Photonic Network Communications (2006) 11:49–64
DOI 10.1007/s11107-006-5323-1

ORIGINAL ARTICLE

On planning of optical networks and representation of their uncertain input parameters

Sofie Verbrugge · Didier Colle · Sophie De Maesschalck · Mario Pickavet · Piet Demeester

Received: 28 June 2004 / Revised: 12 July 2005 / Accepted: 20 July 2005

Abstract This paper studies the use of uncertain inputs in the strategic network planning process. To model uncertain planning inputs three essential parameters are needed: the predicted value expressing for instance an expert's view, the uncertainty level indicating the doubt there is about the predicted value, and the confidence parameter denoting the probability that the output parameter was estimated big enough (compared to the actual output). Several planning approaches that handle uncertain variables are distinguished and their strengths and shortcomings are indicated. This allows to indicate the pitfalls in some common planning practices that use a fixed safety margin to handle uncertainty. It is shown that they can lead to incorrect planning decisions, such as underestimation of the impact of the input uncertainty and overdimensioning in case of inaccurately modelled dimensioning problems. Both a theoretic model and simulation results are shown. A real-life planning problem is studied, including forecasting future network traffic from uncertain inputs and dimensioning a network to accommodate an uncertain traffic demand.

Keywords Optical network planning · Uncertain inputs · Traffic prediction · Robustness against demand uncertainty

S. Verbrugge (✉) · D. Colle · S. De Maesschalck · M. Pickavet · P. Demeester
Department of Information Technology (INTEC), Ghent University—IMEC—IBBT, Gaston Crommenlaan 8 bus 201, B-9050 Gent, Belgium
e-mail: sofie.verbrugge@intec.ugent.be

Strategic planning of today's optical telecommunication networks

The network planning process and the changing environment

The telecommunication landscape has undergone dramatic changes in recent years, especially due to the advances in optical networking. The number of customers grows, as well as the required bandwidth per user, causing the so-called *market pull*. This effect is very important for Internet traffic [1, 2] and is not expected to stop shortly [3]. Moreover, the operator's cost to be able to accommodate this demand (equipment cost, etc.) is subject to dynamic changes over time as a consequence of the *technology push.* The latter is apparent from major technological breakthroughs. We are witnessing an important evolution in optical networking technology, moving from point-to-point WDM transmission towards all-optical networks using Optical Path Cross-Connects [4]. Finally, the liberalization in most countries has completely changed the telecommunication environment, confronting network operators with fierce competition [5, 6]. In this dynamically changing environment, intelligent network planning has even become more important than it used to be [7, 8]. Figure 1 gives a graphical overview of the strategic planning process. The overall goal of the process is to provide a network deployment plan optimising the profit of the network operator (trade-off between the expected revenues generated by the customers and the expected costs to realize the network), while taking into account the indirect customer requirements concerning survivability, etc [9–11].

Especially when planning over a longer time horizon, the uncertainty of the considered planning data becomes important. Forecasting future customer demand is a difficult problem [12–14] with an associated unavoidable forecast error,

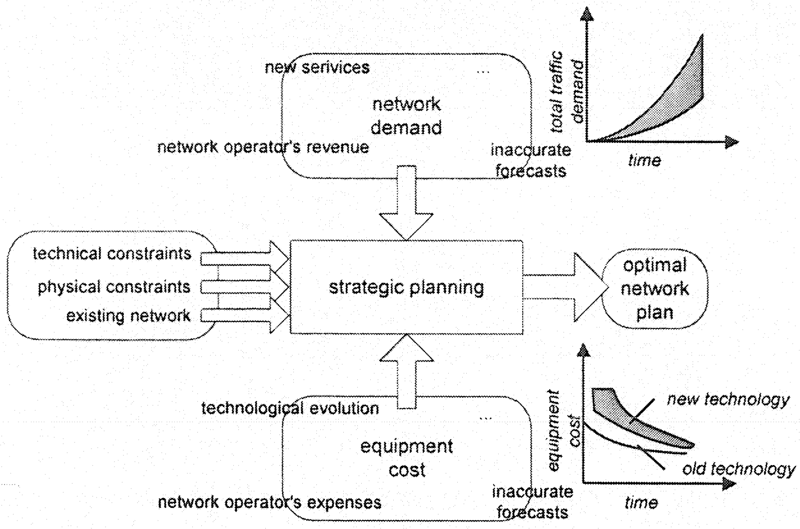


Fig. 1 Strategic planning process

typically increasing exponentially with the horizonal length. Forecast unreliability is problematic to operators, as they are—legitimately—concerned that the investments in network capacity will not match the actual requirements. Insufficient dimensioning may cause a loss of revenue because of displeased customers and penalty costs for unsatisfied Service Level Agreements (SLA). Overdimensioning, on the other hand, results in unused capacity that does not generate any revenue. Using optical technology, the equipment cost per unit is big, e.g., the cost of an OXC is enormous compared to the cost of an IP router line card (a ratio of *40/1* is common). As a consequence, investment decisions for optical networks, dealing with important traffic amounts, are critical for network operators. Understanding the impact of planning input uncertainty is therefore especially relevant in the context of optical networking.

Paper outline and related work

A lot of attention has been paid in the literature to the problem of network planning with uncertain inputs. The fundamentals of the network planning problem were described by Gupta [15] and King [16]. Results of multiple case studies were published: planning approaches using stochastic programming [17, 18], the use of sensitivity analysis when considering uncertain inputs [19] and network dimensioning under traffic uncertainty [20–23]. Furthermore, consulting companies emphasized the importance of handling uncertainty in strategic decisions [24, 25].

In this paper, we focus on the evaluation of some widespread planning practices. The use of a safety margin to handle uncertainty is compared to the representation of uncertain planning inputs as random variables. In Section 'Uncertainity modelling approaches' of the paper, the considered uncertainty handling methods and their parameters are introduced. The safety margin is studied in more detail and a reference scenario is discussed, which allows to determine a useful safety margin value. Section 'Pratical use' compares this reference scenario to what actually happens in practice, revealing some pitfalls of the use of safety margins in real-life situations. Section 'Real-life planning problems using uncertain variables' covers a realistic case study. As a starting point, a traffic model with uncertain inputs is considered. In a second stage, the uncertain traffic forecasts resulting from this model are used for the dimensioning of a pan-European IP-over-Optical network. The last section ends the paper with some conclusions.

Uncertainty modelling approaches

In this section, we indicate where uncertainty comes into play in the long-term network planning process and explain the approach we follow to model this uncertainty.

Our model

Roughly speaking, an uncertain variable consists of a predicted value (sharp number) and some description of the inherent uncertainty. The predicted value of a planning input variable (future demand, cost, ...) will for instance be obtained by hiring an expert for the considered planning domain. However, the real future value of the considered planning input will probably not equal the predicted value, even if made by the best expert. An uncertain variable therefore has some inherent uncertainty which should be taken into account during the planning process and can be represented in several ways (see Fig. 2). A-priori and a-posteriori adjustment are straightforward approaches: they make use of a so-called safety margin, added to the predicted values before or after the planning calculations. The *probabilistic approach* is more formal as it uses probability distributions to model the uncertain character. All parameters needed to describe our model are introduced below.

