^*}{N + N^* + 1 - A^* + \hat{M}}\right). \tag{3.92}$$

The ERT method is iterative since it can be used to compute the blocking of the overflow traffic $(\hat{M}, \hat{V})$, which can also be peaked. Note, however, that the method is not iterative for the carried traffic, which can be smooth, and that it works only for peaked offered traffic. Thus the ERT method must be extended in order to be used for computing the overflow and carried parameters of carried traffic.

ERT for Smooth Traffic

In the case where $Z < 1$, the equations that define the equivalent system (3.87) and (3.88) generally have a negative solution for N^*. Possible extensions to standard ERT can be made along three directions by relaxing various parameters of the model: (1) use an equivalent traffic source that is smoother than Poisson, and match the moments of its overflow with the real parameters, (2) extend the Erlang B formula to the full domain of real N, or (3) generate the real smooth traffic by a mechanism other than overflow.

The first technique, although not widely used, falls within the framework of equivalent renewal processes and is dealt with below. The other methods have been proposed and used for network algorithms, and thus are described briefly here.

The second technique — the possibility of using an extended Erlang B function to construct an equivalent system for smooth traffic — was proposed first by Bretschneider [17]. The idea was extended by Nightingale [18] under the name of ERT-N, based on the analytic continuation of the Erlang B function in the complete (A, N) plane, as shown in [5]. For this case, only the extension to negative N is used; the extension to negative A has not been used in any practical case. The application of the method is straightforward. Given the parameters of the real traffic, the equivalent system is computed by solving the standard ERT equations (3.87) and (3.88); the parameters of the overflow are computed via Eqs. (3.91) and (3.92), where now N^* is negative if the traffic is smooth. Note that the equivalent system must still have nonnegative A. This, however, is not always possible. The derivation of the conditions in which the equivalent traffic is positive is left as an exercise (Problem 3.18).

The third way to extend ERT to smooth traffic is to assume that the mechanism producing the real traffic is something other than overflow. Such a model has been proposed by Katz [19] in which the real offered traffic is produced by the portion of the equivalent traffic that is *carried* in the equivalent group. The Laplace transform of the distribution of the interarrival time is given by Eq. (3.55). We know, however, that this process is not renewal and that its parameters cannot be computed using the general renewal method. The parameters of the equivalent system are given implicitly by the system:

$$M = A^* \left[1 - E(A^*, N^*)\right] \tag{3.93}$$

$$V = M - AB(N^* - M) \tag{3.94}$$

Having determined the value of the equivalent system, the parameters of the traffic overflowing the real group are computed as follows:

1. If $N^* \leq N$, then all calls carried in the equivalent system are also carried in the real group. Thus, we get $\hat{M} = \hat{V} = 0$; that is, there is no overflow from the real group.

2. If $N^* > N$, then some of the calls carried in the equivalent group are blocked in the real group. Assuming that the traffic offered to the real group is effectively A^*, we have for the parameters of the traffic carried in the real system

$$\overline{M} = A^* \left[1 - E(A^*, N)\right]$$
$$\overline{V} = \overline{M} + (\overline{M} - A^*)(N - \overline{M})$$

The parameters of the overflow traffic are computed by

$$\hat{M} = A^* - \overline{M}$$
$$\hat{V} = V - \overline{V}$$

Note that computing the variance in this way neglects the correlation that exists between the carried and overflow traffic on the real system.

It is claimed that in practice the accuracy of the method is comparable to the accuracy of the equivalent random method for peaked traffic and thus is adequate to evaluate network performance.

Generalized ERT

As mentioned in the preceding section, ERT (now standing for equivalent *renewal* theory) can be generalized to an arbitrary renewal input process by replacing the Poisson input by some arbitrary renewal stream. The moment matching is done between the traffic parameters available, on the one hand, and the moments of the renewal stream and the size of the equivalent group N^*, on the other. The blocking probabilities and overflow and carried traffic parameters are then computed for the equivalent renewal stream offered to a group of size $N + N^*$.

