· [HeSt04] *Oliver Heckmann, Ralf Steinmetz;* **On the Elasticity of Traffic Matrices and the Impact on Capacity Expansion;** 11th International Telecommunications Network Strategy and Planning Symposium, Wien, Juni 2004, S. 223-230.

On the Elasticity of Traffic Matrices and the Impact on Capacity Expansion

Oliver Heckmann, Ralf Steinmetz
Multimedia Communications Lab (KOM), Darmstadt University of Technology
Merckstr. 25, 64293 Darmstadt, Germany Phone: +49-6151-16-5188, Fax: +49-6151-16-6152
Email: {heckmann, steinmetz}@kom.tu-darmstadt.de

Abstract

Traffic matrices are fundamental for network design, capacity expansion and traffic engineering. A Traffic matrix describes the rate or volume transfered between the ingress and egress nodes of a network. Internet traffic is dominantly TCP traffic and thus adapts to changing network conditions like routing or capacity. This effect is systematically neglected when using normal traffic matrices. This paper investigates the effect using three analytical models. It also shows how to use elastic traffic matrices for capacity expansion problems that do not neglect this effect.

1 Introduction

A Traffic Matrix M describes the average rate r_{ij} for a given time interval between the ingress nodes i and egress nodes j of a network A.

$$M = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & r_{ij-1} & \cdots & \cdots \\ \cdots & r_{ij} & r_{i+1j} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

Traffic Matrices are fundamental for network design and traffic engineering problems. Normally, the traffic matrix entry r_{ij} is expressed statically as a scalar (we call a traffic matrix with static predictions r_{ij} a *static traffic matrix*). However, Internet traffic is dominantly TCP traffic that adapts to changing network conditions like routing or the link capacity, this effect is systematically neglected when using static traffic matrices. The effect of capacity changes was probably neglectible in times when the Internet was dominated by web traffic that consisted of huge numbers of short lived TCP connections dominated by the slow start and not the elastic congestion avoidance phase. Traffic matrix entries at these times mainly increased if the customer base or browsing behavior changed.

Nowadays, however, most of the traffic [1, 2] is generated by peer-to-peer (P2P) applications. These applications are first of all more bandwidth greedy and second they generate more long-lived and therefore reactive TCP connections over which the dominating part of traffic is exchanged: To support this claim we did some measurements in the Edonkey network. It is with 52% of the generated filesharing traffic [1] the most successful P2P filesharing network in Germany. Our measurements [3] show that an average Edonkey user is sharing 57.8 files with an average size of 217 MB, a large proportion of those files being movies. An average active TCP connection between two clients is with almost 30 minutes definitely long-lived. During this time on average 4 MB are transferred, this volume is mostly limited by the ADSL upload capacity that is typically almost fully used by the P2P application. This supports the assumption of this paper, that long-lived reactive TCP connections start dominating the Internet traffic.

Besides P2P traffic, future multimedia Internet traffic like streaming videos can also be expected to be TCP friendly and therefore show similar reactive effects as long lived TCP connections that we are looking at in this paper [4].

Because of this it is time to investigate the effect of the elasticity of long-lived TCP connections in their congestion avoidance phase on traffic matrices used as input for *network design* and *capacity expansion problems*. Network design like [5, 6] is concerned with creating a new topology which typically includes a capacity assignment subproblem. A *capacity assignment* problem describes the problem of assigning capacity - most importantly link bandwidth - to a network; a *capacity expansion* problem is a special form of a capacity assignment problem, where the structure of the network and the initial capacity are given, the capacity can only be increased in limited ways (e.g. doubled or in discrete steps). All these problems are based on a traffic matrix and usually use static traffic matrices as input and ignore the effect that the new capacity (or capacity change) has on the traffic matrix itself. We use the term *elastic traffic matrix* for a traffic matrix M with entries $r_{ij} = f(...)$ that capture the elasticity of the TCP traffic and investigate the use of these elastic matrices in this paper.

Our paper is structured as follows. After this introduction we discuss related work. We then present three different network models in Section 3 that we use as an analytical foundation for our further analysis. In a set of experiments we use our network models to analyze the effect of the network capacity on the traffic matrix entries r_{ij}. The results show that in many cases the elasticity of the traffic matrix should not be neglected. Therefore, in Section 5 we extend the typical capacity expansion problem to account for the elasticity of the traffic matrix and test the model in a simulation. We conclude with a short summary.