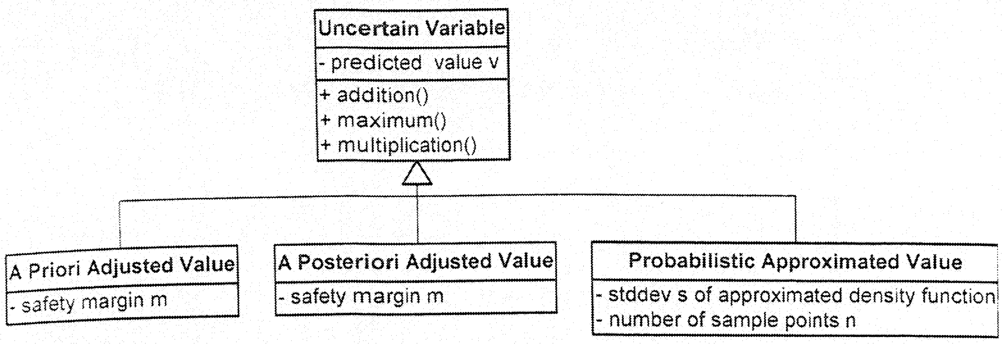


Fig. 2 Uncertain variable: Proposed models

- The *predicted value* v is the expected value for the uncertain variable under consideration. It is the forecasted value proposed by an expert, an extrapolation model, or the average of multiple forecasted values proposed by several experts. For instance, the predicted value for next year's traffic may be obtained by an extrapolation of the traffic measured during previous years and the current year.
- In the a-priori and a-posteriori adjustment method, a relative *safety margin m* is used to take the uncertain character of the forecasts into account in a very straightforward way. In a-priori adjustment, a margin is added to the predicted value of every input variable before the start of the calculations. It indicates the uncertainty inherent to these inputs and may be different for every input variable. For instance, we could add a safety margin of 10% to all demand matrix entries so that they represent some kind of upper limit on the expected traffic. In a-posteriori adjustment, on the other hand, the margin is added to the sharp calculated result at the end of all calculations: only a single safety margin is used, independently of the number of inputs. For instance, if the predicted traffic for next year on a certain network link is 3000 Mbps, one might take a safety margin of 10% and thus decide to foresee 3300 Mbps to be able to cope with the real demand that might be bigger than expected.[1] As an illustration, the results found by a-priori or a-posteriori adjustment for some operations are given in Table 1. Remark that even when using a-posteriori adjustment the intention of the safety margin is still to incorporate the uncertain character of the input variables. The a-posteriori added margin should show how the input uncertainty is reflected in the output variables. It is possible to indicate the a-posteriori margin m as a function of the a-priori margins m_1 and m_2, giving the same result, e.g., for the sum of two uncertain variables: $m = (m_1 \cdot v_1 + m_2 \cdot v_2)/(v_1 + v_2)$. If the predicted values of both inputs are equal to each other, the appropriate a-posteriori margin equals the mean of the a-priori margins.
- In the probabilistic model, an uncertain variable is represented as a random variable with an associated probability distribution. The predicted value v is used as the mean value for this distribution. The standard deviation of a probability distribution can be seen as an indication of the uncertainty associated with this mean value. As it is likely that this uncertainty will grow with the magnitude of the forecasted value, the standard deviation s is chosen to be a percentage of the predicted value: $s = p \cdot v/100$ with p being the *percentual standard deviation*. In this model, the operations on uncertain variables are performed using the common formulae for continuous random variables on piecewise linear approximations of the considered density functions. Piecewise linear functions allow to approximate any probability distribution, even if it cannot easily be described analytically. For scalability reasons,

Table 1 A-priori and a-posteriori adjustment methods

	A-priori adjustment	A-posteriori adjustment
Operand 1	Aprio (v_1, m_1)	Apost (v_1, m)
Operand 2	Aprio (v_2, m_2)	Apost (v_2, m)
Sum	$v_1 \frac{(100+m_1)}{100} + v_2 \frac{(100+m_2)}{100}$	$(v_1 + v_2) \frac{(100+m)}{100}$
Maximum	$MAX\left[v_1 \frac{(100+m_1)}{100}, v_2 \frac{(100+m_2)}{100}\right]$	$MAX\,[v_1, v_2]\, \frac{(100+m)}{100}$
Product	$v_1 \frac{(100+m_1)}{100} \cdot v_2 \frac{(100+m_2)}{100}$	$(v_1 \cdot v_2) \frac{(100+m)}{100}$

[1] From this example it is clear that the safety margins we consider in this paper always indicate positive margins: we add a safety margin m to the sharp calculated result v, so that the actual result (taking into account uncertainty) equals $v+m$. For the planning problems considered here, a negative safety margin (resulting in $v-m$) would be meaningless. However, cases where negative margins would make sense (e.g., when considering the minimal demand needed for the minimal expected revenue) could be studied in a completely similar way.

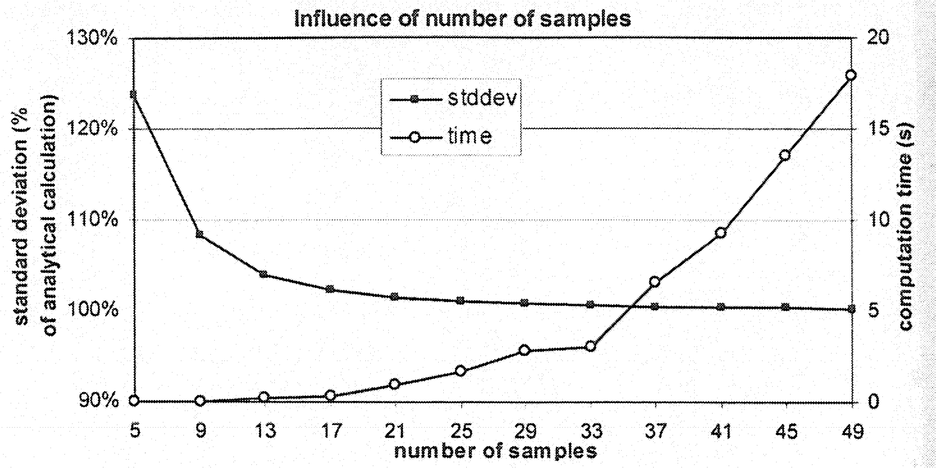


Fig. 3 Influence of number of samples (n) on results found for the sum of probabilistic approximated uncertain variables ($v_1 = 10$, $v_2 = 1000$, $p = 10\%$)

the result of each operation is resampled to have the same number of 'pieces' (=number of samples −1) as its operands. The accuracy of the approximation grows with the amount of samples, as illustrated in Fig. 3 for the standard deviation of a sum (where the exact result can easily be determined analytically). A complexity analysis showed that the computational complexity of the addition of probabilistic approximated variables grows with the third power of the number of samples, which is also apparent in the figure. In contrast to the traditional discrete sampling approach, our model uses continuous probability distributions and imposes lower memory requirements (see the Appendixs for more details on this model).

- The last important parameter is the so-called *confidence parameter* c. For example, if a network planner is interested in the amount of capacity that has to be foreseen to accommodate the future traffic on the network with some (given) chance, this chance is indicated by the confidence parameter. Mathematically speaking, we want to indicate a limit value y_{limit} that will be greater than or equal to the real future outcome of the uncertain value y with a probability c:

$$y_{\text{limit}} = y_{\text{predicted}} + y_{\text{margin}}$$
$$Pr[y \leq y_{\text{limit}}] = c$$

where y_{limit} is the desired limit value, obtained as the end result of planning calculations using one of the proposed uncertainty models, $y_{\text{predicted}}$ is the predicted value of the result obtained by neglecting uncertainty and y_{margin} is determined by the used uncertainty model. It can be different for all three implementations (a-priori, a-posteriori adjustment and probabilistic adjustment). The percentual confidence parameter value indicates the confidence interval, e.g., if we want a network link to be robust to uncertain traffic demand with a probability of 95%, c is set to 95%. Note that this can be an important factor in SLAs. The penalty costs to be paid because of an underdimensioned network can be enormous. Therefore, it is very important to be prepared for uncertain traffic changes.

How to choose an appropriate safety sargin value

A-priori and a-posteriori adjustment are popular planning approaches because of their ease of use (both conceptual and computational). When using those methods however, it is crucial to determine a suitable safety margin value. Below, we indicate a possible way of working to obtain a safety margin that allows to approximate the analytical value as closely as possible.

We find a *reference scenario* in the addition of two uncertain planning inputs x_1 and x_2, represented by normally distributed variables. They have v_1 and v_2 as mean values, respectively. The percentual standard deviation p is equal for both. We know from probability theory that their sum y is normally distributed with as standard deviation the square root of the sum of the variance of both inputs. From the 68%, 95% and 99.7% -rules for the normal distribution (Fig. 4), it is clear that approximately 68% of the observations fall within one standard deviation from the mean.

For a confidence parameter of 84%, the limit value y_{limit}^{2} can thus be expressed analytically by

$$y_{\text{limit}} = y_{\text{predicted}} + y_{\text{margin}}$$
$$\Pr[y \leq y_{\text{limit}}] = \Pr[y \leq v_y + s_y] = 84\%$$
$$y_{\text{margin}} = s_y = \sqrt{(pv_1/100)^2 + (pv_2/100)^2}$$

where

$$x_1 : N(v_1, s_1) = N(v_1, pv_1/100)$$
$$x_2 : N(v_2, s_2) = N(v_2, pv_2/100)$$
$$y : N(v_y, s_y) = N(v_1 + v_2, \sqrt{(pv_1/100)^2 + (pv_2/100)^2})$$

[2] Remark that we focus on the limit value y_{limit} because this value has most practical importance: it is used in planning decisions where it indicates for instance actual link capacities to foresee for the future.

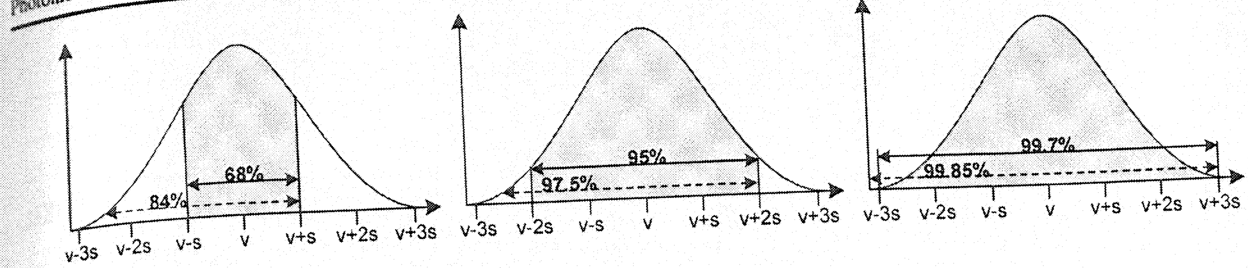


Fig. 4 The 68, 95 and 99.7% rules for normal distribution

$x : N(v, s)$ denotes a normal distributed random variable x with mean value v and standard deviation s. As long as the safety margins considered in a-priori and a-posteriori adjustment are equal, both approaches yield an equal sum. In this case, $y = x_1 + x_2 = \text{Aprio}(v_1, m) + \text{Aprio}(v_2, m) = \text{Apost}(v_1, m) + \text{Apost}(v_2, m)$. We recall the following formulae from above:

$$y_{\text{limit}} = y_{\text{predicted}} + y_{\text{margin}}$$
$$\Pr[y \le y_{\text{limit}}] = \Pr[y \le (v_1 + v_2)(100 + m)/100] = c$$
$$y_{\text{margin}} = (v_1 + v_2)m/100.$$

