University of Würzburg
Institute of Computer Science
Research Report Series

Towards an optimization of the routing parameters for IP networks

Dirk Staehle[1], Stefan Köhler[2], and Ute Kohlhaas[3]

[1] Department
of Computer Science,
University of Würzburg
Am Hubland, D-97074 Würzburg, Germany
staehle@informatik.uni-wuerzburg.de

[2] Department
of Computer Science,
University of Würzburg
Am Hubland, D-97074 Würzburg, Germany
koehler@informatik.uni-wuerzburg.de

[3] Department of Stochastics,
Aachen, University of technology
D-52056 Aachen, Germany
u.kohlhaas@stochastik.rwth-aachen.de

Towards an optimization of the routing parameters for IP networks

Dirk Staehle
Department
of Computer Science,
University of Würzburg
Am Hubland, D-97074 Würzburg, Germany
staehle@informatik.uni-wuerzburg.de

Stefan Köhler
Department
of Computer Science,
University of Würzburg
Am Hubland, D-97074 Würzburg, Germany
koehler@informatik.uni-wuerzburg.de

Ute Kohlhaas
Department of Stochastics,
Aachen, University of technology
D-52056 Aachen, Germany
u.kohlhaas@stochastik.rwth-aachen.de

Abstract

Routing is one of the key issues in IP networks. However, few methods exist to optimize the routing for a particular network. Most effort is invested to improve the routing protocols itself. In this work a possibility to specify appropriate values for the link costs of a given network with linear programs is presented. The obtained link costs can be directly translated into values suitable for the metrics of the two currently most important routing protocols EIGRP and OSPF in today's Internet. With this method a homogeneous distribution of traffic in IP-based networks can be achieved.

1 Introduction

We consider the problem of introducing optimal routing methods for IP packet traffic. IP datagrams are currently used to transmit different types of network services. The rapid explosion in the use of the Internet for Web browsing, telephony and video services, as well as more traditional services such as telnet and ftp, has resulted in a massive increase in traffic load. With this increase in traffic volume, there has been a corresponding significant increase in congestion due to the lack of network resources. Efficient routing of IP packets is becoming a crucial issue both from the point of view of the providers, and of the users of the network.

The most effort concerning routing optimization is focused on improvements of the routing protocols itself [5, 8, 11]. However, few methods deal with the optimization of the routing of a particular network. Those methods are mainly known from telephone networks with fixed connections [1, 2, 4]. In IP networks the possibilities to optimize the routing are more restricted if the existing routing protocols are taken as a given fact. Since IP is connectionless a new routing decision is made at each router in separation, and all IP packets with the same destination are routed on the same path independent of the source they come from. Inside an intranet or inside the network of an internet service provider (ISP), the chosen path is always the shortest path following a certain metric specified by the routing protocols. In the next section an overview of the metrics of IP routing protocols is given. It is described how the metric can be reduced to consist of the sum of link costs which can be set in the routers database.
In the following sections two linear problems are outlined. The first one describes a method to specify paths for given flows in a given network. These paths satisfy the condition that link costs can be determined such that the received paths are the shortest paths. The link costs are computed by a second linear program. Finally, the results for some example networks are presented.

2 Routing in IP networks

Up to now there have been a number of different unicast IP routing strategies employed by router manufacturers or the IETF (Internet Engineering Task Force). An overview is given in Figure 1. They can be broadly classified as:

- Interior Gateway Protocols (IGP) and
- Exterior Gateway Protocols (EGP).

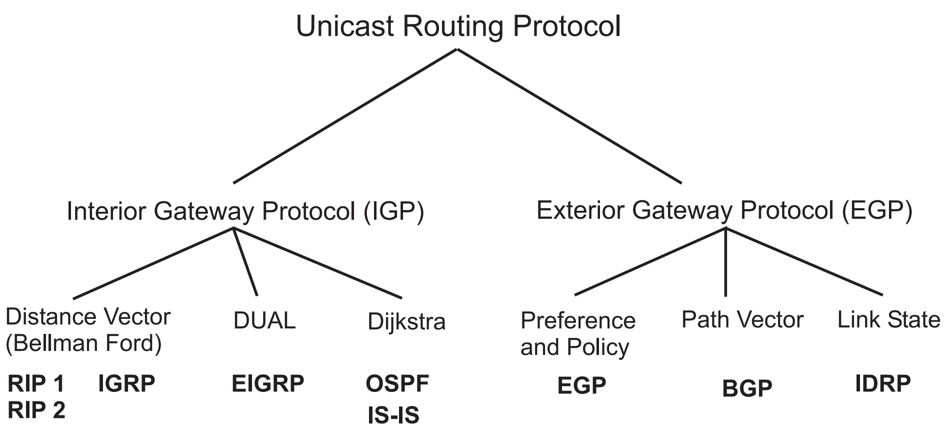

Figure 1: Overview of Unicast Routing Protocols

The IETF defines protocols as interior gateway protocols if they are used for "routing networks that are under a common network administration". This common administration is frequently referred to as an *Autonomous System (AS)*. The most popular interior routing protocols are the following:

- Routing Information Protocol (RIP)
- Enhanced Internet Gateway Routing Protocol (EIGRP)
- Open Shortest Path First (OSPF) Protocol

More information about the interior routing protocols can be found in [6, 10, 7]. Exterior gateway protocols as defined by CISCO, "exchange routing information between networks that do not share a common administration." More about exterior gateway protocols can be found in e.g. [7].