2 Related Work

2.1 Traffic Matrices

[7] gives a good overview over the two distinct approaches to measuring a traffic matrix: The *direct measurement* approach as advocated by [8] uses e.g. Netflow to collect flow information. This information is evaluated offline to derive the traffic matrix using the routing tables active at the measurement time that also have to be recorded. This approach is storage space and router-CPU intensive but contrary to other approaches allows to derive the point-to-multipoint traffic matrix. A point-to-point traffic matrix M models the traffic between ingress node i and egress j while the point-to-multipoint traffic matrix $\tilde{M}$ models the traffic and ingress node i and captures the fact that this traffic can exit at more than one egress j.

Most of the other works favour *deriving the traffic matrix from link measurements* as they are more readily available for all router interfaces via SNMP (simple network management protocol) in production networks. The problem with this approach is that estimating the traffic matrix is an ill-posed inverse linear problem: In a network with N ingress/egress nodes the traffic matrix size is $O(N^2)$ while there are only $O(N)$ measurements - the problem becomes massively underconstrained for large N. To solve this problem additional assumptions e.g. about the traffic and the routing have to be made. Approaches to this problem can be classified into statistical tomographic methods, optimization-based tomographic methods and other methods.

Statistical tomographic methods use higher order statistics of the link load data like the covariance between two loads to create additional constraints. Examples are [9, 10, 11]. [9] and [10] assume a Poisson traffic model. [11] assumes a Gaussian traffic model

Optimization-based tomographic methods select a solution out of the solution space of the underconstrained problem that optimizes a certain objective function using methods like linear or quadratic programming. [12] is a simple example for this approach.

Classified as *other methods* are approaches that combine the tomographic methods with other methods like gravity or choice models. [13] that use a logit choice model that captures the choices of users (where to download from) and network designers (how to interconnect the point of presences/POPs). The decision process is modeled as a utility maximization problem.

[14] combines a optimization-based tomographic methods with a generalized gravity model. A gravity model can for example be used to estimate the traffic between edge links by assuming that the traffic between i and j is proportional to the total traffic entering at i multiplied with the total traffic exiting at j

[15] uses an information theoretic approach that chooses the traffic matrix consistent with the measured data that is as close as possible to a model in which the source and destination pairs are independent and therefore the conditional probability $p(j|i)$ that source i sends traffic to j is equal to the probability $p(j)$ that the whole network sends traffic to j.

All the works mentioned above derive static traffic matrices from measurement data. Traditional network design problems are placing links, nodes and assigning link and node capacities, typical examples are [5, 6]. They use static traffic matrices. Contrary to that in this paper, we assume that have estimated the static traffic matrix as a starting point using one of the methods above and try to incorporate the elasticity of the traffic into the matrix.

While we focus on changes of the traffic matrix because the capacity of the network has changed in the context of network design and capacity expansion, [16] focuses on traffic engineering and investigates how the routing performance in form of network utilization is affected by the fact that traffic matrices change in time and cannot be predicted exactly.

[17] presents an OSPF traffic engineering approach that also takes into account that traffic matrices change over time (e.g. the course of a day) and tries to minimize OSPF weight changes. Their approach is to optimize over multiple traffic matrices.

2.2 Related Network Models

Some works use network models similar to our models of Section 3. The performance models of [18, 19, 20] are used to analyse quality of service (QoS) in DiffServ [21] 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 [20] combines the DiffServ resource models with the TCP formula. We are using a similar approach but we also investigate non-exponential service times and non-Poisson arrivals. Also, we investigate performance in the context of network design and capacity expansion and not QoS and therefore do not use different service classes.

[22] presents an analytical TCP model for multiple flows and verifies it against NS2 simulations. Similar to our model they use a TCP and a network submodel and calculate the fixed point of the two models. Their TCP submodel however, is more fine-grained and complex than our TCP formula based TCP submodel. This however, comes at the cost of loosing a closed form formulation of the whole model. The authors investigate different network submodels and find that the simple $M/M/1/B$ gives sufficiently accurate results.

[23] introduces a queuing model that is based on multiple ON/OFF arrival processes which allows to account for long range dependency. It is extended to be reactive to congestion by slowing down the rate similar to the way TCP is reacting and can thus be used for performance analysis of TCP generated bursty traffic. Contrary to this approach we combine the TCP formula with standard queueing theory.