Identification of y_{margin} calculated analytically and y_{margin} found by a-priori or a-posteriori adjusted values leads to the following expression for the safety margin m as a function of the percentual standard deviation p, the relative magnitude of the predicted values $r = v_1/v_2$ and the confidence parameter c : $m(p, r, c)$.

$$(v_1+v_2)^* m(p, r, 84\%)/100 = \sqrt{(pv_1/100)^2 + (pv_2/100)^2}$$
$$m(p, r, 84\%) = \sqrt{(pv_1)^2 + (pv_2)^2}/(v_1 + v_2)$$
$$m(p, r, 84\%) = p\sqrt{r^2 + 1}/(r + 1)$$

This last equation learns that the sum of normally distributed variables (where the standard deviation is a percentage of the mean value) can perfectly be approximated by a-priori or a-posteriori adjustment if the confidence parameter equals 84%. The needed safety margin is proportional to the percentual standard deviation p and the proportion factor is a function of r, the ratio between the operands' magnitudes. For the special case where both operands are equal ($v_1 = v_2$, $r = 1$) this leads to $m(p, 1, 84\%) = p/\sqrt{2}$. If one operand is infinitely small compared to the other ($v_1 \gg v_2$, $r \gg 1$), we find $m(p, +\infty, 84\%) = p$. Note that exactly the same result is obtained when $v_1 \ll v_2$, ($r = 0$), $m(p, 0, 84\%) = p$. In what follows we will always assume that $v_1 > v_2$.

Following the 68%, 95% and 99.7%-rules we can make a similar calculation for a confidence parameter of 97.5% or 99.85%, resulting in $m(p, 1, 97.5\%) = p \cdot \sqrt{2}$, $m(p, +\infty, 97.5\%) = 2p$, in $m(p, 1, 99.85\%) = 3p/\sqrt{2}$ and $m(p, +\infty, 99.85\%) = 3p$. A summary of the obtained results can be found in Fig. 5.

To determine the safety margin for other confidence parameters, we need numeric values for the cumulative distribution of normal distributed variables. This information is widely spread in the form of distribution tables for the normalised $x : N(0, 1)$ -distribution [26]. Note that via the transformation $z = (x - v)/s$ these tables provide information for all normal distributions $z : N(v, s)$.

General observations concerning the safety margin

In this paragraph, some general properties of the margins $m(p, r, c)$ obtained from cumulative distributions (as indicated in the previous paragraph) are discussed.

The influence of the ratio r of the predicted values on the choice of the safety margin value is examined in Fig. 6, which plots the limit value y_{limit} for the sum found by a-priori adjustment relative to the analytical value (analytical value is shown by the 100%-line in the figure) for a confidence parameters of 84% and 97.5%. Leaving all the rest unchanged, a growing ratio r of the operands causes the appropriate safety margin to decrease (for small values of the ratio), as could be expected by the obtained formulae for $m(p, r, c)$. Remark that unexpected results can be obtained if the wrong m-value is used. For instance, naively setting the safety margin to $p \cdot \sqrt{2}$ when adding uncertain future traffic demands in a dimensioning problem where the confidence parameter equals 97.5 %, may result in serious capacity shortage if the actual ratio of the predicted values turns out to be 100 or more (which might be the demand ratio between a busy and a quiet link).

In Fig. 6 the distance between the *lines* $m(p, 1, c)$ and $m(p, +\infty, c)$, and thus the difference between the appropriate safety margins for operands in the same order of magnitude and in a completely different order of magnitude, grows with an increasing confidence parameter c (p is fixed). An intelligent choice (taking into account the ratio of the parameters) for the safety margin is thus even more important when the confidence parameter grows. Remark in this context that confidence parameters in the range of 95% (or higher) are common in realistic planning problems.

So far we have considered the addition of only two uncertain variables. It is known, however, that the relative standard deviation (std dev/mean) of the sum of positive uncertain variables decreases as we add more variables. As a consequence, the appropriate safety margin value decreases with the number of operands. In real network dimensioning prob-

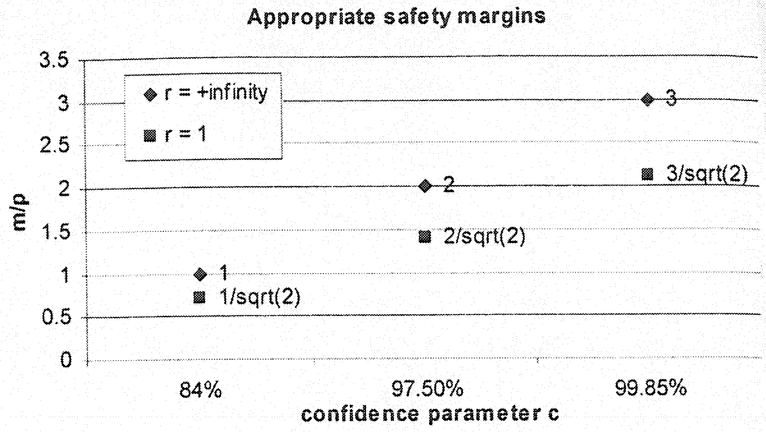


Fig. 5 Appropriate safety margin value (m) as a function of the confidence parameter (c)

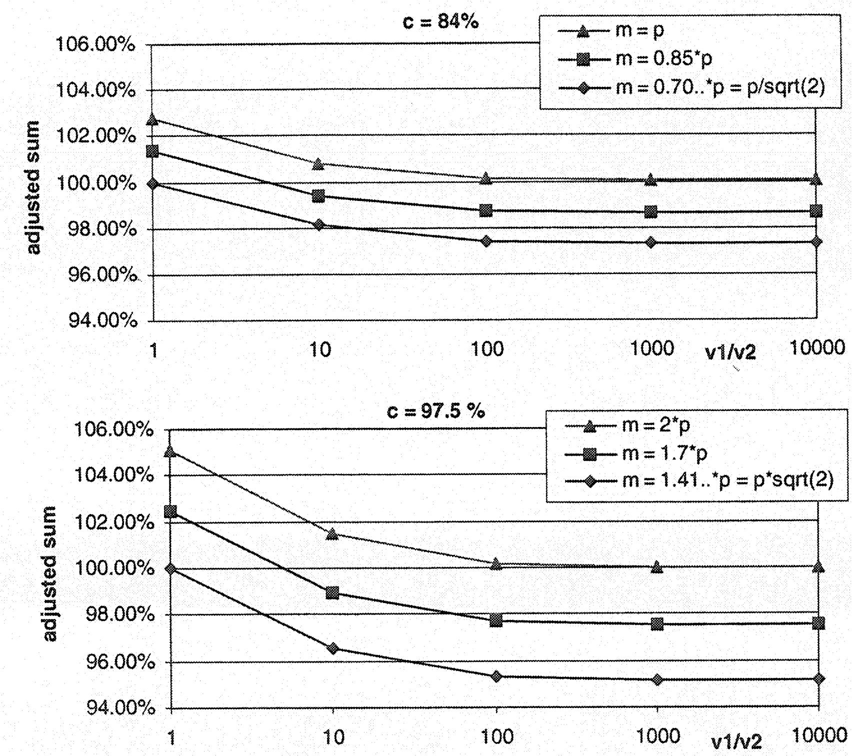


Fig. 6 Influence of the predicted values ratio (r)

lems, multiple (>2) uncertain demands might be added. Using the safety margin for two values will in this case lead to results that are much larger than the analytically calculated ones. This problem can be overcome by adding the variables two-by-two and recalculating the appropriate margin in every step or by adjusting the margin in the following way (for the sum of n uncertain variables with equal predicted values v and percentual standard deviations p)

$$\sqrt{n(vp/100)^2} = nv * m(p,\ 1,\ 84\%)/100$$
$$p/\sqrt{n} = m(p,\ 1,\ 84\%).$$

Practical use

So far we have shown how a safety margin value can be determined for the a-priori or a-posteriori adjustment approach which allows to approximate the analytical results. But what happens if the result of the operation under study no longer has a known distribution (Gaussian in the considered case)? If there is no information available concerning the cumulative distribution of the expected result, a safety margin value cannot be obtained in the way indicated by the reference scenario.

