

# The stochastic single resource service-provision problem

**Citation for published version (APA):**

Dye, S., Stougie, L., & Tomasgard, A. (2003). *The stochastic single resource service-provision problem*. (SPOR-Report : reports in statistics, probability and operations research; Vol. 200313). Technische Universiteit Eindhoven.

**Document status and date:**

Published: 01/01/2003

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

SPOR-Report 2003-13

**The stochastic single resource  
service-provision problem**

S. Dye  
L. Stougie  
A. Tomasgard

SPOR-Report  
Reports in Statistics, Probability and Operations Research

Eindhoven, April 2003  
The Netherlands

SPOR-Report  
Reports in Statistics, Probability and Operations Research

Eindhoven University of Technology  
Department of Mathematics and Computing Science  
Probability theory, Statistics and Operations research  
P.O. Box 513  
5600 MB Eindhoven - The Netherlands

Secretariat: Main Building 9.10  
Telephone: + 31 40 247 3130  
E-mail: [wscosor@win.tue.nl](mailto:wscosor@win.tue.nl)  
Internet: <http://www.win.tue.nl/math/bs/cosor.html>

ISSN 1567-5211

# The stochastic single resource service-provision problem\*

Shane Dye<sup>†</sup>  
Leen Stougie<sup>‡</sup>  
Asgeir Tomasgard<sup>§</sup>

December 17, 2002

## Abstract

The service-provision problem described in this paper comes from an application of distributed processing in telecommunications networks. The objective is to maximize a service provider's profit from offering computational based services to customers. The service provider has limited capacity and must choose which of a set of software applications he would like to offer. This can be done dynamically, taking into consideration that demand for the different services is uncertain. The problem is examined in the framework of stochastic integer programming.

Approximations and complexity are examined for the case when demand is described by a discrete probability distribution. For the deterministic counterpart a fully polynomial approximation scheme is known [2]. We show that introduction of stochasticity makes the problem strongly NP-hard, implying that the existence of such a scheme for the stochastic problem is highly unlikely. For the general case a heuristic with a worst-case performance ratio that increases in the number of scenarios is presented. Restricting the class of problem instances in a way that many reasonable practical problem instances satisfy, allows for the derivation of a heuristic with a constant worst-case performance ratio. Worst-case performance analysis of approximation algorithms is classical in the field of combinatorial optimization, but in stochastic programming the authors are not aware of any previous results in this direction.

**Keywords:** distributed processing; telecommunications; service-provision; stochastic (integer) programming; strong NP-hardness; approximation; worst-case analysis

---

\*For this work, financial support has been received from Telenor and Leonardo Da Vinci. Dye and Tomasgard were employees at The Norwegian University of Science and Technology in Trondheim, Norway when the main part of the work was done.

<sup>†</sup>University of Canterbury, New Zealand

<sup>‡</sup>Eindhoven University of Technology and CWI Amsterdam, The Netherlands

<sup>§</sup>SINTEF Industrial Management, Norway

# 1 Introduction

The problem we study concerns the provision of a set of services from a network of resources. For each service there are requests from customers. In order to meet a request for a service, the service has to be installed and, once installed, the request served. Both installation and provision of a service require capacity from the same resource. Each resource has a limited capacity. Every request served yields a given profit. The problem is to select a subset of the services to be installed and to decide which customer requests to serve, so as to maximize the total profit from serving requests.

Deterministic versions of the above, in which all demands are known in advance, have been studied previously. The decision problem with multiple resources that asks whether there is a solution for which all demand is met, is strongly NP-complete [3]. The deterministic optimization problem with multiple resources is strongly NP-hard even when the number of resources is fixed, [2]. The deterministic optimization problem with a single resource is NP-hard *in the ordinary sense* [2]. The same paper presents a fully-polynomial time approximation scheme for this problem.

We study the single resource version of the problem with uncertain demand for services. The uncertainty is represented by a discrete probability distribution over the demands. The mass points of the distribution are called scenarios in stochastic programming literature (see e.g. [5]). The problem is to select services to be installed so as to maximize the expected profit from serving requests. This is a two-stage stochastic integer programming problem. We show that this problem is *strongly* NP-hard. Thus, the complexity of the problem increases by introducing stochasticity. If the number of scenarios describing the probability distribution on the demand is constant, then a pseudo-polynomial time dynamic programming algorithm exists.

We analyze the performance of approximation algorithms for this problem from a worst-case perspective. This is commonly done in combinatorial optimization, but, to the best knowledge of the authors, it is the first result of this type in stochastic programming.

The motivation for this problem comes from an application in telecommunication that considers the installation of different *processing based services* on a set of computer nodes in a network with distributed processing capabilities. The services, if installed, are provided on demand. In the practical problem, services are constructed from more elementary building blocks, so-called *sub-services*. These are software applications, which run in a distributed manner throughout the network. The computers typically have limited capacity on memory, processing power and storage. The service provider must decide how to allocate the computational resources to the subservices in order to meet customer demand for services. Because of the limited capacities, at times it may be necessary to reject some customers. The distributed processing capabilities of the network enable the subservice demand to be considered independently of the service which generated it.

The single resource version of the problem can arise when the service provider rents capacity from a network provider, see [7, 8].

In the practical problem, demand for services is uncertain and dynamic. At various times the demand for a single service peaks, affecting the demand for all subservices used by that service. Before the peak actually occurs, a signal is

detected through deviations from the normal subservice demand patterns. The signals can be ambiguous but point to a number of possible services that might peak. For any signal there is some idea about the possible situations that are about to occur in terms of subservice demand. Therefore, given a signal, the number of scenarios required to describe the possible events which may occur, is small relative to the overall number of subservices.