Each of these protocols needs to be configured for use in its respective network environment. Router manufacturers publish details for their customers on how to configure their routers for these protocols, although they do not provide guidance on how to choose these settings. All interior routing protocols are working after the same principle. They define the cost of a link with a protocol depended metric and determine with these cost values the shortest path.

The world largest router manufacturer, CISCO, defines the metric for OSPF and EIGRP as follows:

Protocol	Metric	Range
EIGRP	$M = \left[K_1 \cdot \frac{1}{min_i(C_i)} + \frac{K_2 \cdot \frac{1}{min_i(C)}}{256 - \text{load}} + K_3 \cdot \sum_i delay_i\right] \cdot \frac{K_5}{\text{R} + K_4}$ if $K_4 \neq 0$ and $K_5 \neq 0$	$\left[0;\ \left(2^{32} - 1\right)\right]$
OSPF	$M = \sum_i \frac{100{,}000{,}000\text{bps}}{C_i}$	$[1;\ 65535]$

Table 1: Definition of the metric of EIGRP and OSPF.

C_i is the capacity of the link i, R is the reliability of the path, K_1 to K_5 are scaling parameters and $delay_i$ is the physical delay of link i.

In practice the metric of EIGRP is reduced to $M = K_1 \cdot \frac{1}{min_i(C_i)} + K_3 \cdot \sum_i delay_i$. The metric in OSPF is not defined in an RFC. Thus, it is possible for every router manufacturer to define his own metric. The only constraint is, that the range of link cost must be between 1 and 65535.

As could be seen in Table 1 the metric for both routing protocols is based on time independent parameters, e.g the delay or the capacity of the link [1]. All the parameters could be independently configured in the router database. The idea behind the paper is that for a measured end-to-end traffic matrix the parameters are chosen in such a way, that the traffic is uniformly distributed over the network. The results of our optimization are integer values which represent the cost of the links. It is possible

[1] The full EIGRP metric is time dependent, but the reduced metric which is used in practice is time independent.

to translate the integer values to the routing parameters and set theses values in the router database for an optimized routing.

3 Problem definition

The aim of the routing optimization in this paper is to achieve a routing that uniformizes the link utilizations in an IP network. A network with N routers is defined by a capacity matrix C of size $N \times N$ which comprises the link capacities c_{ij} between each two routers i and j. If no link exists between two routers the entry in the capacity matrix is set to zero. The matrix is not restricted to symmetry since for example ADSL links are asymmetric.
The optimization is performed assuming static flows. Therefore routing for the rush hour case is done and consequently the flows are assumed as the maximum flows. Like the capacities they are defined using an end-to-end traffic matrix F which comprises an entry f_{ij} for each two nodes i and j. The matrix specifies the volume of the data stream that has to be transmitted from i to j.
Furthermore, a third matrix D is introduced which describes the physical delays d_{ij} of the links. These physical delays restrict the set of routing possibilities between two nodes. Of course, one would not send packets on a path which due to its physical delay is many times longer than the shortest path.
So the objective of the routing optimization is to

1. minimize the maximum link utilization
2. minimize all link utilizations
3. keep the physical delays in a certain range

This problem can be formulated as a linear optimization problem.

4 Formulation as Linear Optimization problem

The formulation of a routing optimization problem as a linear optimization program is well-known [9, 3]. A linear problem consists of two parts, the objective function and the constraints. The constraints define a multidimensional solution space, wherein an optimal solution has to be determined. An optimal solution is an element of the solution space producing the maximal resp. minimal objective function value.
In the case of an IP routing optimization the solutions have to fulfil several constraints. First, for each traffic flow a path from the source to the destination has to be found over which the flow is routed. In Section 2 the routing possibilities were restricted to single path routing. So in a possible solution exactly one path has to be defined between each two routers. The second condition for a solution is that the amount of data flowing over a certain interface does not exceed the link's capacity. In contrast to traditional telephone networks the traffic in IP networks is not restricted to link's