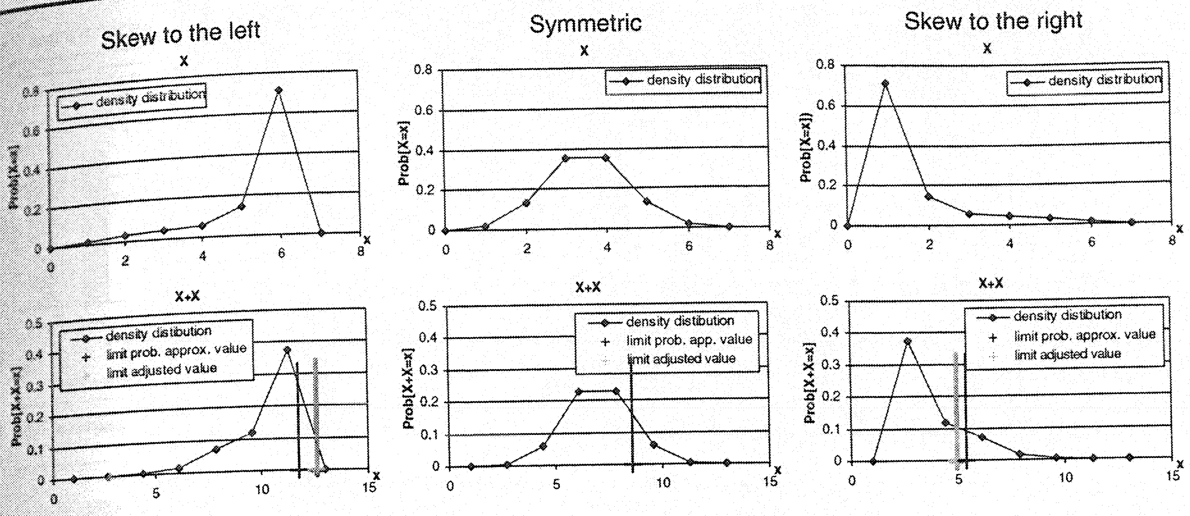


Fig. 7 Influence of the skewness of the probability distribution on the performance of the models

Why the reference scenario cannot always be followed

First of all, the inputs can cause the result not to be normally distributed anymore. Figure 7 shows the influence of the input distribution skewness on the limit value y_{limit} for the sum of two equal uncertain variables. For a symmetric input distribution a-priori or a-posteriori adjustments give the same results as the approximated probabilistic model (when following the reference scenario approach). For a positively skewed distribution (tail to the right) however, the limit value obtained by a-priori or a-posteriori adjustment underestimates the actual limit value. For a negatively skewed distribution, the value is overestimated. The probabilistic approximated value technique is far best suited to model uncertainty in this case, because it calculates the limit value based on the area underneath the density curve.

Also the actual operation performed on the uncertain variables has its influence. If the input variables are represented as normally distributed variables, the sum will be normally distributed as well. Maximum and product will have different distributions. Furthermore, from the information of Table 1 it is clear that the sum and maximum of uncertain variables is equal for both a-priori and a-posteriori adjustment (as long as all considered safety margins are the same). The product obtained by a-priori adjustment, on the other hand, will always be m bigger than that obtained by a-posteriori adjustment:

$$\frac{\text{a-priori adjusted product}}{\text{a-posteriori adjusted product}} = \frac{\frac{v_1(100+m)}{100} \cdot \frac{v_2(100+m)}{100}}{\frac{v_1 \cdot v_2(100+m)}{100}} = \frac{100+m}{100}$$

Pitfalls of common practices

A-priori and a-posteriori adjustments are common practices today. Despite the insights of the previous paragraphs, however, the safety margin value is often chosen in an ad-hoc way. A particular value is chosen at the start of the study and then used throughout all calculations.

In Fig. 8, we illustrate that the use of a fixed safety margin value is a dangerous practice. It may, for example, lead to inaccurate dimensioning decisions if uncertain demands are involved. In the figure, the safety margin m is set to $p\sqrt{2}$, irrespective of the ratio of the operands r. The confidence parameter c equals 97.5%. The results are shown relative to the solution obtained by the probabilistic approximation case, which approximates the analytical value well: the discrepancy between analytical and approximated values obtained when adding two uncertain variables is always below 0.5%.

The uppermost part of Fig. 8 shows that, for equal operands, the adjusted sum (equal for a-priori and a-posteriori adjustments) is exactly the same as the sum found by probabilistic approximation, independently of the uncertainty level value. This is what we expected with the used value for the safety margin (reference scenario). When the operands are not equal anymore, there is a difference between both results which grows with the uncertainty level and with the ratio of the operands' magnitudes. The fact that the adjusted value is always smaller than the probabilistic approximated value can be understood from the influence of the ratio $r = v_1/v_2$. The bigger this ratio, the more the ideal value for m moves away from $p \cdot \sqrt{2}$ towards $2p$ and thus the bigger the fault made by choosing $m = p \cdot \sqrt{2}$. The fault grows linearly with the uncertainty level value.

The middle part of Fig. 8 studies the maximum of uncertain variables. The adjusted value (again, equal for a-priori and a-posteriori adjustments) is always smaller than the value

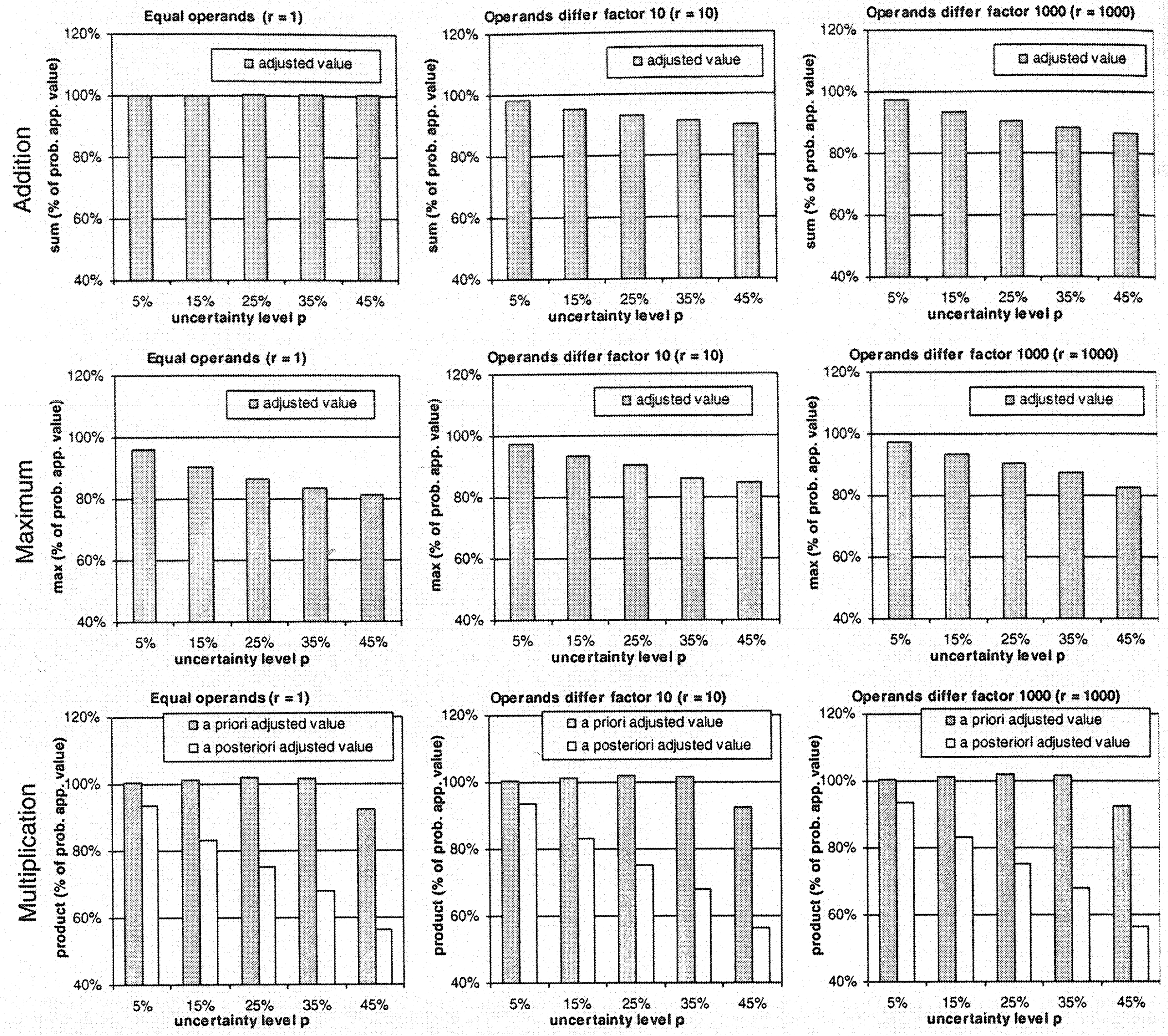


Fig. 8 Influence of uncertainty level ($p = s/v$) on sum, maximum, and product

found by probabilistic approximation, which is understood from the density curve of the maximum of two normally distributed variables. This curve will be higher on the right hand side, so that the abcissa of the limit value having 97.5% of the area underneath the curve to its left will be bigger than the corresponding abcissa on a Gaussian curve with the same mean value. This discrepancy between the adjusted value and the probabilistic approximated value grows with the uncertainty level. It decreases with the ratio of the operands.

The product of uncertain variables is given in the bottom part of Fig. 8. A-priori adjustment and probabilistic approximation lead to almost identical results for most considered cases. The inaccuracy observed for very large uncertainty levels (45%) could be avoided by using more sample points. However, in contrast to the case for addition and maximum, the operands' ratio does not influence the performance at all. The three bottommost parts of Fig. 8 are identical despite the fact that they represent different ratios of the predicted values. This is because the relative fault on the product of two variables always equals the sum of the relative faults of the operands, independently of their absolute values. The discrepancy between the a-priori and a-posteriori adjusted product is striking. This is the ratio $(100 + m)/100$ that was explained above. Because the chosen safety margin m is a linear function of the uncertainty level p, the discrepancy grows with this p. These observations make clear that a-posteriori adjustment will often not lead to the expected results, therefore planners should definitely prefer a-priori adjustment over a-posteriori adjustment when multiplicative operations are involved.