The subservices typically take time and resources for start-up and shutdown. The configuration of subservices cannot react to changes in demand instantaneously. When a signal gives just enough time to re-configure the network before the peak occurs, a two-stage decision situation naturally emerges. In the first stage the decision is which subservices to install given only probabilistic information on demand. During the set-up time uncertainty resolves itself. The only possible recourse action in the second-stage concerns what demand to meet using the subservices installed in the first stage. The available capacity is restricted by the first stage decision. More information on the telecommunication problem and the model can be found in [8]. The subservices in this practical problem are the services in the general definition of the service-provision problem. Thus, the services referred to in the remainder of the paper, relate to the subservices of the practical problem.

In Section 2 we describe the stochastic single resource service-provision problem in a formal way, which will lead to a stochastic integer programming formulation. Its so-called deterministic equivalent formulation is presented in the same section. The complexity of the problem is studied in Section 3. Approximation algorithms are designed and analyzed in Sections 4 and 5. Both design and analysis of these algorithms are based on a study of the LP relaxation of the deterministic equivalent formulation. The approximation algorithm in Section 4 has a worst-case performance ratio which is a function of input parameters of the problem, a highly undesirable property. In Section 5 we restrict the class of problems by introducing the assumption that there is sufficient resource capacity to install all services concurrently. For the practical problem this is reasonable as the service provider would not reject demand when there is no peak in demand. If such rejection was common, it would be in the service provider's interest to invest in more resource capacity. This assumption allows the design of an approximation algorithm with constant worst-case performance ratio.

Computational testing of these approximation algorithms is reported in Section 6. The numerical experiments were performed on randomly drawn problem instances of a size which is reasonable for the practical telecommunication problem mentioned above. These experiments reveal that using standard software to solve the instances requires an unreasonable running time. In contrast, the approximation algorithms yield fast solutions which are within a factor of 1.08 on average.

## 2 A mathematical formulation

To state the problem we first introduce some notation. Let  $n$  be the number of services and  $s$  the capacity of the single resource. Let  $q_j$  be the profit obtained from allocating one resource unit to meeting demand for service  $j$ . Each service,  $j$ , uses a resource capacity of  $r_j$  while installed, which is independent of the

demand met. Demand is denoted by the random vector  $\delta \in R^n$ , with  $\delta_j$  denoting the demand for service  $j$ .

We introduce binary decision variables  $z_j$  that indicate whether service  $j$  is installed ( $z_j = 1$ ), or not ( $z_j = 0$ ),  $j = 1, \dots, n$ . Decision variable  $x_j$  gives the amount of resource used to meet demand for service  $j$ . The stochastic, single resource service-provision problem (SSRP) has a two-stage stochastic programming formulation:

$$\begin{aligned} \max \quad & E[Q(z, \delta)] \\ \text{s.t.} \quad & \sum_{j=1}^n r_j z_j \leq s \\ & z_j \in \{0, 1\} \quad j = 1, \dots, n, \end{aligned} \quad (1)$$

with  $Q(z, \delta)$  the optimal objective value of the second-stage linear program:

$$\begin{aligned} Q(z, \delta) = \max \quad & \sum_{j=1}^n q_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^n x_j \leq s - \sum_{j=1}^n r_j z_j \\ & x_j \leq \delta_j z_j \quad j = 1, \dots, n, \\ & x_j \geq 0 \quad j = 1, \dots, n. \end{aligned} \quad (2)$$

The second-stage problem is to set the values of the variables  $x_j$  under two constraints: the *capacity constraint* ensuring that resource capacity is not exceeded and the *demand constraint* ensuring that demand is not exceeded and met only for services that have been installed. The constraint in the first stage ensures relatively complete recourse (see e.g. [5]).

The probability distribution on demand is described, in stochastic programming terminology, in terms of scenarios [5]. Here a scenario is a vector of demands. A probability is assigned to each scenario. Denote by  $K$  the number of demand scenarios and let  $p_k$  be the probability that scenario  $k$  occurs. Then,  $\delta_{jk}$  is demand for service  $j$  in scenario  $k$ .

Given the scenarios, the stochastic program has a deterministic equivalent formulation (see [5], also called the extensive-form formulation [1]), which is a linear mixed integer programming model. We use  $x_{jk}$  for the second-stage variables, denoting the resource allocated to providing service  $j$  in scenario  $k$ .

$$\begin{aligned} \max \quad & \sum_{k=1}^K p_k \sum_{j=1}^n q_j x_{jk} \\ \text{s.t.} \quad & \sum_{j=1}^n (r_j z_j + x_{jk}) \leq s \quad k = 1, \dots, K, \\ & \delta_{jk} z_j - x_{jk} \geq 0 \quad j = 1, \dots, n, \quad k = 1, \dots, K, \\ & z_j \in \{0, 1\}, \quad x_{jk} \geq 0 \quad j = 1, \dots, n, \quad k = 1, \dots, K. \end{aligned} \quad (3)$$

The mathematical program of interest is a stochastic integer program. The integrality is purely in the first stage. In fact one may argue about the continuity assumption on the  $x$ -variables. However, if the data (resource capacity, installation requirements, and demands) are integral, the second stage will have an integer solution in every scenario.

To facilitate the exposition the assumption is made that no demand is higher than the resource capacity minus the corresponding installation requirement. If necessary, this can be ensured by preprocessing: just cutting off each demand above a level that could, whatsoever, never be met in any feasible solution.

**Assumption 1** For any service  $j$  in any scenario  $k$ ,  $0 \leq \delta_{jk} \leq s - r_j$ .

As a consequence, denoting the optimal value of (3) by  $\pi^{\text{OPT}}$ ,

$$\pi^{\text{OPT}} \geq q_j E[\delta_j], \quad j = 1, \dots, n. \quad (4)$$

### 3 Computational complexity

As indicated in the introduction, the deterministic counterpart of the problem admits a fully polynomial approximation scheme for its solution. Here we show that, unless  $P=NP$ , this is impossible for the stochastic problem.