Such a generalized ERT was proposed by Potter [20], who suggested using input streams with various Erlang$_k$ interarrival distributions as input to the primary group for modeling smooth traffic. The equivalent streams have an interarrival time distribution of the E_n type, where $n = 1, 2, 3, 6, 10, \infty$. These streams have some interesting properties. As k increases, the peakedness of the overflow traffic decreases, with its smallest value attained for the deterministic stream. In fact, it has been shown [13] that $\hat{Z} \geq 0.5$ for any k. Thus these input processes cannot be used to model traffic whose peakedness is below 0.5. For a given phase k, the equivalent system has only two parameters: (1) the intensity of the Erlang stream and (2) the size of the equivalent group. Thus the model can be used directly for a two-moment representation of the traffic. Tables have been produced that relate the parameters of the overflow traffic of the equivalent group to these two parameters, allowing the equivalent system to be sized. No practical use of the technique, however, seems to exist in real network applications.

Equivalent IPP Model for Peaked Traffic

Consider again a peaked traffic (M, V). Such a traffic could be generated by an interrupted Poisson process [9,21], suitably chosen as to reproduce the values of M and V. Note, however, that the IPP is a three-parameter model, which means that the traffic should be characterized by a third parameter in addition to the mean and variance.

Let us first consider a somewhat simpler problem. Assume for the moment that the IPP is in fact produced by the overflow of some Poisson traffic A^* from a finite group of size N^*. In this case, the parameters of the IPP can be expressed as a function of the parameters of the equivalent system as

$$\lambda = A^* \frac{\delta_2(\delta_1 - \delta_0) - \delta_0(\delta_2 - \delta_1)}{(\delta_1 - \delta_0) - (\delta_2 - \delta_1)} \tag{3.95}$$

$$\omega = \frac{\delta_0}{\lambda}\left(\frac{\lambda - A^*\delta_1}{\delta_1 - \delta_0}\right) \tag{3.96}$$

$$\gamma = \frac{\omega}{A^*}\left(\frac{\lambda - A^*\delta_0}{\delta_0}\right) \tag{3.97}$$

where

$$\delta_n = \frac{1}{A^*} \frac{M_{n+1}}{M_n}, \tag{3.98}$$

where the factorial moments M_n can be computed by the Kosten model of Section 3.1. Given an arbitrary value for A^* and N^*, there is no proof that the corresponding λ, ω, γ are not negative (although this seems to be the case in practice). The same question can be posed for an arbitrary set of three moments with $Z \geq 1$.

This relation between the moments of the IPP process and the overflow of a Poisson stream is useful when the real traffic is described by its first two moments only. The difficulty is that the IPP is a three-parameter process, and that there are only two quantities with which to match it. In this case, the IPP model can be used by making the following approximations. First, given M and V, compute $A^*$and N^* by the standard ERT. Then use these values to obtain the parameters of the IPP system by Eqs. (3.95–3.97), and use Eq. (3.72) to obtain the mean and variance of the overflow. Another possibility is to use the equivalent Poisson traffic computed from ERT as the value for λ. In this case, only two parameters are left to compute, given by

$$\omega = \frac{M}{A^*}\left(\frac{A^* - M}{Z - 1} - 1\right) \tag{3.99}$$

$$\gamma = \omega\left(\frac{A^*}{M} - 1\right)$$

In general, when a two-moment match is desired, one parameter of the IPP must be chosen arbitrarily. In practice, this choice is constrained by the requirement that the parameters of the equivalent IPP system should be non-negative, which imposes the requirement that $\lambda > M$. Finally, the method is recursive, since overflow traffic is always peaked, and the IPP method can be used again on this traffic. This is not the case for the carried traffic, which is generally smooth; some other method must be used to calculate its overflow.