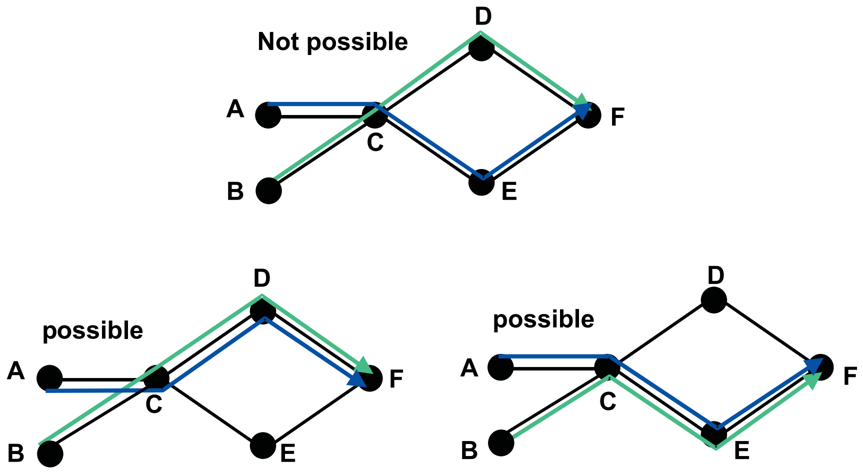

Figure 2: IP Routing restriction

capacity, it may be somewhat higher as well. In that case IP packets will be stored or dropped in the router. At least two different modeling approaches can be used to formulate those conditions in terms of constraints for a linear program. The first one is path-oriented, the second one is link flow-oriented. Due to the smaller number of variables in the flow-oriented approach, this one was chosen.
Furthermore, the routing algorithm specified by the applied routing protocol has to be considered. As mentioned above both routing protocols, OSPF and EIGRP, currently use shortest path routing, however with different metrics and different algorithms. This implies that the routing between each two routers is identical for all flows unless multipath routing is considered. However, the influence is neglected in this optimization. In Fig. 2 a simple example for this restriction is shown. The shortest path - independent of the used metric - between router C and router F is either over D or over E. Consequently, both flows A-F and B-F are routed first over C and then either over D or over E, but not one over D and one over E.

4.1 Variables

Before formulating the different constraints and the objective function the used variables have to be specified. Since a flow-oriented approach is used for each flow with $f_{ij} \geq 0$ and for each link with $c_{ij} > 0$ a boolean variable x_{ij}^{uv} is introduced. This variable is set to one if flow uv is routed over link ij and otherwise set to zero. The variable t is an upper bound for the utilization of all links.

4.2 Objective function

The aim of this routing optimization is to receive a traffic distribution as homogeneous as possible in the entire network; this means that all links should be utilized equally at a level as low as possible. This is obtained by minimizing the maximum link utilization as far as possible. But if this maximum utilization is found for a certain link the traffic of all other links shall be minimized without increasing the found maximum value. Therefore, the objective function comprises two additive parts that both have to be minimized:

$$a_t * t + \sum_{ij} \sum_{uv} \frac{f_{uv} x_{ij}^{uv}}{c_{ij}}. \tag{1}$$

In the first part the maximum link utilization is minimized. As we will see in Equ. 7 a constraint is formulated that forces all utilizations below the value of t. The second part reduces the average link utilization.
The parameter a_t defines the importance of a small maximum utilization versus the importance of a small average utilization. If it is set sufficiently large the prime aim of reducing the maximum utilization by directing traffic onto less utilized links is achieved.

4.3 Constraints

4.3.1 Transport Constraints

The transport constraints guarantee that for each flow, say uv, exactly one loopfree path from router u to router v is specified by the resulting values of the variables x_{ij}^{uv}. This is obtained with four constraints for each flow.

1. only one link leading out of a router i may carry traffic of flow uv.

$$\sum_{j=1,\ c_{ij}>0}^{N} x_{ij}^{uv} \leq 1, \text{ for all flows } uv \text{ and all routers } i \tag{2}$$

2. exactly one link ui from router u to another router i has to carry flow uv.

$$\sum_{i=1,\ c_{ui}>0}^{N} x_{ui}^{uv} - \sum_{i=1,\ c_{iu}>0}^{N} x_{iu}^{uv} \geq 1, \text{ for all flows } uv \tag{3}$$

3. for all flows uv and for each router $i \notin \{u, v\}$ the sum of incoming links used by flow uv equals the sum of outgoing uv carrying links and because of Equ. 2 equals one.

$$\sum_{j=1,\ c_{ij}>0}^{N} x_{ij}^{uv} - \sum_{j=1,\ c_{ji}>0}^{N} x_{ji}^{uv} = 0, \text{ for all flows } uv \text{ and all routers } i \notin \{u, v\} \tag{4}$$

4. exactly one link running into router v has to carry the traffic of flow uv.

$$\sum_{i=1,\ c_{iv}>0}^{N} x_{iv}^{uv} - \sum_{i=1,\ c_{vi}>0}^{N} x_{vi}^{uv} \geq 1, \text{ for all flows } uv \tag{5}$$