3 Network Models

In this section we present the analytical foundation of our analysis. 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.

3.1 Basic Model

We model a subnetwork Λ of the Internet consisting of N nodes and L directed links. The traffic through the network consists of long-lived greedy TCP connections and is represented by TCP *macro-flows*. A TCP macro-flow represents a number of TCP connections that have the same ingress node i and egress node j of Λ. We assume that the connections of a macro-flow experience on average the same loss $\bar{p}$ and delay $\bar{q}$ when traversing the other networks that are not modeled in detail with this model from their source to their destination. We assume that the macro-flows modeled in our model are small compared to the other flows flowing through the external networks and that because of this the external loss $\bar{p}$ and delay $\bar{q}$ are independent of the rate of the macro-flows. The macro-flows are elastic, their rate is described by a TCP-Formula and adapt to the network conditions of Λ. There are a number of works about predicting the average TCP throughput depending on the loss and delay properties of a flow [24, 25, 26, 27, 28]. As we are not interested in details like the duration of the connection establishment etc. we use the rather simple square-root formula ([24, 25]) in this work.

An output queue is attached to each link. In the basic model we model the queues as *M/M/1/B* queues [29, 30]. This is not the most realistic approach: First, because Internet traffic is not described very well by a Poisson arrival process [31]. Second, because packet sizes are not exponentially distributed an exponential service rate is also not realistic [32, 33]. However, the *M/M/1/B* model is one of the simplest queuing models and used in related works like [20, 22]. We will investigate more realistic queueing models later in this section.

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

The total loss probability of a macro-flow ij can be approximated by $p_{ij} = \bar{p} + \sum_{l \in \psi_{ij}} p_l$ for small loss probabilities. Similarly, for small loss probabilities at a link l the utilization (2) can be approximated by $\rho_l = (\sum_{(i,j) | l \in \psi_{ij}} r_{ij}) \cdot \frac{1}{\mu_l}$

These simplification 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 topology [34] with different parameters of t_{ij}, B and μ_l. We solve the non-linear equation system from Fig. 1 using MAPLE [35] and compare the difference in ρ_l. The maximum errors of different settings are listed in Table 1. They are

Indices	
$i, j = 1, \ldots, N$	Node i rsp. j
$l = 1, \ldots, L$	Link rsp. output queue l
Parameters	
ψ_{ij}	Path from node i to node j and back
t_{ij}	Size of macroflow between node pair i, j
μ_l	Service rate of link rsp. queue l
$\bar{q}$	Av. ext. queueing + total prop. delay
$\bar{p}$	Av. ext. loss probability
Variables:	
r_{ij}	Rate (pkts/sec) of macroflow betw. i, j
ρ_l	Utilization of link rsp. queue l
p_l	Loss probability of link rsp. queue l
q_l	Queuing delay of link rsp. queue l

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

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

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

$$q_l = \frac{1 - \rho_l}{\mu_l} \cdot \frac{\rho_l^{-}}{1 - {}^{B+1}} (B + 1) \qquad \forall l \quad (4)$$

Figure 1: The Basic Model

Approximation	Maximal Error [%]
for p_{ij}	0.0004795
for ρ_l	0.0009097

Table 1: Assessment of the Approximations

extremely small and can be neglected.

Next we discuss the possible extensions of the basic model.

3.2 Discrete Service Times

We first investigate how we can extend the basic model from Section 3.1 to account for more realistic service times. IP packets can differ drastically in their size (40 to 1500 Bytes) [32, 33]. 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 $pdf(x) = \sum_c h_c \cdot \delta(x - \frac{si_c}{sp_l})$ where $\delta(x)$ is the Dirac impulse $\delta(x) = 1$ for $x = 0$ and 0 otherwise. The probability distribution function is $PDF(x) = \sum_c h_c \cdot u(x - \frac{si_c}{sp_l})$ 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$q_l = E(x) \cdot (1 + \frac{1 + C_v^2}{2} \frac{\rho_l}{1 - \rho_l})$ with the expected service time[1] $E(x) = \frac{1}{\mu} = \int_{-\infty}^{\infty} x \cdot pdf(x)\,dx = \sum_c h_c \cdot \frac{s_{1c}}{sp_l}$ and the square of the coefficient of variation $C_v^2 = \frac{Var(x)}{E(x)^2} = \frac{\int_{-\infty}^{\infty}(x-E(x))^2 \cdot pdf(x)\,dx}{E(x)^2}$

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 according $M/G/1$ queue l.