Probabilistic approximation is conceptually more complicated than a-priori or a-posteriori adjustment and has a higher computational complexity as well, but will always give a good approximation of the analytically calculated results. If there is no information available concerning the cumulative distribution of the expected results (and there-

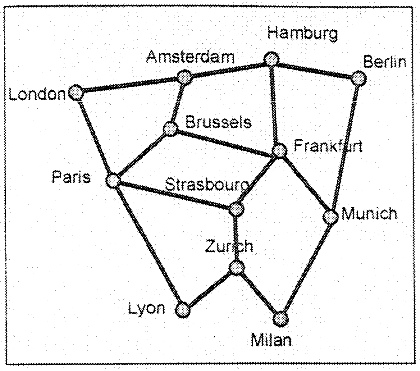


Fig. 9 Pan-European network topology under study

fore it is infeasible to determine an appropriate safety margin value for a-priori or a-posteriori adjustment), probabilistic approximation certainly is the best uncertainty handling technique.

Real-life planning problems using uncertain variables

In this section, a realistic planning problem is studied. Based on the uncertain traffic forecasts resulting from a traffic prediction model, the dimensioning of the pan-European IP-over-Optical network of Fig. 9 (similar to the topologies described in [27, 28]) is performed. This happens in two steps. Section 'Forecasting future traffic' determines the expected future traffic from a traffic model with uncertain inputs (based on the forecasts of [28]), Section 'Network dimensioning based on uncertain traffic predictions' performs the dimensioning of the network links based on the obtained traffic matrix.

Forecasting future traffic

The considered traffic model, proposed by Refs. [12–14],[3] distinguishes voice, transaction data, and IP traffic. It is based on the population, the number of non-production business employees, and the number of internet hosts in the considered cities, respectively. Moreover, the three traffic types have different distance dependencies. The total traffic is given as the sum of those three types, as shown in Fig. 10. In this figure, uncertain variables are shown in grey boxes, sharp values in white boxes. Consider for example the prediction of IP traffic between city i and city j. The number of hosts in those cities are estimations and thus the resulting amount of IP traffic will be an estimation as well. The planner's goal is to determine a realistic upper bound on the expected IP traffic.

The uncertain inputs are represented as uncertain variables and the upperbounds for the considered traffic types (and the total traffic) can be calculated according to the three uncertainty handling approaches. When using the probabilistic approximated case, uncertain variables are represented as normally distributed variables, approximated by a piecewise linear function with 19 sample points. Although this model can deal with all kinds of distributions, we use normal distributions for our planning inputs because we suppose them to be obtained from multiple experts' forecasts. According to the central limit theorem, the average of (infinitely) many forecasts provided by independent experts will in most cases be approximately normally distributed.

First, we consider the *immediate introduction of uncertainty*: the inputs of Fig. 10 are considered as uncertain variables. It is our impression that the uncertainty level associated with the three kinds of uncertain inputs is different. The population (P) for the next year in a certain city region is probably known rather accurately. The number of non-production business employees (E) might be more difficult to forecast and the prediction for the future number of hosts (H) is probably most uncertain. For this reason, we set different p-values for the different inputs: $p(P) = 5\%$, $p(E) = 10\%$, and $p(H) = 20\%$. For a confidence parameter of 97.5%, we have chosen the safety margin value as $m = p \cdot \sqrt{2}$. As we want to illustrate the practical use of the safety margin, the operands' ratio is not taken into account here (this is the so-called *use of fixed safety margin value*). Fig. 11 shows results for the traffic to and from a single source node (London). In the considered case, the forecast for the transaction data traffic is about four times as big as the voice traffic. IP traffic is approximately twice as big as voice traffic, because of the used reference information for the considered year 2003. Using realistic forecasts for future years would probably result in an important growth of the expected IP traffic. A-priori adjustment and probabilistic approximation give similar results, the used safety margin value was a good choice because the relative magnitude of the predicted values does not influence the choice of the safety margin in case of multiplication. This is clear for all three traffic types. A-posteriori adjustment results in smaller values, which was expected as well. The difference between the a-priori and the a-posteriori adjusted values equals $m\%$ (which was indicated theoretically before) and therefore is smallest for voice traffic, bigger for transaction data traffic and, the biggest for IP traffic (reflecting the growth in the uncertainty level of the input values for these traffic types).

The total expected traffic per link is obtained by summing up the above amounts of traffic for the three considered types (voice, transaction data, and IP traffic). We find the results in Fig. 12, indicated by the black points. The uncertainty has been introduced in the first step of the model (Fig. 10), which justifies the name *immediate introduction of uncertainty*. Remark that, from hereon, we will only consider a-priori adjusted values and probabilistic approximated values. A-posteriori adjustment of uncertain variables indeed can-

[3] In the presented simulation results, the formulae of [14] are used.

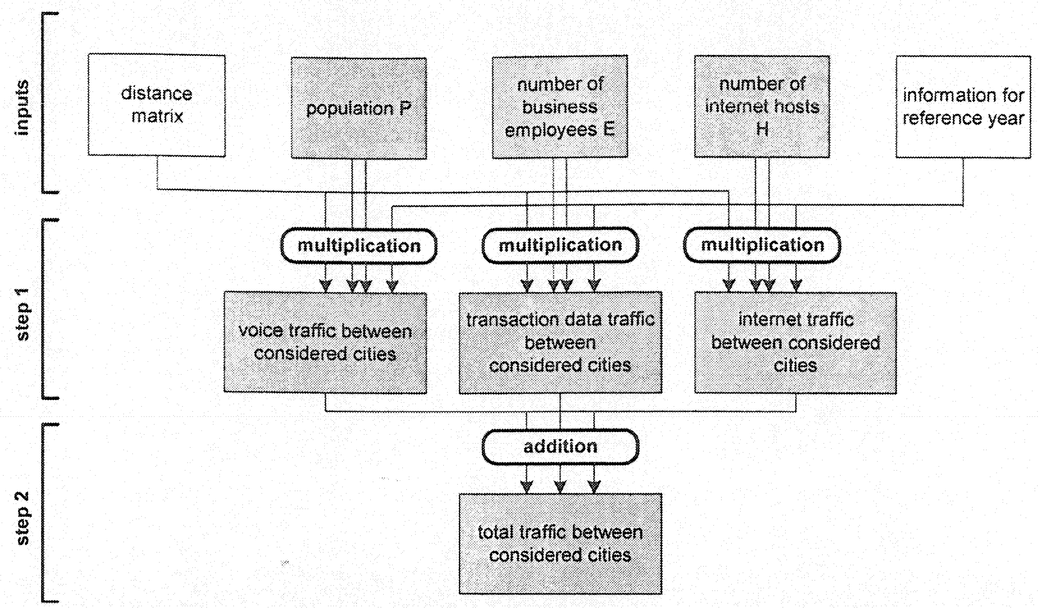


Fig. 10 Traffic prediction model proposed in [12–14]

not handle the different safety margins used for the different traffic types in a straightforward way (see remark in Section Our model). Moreover, for equal safety margins a-priori and a-posteriori adjustments would give the same result, indicated by the square black points.

It is apparent that a-priori adjustment and probabilistic approximation no longer give equal results, despite the chosen values for c and m. This example shows that it is very difficult to find a reliable safety margin in cases where several kinds of operations are involved and a wide range of predicted values is considered. Therefore, the a-priori adjustment method should be used with care in real-life situations.

In practice, it can happen that the network planner does not know the traffic prediction model (the first step of Fig. 10), but only disposes of its outcome (amount of voice, transaction data, and IP traffic). This means that he knows forecasted values for the three traffic types as sharp numbers without knowing where they come from. However, he still wants to incorporate the inherent uncertain character of those forecasts and therefore models the amounts of voice, transaction data, and IP traffic as uncertain inputs for the determination of the total amount of traffic. In other words, the *introduction of uncertainty is delayed* till the second step of the traffic model of Fig. 10. The transparent points of Fig. 12 show results for this case when $c = 97.5\%$, $p(voice\ traffic) = 5\%$, $p(transaction\ datatraffic) = 10\%$, $p(IPtraffic) = 20\%$ and $m = p \cdot \sqrt{2}$. The results obtained by a-priori adjustment and probabilistic approximation are similar. If the uncertainty level on all inputs would have been the same, both results would have been identical. When we compare the two ways of calculating the total expected traffic on our pan-European network (immediate introduction versus delayed introduction of uncertainty), we notice that delayed introduction leads to smaller limit values for the total traffic in the considered case. This can be explained using the information of Table 2. This table shows that the uncertain variables obtained for the different traffic types as intermediate results in case of immediate introduction of uncertainty have bigger percentual standard deviations than the values that are introduced in the case of delayed uncertainty introduction. A bigger standard deviation means a wider density curve and thus a bigger abscissa value for the limit having $c\%$ of the area below the density curve to its left. The method of delayed uncertainty introduction therefore may lead to an underestimation of the total future traffic.