**Theorem 1** *SSRP is strongly NP-hard, even when  $\sum_{j=1}^n r_j \leq s$ .*

**Proof** The natural recognition version of this problem, obtained by introducing a number and asking if there is a feasible solution giving profit at least that number, is obviously in NP. This can be seen because the representation of the probabilistic input in scenarios allows the formulation of a deterministic equivalent mixed-integer programming problem. To see that the recognition version is strongly NP-Complete consider a reduction from the well-known, strongly NP-Complete vertex cover problem (see [4]):

Given a graph  $G = (V, E)$  with  $|V|$  vertices and  $|E|$  edges and a constant  $\kappa$ , does there exist a subset  $V'$  of the vertices, such that each edge in  $E$  is incident to at least one vertex in  $V'$ , and such that  $|V'| \leq \kappa$ ?

For every vertex  $j \in V$  introduce a service,  $j$ , with installation requirement  $\alpha = \frac{1}{\kappa|E|}$ . For every edge, introduce a scenario with demand 1 for the two services incident to it and demand 0 for all other services. Let  $q_j = |E| \forall j \in V$ , and let all scenarios have a probability  $\frac{1}{|E|}$  of occurring. Then the expected profit from meeting one unit of demand in a single scenario is 1. Take  $\kappa\alpha + 1$  as capacity of the resource in SSRP. The question is whether there is a solution to this instance of SSRP with total expected profit at least  $|E|$ .

This transformation is obviously polynomial. In case there exists a vertex cover of size at most  $\kappa$  then there is a service-provision with total expected profit at least  $|E|$ . Install the services corresponding to the vertices in the vertex cover. Then for each scenario (edge) at least one of the services with demand 1 is installed. The total capacity used by the installation of the services is at most  $\kappa\alpha$  leaving at least capacity 1 to fill with the demands for each scenario.

The other direction is a bit more complicated. Suppose there does not exist a vertex cover of size  $\kappa$  or less. Then installing all services corresponding to a vertex cover would use resource capacity strictly greater than  $\kappa\alpha$  leaving strictly less than 1 for meeting demand in each of the  $|E|$  scenarios, making a total expected profit of at least  $|E|$  unattainable. Installing any set of services of size  $L \leq \kappa$  would leave  $(\kappa - L)\alpha + 1$  resource capacity for meeting demand in



each scenario. However, at least one edge will remain uncovered, implying that there is at least one scenario in which both services with a positive demand are not installed. With at most  $|E| - 1$  scenarios the expected profit will be at most  $(|E| - 1)((\kappa - L)\alpha + 1) \leq (|E| - 1)(\kappa\alpha + 1) = (|E| - 1)(\frac{1}{|E|} + 1) < |E|$ . For the transformed problem we have  $\sum_{j=1}^n r_j = |E|\alpha = \frac{1}{\kappa} \leq 1 < s$ .  $\square$

Below we will show that if the number of scenarios is fixed, the problem can be solved using a dynamic programming algorithm.

**Theorem 2** *The stochastic, single resource service-provision problem with a fixed number of scenarios can be solved in pseudo-polynomial time.*

**Proof** Consider the following DP that has the services as its stages. A state,  $S \in \mathbb{Z}_+^K$ , gives the capacity used in each scenario. Define  $f_j(S)$  as the maximum profit that can be achieved from scenario capacities  $S = (S_1, \dots, S_K)$  using the services  $1, \dots, j$ . Each  $S_k$  may take a value between 0 and  $s$ , hence there are at most  $(s + 1)^K$  states per stage. There are two types of transitions in every stage, either the service is not installed, or it is installed and some demand is met. There are fewer than  $s + 1$  possible choices concerning the demand to meet in each scenario, and overall there are then fewer than  $(s + 1)^K$  different feasible decisions in a state. The initial settings are

$$f_0(S_1, \dots, S_K) = \begin{cases} 0 & \text{if } 0 \leq S_i \leq s, \quad i = 1, \dots, K, \\ -\infty & \text{otherwise.} \end{cases}$$

The recurrence is given by

$$f_j(S_1, \dots, S_K) = \max_{0 \leq x_k \leq \delta_{jk}} \left\{ f_{j-1}(S_1 - r_j - x_1, \dots, S_K - r_j - x_K) + q_j \sum_{k=1}^K p_k x_k, f_{j-1}(S_1, \dots, S_K) \right\}.$$

From each state there are at most  $(s + 1)^K + 1$  possible transitions. At each of the  $n$  stages there are at most  $(s + 1)^K$  states. The running time of the DP is therefore at most  $O(ns^{2K})$ , which implies the theorem.  $\square$

Thus, the conclusion is that the problem with a fixed number of scenarios is not strongly NP-hard. This suggests also the existence of a fully polynomial approximation scheme for the problem, a nice subject for future investigations. This subclass of problems is still NP-hard since the problem with only one scenario (the deterministic problem) is NP-hard [2]. We argued in the introduction that this problem is not only of academic interest: in the telecommunication problem the number of scenarios to describe the random peak event is very small relative to the total number of services.

## 4 The LP bound and a heuristic

Given strong NP-hardness we study approximation algorithms. As a preliminary we concentrate on the LP relaxation of SSRP in which the restrictions  $z_j \in \{0, 1\}$ ,  $j = 1, \dots, n$ , in (3) are replaced by  $0 \leq z_j \leq 1$ ,  $j = 1, \dots, n$ . This section describes an optimal basis for the LP relaxation of SSRP and uses it to give an upper bound on the ratio of the LP versus the optimal solution. A heuristic based on the bound is given subsequently in Subsection 4.2.

## 4.1 The LP bound

The following theorem bounds the number of variables in an optimal LP solution that have a value different from their lower and (in some cases variable) upper bound, i.e., variables  $z_j$  with a value  $0 < z_j < 1$ , and variables  $x_{jk}$  with  $0 < x_{jk} < \delta_{jk}z_j$ . We will call these variables *interior*.