The Two-Stage Cox Model

The Cox model of Section 3.3 can be used to represent a stream of traffic whose first three moments are known, provided the system of Eqs. (3.81) and (3.82) can be inverted. This question is discussed in detail by Guérineau and Labetoulle in [22], in which the following conditions are given. Let $P = \alpha\beta$ and $S = \alpha + \beta$. These can be written in terms of the coefficients δ_0, δ_1 and δ_2 as

$$P = \frac{2\mu^2 \delta_0(\delta_2 - \delta_1)}{2\delta_1 - \delta_0 - \delta_2}$$

$$S = \frac{\mu\left[2(\delta_1 - \delta_0 + 1)(\delta_2 - \delta_1) + \delta_1(2\delta_1 - \delta_0 - \delta_2)\right]}{2\delta_1 - \delta_0 - \delta_2}$$

and of course we have

$$\alpha = \frac{S + \sqrt{S^2 - 4P}}{2}$$

$$\beta = \frac{S - \sqrt{S^2 - 4P}}{2}$$

$$b = \frac{P}{\mu\delta_0\alpha}$$

Obviously, there exist solutions if and only if

$$P > 0,\ S > 0,\ S^2 - 4P > 0,\ a > 0.$$

For peaky traffic, a necessary and sufficient condition is that $\delta_0 < \delta_1 < \delta_2 < 2\delta_1 - \delta_0$. In the case of smooth traffic, the condition $2\delta_1 - \delta_0 < \delta_2 < \delta_1 < \delta_0$ is necessary but not sufficient.

Because the parameters of the real offered traffic are computed from the superposition of various overflow and carried traffic streams, quite possibly the set of three moments does not meet the required conditions. In such a case, some further approximation is required. If the traffic is peaky, the first two moments can always be matched. Since the equivalent system is a three-parameter model, one parameter is arbitrary. Guérineau and Labetoulle suggest choosing $\mu\delta_1 < S < \infty$, with a recommended value of $S = 2\mu\delta_1$. Having chosen S, the parameters for the equivalent Cox generator are given by

$$\alpha = \frac{S + \sqrt{S^2 - 4P}}{2}$$

$$\beta = \frac{S - \sqrt{S^2 - 4P}}{2}$$

$$b = \frac{1}{\alpha}(S - P/\mu\delta_0)$$

where

$$P = \frac{\mu\delta_0}{\delta_1 - \delta_0 + 1}(S - \mu\delta_1).$$

In the case of smooth traffic, one first tries the same values as for peaky traffic. If the parameters of the generator are positive, then a solution has been obtained. If not, only the first two moments are matched. In order to do this, α and β are kept arbitrary, and a is fixed to 1. The parameters are then

$$\alpha = \frac{S + \sqrt{S^2 - 4P}}{2}$$

$$\beta = \frac{S - \sqrt{S^2 - 4P}}{2}$$

which is valid if and only if

$$\delta_1 \geq \frac{4\delta_0^2}{1 + 4\delta_0}.$$

This means that smooth traffic can be modeled by the Cox system if it is no smoother than Erlang_2. When this is not so, the authors suggest replacing it with an Erlang_2 traffic and using the Cox model. In this case, we have $\alpha = \beta = 2\mu\delta_0$.

After constructing the equivalent system, we must calculate the parameters of the carried and overflow traffics. This can be done by means of the transform of the overflow process from the primary group, using either the method proposed in [22] or the formulas derived in Section 3.3. It is not clear which of these two methods is preferable from the point of view of numerical stability and computation time.

The Bernouilli-Poisson-Pascal Process

As is well known [6], the overflow traffic offered to a large group has a busy-circuit distribution that is very close to a Pascal distribution. Given the mean and variance of a peaked traffic, the parameters of the corresponding Pascal distribution can be computed analytically from Eqs. (3.27) and (3.28). From the values of the parameters, the overflow could be analyzed just as for the Kosten system, with the arrival rate corresponding to a Pascal process instead of the constant rate of the Poisson process. Note, however, that the analysis would become quite complex — probably the reason why it has not been carried out.