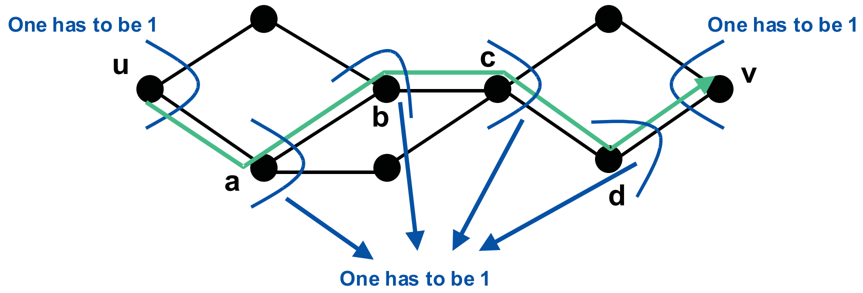

Figure 3: Effect of the transport constraints

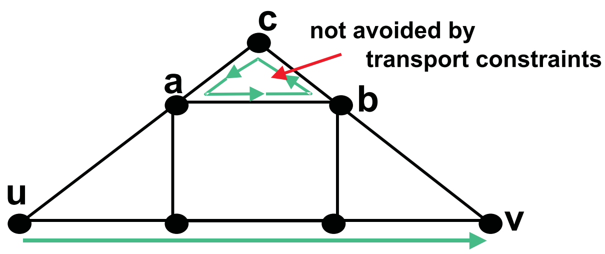

Figure 4: Loops despite Transport Constraints

Together, the constraints satisfy that exactly one loopfree path from u to v is specified in each solution. Fig. 3 shows by means of an example network how a coherent path is received by the four conditions. The blue lines indicate the points where the constraints require a link to carry flow uv. Following Equ. 3, from the links out of u only the link ua carries flow uv. At routers a, b, and c one link runs into these nodes, hence conditioned by Equation 4 one link out of them has to be used by uv as well. Finally by Equation 5 the path runs over link cv and ends in v.
Nevertheless, it is still possible that a loop exists beside the path, as shown is Figure 4. We can see that no constraint contradicts the existence of such a loop. However, this absurd solution is avoided by the minimization of the average link utilization in the objective function.

4.3.2 Capacity Constraints

The capacity constraints guarantee that the traffic over a link does not exceed certain limits. Two constraints are required for each link ij. The first one achieves that the link utilization does stay below a fixed limit given by the parameter a_c. This is obtained by the following constraint:

$$\sum_{uv} x_{ij}^{uv} f_{uv} \leq a_c c_{ij}, \text{ for all links } ij. \tag{6}$$

As mentioned above, within IP networks it is possible that links are offered more traffic than they are able to handle. Therefore a_c is not restricted to values between zero and

one. Nevertheless, by default a_c is set to one.
Furthermore the value of the variable t is another upper bound for the link utilizations. However, it is a variable that shall be minimized whereas the parameter a_c is a fixed value. All link utilizations have to be less than t. This is realized by the following constraint:

$$\sum_{uv} x_{ij}^{uv} f_{uv} \leq \frac{t}{100} c_{ij}, \text{ for all links } ij. \tag{7}$$

4.3.3 Routing Constraints

Up to now the constraints specify that each found solution provides a path for each flow while the given link utilizations are not exceeded. The routing in an IP network is more restricted due to the functionality of the routing protocols. The crucial point thereby is that all IP packets are always routed on the shortest path. The shortest path is determined following a certain metric for each link. This metric depends on the used routing protocol.
As a consequence all flows with routers i and j in the same order on their path have to be routed over the same way between i and j. This restriction was further illustrated in Figure 2. The IP conform routing is achieved by adding the following constraints to the linear program:

$$x_{ui}^{uv} + x_{st}^{iv} - x_{st}^{uv} \leq 1, \qquad \text{for all flows } uv, \text{ routers } i \notin \{u, v\} \text{ and links } st \tag{8}$$

$$x_{jv}^{uv} + x_{st}^{uj} - x_{st}^{uv} \leq 1, \qquad \text{for all flows } uv, \text{ routers } j \notin \{u, v\} \text{ and links } st. \tag{9}$$

Equ. 8 can be interpreted in the following way:

$$x_{ui}^{uv} = 1 \Rightarrow x_{st}^{iv} \leq x_{st}^{uv}$$

If the first hop of flow uv leads to router i, all other links st used by flow iv have to be used by flow uv as well. So uv and iv are routed on the same path from router i to router v.
Equ. 9 can be interpreted similarly. If the last hop of flow uv goes out of router i, all other links st used by flow uj have to be used by flow uv as well. So uv and uj are routed on the same path from router u to router j.
When we conceive an iterated application of these equations we notice that for each two routers i and j the routing between these routers is identical for all flows through these nodes. And so the found solutions correspond to IP routing implementations.
In Fig. 5 the meaning of the IP constraints is illustrated. It is assumed that flows ub, ac, and cv are routed as shown by the red arrows. Due to Equ. 8 and 9 the flow uv has to be routed as shown by the green arrow.
If we look at flow uc first, two other paths different from the green one exist, either