**Theorem 3** *Any basic optimal solution to the LP relaxation of SSRP with  $K$  scenarios has at most  $K$  interior  $z$  and  $x$  variables.*

**Proof** Let  $(z^{\text{LP}}, x^{\text{LP}})$  be an optimal basic solution to the LP relaxation of SSRP. Reduce the problem by eliminating all services for which  $z_j^{\text{LP}} = 0$  from the original problem. Clearly, the optimal solution of the LP relaxation of the reduced problem has the same number of interior  $x$  and  $z$  variables, and  $z^{\text{LP}} > 0$ .

Introducing slack variables  $t_k$  for the capacity constraints and  $u_{jk}$  for the demand constraints, results in the following reformulation of the LP relaxation.

$$\max \sum_{j=1}^n \sum_{k=1}^K p_k q_j x_{jk} \quad (5)$$

s.t.

$$\sum_{j=1}^n r_j z_j + \sum_{j=1}^n x_{jk} + t_k = s \quad k = 1, \dots, K, \quad (6)$$

$$\delta_{jk} z_j - x_{jk} - u_{jk} = 0 \quad j = 1, \dots, n, \quad k = 1, \dots, K, \quad (7)$$

$$0 \leq z_j \leq 1, \quad x_{jk}, u_{jk}, t_k \geq 0 \quad j = 1, \dots, n, \quad k = 1, \dots, K. \quad (8)$$

This LP has  $K + nK$  functional constraints, so that at any basic solution there are at most  $K + nK$  basic variables. Let  $(t^{\text{LP}}, u^{\text{LP}}, x^{\text{LP}}, z^{\text{LP}})$  be a basic optimal solution. We count the number of basic variables, i.e., variables with strictly positive values and, in the case of  $z$  variables, with value strictly less than 1.

Remember that we consider a problem with  $z_j^{\text{LP}} > 0$ ,  $j = 1, \dots, n$ . Thus, constraints (7) imply that at least one of  $x_{jk}^{\text{LP}}$  or  $u_{jk}^{\text{LP}}$  will be positive for each pair  $(j, k)$ ,  $j = 1, \dots, n$ ,  $k = 1, \dots, K$ , accounting for at least  $nK$  basic variables.

Define the following sets

$$\mathcal{Z} = \{j \mid z_j^{\text{LP}} = 1\},$$

$$\mathcal{F} = \{j \mid 0 < z_j^{\text{LP}} < 1\},$$

$$\mathcal{U} = \{(j, k) \mid x_{jk}^{\text{LP}} > 0 \text{ and } u_{jk}^{\text{LP}} > 0\}$$

and

$$\mathcal{T} = \{k \mid t_k^{\text{LP}} > 0\}.$$

Notice that  $\mathcal{U}$  is exactly the set of indices for which  $x_{jk}^{\text{LP}}$  are interior because they are positive but not equal to  $\delta_{jk}z_j^{\text{LP}}$ .

The number of interior  $z^{\text{LP}}$  and  $x^{\text{LP}}$  is  $|\mathcal{F}| + |\mathcal{U}|$  and the number of basic variables is  $|\mathcal{F}| + |\mathcal{T}| + |\mathcal{U}| + nK$ . From the above this is no greater than  $K + nK$ , implying  $|\mathcal{F}| + |\mathcal{T}| + |\mathcal{U}| \leq K$ .  $\square$

The optimal solution value,  $\pi^{\text{LP}}$ , can be written as

$$\pi^{\text{LP}} = \sum_{j \in \mathcal{Z}} \sum_{k=1}^K p_k q_j x_{jk}^{\text{LP}} + \sum_{j \in \mathcal{F}} \sum_{k=1}^K p_k q_j x_{jk}^{\text{LP}}. \quad (9)$$

In particular,  $|\mathcal{F}| \leq \min\{K, n\}$ . Under Assumption 1, the above theorem immediately yields the following corollary, in which  $\pi^{\text{OPT}}$  denotes the optimal solution value of SSRP.

**Corollary 1** *Under Assumption 1,  $\pi^{\text{LP}} \leq \min\{K + 1, n\} \pi^{\text{OPT}}$ .*

**Proof** Apply  $x_{jk}^{\text{LP}} \leq \delta_{jk}$  to (9):

$$\pi^{\text{LP}} \leq \sum_{j \in \mathcal{Z}} \sum_{k=1}^K p_k q_j x_{jk}^{\text{LP}} + \sum_{j \in \mathcal{F}} q_j \sum_{k=1}^K p_k \delta_{jk}. \quad (10)$$

Since  $\sum_{j \in \mathcal{Z}} \sum_{k=1}^K p_k q_j x_{jk}^{\text{LP}}$  is the value of a feasible solution it is no greater than  $\pi^{\text{OPT}}$ . Assumption 1 implies that (4) holds. Inserted in (10) the two observations yield

$$\pi^{\text{LP}} \leq \pi^{\text{OPT}} + \sum_{j \in \mathcal{F}} \pi^{\text{OPT}} \leq (\min\{K, n\} + 1) \pi^{\text{OPT}},$$

since  $|\mathcal{F}| \leq \min\{K, n\}$ . Finally, note that if  $|\mathcal{F}| = n$ , then  $\mathcal{Z} = \emptyset$  and (10) implies  $\pi^{\text{LP}} \leq n \pi^{\text{OPT}}$ .  $\square$

We have no example that shows tightness of this bound. The worst example found so far has ratio  $\pi^{\text{LP}}/\pi^{\text{OPT}} = 4$ .

Consider the following instance for an odd integer  $t \geq 3$ . Define the parameters  $s = 1$ ,  $K = n = t^2$ ,  $r_j = 1/(t + 1)$ ,  $q_j = 1$ ,  $p_j = 1/t^2$ ,  $\delta_{jj} = t/(t + 1)$ , for  $j = 1, \dots, t^2$ , and  $\delta_{jk} = 0$ , for  $j \neq k$ ,  $j, k = 1, \dots, t^2$ .

