# D

# Elasticity of Traffic Matrices – Network Models*

Here, we present the analytical foundation for the analysis of elastic traffic matrices described in Section 13.3. Several network models of increasing complexity that describe the behaviour of the traffic flows through a network with respect to the capacity of the links and nodes of that network are described.

## D.1 Basic Model

We model a subnetwork $\Lambda$ of the Internet consisting of $N$ nodes and $L$ directed links. The traffic through the network consists of long-lived greedy Transmission Control Protocol (TCP) connections and is represented by TCP *macroflows*. A TCP macroflow represents a number of TCP connections that have the same ingress node $i$ and egress node $j$ of $\Lambda$. We assume that the connections of a macroflow experience on average the same loss $\tilde{p}$ and delay $\tilde{q}$ when traversing the other networks that are not modelled in detail with this model. The macroflows are assumed to be small compared to the other flows flowing through the external networks; therefore, the external loss $\tilde{p}$ and delay $\tilde{q}$ are independent of the rate of the macroflows. The macroflows are elastic traffic; their rate is described by a TCP formula and adapts to the network conditions of $\Lambda$. There are a number of works about predicting the average TCP throughput depending on the loss and delay properties of a flow, see Section 4.1.3. As we are not interested in details such as the duration of the connection establishment etc., we use the rather simple square-root formula (4.2) here.

An output queue is attached to each link. In the basic model, we describe the queues as *M/M/1/B* queues. This is not the most realistic approach: First, because Internet traffic is not described very well by a Poisson arrival process, see Paxson and Floyd (1995). Second, since packet sizes are not exponentially distributed, an exponential service rate is also not realistic, see AIX – NASA Ames Internet Exchange (2000); Claffy *et al.* (1998). However, the *M/M/1/B* model is one of the simplest queueing models and is used in related works like Garetto *et al.* (2001); Gibbens *et al.* (2000). We will investigate more realistic queueing models later in this section.

* Reproduced by permission of VDE Verlag GMBH

*The Competitive Internet Service Provider: Network Architecture, Interconnection, Traffic Engineering and Network Design*
Oliver Heckmann  ISBN: 0-470-01293-5

The basic network model with elastic traffic is described by the non-linear equation system in Model D.1.

**Model D.1** Basic Network Model for Elastic Traffic Matrices

Indices:

| | |
|---|---|
| $i, j = 1, ..., N$ | Node $i$ or $j$ |
| $l = 1, ..., L$ | Link or output queue $l$ |

Parameters:

| | |
|---|---|
| $\psi_{ij}$ | Path from node $i$ to node $j$ and back (set of links) |
| $t_{ij}$ | Size of macroflow between node pair $i, j$ [pkts] |
| $\mu_l$ | Service rate of link or queue $l$ [pkts/s] |
| $B$ | Buffer size [pkts] |
| $\tilde{q}$ | Average external queueing and total propagation delay [s] |
| $\tilde{p}$ | Average external loss probability |

Variables:

| | |
|---|---|
| $r_{ij}$ | Rate of macroflow between node pair $i, j$ [pkts/s] |
| $\rho_l$ | Utilisation of link or queue $l$ |
| $p_l$ | Loss probability of link or queue $l$ |
| $q_l$ | Queueing delay of link or queue $l$ [s] |

$$r_{ij} = \frac{t_{ij}}{[(\sum_{l \in \psi_{ij}} q_l) + \tilde{q}] \cdot \sqrt{\frac{2}{3}} \cdot \sqrt{1 - [\prod_{l \in \psi_{ij}} (1 - p_l)] \cdot (1 - \tilde{p})}} \qquad \forall i, j \,|\, i \neq j \tag{D.1}$$

$$\rho_l = (\sum_{(i,j) \,|\, l \in \psi_{ij}} r_{ij}) \cdot \frac{1}{1 - p_l} \cdot \frac{1}{\mu_l} \qquad \forall l \tag{D.2}$$

$$p_l = (1 - \rho_l) \cdot \frac{\rho_l^B}{1 - \rho_l^{B+1}} \qquad \forall l \tag{D.3}$$

$$q_l = \frac{1 + B\rho^{B+1} - (B+1)\rho^B}{\mu_l (1 - \rho_l)(1 - \rho_l^B)} \qquad \forall l \tag{D.4}$$