Table 2 Values for relative standard deviation (p) of different traffic types

Traffic type	Std dev (obtained from step 1) in case of immediate uncertainty introduction	Std dev (introduced) in case of delayed uncertainty introduction
Voice	7.21%	5%
Transaction data	14.47%	10%
IP	29.27%	20%

Network dimensioning based on uncertain traffic predictions

Knowing the expected future traffic on the network, the network planner can take dimensioning decisions to

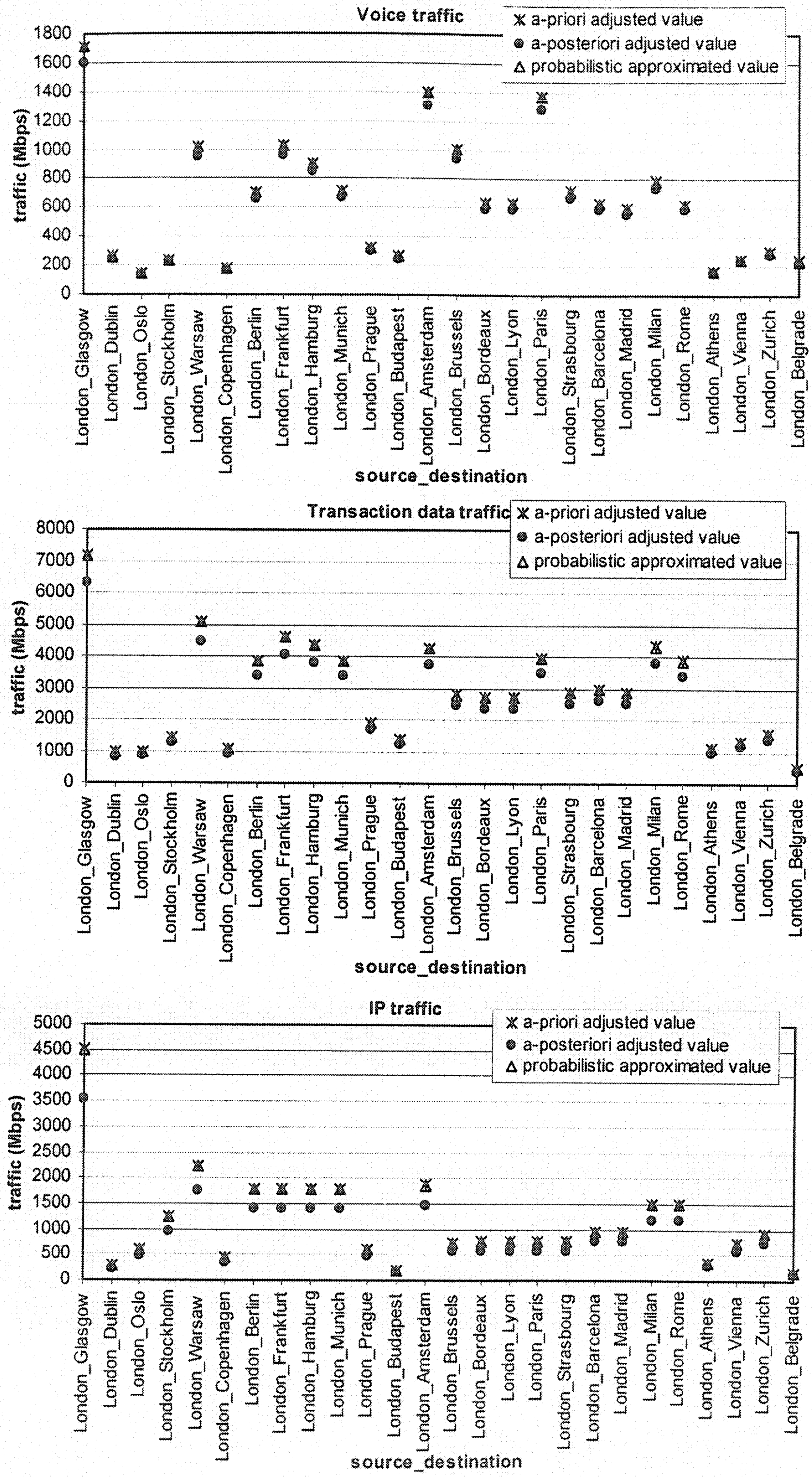


Fig. 11 Limit values (y_{limit}) for expected voice, transaction data, and IP traffic demand

accommodate this future traffic. Resilience questions come up at this point in time. Do we want to protect this traffic? Do all traffic types have to be protected in the same way? Also in this planning phase, uncertainty has to be taken into account. Below we study the dimensioning of the considered pan-European network under traffic uncertainty.

If we send all traffic unprotected, the total required capacity on each link l can be calculated as the sum of all voice, transaction data, and IP traffic over that link:

$$\text{required capacity } (l) = \sum_g [V_g : l \in P_g] + \sum_g [TD_g : l \in P_g] + \sum_g [IP_g : l \in P_g]$$

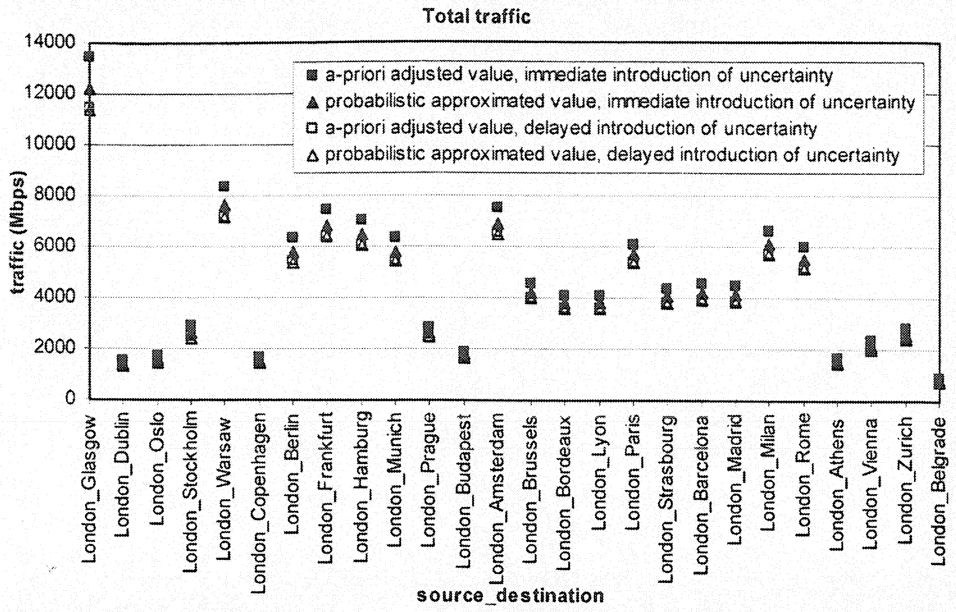


Fig. 12 Limit values (y_{limit}) for expected total traffic demand

with P_g being the path of the traffic for source-destination pair g; V_g, TD_g, IP_g, the amounts of voice, transaction data and IP traffic for source-destination pair g. The amounts of traffic are not known exactly, they are modelled as uncertain variables with an uncertainty level of 10%. The first part of Fig. 13 plots the results for a confidence parameter of 97.5%. The required capacity on every link is shown relative to the amount determined by the approximated probabilistic method. The latter method closely approximates the actual required capacity to be robust to an uncertain demand with a chance of 97.5%. For a-priori and a-posteriori adjustments several m-values are considered. The appropriate value for the addition of two uncertain variables in the considered case ($m = p \cdot \sqrt{2}$) leads to an important overprovisioning on most links, compared to the probabilistic case. This can be understood by the decreasing appropriate safety margin with a growing number of operands and the fact that the safety margin is not recalculated here for every single operation (fixed safety margin value). A better approximation of the 100% -line is obtained for smaller values of the safety margin (e.g., $m = p/\sqrt{2}$). On some links on the other hand (e.g., the link Vienna-Rome) we notice a capacity shortage, even for rather big values of the safety margin. In this regard, remark that a part of the capacity shortage could be solved by performing some form of traffic engineering (which is not considered here). It is clear from the observations in this paragraph that the usage of a fixed safety margin to incorporate uncertainty leads to unexpected dimensioning decisions.

Another approach is to differentiate the protection schemes for the considered traffic types, as they may have different survivability requirements. Voice and transaction data traffic can be protected by a pre-calculated link restoration scheme, while IP traffic is sent unprotected. Restoration is especially important in optical networks. Because of the huge traffic amounts a single failure may have a dramatic impact, while the cost of spare resources (optical equipment) is substantial. Sharing resources among several paths is a good option in this case. The differentiated restoration scheme, we use here implies that for IP traffic the formula from the previous section still holds, while for voice and transaction data traffic, the value stated there has to be augmented by the maximum of all traffic that can be affected by a single link failure, resulting in

$$\begin{aligned}&\text{required capacity } (l)\\ &= \textstyle\sum_g [V_g : l \in P_g] + \underset{k \neq l}{\text{MAX}}\left[\textstyle\sum_g [V_g : k \in P_g \wedge l \in S_g]\right]\\ &\quad + \textstyle\sum_g [\text{TD}_g : l \in P_g] + \underset{k \neq l}{\text{MAX}}\Big[\textstyle\sum_g [\text{TD}_g : k \in P_g \wedge\\ &\qquad l \in S_g]\Big] + \textstyle\sum_g [IP_g : l \in P_g]\end{aligned}$$

with

P_g and S_g being the working and backup path for the traffic for source-destination pair g; V_g, TD_g, IP_g, the amounts of voice, transaction data, and IP traffic for source-destination pair g.