For this instance  $r_j + \delta_{jk} \leq s$  for  $j, k = 1, \dots, t^2$  and

$$\sum_{j=1}^n r_j = \frac{t^2}{t + 1} > 1.$$

In an optimal solution of the LP relaxation, the constraints of the form  $\delta_{jj} z_j - x_{jj} \geq 0$  will be binding. Thus, the LP relaxation may be re-written as

$$\begin{aligned} \max \quad & \sum_{j=1}^{t^2} \frac{z_j}{t(t + 1)} \\ \text{s.t.} \quad & \sum_{j=1}^{t^2} \frac{1}{t + 1} z_j + \frac{t}{t + 1} z_k \leq 1 \quad k = 1, \dots, t^2, \\ & 0 \leq z_j \leq 1 \quad j = 1, \dots, t^2. \end{aligned} \quad (11)$$

As the problem is symmetric in  $z_j$  there is an optimal solution for which all  $z_j$  are equal to  $\alpha$ , say. For this solution, the first set of constraints will be binding yielding  $\alpha = \frac{1}{t} < 1$ . The optimal LP objective value is  $\pi^{\text{LP}} = \frac{1}{t+1}$ .

For the optimal IP solution, the formulation is symmetric so it suffices to know how many  $z_j = 1$ . Assume that it is the first  $m$  of them, then for  $j =$

$1, \dots, m$ ,  $x_{jj} = 1 - \frac{m}{t+1}$ , otherwise  $x_{jk} = 0$ . The objective value is  $\frac{m}{t^2}(1 - \frac{m}{t+1})$  which is maximized when  $m = (t+1)/2$ . As  $t$  is odd,  $m$  is an integer, the optimal IP value is  $\pi^{OPT} = \frac{t+1}{4t^2}$  and

$$\pi^{LP}/\pi^{OPT} = \frac{4t^2}{(t+1)^2} \rightarrow 4 \quad \text{as} \quad t \rightarrow \infty.$$

## 4.2 The LP round-down heuristic

We present an approximation algorithm which amounts to rounding down the optimal solution of the LP relaxation of SSRP. Its worst-case performance analysis is related to the analysis of the greedy heuristic for the knapsack problem (see e.g. [6, Subsection 2.4]).

In the deterministic case the knapsack LP solution can be found in  $O(n)$  time by a median finding algorithm using the price per unit criterion [2]. Here a similar approach is not known.

Our algorithm installs service  $j$  if and only if  $z_j^{LP} = 1$ ; i.e., installs all services  $j \in \mathcal{Z}$ . The remaining capacity is then allocated to serve demand for the installed services in a greedy way, in order of non-increasing  $q_j$  values. We call this algorithm  $G_1$ , denote its solution by  $(z^{G_1}, x^{G_1})$  and its value by  $\pi^{G_1}$ .

Since  $G_1$  installs only services in  $\mathcal{Z}$ , the following lower bound is trivial.

### Proposition 1

$$\pi^{G_1} \geq \sum_{j \in \mathcal{Z}} \sum_{k=1}^K p_k q_j x_{jk}^{LP}.$$

That  $\pi^{OPT}/\pi^{G_1}$  can be arbitrarily high is seen from the example at the end of Section 4.1 where  $0 < z_j^{LP} < 1, \forall j$  and hence  $\pi^{G_1} = 0$ . Therefore we modify the algorithm: compute for each service  $j \in \mathcal{F}$  the objective value if  $j$  is the only service installed. Choose from the best of these values and  $\pi^{G_1}$  as output for the algorithm. Call this approximation algorithm  $G_2$ , then,

$$\pi^{G_2} = \max\{\pi^{G_1}, \max_{j \in \mathcal{F}} q_j E[\delta_j]\}.$$

A pseudo-code of the algorithm is shown in Figure 1, in which services are assumed to be ordered according to decreasing profits.

### Theorem 4 $G_2$ has worst case performance ratio

$$\pi^{OPT}/\pi^{G_2} \leq \min\{K+1, n\},$$

and this ratio is tight.

### Proof

$$\begin{aligned} \pi^{OPT} &\leq \pi^{LP} = \sum_{j \in \mathcal{Z}} \sum_{k=1}^K p_k q_j x_{jk}^{LP} + \sum_{j \in \mathcal{F}} \sum_{k=1}^K p_k q_j x_{jk}^{LP} \\ &\leq \pi^{G_1} + \sum_{j \in \mathcal{F}} q_j E[\delta_j] \\ &\leq \pi^{G_2} + |\mathcal{F}| \pi^{G_2} \leq (\min\{K, n\} + 1) \pi^{G_2}. \end{aligned}$$

---

Solve the LP relaxation, set  $s' = s$ ,  $\pi^{G_1} = 0$ .

**For each service  $j = 1, \dots, n$ :**

**If  $z_j^{LP} < 1$  then set  $z_j^{G_1} = 0$  and  $x_{jk}^{G_1} = 0$  for all  $k = 1, \dots, K$ ,**

**Otherwise set  $z_j^{G_1} = 1$  and  $s' = s' - r_j$ .**

**For each scenario  $k = 1, \dots, K$ :**

**Set  $s'' = s'$ .**

**For each service  $j = 1, \dots, n$ :**

**If  $z_j^{G_1} = 1$ :**

**Set  $x_{jk} = \min\{s'', \delta_{jk}\}$ ,**

**Set  $s'' = \max\{0, s'' - \delta_{jk}\}$ .**

**Set  $\pi^{G_2} = \pi^{G_1}$ ,  $z^{G_2} = z^{G_1}$  and  $x^{G_2} = x^{G_1}$ .**

**For each service  $j = 1, \dots, n$ :**

**If  $0 < z_j^{LP} < 1$  and  $\pi^{G_2} < q_j E[\delta_j]$ :**

**Set  $z_i^{G_2} = 0$  and  $x_{ik}^{G_2} = 0$  for all  $i = 1, \dots, n$  and  $k = 1, \dots, K$ ,**

**Set  $\pi^{G_2} = q_j E[\delta_j]$ ,  $z_j^{G_2} = 1$  and  $x_{jk}^{G_2} = \delta_{jk}$  for all  $k = 1, \dots, K$ .**

---

Figure 1: Approximation algorithm  $G_2$ .