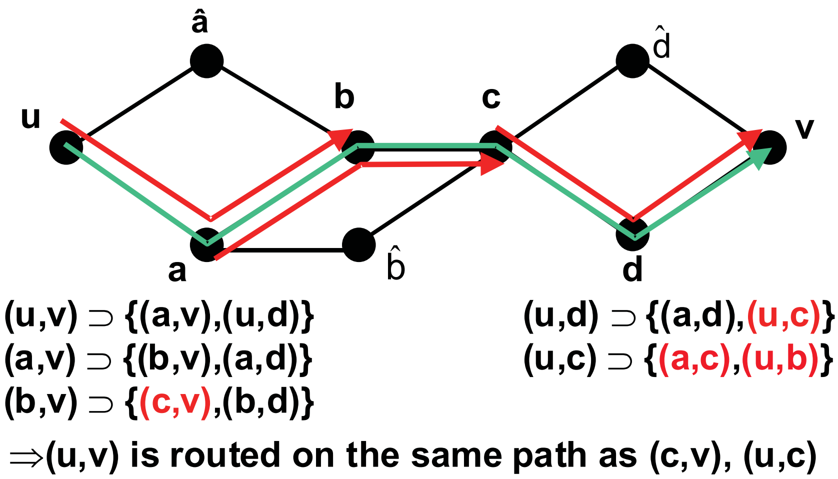

Figure 5: IP Constraints

over $\hat{a}$ or over $\hat{b}$. Assumed uc runs over $\hat{b}$ link ua is used. But then due to Equ. 8 between a and c flows uc and ac have to use the same path. For the other path the case is analogous but with Equ. 9. Thus, flow uc is routed along the green arrow. Again, for flow bv an alternative path over $\hat{d}$ exists but it can not be used due to Equ. 8. With both routing constraints applied repeatedly uv has to use the same path like flow uc between u and c and also the same path like flow bv between b and v. Altogether, flow uv has to use the path indicated by the green arrow.

4.3.4 Physical Delay Constraints

With the constraints and the objective function described above a shortest path conform routing with evenly distributed traffic is achieved. Thus, only the third aim to keep the physical delays in a certain range is not fulfilled yet. The range of the possible physical delays is defined over the parameter a_r and the lowest possible physical delay for a flow d_{uv}^{min}. Let P_{uv} be the set of all loopfree paths from u to v and let d_p be the sum of the physical delays of the interfaces of path p. Then the minimum physical delay d_{uv}^{min} is the delay of the path p from u to v with the smallest delay d_p:

$$d_{uv}^{min} = min_{p \in P_{uv}}(d_p). \tag{10}$$

This value can be found out by computing the physical delays of all possible loopfree paths from u to v.
With these two values a_r and d_{uv}^{min} for each flow a constraint can be specified that keeps the physical delay below or equal $a_r d_{uv}^{min}$:

$$\sum_{ij} x_{ij}^{uv} d_{ij} \leq a_r d_{uv}^{min}, \text{ for all flows } uv. \tag{11}$$

4.4 Reduction of the complexity of the linear program

In the above formulation of the problem the number of variables is in the magnitude of N^2M^2 where N is the number of routers and M is the number of interfaces. Furthermore the number of constraints is in the magnitude of N^3M^2, determined by the routing constraints. Accordingly, the problem can be solved for rather small networks only. However, if one takes a closer look at greater intranets the routing decision is only interesting for a few competing paths. Therefore, many variables can be presolved by considering only relevant paths.

Whether a path is relevant or not is defined over its physical delay. A path $p \in P_{uv}$ is considered as not relevant if its physical delay d_p is more than a_r times higher than the minimal physical delay d_{uv}^{min}. Then $\hat{P}_{uv}$ denotes the set of all relevant paths:

$$\hat{P}_{uv} = \{p \in P_{uv} \mid d_p \leq a_r d_{uv}^{min}\} \tag{12}$$

With the set of all relevant paths the value of certain variables can be determined:

$$x_{ij}^{uv} = \begin{cases} 0 & : \forall p \in \hat{P}_{uv} : \delta_{ij}(p) = 0 \\ 1 & : \forall p \in \hat{P}_{uv} : \delta_{ij}(p) = 1 \\ unspecified & : else, \end{cases} \tag{13}$$

where $\delta_{ij}(p)$ is one if p uses link ij and zero otherwise. Here, unspecified means that the link ij is used by some but not all of the relvant paths in P_{uv}. The fixed variable values can be substituted into the above program formulation and the constraints which are always complied with these values can be omitted.

Another possible simplification is to fix the path between neighbor routers to the link connecting them. Then for each link ij the value of x_{ij}^{ij} is one. This assumption makes sense since if the traffic of the neighbor routers is not transmitted over the direct link no traffic at all will be transmitted over it due to the routing constraints. The shortest path principle would force all other flows between i and j to use the same path as the flow ij. Hence, either x_{ij}^{ij} is set to one or otherwise link ij is omitted completely.

The previous possibilities to reduce the complexity of the linear problem both presolved the values of certain variables by a restriction of the solution space. The third simplification does reduce the complexity of the problem itself. The treated test networks revealed that the simultaneous minimization of the maximum utilization and the average utilization together increases the problem complexity. However, it is possible to omit the variable t from the objective function and also the constraints of Equation 7 which force the maximum utilization below t. The reduction of the maximum utilization is now obtained by the reduction of parameter a_c. A first solution is found with a_c set to one or an even greater value. With this solution the maximum link utilization can be computed and the linear program can be formulated again but with a_c set to a value smaller than the received maximum utilization. If this is repeated until either the problem is identified as infeasible or the solver can not handle the problem due to its complexity, the optimal or at least a good solution is found. In Section 6 results with and without simplifications are compared.