[30, 36] 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 way 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 $b_l^*(s) = \sum_c h_c \cdot e^{-s\frac{s_{1c}}{sp_l}}$ [29, 36] show that the transformed state probabilities follow the Pollaczek-Khinchin transform formula for the queue length$Q_l(z) = (1-\rho_l)\frac{b_l^*(\lambda-\lambda z)}{b_l^*(\lambda-\lambda z)-z}(1-z)$. 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 quite complex term $Q_l(z)$ back:$\pi_{lk}^{(\infty)} = \frac{1}{k!}\frac{d^k}{dz^n}Q(z)\,|_{z=0}$. The loss probability of the related $M/G/1/B$ queue is now given as $p_l = 1 - \frac{1}{\rho_l + \pi_{l0}^{(B)}}$ using the state probability $\pi_{l0}^{(B)}$ of the finite queue $\pi_{l0}^{(B)} = \frac{\pi_{l0}^{(\infty)}}{\sum_{j=0}^{B-1}\pi_{lj}^{(\infty)}}$ [36]. 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.

3.3 Self-Similar Traffic

Internet traffic measurements show self-similar, heavy-tailed and long-range dependent properties [31]. 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 [37, 38]:

$$p_l = \frac{C}{\frac{B-\lambda\hat{t}}{\sigma_{\hat{t}}^2}\sqrt{2\pi\sigma_{\hat{t}}^2}} \cdot e^{-\inf_{t\in\Re^+}\frac{(B+(\mu-\lambda)\cdot t)^2}{2\cdot\sigma_t^2}} \quad (5)$$

$\hat{t}$ is the optimizer from the infimum condition, t is the timescale, B is the buffer size, λ and μ are the arrival rsp. service rate. For a given Hurst parameter σ_t^2 is given as $\sigma_t^2 = \sigma^2 \cdot t^{2H}$.

[1]We keep using μ for the inverse of the expected service time as we did with the $M/M/1/B$ queue.

4 Elasticity of Traffic Matrices

The influence that the elasticity of a traffic matrix has when the capacity of the network changes while all other conditions remain the same (ceteris paribus) is being analyzed in this section. The effects described in this section are normally neglected when static traffic matrices are used. We base our analysis on the different network models of the previous section.

4.1 Single-Link Experiments

We start our analysis with an extensive series of experiments on a single link. The Figure 2 shows the rate increase $\frac{r_{ij}^{new} - r_{ij}^{old}}{r_{ij}^{old}}$ of the symmetrical macroflows over the single-link topology for different queue lengths B (measured in packets) and different values for the external loss $\tilde{p}$ and delay $\tilde{q}$ when the link capacity μ_l is doubled $\mu_l^{new} = 2 \cdot \mu_l^{old}$. Figure 2 (a) lists the results for the basic model of Section 3.1, (b) shows the results for the model with discrete service times of Section 3.2. We used two different service time distributions, distribution A consists of 50% packets with a size of 40 and 50% packets with a size of 1500 Bytes. Distribution B consists of packets of size 1000 Bytes only. We assumed a line rate of 1 Mbps and had to use a rather low queue length of $B = 10$ packets because the loss probability formula gets too complicated for larger values of B to be handled analytically.

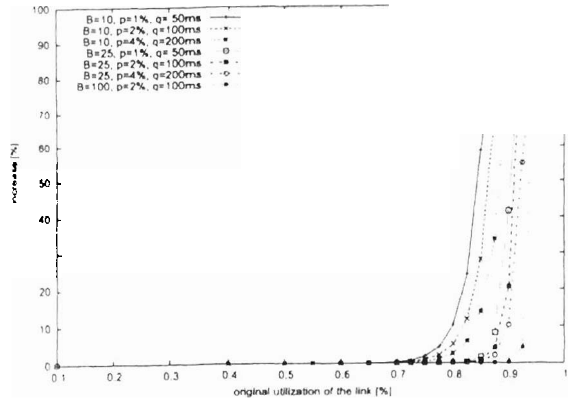
Figure 2 (c) shows the results we obtain if we assume the loss probability of the model for self-similar traffic from Section 3.3. We used a Hurst parameter of $H = 0.75$, a line rate of 1 Mbps, an average service packets size of 1000 Bytes and the according average service time.