Instead, the Bernouilli-Poisson-Pascal (BPP) method can be used to approximate the overflow system by making a second assumption [23]. Given an offered traffic with arbitrary mean and variance, *the busy-circuit distribution in a group of size N is given by the same distribution used to describe the offered traffic*, that is, the Pascal distribution (or Bernouilli distribution for smooth traffic). For a group of finite size N, the distribution (3.25) must be truncated; with $\mu = 1$, we have the truncated Pascal distribution, denoted by q_j:

$$q_j = \frac{p_j}{\sum_{i=0}^{N} p_i}, \quad j = 0, \ldots N$$

The time congestion for a group of size N is given by

$$E_N \triangleq q_N = \frac{\binom{-r}{N}(-\beta)^N}{\sum_{j=0}^{N} \binom{-r}{j}(-\beta)^j} \tag{3.100}$$

and the call congestion by

$$B = E\left[1 + \frac{N}{M}(Z - 1)\right]. \tag{3.101}$$

Equation (3.101) should be compared with the general relation (3.48) between time and call congestion in the case of a general renewal input. The time congestion obeys the interesting recurrence

$$
\begin{aligned}
E_k &= \frac{q_{r+k-1}E_{k-1}}{k + q_{r+k-1}E_{k-1}} \qquad k = 1, \ldots \\
E_0 &= 1
\end{aligned}
\tag{3.102}
$$

which is a generalization of the well-known recurrence for the Erlang B function. Similarly, under this approximation, the mean and variance of the traffic carried in the group are given by

$$
\overline{M} = \frac{nq - (n + N)qE}{1 - q} \tag{3.103}
$$

$$
\overline{V} = \frac{nq + \overline{M}(n + 1)q - (N + 1)(n + N)qE}{1 - q} - \overline{M}^2 \tag{3.104}
$$

From these values, we can get $\hat{M} = M - \overline{M}$ and, neglecting the correlation, $\hat{V} = V - \overline{V}$. It is easy to show that these quantities reduce to the values already calculated for the Kosten system, that is, when $q \to 0$ and $n \to \infty$ while maintaining $Z = 1$.

As indicated in Problem 3.4, the BPP distribution is inadequate to describe smooth offered traffic because the probabilities given by the binomial distribution could become negative. The situation is better in the case of a finite group since a sufficient condition exists to guarantee that the probabilities do not become negative in the range of interest. That is, for $j \leq N$, $N \leq M/(1 - Z)$, a condition generally met in practical situations unless Z is very low. Although the method cannot represent all potential situations, the advantages of simplicity and uniformity in representing both peaked and smooth traffics make it attractive for modeling traffic in network studies.

Finally, let us comment on the accuracy of the method. The accuracy was checked in [23] by comparing the parameters of the carried traffic for a hyperexponential and a gamma distribution of the interarrival times. It was shown numerically that the values were similar for the two cases and were quite good. A strong assumption is needed for the BPP model, stating that the busy-circuit distribution is BPP for any arbitrary arrival process. In fact, the situation is even worse from a theoretical point of view, since it can be shown that there is *no* renewal input process such that its overflow has the Pascal distribution [13]. These two facts would seem to make the BPP model a very poor candidate indeed for the analysis of overflow. The accuracy of this model, however, is at least as good as that of the other models. Let us explain this apparent anomaly in the more general context of renewal arrival processes and moment-matching methods.

General Renewal Input

In general, when a traffic is represented by a renewal process, specifying a number of moments of the real traffic constrains the input process and to some degree determines the blocking experienced by this process. Moment-matching methods would be expected to work well if specifying the moments of the real traffic limits the equivalent renewal process to a small range. In this case, we would expect the moment-matching technique to be "forced" to choose the right equivalent process, which would explain the good accuracy obtained by moment matching.

From Eqs. (3.53) and (3.54), the values of M and V uniquely determine $\Phi'(0)$ and $\Phi(\mu)$ for *any* renewal input. In fact, these values constrain all of the $\Phi(\mu)$ since they must satisfy Eq. (3.52), where $\Phi(\mu)$ appears in all the terms through h_n. Bounds have been computed in [24] for $\Phi(x), x \geq 1$ as

$$[\Phi(\mu)]^x \leq \Phi(x) \leq \Phi_m(x) \triangleq p_1 + p_2 e^{-bx},$$

where b is defined implicitly by the equation

$$b = \frac{A(1 - e^{-b})}{1 - \Phi(\mu)}$$

and

$$p_2 = \frac{A}{b}$$
$$p_1 = 1 - p_2$$

Furthermore, from Eq. (3.46), we see that $\Phi(\mu)$ is the leading term in the expression for the call congestion B although B also depends on all the $\Phi(\mu)$s. Using the bounds for Φ, it is possible to write bounds for the blocking experienced by an arbitrary renewal stream $B_{min} \leq B \leq B_{max}$ where