5 Specification of the link costs

In the previous chapter a routing scenario was obtained which minimizes the maximum and average link utilizations and hence the IP packet transmission delays. Additionally, it is suitable for shortest path routing. However, the costs for the interfaces with which the found out routing is the shortest path routing as well, have still to be specified. Again this can be achieved by a linear program. In this case the objective function is rather unimportant. The linear problem solver is used only as a solver for an inequality system. The objective function is hereby used to minimize the obtained link costs. The important part of the linear program is the constraint formulation. They restrict the solution space to contain only solutions where the shortest path routing is identical with the optimized routing. The variables m_{ij} of the program represent the cost of interface ij. The shortest path from u to v is always the path with the least cost. The cost of a path is the sum of the cost values of the single interfaces the path comprises. For flow uv the cost m^{uv} is given by

$$m^{uv} = \sum_{ij} m_{ij} x_{ij}^{uv}, \tag{14}$$

where x_{ij}^{uv} are the values of the variables received by the routing optimization. If the optimized path and shortest path have to be identical m^{uv} has to be the minimum cost of all possible paths from u to v. Therefore, for each loopfree path p from u to v different from the optimized path a constraint has to be specified that forces the cost of the optimized path to be the smaller one:

$$m^{uv} < \sum_{ij \in p} m_{ij}, \tag{15}$$

where $ij \in p$ if path p runs over interface ij.
However, to include all such constraints for all paths would be too complex for greater networks. Therefore only the necessary constraints are added to the problem. They are identified by Algorithm 5.1.

6 Results

In this Section the results of the path optimization and the link cost specification are presented. They are demonstrated with three networks of different size. These networks are shown in Figure 6. The network on the left with only six routers was chosen because its size allows to depict the resulting paths. The networks with eight and fourteen routers were selected because of their complex structure which makes many different path choices possible. Hence, the routing optimization is rather complex. For all three networks between each two nodes a flow exists. Consequently the flow matrix is filled with exception of the diagonal. The flow matrix for the six router network is given in Table 3. In Table 2 the link capacities for this network are shown. As physical

Algorithm 5.1 (Algorithm to identify the delay constraints)

variables:

c_{ui}	*capacity of link ui*
P_{uv}	*set of links used by flow uv*
M_{ij}	*cost of link ij*

algorithm:

foreach flow uv
 foreach router i with $c_{ui} > 0$ and $x_{ui}^{uv} = 0$
  if $\forall j \in P_{uv}, j \neq v : j \notin P_{iv}$ /* where $j \in P \Leftrightarrow \exists$ jk$\in P$ */
   add $\sum_{ij \in P_{uv}} m_{ij} < \sum_{ij \in P_{iv}} m_{ij} + m_{ui}$ to constraints

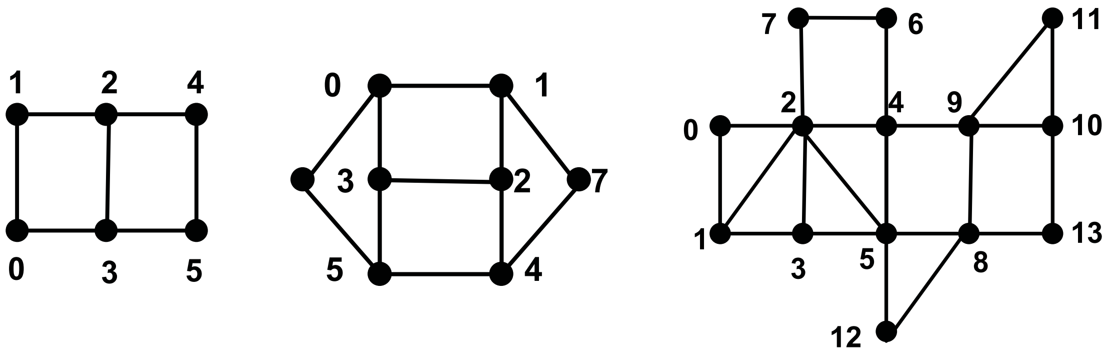

Figure 6: Example networks with six, eight and fourteen routers

delays the hopcounts were taken for all three networks such that for each link the physical delay d_{ij} was set to one. Furthermore, in every optimization run the parameter a_r was three. So the physical delay constraint does actually restrict the solution space hardly for the six and eight router network. Nevertheless, the routing optimization was done with the simplification of presolved variables as described above. For the fourteen router network 4714 of the 8008 variables were presolved and set to zero. Furthermore, the number of constraints was reduced to less than the half. Additionally, the path of the flows for neighbored routers was fixed, too.

For all networks optimizations with several parameter settings were performed. First, the optimization was performed without an effective upper bound for the link utilizations. This was achieved by omitting the constraints that kept the link utilizations below the value of t and setting the parameter a_c to a value of 10. Thus, the load of the links was permitted to be ten times greater than the link's capacity.

The resulting routing scenario is depicted in Figure 7. The link with the highest utilization of 42.9% is $4-3$. The value received for the minimized average utilization is 22.4%. As described above the value of parameter a_c was then repeatedly decreased

	0	1	2	3	4	5
0	0	120	0	100	0	0
1	110	0	80	0	0	0
2	0	90	0	100	0	70
3	130	0	80	0	90	0
4	0	0	0	70	0	120
5	0	0	75	0	125	0

Table 2: Capacity Matrix for the six router network

	0	1	2	3	4	5
0	0	5	6	7	4	7
1	4	0	6	4	8	5
2	6	8	0	4	6	4
3	8	9	4	0	5	6
4	5	7	8	5	0	3
5	5	7	6	8	4	0

Table 3: Flow matrix for the six router network

below the previous obtained utilizations. The settings for a_c were 0.4, 0.375 and 0.36. The result for $a_c = 0.36$ was identical with the result of the default optimization.
In the default optimization the parameter a_t was set to 1000. With this setting the maximum link utilization was given much more importance than the average utilization. The result of this optimization for the six router network is shown in Figure 8. Here the most utilized link is still $4-3$, however it's utilization was decreased to 35.7%. As a compensation, the average utilization increased to a value of 22.7%.