Looking at the results we notice that for all three different network models and most parameters the general behaviour of the traffic is the same. Up to a certain utilization threshold of the analyzed link the traffic largely unaffected by the increase in capacity. Then the traffic increases very quickly. If the initial utilization of the link is high enough the analyzed link forms a strong bottleneck and all additional capacity is used up completely by a rate increase of 100%.

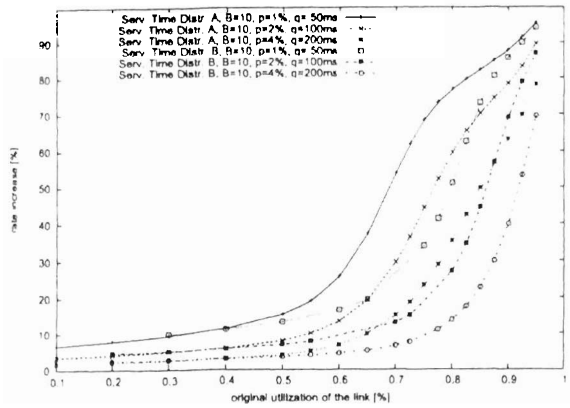
The step is steeper for the $M/M/1/B$ network model than for the other two models that can be deemed more realistic.

4.2 Different Topologies

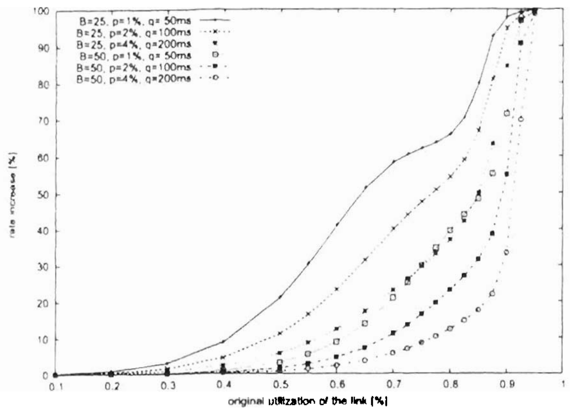
We now analyze the elasticity in form of the rate increase for more complex topologies than the single link topology of the previous experiments. 3 summarizes the results for three different topologies, the backbone of the Deutsche Telekom [34], a dumbbell topology with a single bottleneck link and three nodes on each side of the bottleneck and a star topology with one internal and 4 external nodes. We varied the value of t_{ij} between 10^{-1} and 10^2, doubled the network capacity for each t_{ij} and recorded the rate increase. As



(a) Results for the Basic Network Model



(b) Results for Discrete Service Times



(c) Results for Self-Similar Traffic

Figure 2: Single-Link Experiment Results

one can see, the different topologies lead to similar results. While most of the rate increases are very small (more than 50% of the times the rate increase was below 10%), there are a significant number of times where the rate increase was very high. Because of the different paths the different flows take through the topology the rate increase can be higher than 100% if a series of links is doubled in capacity for a flow.

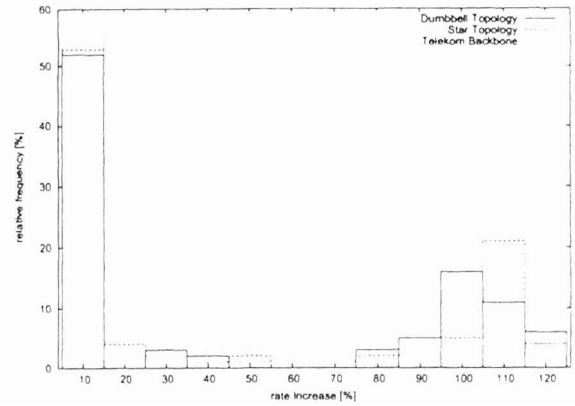


Figure 3: Rate Increase for Different Topologies

Conclusion If a traffic matrix is used in the context of network design or capacity expansion the elasticity of the traffic can be quite important as our results show. The elasticity of the traffic can only be neglected up to a certain threshold of the utilization of a link. Once that threshold is passed, the error can be quite significant.