$$B_{min} = \left\{ 1 + \binom{N}{1} \frac{1 - \Phi(\mu)}{\Phi(\mu)} + \ldots + \binom{N}{N} \frac{[1 - \Phi(\mu)] \ldots [1 - \Phi^N(\mu)]}{[\Phi(\mu)]^{N(N+1)/2}} \right\}^{-1}$$

$$B_{max} = \left\{ 1 + \binom{N}{1} \frac{1 - \Phi_m(\mu)}{\Phi_m(\mu)} + \ldots + \binom{N}{N} \frac{[1 - \Phi_m(\mu)] \ldots [1 - \Phi_m(N\mu)]}{\Phi_m(\mu) \ldots \Phi_m(N\mu)} \right\}^{-1}$$

The point is that these bounds are sharp, and it is possible to identify renewal processes that reach them. The lower bound is given by a renewal input with constant interarrival times m', given by $\exp\{-m'\} = \Phi(\mu)$. Similarly, a renewal process with a step at time $t \to 0$ and another at $t \to \infty$ has a blocking probability arbitrarily close to 1. The upper bound can be viewed as such a process, but constrained by V to have the second step at finite t. The point is that the two bounds can be widely separated, as shown in [24]. This means that the blind selection of a renewal process to represent a traffic specified by only two moments cannot be expected to yield an accurate value for the blocking.

The excellent accuracy of moment matching has been explained in the case of the BPP method [23]. Assume that the system is divided into K subsystems of infinite size, and let p_i be the probability that an incoming call is served by the i^{th} subsystem. Let X_i be the number of calls in subsystem i, and M_i and V_i its mean and variance. We then have

$$X = \sum_i X_i, \; M_i = Mp_i$$

$$V = M_i p_i (1 + p_i(Z - 1))$$

$$\operatorname{cov}(X_i, X - X_i) = Mp_i(1 - p_i)(Z - 1)$$

If A_i is the interarrival time distribution at subsystem i, its Laplace-Stieltjes transform is given by

$$a_i(s) = \frac{a(s)\,(1 - (1 - p_i))}{1 - (1 - p_i)a(s)}.$$

Choosing $p_i = 1/K$, we can choose K such that M_i/V_i is as close to 1 as desired. We also note that $a_i(s)$ tends to a negative exponential distribution as $p_i \to 0$. The mean carried traffic can then be viewed as the sum of a large number of infinitesimal quasi-regular traffics; this property is independent of the precise nature of the arrival renewal process. Because this microscopic structure of the stream is the important factor in the particular value of the mean, we can expect the BPP model to adequately describe the carried traffic in an infinite group, irrespective of the particular nature of the arrival renewal process. This is indeed the case.

More generally, equivalent methods are found to work well when used to evaluate overflow systems. The reason for this success is found not in the constraints imposed on the equivalent process by the prescribed values for the mean and variance, but in the additional information used in selecting the equivalent process. In other words, equivalent theory works because, when it is applied to overflow systems, the fact that the M, V traffic is produced by an overflow mechanism is used to select the equivalent renewal process. This is obvious when considering that the most successful equivalent systems for peaked traffic are the Poisson overflow and the IPP. Similarly, the success of the BPP method is also due to the choice of the correct distribution to represent carried traffic; it would probably be not as successful if an arbitrary renewal input stream with the correct M, V parameters were used. This is only a conjecture, however, and for now we only note that, for the problems arising in circuit-switched studies, the renewal equivalent method is generally sufficient.

Summary

Moment-matching techniques are widely used in analyzing networks. In fact, it is fair to say that these techniques are the only tools available that are rapid enough to be used in this context. Even so, the situation is not completely satisfactory, at least for certain problems. Network analysis really requires a technique that can represent *any* kind of traffic, smooth or peaked, within the same model, and for *arbitrary* values of the parameters. None of the methods we have discussed here meets these requirements; all are designed for a particular type of traffic, smooth or peaked — or, if they apply to both, do so by using different models for different ranges of peakedness.