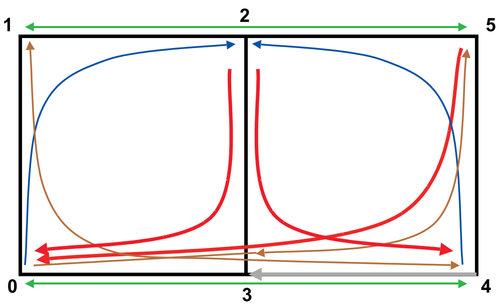

Figure 7: Routing with minimized average utilization

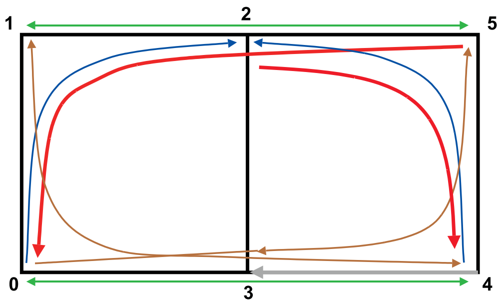

Figure 8: Routing with minimized maximum utilization

In Figures 7 and 8 the thick red lines show the paths that have changed with the reduction of the maximum link utilization. We see that flow $5-0$ was taken away from link $4-3$ and is now routed over routers 2 and 1. Consequently, the path of flow $2-0$ was changed as well. With only these changes link $3-4$ would have been utilized with 40%, so flow $2-4$ was routed over router 5 instead of router 3.
With the second linear program the link costs as small as possible were specified such that the shortest path routing is equivalent to the routing defined by the first linear program. In Figures 9 and 10 the network with the received costs is depicted. The costs are drawn in the gray boxes.
Additionally, the pictures show the path for flow $5-0$ represented with the green arrow

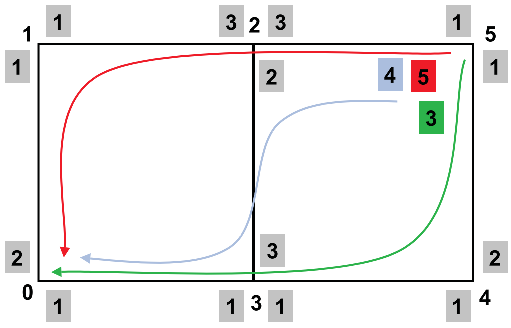

Figure 9: Link costs with minimized average utilization

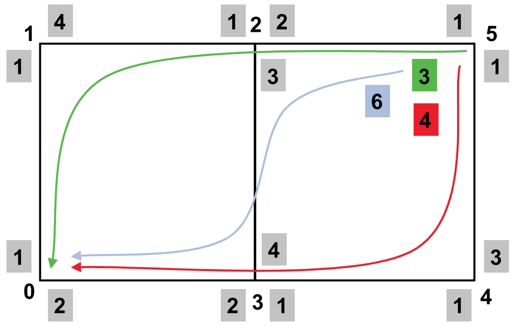

Figure 10: Link costs with minimized maximum utilization

and the other possible paths with a light blue and a red arrow. The boxes show the total link costs of the path. The color of the arrow representing the path and the color of the box with the delay belong together. We can see that in Figure 9 the total cost for flow $5-0$ is 3. The other possible paths are either over routers 2 and 1 or over routers 2 and 3. The total delays of these paths are 5 and 4 respectively. Thus, the shortest path is the same as the desired path. This is also true for the other flows.
In Figure 10 the delay for the path of flow $5-0$ is again 3. The delays of the other possible paths over 4 and 3 or 2 and 3 are 4 and 6 respectively. The shortest path is here identical with the desired path as well.

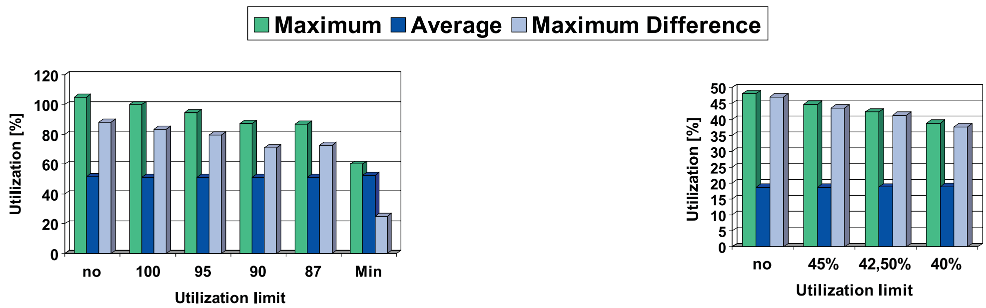

Figure 11: Utilization changes for the 8 router network

Figure 12: Utilization changes for the 14 router network

For the other two larger networks the procedure was similar. First, the average utilization was minimized without restriction for the maximum utilization. Then the maximum utilization was decreased iteratively. At the end the optimization was performed with the objective to obtain the minimal possible maximum utilization. However, the network with 14 routers proved to be too complex to achieve an optimal value. The maximum utilization was reduced to 38.8%. The minimal value for the maximum utilization is not known. However, it has to be greater than 32% since for this value the linear problem proved to be infeasible.