The total loss probability of a macroflow $ij$ can be approximated by

$$p_{ij} = \tilde{p} + \sum_{l \in \psi_{ij}} p_l \tag{D.5}$$

**Table D.1** Assessment of the Approximations

| Approximation | Maximal Error [%] |
|---|---|
| For $p_{ij}$ | 0.0004795 |
| For $\rho_l$ | 0.0009097 |

for small loss probabilities. Similarly, for small loss probabilities at a link $l$ the utilisation (D.2) can be approximated as

$$\rho_l = \Big( \sum_{(i,j)\,|\,l \in \psi_{ij}} r_{ij} \Big) \cdot \frac{1}{\mu_l} \tag{D.6}$$

These simplifications can reduce the computational effort to solve the resulting non-linear equation system by up to 25%. In order to assess the systematic error of these approximations we ran a number of experiments on the Deutsche Telekom backbone topology (see Appendix A) with different parameters of $t_{ij}$, $B$ and $\mu_l$. We solve the non-linear equation system from Model D.1 using Maplesoft (2004) and compare the difference in $\rho_l$. The maximum errors of 25 different settings are listed in Table D.1. They are extremely small and can be neglected.

Next we discuss the possible extensions of the basic model.

## D.2 Discrete Service Times

We first investigate how the basic model from Section D.1 can be extended to account for more realistic service times. IP packets can differ drastically in their size (40 to 1500 bytes), see AIX – NASA Ames Internet Exchange (2000); Claffy *et al.* (1998). We assume a service time proportional to the packet size and use a discrete distribution with $c = 1, \ldots, C$ classes of differently sized packets to model the service time; $si_c$ is the packet size of class $c$ and $h_c$ the relative frequency of class $c$ with $\sum_c h_c = 1$. Using $sp_l$ as the line speed of link $l$, the probability density function of the service time distribution is given as

$$pdf(x) = \sum_c h_c \cdot \delta\Big(x - \frac{si_c}{sp_l}\Big) \tag{D.7}$$

where $\delta(x)$ is the Dirac impulse $\delta(x) = 1$ for $x = 0$ and 0 otherwise. The probability distribution function is

$$PDF(x) = \sum_c h_c \cdot u\Big(x - \frac{si_c}{sp_l}\Big) \tag{D.8}$$

where $u(x)$ is the unit function $u(x) = 1$ for $x \geq 0$ and 0 otherwise. In order to model the *queueing delay*, we use the Pollaczek–Khinchin formula for the queueing delay of an *M/G/1* queue

$$q_l = E(x) \cdot \Big(1 + \frac{1 + C_v^2}{2} \frac{\rho_l}{1 - \rho_l}\Big) \tag{D.9}$$

with the expected service time[1]

$$E(x) = \frac{1}{\mu} = \int_{-\infty}^{\infty} x \cdot pdf(x)\,dx = \sum_c h_c \cdot \frac{si_c}{sp_l} \tag{D.10}$$

and the square of the coefficient of variation

$$C_v^2 = \frac{Var(x)}{E(x)^2} = \frac{\int_{-\infty}^{\infty} (x - E(x))^2 \cdot pdf(x)\,dx}{E(x)^2} \tag{D.11}$$

For the *loss probability* $p_l$ we turn to the *M/G/1/B* queue. There is no general closed form for the loss probability of the *M/G/1/B* or the queue length distribution of the *M/G/1* queue. We can derive the loss probability of the *M/G/1/B* queue exactly if we know the state probabilities $\pi_{lk}^{(\infty)}$ for queue length $k$ of the corresponding *M/G/1* queue $l$.