Again, the case where  $|\mathcal{F}| = n$  strengthens the bound to  $\min\{K + 1, n\}\pi^{G_2}$ .

For the following instance the bound is tight. It has  $v + 1$  services and  $v$  scenarios,  $v \geq 2$ ,  $q_j = v - \epsilon$ ,  $r_j = \epsilon$ ,  $j = 1, \dots, v$ ,  $q_{v+1} = v/\epsilon$  and  $r_{v+1} = 0$  where  $0 < \epsilon < 1$ . The resource capacity is  $s = 1 + v\epsilon$  and all scenarios are equally likely. Demand for service  $v + 1$ , is the same for all scenarios:  $\delta_{v+1,k} = \epsilon/v$ ,  $k = 1, \dots, v$ . For each service  $j$ , with  $j = 1, \dots, v$ , there is positive demand only in scenario  $j$ :  $\delta_{jj} = 1$  and  $\delta_{jk} = 0$  when  $k \neq j$ .

The optimal solution is to install all services. Demand for service  $v + 1$  is always met completely, while in scenarios  $j = 1, \dots, v$  the optimal solution has  $x_{jj}^{\text{OPT}} = 1 - \frac{\epsilon}{v}$ . The profit from this is  $\pi^{\text{OPT}} = 1 + v - 2\epsilon + \frac{\epsilon^2}{v}$ .

The optimal solution of the LP relaxation is  $z_{v+1}^{\text{LP}} = 1$ ,  $x_{v+1,k}^{\text{LP}} = \epsilon/v$ ,  $\forall k$ ,  $x_{jj}^{\text{LP}} = z_j^{\text{LP}} = \frac{1+v\epsilon-\epsilon}{1+v\epsilon}$ ,  $\forall j$ ,  $x_{jk}^{\text{LP}} = 0$ ,  $k \neq j$ . This solution has  $K = v$  interior  $z$ -values and no interior  $x$ -values. Thus,  $\pi^{G_2} = \max\{1, \frac{1}{v}(v - \epsilon)\} = 1$ . As  $\epsilon$  gets arbitrarily small,

$$\pi^{\text{OPT}}/\pi^{G_2} = 1 + v - 2\epsilon + \frac{\epsilon^2}{v}$$

gets arbitrarily close to  $v + 1$ . □

Notice that the given bound on the performance ratio holds for any possible discrete distribution defined in terms of scenarios. It is increasing in the number of scenarios and if the number of scenarios is greater than the number of services the bound is linear in the number of services, which, in general, is not a very favorable situation. For the telecommunication problem described in the introduction the number of scenarios is typically small, so the algorithm might still be useful. Yet, it would be preferable to have an algorithm with a constant performance ratio. In the next section we derive such a bound by restricting the problem class slightly.

## 5 A constant bound

Here, we consider the class of problems for which it is feasible (but not necessarily optimal) to install all services concurrently; i.e., the sum of the installation requirements is less than the resource capacity. As discussed in the introduction, this assumption is reasonable in the telecommunication problem setting.

**Assumption 2** For the class considered:  $\sum_{j=1}^n r_j \leq s$ .

In order to facilitate the exposition, the resource capacity is scaled to 1:  $s = 1$ .

### 5.1 The LP bound

Again, we start by studying the ratio between the optimal solution value of the LP-relaxation with that of the original problem. Let  $(z^{\text{LP}}, x^{\text{LP}})$  be a basic optimal LP relaxation solution. Let  $\ell$  be the number of interior  $z_j^{\text{LP}}$  and assume that  $\ell_w$  of these services have  $r_j \leq w$  for some  $0 < w < 1$  to be chosen later. These services will be installed in groups while those with  $r_j > w$  will be installed separately as before. Again, let  $\mathcal{Z} = \{j \mid z_j^{\text{LP}} = 1\}$ . Without loss of generality let  $0 < z_j^{\text{LP}} < 1$  and  $r_j \leq w$  for  $j = 1, \dots, \ell_w$ , and  $0 < z_j^{\text{LP}} < 1$  and  $r_j > w$  for  $j = \ell_w + 1, \dots, \ell$ . Write the optimal LP value as

$$\pi^{\text{LP}} = \pi_0^{\text{LP}} + \pi_1^{\text{LP}} + \pi_2^{\text{LP}}, \quad (12)$$

where

$$\pi_0^{\text{LP}} = \sum_{j \in \mathcal{Z}} \sum_{k=1}^K p_k q_j x_{jk}^{\text{LP}},$$

$$\pi_1^{\text{LP}} = \sum_{j=1}^{\ell_w} \sum_{k=1}^K p_k q_j x_{jk}^{\text{LP}}$$

and

$$\pi_2^{\text{LP}} = \sum_{j=\ell_w+1}^{\ell} \sum_{k=1}^K p_k q_j x_{jk}^{\text{LP}}.$$

Feasible solutions generated from the LP solution will be used to bound parts of (12). Proposition 1 implies

$$\pi^{\text{OPT}} \geq \pi^{G_1} \geq \pi_0^{\text{LP}}. \quad (13)$$

To bound  $\pi_1^{\text{LP}}$  define  $A = \sum_{j=1}^{\ell_w} r_j z_j^{\text{LP}}$  and note that  $\sum_{j=1}^{\ell_w} x_{jk}^{\text{LP}} \leq 1 - A$  for each  $k = 1, \dots, K$ . Partition the set  $\{1, \dots, \ell_w\}$  into  $I$  subsets,  $\{S_i\}_{i=1}^I$ , such that

$$\sum_{j \in S_i} r_j \leq \beta + w \quad i = 1, \dots, I$$

and

$$\sum_{j \in S_i} r_j \geq \beta \quad i = 1, \dots, I - 1, \quad (14)$$

for some constant  $\beta > 0$  to be chosen later, such that  $\beta + w < 1$ . Notice that  $\sum_{j \in S_i} r_j$  is allowed to be smaller than  $\beta$ . In the optimal solution of the

LP relaxation at most  $1 - A$  units of capacity are available for the  $x$  variables. Installing only the services in one of the sets  $S_i$  will leave at least  $1 - \beta - w$  units of capacity available. The  $x$ -variable values from the LP relaxation solution corresponding to services in  $S_i$  may be scaled down, if necessary, to use a total of no more than  $1 - \beta - w$  units of capacity in each scenario.