In Figures 11 and 12 the results for the path optimizations with the chosen upper bound for the link utilization can be seen. The green bars show the maximum utilizations and the dark blue bars the average utilization. The maximum utilization is decreased conspicuously, whereas the average utilization stays almost unchanged. The bright blue lines show the difference between the utilizations of the most and the least utilized link. This difference indicates whether the traffic is evenly distributed over the network or not. The graphs show that the traffic distribution becomes more homogeneous with a stricter upper bound for the link utilizations. In particular for the eight-router network the difference is reduced to almost a quarter. For the fourteen-router network the effect is less clear as there exist links with an utilization of only about 1% independent of the upper bound.

Nevertheless, since the highest utilization was decreased some utilizations have to be increased as well. This can be seen in Figure 13. The utilization of selected links is shown for the different upper bounds. We can see that the utilization of the links that are initially highly utilized is decreasing with a reduction of the upper bound, like with links $9-4$, $4-9$ and $4-2$. And as compensation the utilization of links $8-5$, $5-8$ and $3-1$ is increasing.

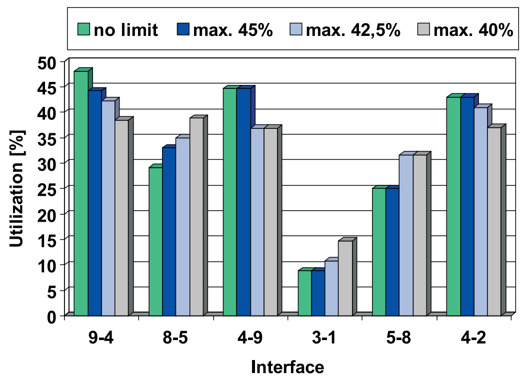

Figure 13: Selected interfaces of the 14 router network

7 Conclusions

In the last section we presented results for networks of different size. Though the optimum for the largest network could not be identified with the standard linear program solver CPLEX, at least a near optimal solution was found. If we look at the structure of the fourteen router network one has to notice that the structure is quite complex with regard to routing optimization. Between most pairs of nodes a lot of possible paths with the same or a similar physical delay exist. Of course actual intranets mostly comprise more that fourteen routers. However, in many cases a main part of the network can be made out which is relevant for the routing and consists of a number of routers

which is in a magnitude the linear programs are able to deal with. Nevertheless, some effort still has to be invested to reduce the problem's complexity further or to identify better problem solving algorithms.
The major issue of this work was to find a possibility to optimize the routing for existing routing protocols. It was shown that the results obtained by the two linear programs are applicable in the currently most important routing protocols OSPF and EIGRP.
The major lack of this approach is that only the instantaneous or peak traffic can be handled whereas the IP traffic is actually time-dependent. Another problem is to quantify the quality of a found routing decision. In this work two routing decisions were compared by the average and maximum link utilizations. However, the effects on the actual transmission delays of the packets in the network are not compared. As well, the influence of this delays on TCP connections has to be considered.

References

[1] G.R. Ash, R.H. Cardwell, R.P. Murray, "Design and Optimisation of Networks with Dynamic Routing", Bell System Technical Journal, Vol. 60, No.8, Oct. 1981.

[2] G.R. Ash, A.H. Kafker, K.R. Krishnan, "Servicing and Real Time Control of Networks with Dynamic Routing", Bell System Technical Journal, Vol. 60, No.8, Oct. 1981.

[3] Barahona F., "Network Design Using Cut Inequalities", SIAM J. Optimization

[4] D. Bertsekas, R. Gallager, "Data Networks", Second Edition, Prentice Hall, 1992

[5] E. Crawley, R. Nair, B. Rajagopalan, H. Sandick "A Framework for QoS-based Routing in the Internet" RFC2386, Aug, 1998

[6] R. Harris, S. Köhler, "Possibilities for QoS in existing Internet routing Protocols", University of Würzburg, Research Report Series, Report No. 228, May, 1999

[7] C. Huitema, "Routing in the Internet", Prentice Hall, 1995

[8] D. Lorenz, A. Orda, "QoS Routing in Networks with Uncertain Parameters" Proc. of INFOCOM'98, 1998

[9] Magnanti, T. et al, "Modeling and Solving the Two-Facility Capacity Network Loading Problem", Operations Research, Vol.43, No.1, Jan-Feb, 1995

[10] J. Moy, "OSPF Version 2" RFC 2328, April, 1998

[11] G. Apostolopoulos et al, "QoS Routing Mechanisms and OSPF Extensions", draft-guerin-qos-routing-ospf-04, December, 1998