5 Capacity Expansion

Capacity expansion of an existing network topology are an important problem for ISPs. In this section, we focus on a special type of capacity expansion problem with elastic traffic:

The provider measures and evaluates the link utilization (as available e.g. from SNMP data) every period. If the utilization ρ_l exceeds a certain threshold ξ on a link l the capacity expansion for that link is triggered. The motivation for this approach is that the quality of service of a best-effort network is only good if certain utilization thresholds are not exceeded [39].

The capacity of a link is expanded by doubling the bandwidth of that link as new line cards are added to the routers connected by the link. The expansion will take some time, we assume that the link capacity is effectively doubled to the beginning of the next period after the one that triggered the expansion.

Traffic is given in form of the parameter t_{ij} of equation 1. The actual traffic volume passed through the network is elastic and will react to the capacity expansion.

In "classical" network design and capacity expansion algorithms the elasticity of the traffic is ignored. The problem is that by increasing the capacity of a link the traffic flows through that link will increase their rate and therefore the utilization also of the other links they are flowing through. This can lead to the situation (a) that immediately after the expansion the threshold ξ on other links will be exceeded which is not predicted by the classical model with static traffic matrices and

Telekom Backbone	how often avoided	in percent of total number of link expansions
Effect (a)	23	21.5%
Effect (b)	-	-

Table 2: Simulation Results

it will take an additional period until they can be expanded, too. Furthermore, if a link is an extreme bottleneck for some flows it is possible that the utilization will not significantly decrease if the link is doubled. This effect (b) can also not be predicted with static matrices. This effect was for example observed when the UK ISP Rednet quadrupled their DSL access link capacity.

Using the our models of Section 3, we can predict the traffic increase rsp. utilization change of a planned network expansion and avoid the effects (a) and (b). We use the following simulation as a proof of concept:

Using the backbone topology of the Deutsche Telekom [34] we generate a traffic matrix with random entries r_{ij} between 1.0 and 5.0. We use this for the inital parameters t_{ij}. We choose a starting line-rate of 1 Mbps for all links that is doubled for each link before the actual simulation until all link utilizations are below 70%. We then simulate 10 periods, at the beginning of each period each traffic matrix entry is increased randomly between 5 and 20%. Our basic model of Section 1 is used to calculate the link utilizations - we assume that the result of these initial calculations represents the SNMP data collected by the provider. We used an external loss of 2% and delay of 100ms. This results in a not too aggressive behavior of TCP. The expansion of a link l is triggered if it has a utilization of $\rho_l \geq \xi = 0.75$.

In order to capture the elasticity of the traffic matrix we can again use our basic model to predict the effect of the triggered capacity expansions in order to avoid the effects (a) and (b) described above. We do so and measure how often these effects were avoided. The results are summarized in Table 2. Because we increased the rates only in moderate steps and allowed to increase the capacity in each period effect (b) did not occur in our simulations and could therefore not be avoided by the model. Effect (a) however occured 23 times and could be avoided by using our prediction of the elastic traffic. This shows that our concept works and significantly helps in capacity expansion decisions.

6 Summary and Outlook

In this paper we presented three analytical network models of varying complexity that predict the queueing behaviour (loss and delay) and the traffic behavior of a network. We assume that long-lived TCP connections dominate the traffic behavior and therefore model the elasticity of these connections using the TCP formula

Because of the elasticity of that traffic it will adapt to changes in the network, e.g. to capacity increases. The capacities of operational networks are increased constantly as the traffic demand (e.g. the customer base and the downloading behavior) increases rapidly. With this in mind, we used our models to analyse the reaction of the traffic to an increase in link bandwidth. The results show that the elastic effects can be ignored only if the initial utilization of that link is low. Unfortunately, it is especially the links with a high utilization that are increased in capacity regularly, therefore it is important to take into account the elasticity of the traffic matrix. In the last part of this paper we showed a simple way of incorporating our models into a capacity expansion problem in order to predict the elasticity. A small simulation shows how this approach works.

Acknowledgements The authors would like to thank Richard Gibbens and Damon Wischik of the Cambridge University, Raza Rizvi of the UK ISP Rednet and Gerhard Hasslinger of T-Systems for the valuable and helpful discussions.

References

[1] Sandvine Incorporated, "Regional Characteristics of P2P - File Sharing as a Multi-Application, Multi-National Phenomenon," *White Paper*, 2003.