For each  $i = 1, \dots, I$  define the feasible solution  $(z^{H_i}, x^{H_i})$  with  $z_j^{H_i} = 1$  for  $j \in S_i$ ,  $z_j^{H_i} = 0$  for  $j \notin S_i$ ,  $x_{jk}^{H_i} = \gamma x_{jk}^{\text{LP}}$  for  $j \in S_i$ ,  $k = 1, \dots, K$  and  $x_{jk}^{H_i} = 0$  for  $j \notin S_i$  and all  $k$ , where

$$\gamma = \begin{cases} \frac{1 - \beta - w}{1 - A} & \text{if } \beta + w \geq A, \\ 1 & \text{otherwise.} \end{cases} \quad (15)$$

The objective value of solution  $(z^{H_i}, x^{H_i})$  is

$$\pi^{H_i} = \sum_{j \in S_i} \sum_{k=1}^K p_k q_j x_{jk}^{H_i} = \gamma \sum_{j \in S_i} \sum_{k=1}^K p_k q_j x_{jk}^{\text{LP}}.$$

Hence,

$$\pi_1^{\text{LP}} = \sum_{i=1}^I \sum_{j \in S_i} \sum_{k=1}^K p_k q_j x_{jk}^{\text{LP}} = \frac{1}{\gamma} \sum_{i=1}^I \pi^{H_i} \leq \frac{I}{\gamma} \pi^{\text{OPT}}. \quad (16)$$

By the assumption on the problem class and the definition of the sets  $S_i$  (specifically (14)) we have

$$1 \geq \sum_{j=1}^n r_j \geq \sum_{j=1}^{\ell_w} r_j = \sum_{i=1}^I \sum_{j \in S_i} r_j \geq (I - 1)\beta. \quad (17)$$

Thus,  $I \leq 1 + 1/\beta$  which, inserted in (16), implies that

$$\pi_1^{\text{LP}} \leq \frac{\beta + 1}{\beta\gamma} \pi^{\text{OPT}}. \quad (18)$$

To bound  $\pi_2^{\text{LP}}$  consider installing each service  $j = \ell_w + 1, \dots, \ell$  (having  $r_j \geq w$ ) individually. Using the definition of  $A$ :

$$A = \sum_{j=1}^{\ell} r_j z_j^{\text{LP}} \geq \sum_{j=\ell_w+1}^{\ell} r_j z_j^{\text{LP}} \geq w \sum_{j=\ell_w+1}^{\ell} z_j^{\text{LP}}.$$

Thus,

$$\sum_{j=\ell_w+1}^{\ell} z_j^{\text{LP}} \leq \frac{A}{w}.$$

Just installing service  $j$  has objective value  $q_j E[\delta_j]$ . Satisfying the demand constraints implies that  $\sum_{k=1}^K p_k x_{jk}^{\text{LP}} \leq E[\delta_j] z_j^{\text{LP}}$ . Using Assumption 1, this leads to the following bound.

$$\pi_2^{\text{LP}} = \sum_{j=\ell_w+1}^{\ell} \sum_{k=1}^K p_k q_j x_{jk}^{\text{LP}}$$

$$\begin{aligned}
&\leq \sum_{j=\ell_w+1}^{\ell} q_j E[\delta_j] z_j^{\text{LP}} \\
&\leq \pi^{\text{OPT}} \sum_{j=\ell_w+1}^{\ell} z_j^{\text{LP}} \leq \frac{A}{w} \pi^{\text{OPT}}.
\end{aligned} \tag{19}$$

Combining (13), (18), and (19) gives

$$\pi^{\text{LP}} \leq \left(1 + \frac{\beta+1}{\beta\gamma} + \frac{A}{w}\right) \pi^{\text{OPT}}. \tag{20}$$

**Theorem 5** *Under the assumption that  $\sum_{j=1}^n r_j \leq 1$ ,*

$$\pi^{\text{LP}} \leq (5 + 2\sqrt{3})\pi^{\text{OPT}}.$$

**Proof**  $w$  and  $\beta$  are chosen to ensure the smallest bound in (20). When  $A < \frac{1}{2}$  take  $w = 1 - \frac{1}{2}\sqrt{3}$  and  $\beta = -\frac{1}{2} + \frac{1}{2}\sqrt{3}$  and when  $A \geq \frac{1}{2}$  take  $w = \beta = \frac{1}{2}A$ . In both cases  $w + \beta \geq A$ , and therefore  $\gamma = \frac{1-\beta-w}{1-A}$ . In the former case (20) yields

$$\begin{aligned}
\pi^{\text{LP}} &\leq \left(1 + \frac{2(1+\sqrt{3})(1-A)}{-1+\sqrt{3}} + \frac{A}{1-\frac{1}{2}\sqrt{3}}\right) \pi^{\text{OPT}} \\
&= \left(1 + (1+\sqrt{3})^2(1-A) + 4\left(1 + \frac{1}{2}\sqrt{3}\right)A\right) \pi^{\text{OPT}} = (5 + 2\sqrt{3})\pi^{\text{OPT}}.
\end{aligned}$$

In the latter case (20) leads to

$$\pi^{\text{LP}} \leq \left(4 + \frac{2}{A}\right) \pi^{\text{OPT}} \leq 8\pi^{\text{OPT}} \leq (5 + 2\sqrt{3})\pi^{\text{OPT}}.$$

□

We have also not been able to show tightness for this bound. The worst instance found (with ratio 2) is similar to the instance given at the end of Section 4.1 but with  $r_j = 1/t^2$  and  $\delta_{jj} = (t^2 - 1)/t^2$ .