The dimensioning decisions obtained for an uncertainty level of 10% associated with all traffic types (V, TD, IP) and a confidence parameter of 97.5% are shown in the second part of Fig. 13. In this case the overdimensioning compared to the probabilistic case is even worse than it was for unprotected traffic.[4] A network operator following a-priori

[4] Note that there can be some confusion concerning the word *overdimensioning*. With overdimensioning in the context of traffic uncertainty we mean that there is more capacity installed that what is needed to fulfill the requirements set by the confidence parameter (to be robust to changes in the traffic with a probability of 97.5%). The actual amount of capacity to be foreseen to fulfill this requirement is indicated by the approximated probabilistic case. Overdimensioning in the context of

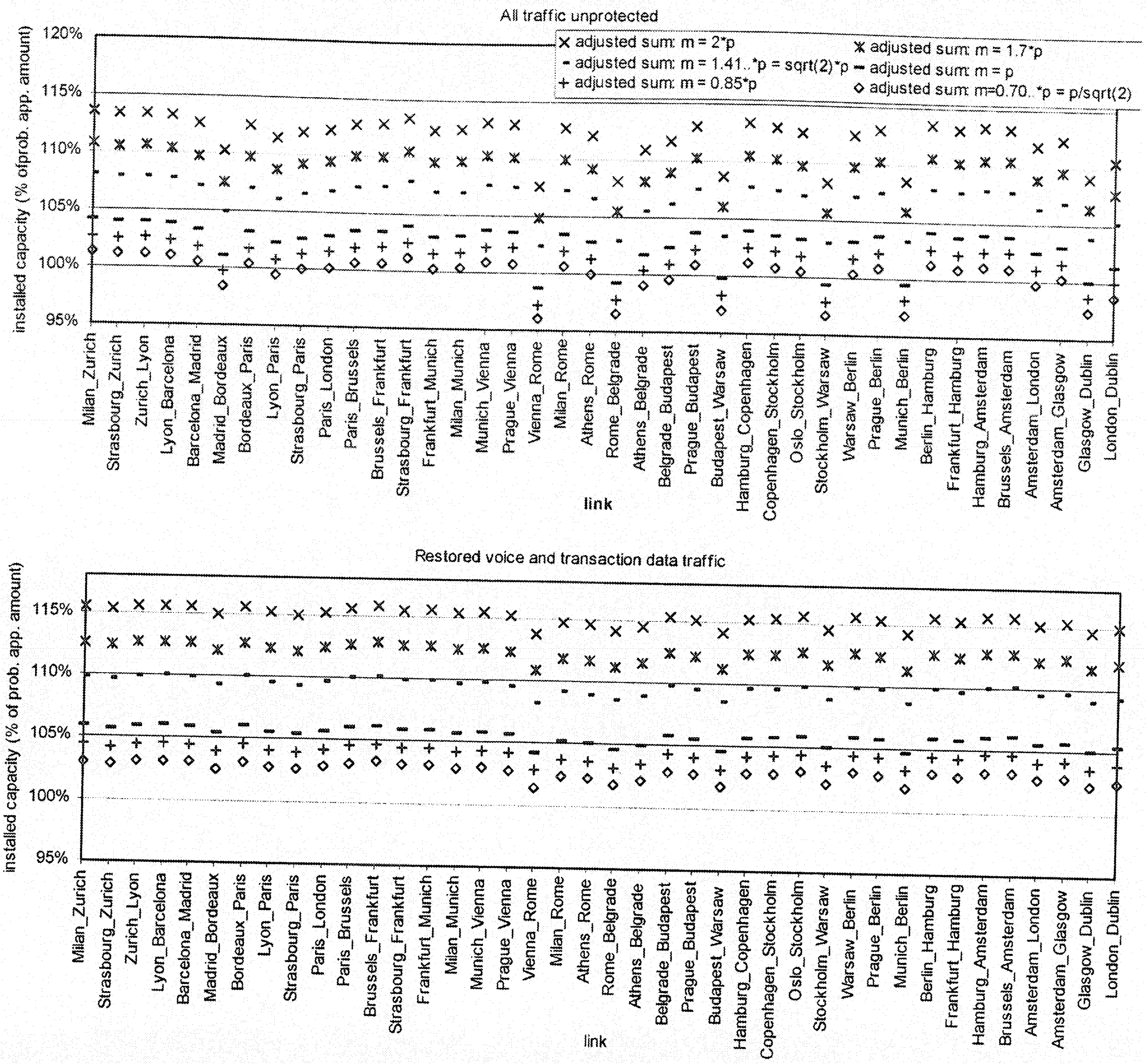


Fig. 13 Limit values (y_{limit}) for link dimensions in the pan-European network

adjustment would make an overinvestment, compared to one following the approximated probabilistic case. Because determination of maxima is involved here, the result is not normally distributed and therefore it is not possible to determine a suitable safety margin in the indicated way. Moreover, the changing number of operands involved complicates the issue even more. For the sake of completeness, we remark that the extra capacity needed to protect voice and transaction data represents 40% of all capacity in this case.

Conclusions

Strategic network planning needs forecasts for all major planning inputs. These forecasts are uncertain in nature and in order to model them as uncertain variables three essential parameters are needed: the predicted value expressing for instance an expert's view, the uncertainty level indicating the doubt there is about the presented predicted value, and the confidence parameter denoting the importance that the estimated output parameter exceeds the actual outcome. We distinguish several planning approaches for handling uncertainty, starting with a-priori and a-posteriori adjustments, which are widely used in practice today. Both use a safety margin to incorporate the effects of uncertainty. In the first case, the margin is added to the inputs of the planning problem, in the latter to the sharply calculated result after all calculations are performed. The popularity of those models can be explained by their conceptual as well as computational simplicity. The third considered approach represents uncertain variables using probability distributions. To be able to approximate all kind of density functions, a scalable piecewise linear approximation method is implemented.

survivability to network faults on the other hand, may indicate the extra capacity we foresee compared to the unprotected case, because of the use of backup paths.

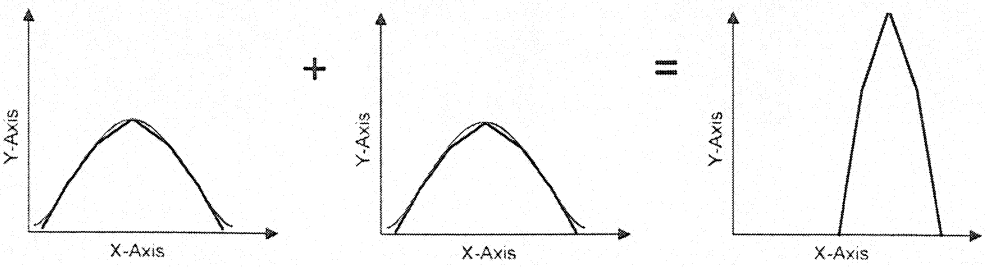


Fig. A1 Piecewise linear approximation of the density curve

We have shown that under some specific conditions it is possible to determine a safety margin value which allows a-priori or a-posteriori adjustment to approximate the probabilistic case. This appropriate value is a function of the confidence parameter, the ratio of the predicted values, and the number of operands. In practice, however, it is often infeasible to determine a useful safety margin value before the start of the planning calculations. This can for instance be the case when the shape of the distribution of the result is not known in advance, because of the skewness of the input distributions, the considered operations, etc. As a consequence, network planners often use a fixed safety margin value. We have shown that this may lead to incorrect planning decisions. Our findings were based on a theoretic model as well as on simulation results for a realistic planning problem.

Appendix A: some implementation details

When using the probabilistic approximated value model, the input distributions are sampled in n equidistant sample points. For normally distributed variables $N(v, s)$, we choose to sample in $v - 4s,\ v + 4s$ (where the density is set to 0) and $n - 2$ equidistant points between those extremes. A piecewise approximation of the original distribution is obtained, which, from this point onwards, is used as the continuous density distribution of the considered uncertain variable.

To determine the sum of two probabilistic approximated uncertain variable x_1 and x_2 (each having n sample points), the following steps are performed:

- Determine n^2 sample points for $Y :\ y = x_1 + x_2$, with x_1 and x_2 sample points for X_1 and X_2. Sort these y-values in increasing order.
- Calculate $F_Y(y)$ for all sample points y by calculating $F_{X_1}(y - x_2) \cdot f_{X_2}(x_2)$ and determining the area below the linear interpolation between those points. Connecting all points for $F_Y(y)$ results in a piecewise linear approximation of F_Y.

$$F_Y(y) = \text{Prob}[Y \le y] = \text{Prob}[X_1 + X_2 \le y]$$
$$= \int_{-\infty}^{+\infty} \text{Prob}[X_1 \le y - x_2 \text{ and } x_2 \le X_2 \le x_2 + dx_2]$$
$$= \int_{-\infty}^{+\infty} F_{X_1}(y - x_2) \cdot f_{X_2}(x_2) \cdot dx_2.$$

- Determine a piecewise linear density function $f(y)$ from this cumulative function $F(Y)$ by the following formula (for two consecutive sample points y_i and y_j:

$$F_Y(y_j) - F_Y(y_i) = (y_j - y_i) \cdot \frac{\lfloor f_Y(y_j) + f_Y(y_i) \rfloor}{2}$$

- Resample $f(y)$ to n sample points and normalise (area below $f(Y) = 1$)

Figure A1 illustrates how the density curve of the operands is approximated by a piecewise linear curve. The result is represented by a piecewise linear curve with the same amount of sample points (due to resampling) to order to serve as an input for another calculation lateron.