[2] C. Fraleigh, S. Moon, B. Lyles, C. Cotton, M. Khan, D. Moll, R. Rockell, and T. Seely, "Packet-Level Traffic Measurements from the Sprint IP Backbone," *IEEE Networks*, vol. 17, no. 6, pp. 6–16, Nov/Dec 2003.

[3] A. Bock, "Measurements in the Edonkey2000 Network," Master's thesis, Multimedia Communications Lab (KOM), Darmstadt University of Technology, 2002.

[4] M. Handley, S. Floyd, J. Pahdye, and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification," RFC 3448, Proposed Standard, Jan 2003.

[5] B. Gavish, "Topological Design of Communication Networks - The Overall Design Problem Overall Design Problem," *European Journal of Operational Research*, vol. 58, pp. 149–172, 1992.

[6] L.T.M.Berry, B. Murtagh, G. McMahon, S. Sugden, and L. Welling, "An integrated GA-LP Approach to Communication Network Design," *Telecommunication Systems Journal*, vol. 12, no. 2-3, pp. 265–280, 1999.

[7] N. Benameur and J. Roberts, "Traffic Matrix Inference in IP Networks," in *Proceedings of Networks 2002*, Jun 2002, pp. 151–156.

[8] A. Feldmann, A. G. Greenberg, C. Lund, N. Reingold, J. Rexford, and F. True, "Deriving Traffic Demands for Operational IP Networks: Methodology and Experience," in *Proceedings of ACM Special Interest Group on Data Communication Conference (SIGCOMM 2000)*, Aug 2000, pp. 257–270.

[9] Y. Vardi, "Network Tomography: Estimating Source-Destination Traffic Intensities from Link Data," *Journal of The American Statistical Association*, vol. 91, pp. 557–576, 1996.

[10] C. Tebaldi and M. West, "Bayesian Inference on Network Traffic using Link Count Data," *Journal of the American Statistical Association*, vol. 93, pp. 557–576, 1998.

[11] J. Cao, D. Davis, S. V. Wiel, and B. Yu, "Time-Varying Network Tomography," *Journal of the American Statistical Association*, vol. 95, pp. 1063–1075, 2000.

[12] O. Goldschmidt, "ISP Backbone Traffic Inference Methods to Support Traffic Engineering," in *Proceedings of Internet Statistics and Metrics Analysis (ISMA) Workshop*, Dec 2000.

[13] A. Medina, N. Taft, K. Salamatian, S. Bhattacharyya, and C. Diot, "Traffic Matrix Estimation: Existing Techniques and New Directions," in *Proceedings of ACM Special Interest Group on Data Communication Conference (SIGCOMM 2002)*, Aug 2002, pp. 161–174.

[14] Y. Zhang, M. Roughan, N. G. Duffield, and A. G. Greenberg, "Fast Accurate Computation of Large-Scale IP Traffic Matrices from Link Loads," in *Proceedings of the International Conference on Measurements and Modeling of Computer Systems (SIGMETRICS 2003)*, Jun 2003, pp. 206–217.

[15] Y. Zhang, M. Roughan, C. Lund, and D. Donoho, "An Information-Theoretic Approach to Traffic Matrix Estimation," in *Proceedings of ACM Special Interest Group on Data Communication Conference (SIGCOMM 2003)*, Aug 2003, pp. 301–311.

[16] D. Applegate and E. Cohen, "Making Intra-Domain Routing Robust to Changing and Uncertain Traffic Demands: Understanding Fundamental Tradeoffs," in *Proceedings of ACM Special Interest Group on Data Communication Conference (SIGCOMM 2003)*, Aug 2003, pp. 313–324.

[17] B. Fortz and M. Thorup, "Optimizing OSPF/IS-IS Weights in a Changing World," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 4, pp. 756–767, May 2002.

[18] M. May, J. C. Bolot, A. Jean-Marie, and C. Diot, "Simple Performance Models of Differentiated Services Schemes for the Internet," *Proceedings of IEEE Conference on Computer Communications and Networking (INFOCOM 1999)*, pp. 1385–1394, Mar 1999.

[19] R. Gibbens, S. Sargood, F. Kelly, H. Azmoodeh, R. Macfadyen, and N. Macfadyen, "An Approach to Service Level Agreements for IP Networks with Differentiated Services," in *Royal Society Discussion Meeting on Network Modelling in the 21st Century*, Dec 1999.