Cooper (1981); Virtamo (2003) list an iterative algorithm based on Markov chains that can be used to numerically derive $\pi_{lk}^{(\infty)}$. We do not want to use this Markov chain algorithm; first, because it does not give us a closed form for the loss probability that we need for our equation system and, second, because for that approach we would have to solve several complex integrals numerically, while we are interested in an analytical form. Therefore, we use a different method to derive the state probabilities $\pi_{lk}^{(\infty)}$ of the *M/G/1* queue: The Laplace transform of the service time distribution $pdf(x)$ is

$$b_l^*(s) = \sum_c h_c \cdot e^{-s\frac{si_c}{sp_l}} \tag{D.12}$$

Kleinrock (1975); Virtamo (2003) show that the transformed state probabilities follow the Pollaczek–Khinchin transform formula for the queue length

$$Q_l(z) = (1 - \rho_l)\frac{b_l^*(\lambda - \lambda z)}{b_l^*(\lambda - \lambda z) - z}(1 - z) \tag{D.13}$$

With the inverse $Z$-transformation on $Q_l(z)$, we can derive the state probabilities $\pi_{lk}^{(\infty)}$ analytically. We can use the Taylor series expansion to analytically transform the somewhat complex term $Q_l(z)$ back:

$$\pi_{lk}^{(\infty)} = \frac{1}{k!}\frac{d^k}{dz^n} Q_l(z)\,|_{z=0} \tag{D.14}$$

The loss probability of the related *M/G/1/B* queue is now given as

$$p_l = 1 - \frac{1}{\rho_l + \pi_{l0}^{(B)}} \tag{D.15}$$

using the state probability $\pi_{l0}^{(B)}$ of the finite queue as in Virtamo (2003)

$$\pi_{l0}^{(B)} = \frac{\pi_{l0}^{(\infty)}}{\sum_{j=0}^{B-1} \pi_{lj}^{(\infty)}} \tag{D.16}$$

This leaves us with closed-form non-linear equations for loss and delay of the *M/G/1/B* queue with a discrete service time distribution.

[1] We continue using $\mu$ for the inverse of the expected service time as we did with the *M/M/1/B* queue.

## D.3 Self-similar Traffic

Internet traffic measurements show self-similar, heavy-tailed and long-range dependent properties as discussed in Section 5.1. The burstiness of Internet traffic on larger timescales can significantly influence the loss probability. To take this effect into account we use the Gaussian approximation of aggregate traffic and the following loss formula based on Addie *et al.* (2002); Mannersalo and Norros (2002):

$$p_l = \frac{C}{\frac{B-\lambda\hat{t}}{\sigma_{\hat{t}}^2}\sqrt{2\pi\sigma_{\hat{t}}^2}} \cdot e^{-\inf\limits_{t\in\Re^+} \frac{(B+(\mu-\lambda)\cdot t)^2}{2\cdot\sigma_t^2}} \tag{D.17}$$

$\hat{t}$ is the optimiser from the infimum condition, $t$ is the timescale, $B$ is the buffer size, $\lambda$ and $\mu$ are the arrival and service rates. For a given Hurst parameter, $\sigma_t^2$ is given as $\sigma_t^2 = \sigma^2 \cdot t^{2H}$.

## D.4 Related Work

Some works use network models similar to our models. The performance models of Gibbens *et al.* (1999, 2000); May *et al.* (1999) are used to analyse quality of service (QoS) in Diffserv IP networks with two service classes. They assume a Poisson arrival process and exponential service times (*M/M/1/B*). The fixed point model of Gibbens *et al.* (2000) combine the Diffserv resource models with the TCP formula. We use a similar approach but we also investigate non-exponential service times and non-Poisson arrivals. Also, we investigate the performance in the context of network design and capacity expansion and do not use different service classes.

Garetto *et al.* (2001) present an analytical TCP model for multiple flows and verifies it against NS2 simulations. Similar to our model, they combine a TCP and a network model and calculate the fixed point of the two models. Their TCP model, however, is more fine grained and complex than our TCP formula–based TCP model. This, however, comes at the cost of losing a closed-form formulation of the whole model. The authors investigate different network models and find that the simple *M/M/1/B* gives sufficiently accurate results.

Schwefel (2001) introduces a queueing model that is based on multiple ON/OFF arrival processes; this allows accounting for long-range dependency. It is extended to be reactive to congestion by slowing down the rate similar to the way TCP reacts and can thus be used for the performance analysis of TCP-generated bursty traffic. Contrary to this approach, we combine the TCP formula with the standard queueing theory.