A similar way of working is followed to determine the maximum and product of uncertain variables. A complexity analysis showed that the determination of a maximum is of the order $O(n \cdot \log(n))$, while addition and multiplication are of the order $O(n^3)$.

Acknowledgements This work has been financed by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen) through a PhD grant for the first author, a postdoctoral grant for the second author and the GBOU-project IWT 010058. It was also supported by Ghent University through the BOF-project RODEO and by the European Commission through IST-NOBEL and ePhoton/One.

References

1. Coffman, K.G., Odlyzko, A.: The size and growth rate of the Internet. First Monday (journal on the Internet) 3(10): online (www.firstmonday.org) (1998)
2. Rosenbush, S., Crockett R.O., Haddad C., Ewing J.: The telecom depression: when will it end? BusinessWeek: online (http://www.businessweek.com), (2002)
3. Beardsley, S., Doman A., Edin P.: Making sense of broadband, McKinsey. quart strat uncertain world, 4, 5–14 (2001)
4. Sato, K.: Key enabling optical technologies for future networks. In: Proceedings of OFC 2003. Atlanta, USA, March, vol. 1, pp. 353–355 (2003)

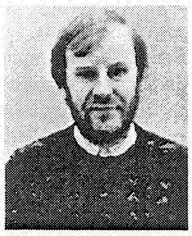
5. DeNardis, L.E.: Market shakeout: Implications for enterprise network planning, Business Commun. Rev. 26–29 (2001)
6. Laffont, J., Rey P., Tirole J.: Competition between telecom operators. Euro. Econ. Rev. **41**, 701–711 (1997)
7. Gryseels, M., Sorbello L., Demeester P.: Network planning in uncertain and dynamic environments, In: Proceedings Of Networks 2000, Toronto, Canada, September 2000
8. Wauters, N. Demeester P.: Design of optical path layer in multiwavelength cross-connected networks. IEEE J. Select Areas Commun. **14**(5), 881–892 (1996)
9. Van Caenegem, B., Van Parys, W., De Turck, F., Demeester, P.: Dimensioning of survivable WDM networks. IEEE J. Select Areas Commun. **16**(7), 1146–1157 (1998)
10. Pickavet, M., Demeester, P.: Multi-period planning of survivable WDM networks. Euro. Trans. Telecommun. Special issue on WDM networks, **11**(1), 7–16 (2000)
11. Demeester, P. et al.: Resilience in multilayer networks, IEEE Commun. Mag. **37**(8), 70–76 (1996)
12. Vaughn, M., Wagner, R.: Metropolitan network traffic demand study. In: Proceedings of LEOS 2000, vol. 1, pp. 102–103 Rio Grande, Puerto Rico, November 2000
13. Vaughn, M., Chia M., Wagner, R.: A bottom-up traffic demand model for long-haul metropolitan optical networks. In: Proceedings of OFC2003, vol. 1, pp. 139–150 MF111, Atlanta, USA, March 2003
14. Dwivedi, A., Wagner R.: Traffic model for USA long-distance optical networks In: Proceedings of OFC2000: vol. 1, pp. 156–158 TuK1-1, Baltimore, USA (2000)
15. Gupta, V.P.: What is network planning? IEEE Commun. Mag. **23**(10), 10–16 (1985)
16. King, W.R., Premkumar, G.: Key issues in Telecommunications Planning. Inform. Manage. **17**, 255–265
17. Carpenter, T., Heyman D., Saniee I.: Studies of random demands on network costs, Telecom. Sys. **10**, 409–421 (1998)
18. Sen, S., Doverspike, D., Cosares, S.: Network planning with random demands. Telecommun. Sys. **3**, 11–30 (1994)
19. Wallace, S.W.: Decision making under uncertainty: is sensitivity analysis of any use? Operat. Res. **48**(1), 20–25 (2000)
20. Mauz, C.: Dimensioning of a transport network being robust to changes in traffic patterns. In: Proceedings of ICTON2002, Warsaw, Poland, April 2002
21. Geary, N. et al.: A framework for optical network planning under traffic uncertainty. In: Proceedings of DRCN2001, pp. 50–56 Budapest, Hungary, February 2001
22. Hallez, A. et al.: Handling uncertainty in network capacity planning: a case study. In: Proceedings of EUROFUSE2002, pp. 87–91 Varenna, Italy, September 2002
23. Verbrugge, S. et al.: Modelling of uncertain demands in optical network planning. In: Proceedings of NOC2002, pp. 29–36 Darmstadt, Germany, June 2002
24. Courtney, H.G., Kirkland, J., Viguerie, P.: Strategy under uncertainty, McKinsey Quart. No. 3 (2000)
25. Metzer, B.: The uncertainty principle, CIO Insight, June 2001
26. Snedecor, G.W., Cochran, W.G.: Statistical methods Iowa State Univ. Press (1989)
27. IBCN: Reference topologies, online (http://www.ibcn.intec.ugent.be/css_design/research/projects/IST_FP5/NRS/index.html)
28. De Maesschalck, S. et al.: Pan-European optical transport networks. Photonic Network Commun. **5**(3), 203–225 (2003)

Sofie Verbrugge received a M.Sc. degree in Computer Science Engineering from the Ghent University (Gent, Belgium) in 2001. She is currently working toward the Ph.D. degree in the IBCN-group of the Department of Information Technology, at the same university, as a researcher for the Institute for the Promotion of Science and Technology in Flanders (IWT-Vlaanderen). Her research interests are situated in the field of long-term planning of optical networks and include techno-economic aspects of telecom network planning, such as migration scenarios, investment decision techniques and cost modelling. She has been involved in the European IST-project LION as well as in the ITEA-TBONES project and the IWT GBOU ONNA and is currently active in the IST-NOBEL.

Didier Colle received a M.Sc. degree in electrotechnical engineering (option: communications) from the Ghent University in 1997. Since then, he has been working at the same university as researcher in the department of Information Technology (INTEC). He is part of the research group INTEC Broadband Communication Networks (IBCN) headed by prof. Piet Demeester. His research lead to a Ph.D degree in February 2002. He was granted a postdoctoral scholarship from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen) in the period 2003–2004. His research deals with design and planning of communication networks. This work is focussing on optical transport networks, to support the next-generation Internet. Up till now, he has actively been involved in several IST projects (LION, OPTIMIST, DAVID, STOLAS, NOBEL and LASAGNE), in the COST-action 266 and 291 and in the ITEA/IWT TBONES project. His work has been published in more than 100 scientific publications in international conferences and journals.

Sophie De Maesschalck received a M.Sc. Degree in electrotechnical engineering (option: communication technology) in 1998 from the Ghent University, Belgium. In 1998, she started working as a research assistant at the Department of Information Technology of the same university. In 2003, she was granted the Ph.D. degree. Her research interests included the design, planning and evaluation of multilayer IP-over-optical transport networks and broadband access networks. She was involved in the ACTS project SONATA, in the IST projects LION, BREAD and MUSE, and in COST action 266. Currently she is working at Proximus, one of the Belgian mobile operators.

Mario Pickavet received a M.Sc. and Ph.D. degree in Electrical Engineering, specialized in telecommunications, from the Ghent University in 1996 and 1999, respectively. Since 2000, he is professor at Ghent University where he is teaching telecommunication networks and algorithm design. His current research interests are related to broadband communication networks (WDM, IP, (G-)MPLS, OPS) and include design, long-term planning and routing of core and access networks. Special attention goes to Operations Research techniques that can be applied for routing and network design. In this context, he is currently involved a.o. in the European IST projects LASAGNE and e-Photon/ONe. He has published over 100 international publications, both in peer-reviewed journals and in proceedings of conferences. He is one of the authors of the book "Network Recovery: Protection and Restoration of Optical, SONET-SDH, IP, and MPLS".



Piet Demeester received the Masters degree in Electro-Technical Engineering and the Ph.D. degree from the Ghent University (Gent, Belgium) in 1984 and 1988, respectively. In 1992 he started a new research activity on broadband communication networks, resulting in the IBCN-group (INTEC broadband communications network group). In 1993, he became a professor at the Ghent University where he is responsible for the research and education on communication networks. The research activities cover various communication networks (IP, ATM, SDH, WDM, access, active, mobile), including network planning, network and service management, telecom software, internetworking, network protocols for QoS support, etc. Piet Demeester is author of more than 250 publications in the area of network design, optimisation and management. He is member of the editorial board of several international journals and has been member of several technical program committees (ECOC, OFC, ICCCN, IZS, ...), he is the founding chair of the Workshop on Design of Reliable Communication Networks (DRCN).

Photonic Network Communications

Volume 11, Number 1, January 2006

ISSN: 1387-974X

Springer