[20] R. Gibbens, S. Sargood, C. V. Eijl, F. Kelly, H. Azmoodeh, R. Macfadyen, and N. Macfadyen, "Fixed-Point Models for the End-to-End Performance Analysis of IP Networks," in *Proceedings of the 13th ITC Specialist Seminar on IP Traffic Management, Modeling and Management*, Sept 2000.

[21] D. Black, S. Blake, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services," 12 1998, RFC 2475.

[22] M. Garetto, R. L. Cigno, M. Meo, and M. A. Marsan, "A Detailed and Accurate Closed Queueing Network Model of Many Interacting TCP Flows," in *Proceedings of IEEE Conference on Computer Communications and Networking (INFOCOM 2001)*, Apr 2001, pp. 1706–1715.

[23] H.-P. Schwefel, "Behavior of TCP-like Elastic Traffic at a Buffered Bottleneck Router," in *Proceedings of IEEE Conference on Computer Communications and Networking (INFOCOM 2001)*, Apr 2001, pp. 1698–1705.

[24] S. Floyd, "Connections with Multiple Congested Gateways in Packet-Switched Networks, Part 1: One-way Traffic," *Computer Communications Review*, vol. 21, no. 5, pp. 30–47, Oct 1991.

[25] T. Lakshman and U. Madhow, "The Performance of TCP/IP for Networks with High Bandwidth-Delay Products and Random Loss," *IEEE/ACM Transactions on Networking*, vol. 5, no. 3, pp. 336–350, 1997.

[26] M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm," *Computer Communication Review*, vol. 27, no. 3, Jul 1997.

[27] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP Throughput: A Simple Model and its Empirical Validation," in *Proceedings of ACM Special Interest Group on Data Communication Conference (SIGCOMM 1998)*, 9 1998, pp. 303–314.

[28] N. Cardwell, S. Savage, and T. Anderson, "Modeling TCP Latency," in *Proceedings of IEEE Conference on Computer Communications and Networking (INFOCOM 2000)*, Mar 2000, pp. 1742–1751.

[29] L. Kleinrock, *Queueing Systems - Theory*. Wiley-Interscience, New York, Vol.1, 1975, iSBN 0471491101.

[30] R. B. Cooper, *Introduction to Queueing Theory*, 2nd ed. Elsevier, North-Holland, 1981, iSBN 0444003797. [Online] Available: www.cse.fau.edu/ bob/ publications/ IntroToQueueingTheory_Cooper.pdf

[31] V. Paxson and S. Floyd, "Wide Area Traffic: the Failure of Poisson Modeling," *IEEE/ACM Transactions on Networking*, vol. 3, no. 3, pp. 226–244, 1995. [Online] Available: http://citeseer.nj.nec.com/paxson95widearea.html

[32] K. Claffy, G. Miller, and K. Thompson, "The Nature of the Beast: Recent Traffic Measurements from an Internet Backbone," in *Proceedings of Internet Society Internet Global Summit (Inet 1998)*, Jul 1998, www.caida.org/Papers/Inet98/index.html.

[33] NASA Ames Internet Exchange (AIX), "Packet Length Distributions," www.caida.org/analysis/AIX/plen_hist/, Feb 2000.

[34] O. Heckmann, "Topologies for ISP level network simulation (website)," http://www.kom.e-technik.tu-darmstadt.de/ heckmann/topologies/.

[35] Maplesoft, "MAPLE: Advanced Mathematical Problem Solving and Programming Environment," http://www.maplesoft.com/.

[36] J. Virtamo, "Queueing Theory," www.netlab.hut.fi/ opetus/ s38143/ luennot/ english.shtml, 2003, lectures notes.

[37] P. Mannersalo and I. Norros, "A Most Probable Path Approach to Queueing Systems with General Gaussian Input," *Computer Networks*, vol. 40, no. 3, pp. 399–412, Sept 2002.

[38] R. Addie, P. Mannersalo, and I. Norros, "Most Probable Paths and Performance Formulae for Buffers with Gaussian Input Traffic," *European Transactions on Telecommunications*, vol. 13, no. 3, pp. 183–196, May/Jun 2002.

[39] G. Hasslinger, "Quality-of-Service Analysis for Statistical Multiplexing with Gaussian Distributed and Autoregressive Input," *Telecommunication Systems Journal*, vol. 16, no. 3-4, pp. 315–334, Mar/Apr 2001.