As a case in point, consider the BPP model. Although probably the most attractive from the point of view of generality, it suffers from some defects. It cannot represents arbitrarily smooth traffic, and it represents smooth and peaky traffics by two separate models. Suppose that we want to use this model for network analysis or synthesis within a computer program. It is true that, in a real network, traffic will not be so smooth that the model cannot apply. This, however, is not the case in the design process, during which a search algorithm may produce some extreme traffic patterns when it is far from its optimal solution. Some care must be taken that these values do not cause a fatal error in the program. Similarly, the nature of the traffic offered to a given link may change, say from peaky to smooth, during this search procedure. Because the BPP model uses two different distributions for these two cases, the parameters of the overflow and carried traffic may not change continuously when this transition occurs — or, more likely, their derivatives may not be smooth. If the search algorithm uses gradients, these discontinuities in the derivative may cause the procedure to fail.

To avoid these difficulties, given that there exists no unified model that is guaranteed to be smooth and robust in terms of the parameters of the offered traffic, caution should be exercised when using moment-matching models in optimization procedures.

3.5 Heterogeneous Blocking

The traffic models described in Sections 3.1 to 3.4 apply to a single stream of offered traffic. Unfortunately, real networks are somewhat more complex since the traffic offered to a group is generally a mixture of streams, each with its own set of moments, and is not of the renewal type. Because of their different statistical characteristics, the streams effectively experience different blocking probabilities.

The heterogeneous blocking problem can thus be formulated as follows. Given n distinct streams of traffic offered to a group of size N, each with a known set of moments, compute the blocking probability (*parcel* blocking), as well as the carried and overflow traffic moments, for each stream. In this general form, the problem has no known solution, even under the assumption that the input streams are independent; only approximate methods are available. The problem with two streams has been solved exactly when one of the streams is Poisson and the other is renewal. The case of two renewal streams has also been analyzed, but simplifications are required to obtain a solution. For this reason, we concentrate here on practical methods that can yield sufficiently accurate results with relatively little computation.

First, we discuss the aggregation of the input streams and the assumption of independence. Then, we describe one case where an exact solution is known, mostly to illustrate the difficulty of the problem. Finally, heuristics usable in network problems are examined, starting with empirical formulas and moving to progressively more complex models based on those of Section 3.4.

Aggregation of Input Streams

The heterogeneous blocking problem arises because calls offered to a group in a network have been either blocked or accepted for service on some other group, but still must be connected through some intermediate group in order to establish a connection to their destination. Depending on the structure of the network and the routing method used, these processes cannot be expected to be independent. The dependence between various streams of traffic offered to a single group was investigated in [25] for simple arrangements of overflowing groups. Such dependence effects are quite difficult to compute; in practice, it is assumed that the streams to be merged are independent. This is a good approximation in large networks, where many different streams overflow on a given link (see Chapter 2).

Assuming this independence, and also assuming that stream i has k^{th} moment α_k^i, the k^{th} moment of the aggregate stream is given by

$$\alpha_k = \sum_{k_1+k_2+\ldots k_m=1} \frac{k!}{k_1!k_2!\ldots k_m!}\alpha^1_{k_1}\alpha^2_{k_2}\ldots\alpha^m_{k_m}. \tag{3.105}$$

In the case of $m = 2$, we have

$$\begin{aligned}
\alpha_1 &= \alpha_1^1 + \alpha_1^2 \\
\alpha_2 &= \alpha_2^1 + 2\alpha_1^1\alpha_1^2 + \alpha_2^2 \\
\alpha_3 &= \alpha_3^1 + 3\alpha_2^1\alpha_1^2 + 3\alpha_1^1\alpha_2^2 + \alpha_3^2
\end{aligned}$$

Given the moments of the aggregate stream, its blocking probability and the parameters of the carried and overflow traffic can be computed by the methods of Section 3.4. The difficult part is determining how much of this blocking is actually experienced by each stream.

Heterogeneous Blocking for Poisson Inputs

The heterogeneous blocking problem has been solved in the case where all the input streams are Poisson. In this case, all streams experience the same