## 5.2 A round and partition heuristic with constant worst-case ratio

Based on the previous LP bound a round and partition heuristic, RP, is developed with a worst-case performance ratio bounded above by  $5 + 2\sqrt{3}$ .

Let  $(z^{\text{LP}}, x^{\text{LP}})$  be a basic optimal LP relaxation solution with the optimal solution value given by (12). Choose the constants  $w$  and  $\beta$  of the previous subsection as in the proof of Theorem 5 and generate a partition  $\{Z, V, W, S_1, \dots, S_I\}$  for some  $I$  in  $\{1, \dots, n\}$  as follows.

$$\begin{aligned}
Z &= \{j \mid z_j^{\text{LP}} = 1\}; \\
V &= \{j \mid z_j^{\text{LP}} = 0\}; \\
W &= \{j \mid 0 < z_j^{\text{LP}} < 1, r_j > w\}.
\end{aligned}$$

The remaining services  $\{j \mid 0 < z_j^{\text{LP}} < 1, r_j \leq w\}$  are partitioned into the sets  $S_1, \dots, S_I$  in the following way. Consider these services in arbitrary order. Start



filling set  $S_1$  with the first services until addition of the next service will raise the sum of the installation requirements above  $w + \beta$ . That service will be the first one to go into the set  $S_2$ . Continue in the same way filling the set  $S_2$  and so on until the last set  $S_I$  is constituted by the last few items. Thus, the sets  $S_1, \dots, S_I$  have the properties

$$\sum_{j \in S_i} r_j \in [\beta, \beta + w], \quad i = 1, \dots, I - 1,$$

$$\sum_{j \in S_I} r_j \leq \beta + w.$$

For each set  $T \in \{Z, V, \{j\}_{j \in W}, S_1, \dots, S_I\}$  heuristic RP generates the solution  $(z^T, x^T)$  with objective value  $\pi^T$ , by setting  $z_j^T = 1$  if  $j \in T$  and  $z_j^T = 0$  if  $j \notin T$  and choosing  $x^T$  to maximize the LP created by fixing  $z$  to  $z^T$ . Then RP chooses the solution with highest objective value. We denote this value by  $\pi^{RP}$ . A pseudo-code of RP is shown in Figure 2.

---

Solve the LP relaxation to get basic optimal solution  $(z^{LP}, x^{LP})$   
Set  $I = 1$  and  $r^* = 0$  and declare sets  $Z, V, W$ , and  $S_1$  open.  
**For each** subservice  $j = 1, \dots, n$  (order arbitrary):  
  **If**  $z_j^{LP} = 1$  **then add**  $j$  to  $Z$ ,  
  **Otherwise if**  $z_j^{LP} = 0$  **then add**  $j$  to  $V$ ,  
  **Otherwise if**  $0 < z_j^{LP} < 1$  **and**  $r_j > w$  **then add**  $j$  to  $W$ ,  
  **Otherwise** (Comment:  $0 < z_j^{LP} < 1$  and  $r_j \leq w$ )  
    **If**  $r^* + r_j \leq w + \beta$  **then**  
      **Add**  $j$  to  $S_I$  and **set**  $r^* = r^* + r_j$ ,  
    **Otherwise**  
      Close  $S_I$  and open  $S_{I+1}$ ,  
      **Add**  $j$  to  $S_{I+1}$  and **set**  $r^* = r_j$ ,  
      **Set**  $I = I + 1$ .  
Close sets  $Z, V, W$ , and  $S_I$ .  
Generate the solutions  $(x^Z, z^Z)$  and  $(x^{\{j\}}, z^{\{j\}})$  for all  $j \in W$ .  
Generate the solutions  $(x^{S_i}, z^{S_i})$  for  $i = 1, \dots, K$ .  
**Set**  $\pi^{RP} = \max\{\pi^Z, \max\{\pi^{\{j\}} \mid j \in W\}, \max\{\pi^{S_i} \mid i = 1, \dots, I\}\}$ .  
**Set**  $(x^{RP}, z^{RP})$  to the solution corresponding to  $\pi^{RP}$ .

---

Figure 2: Approximation Algorithm RP.

**Theorem 6** For SSRP with  $\sum_{j=1}^n r_j \leq s$ , RP has worst-case performance ratio

$$\pi^{\text{OPT}} / \pi^{\text{RP}} \leq (5 + 2\sqrt{3}).$$

**Proof** This follows almost immediately from the proof of the bound for the LP relaxation in Section 5.1. Notice that  $\pi^Z = \pi^{G_1}$  and for any  $j \in W$ ,  $\pi^{\{j\}} = q_j E[\delta_j]$ . Also, for each  $i \in \{1, \dots, I\}$ ,  $\pi^{S_i} \geq \pi^{H_i}$ . Thus,

$$\pi^{\text{RP}} \geq \pi^{G_1}, \quad \pi^{\text{RP}} \geq \pi^{H_i}, \quad \forall i = 1, \dots, I \quad \text{and} \quad \pi^{\text{RP}} \geq q_j E[\delta_j], \quad \forall j \in W.$$

These observations allow  $\pi^{\text{OPT}}$  to be replaced by  $\pi^{\text{RP}}$  in (13), (16) and (19). Following this throughout the proof of the LP bound in Theorem 5 gives

$$\pi^{\text{OPT}} \leq \pi^{\text{LP}} \leq (5 + 2\sqrt{3})\pi^{\text{RP}}.$$

□

We have not been able to show tightness of the bound. The worst instance we have been able to find has a performance ratio of 2. When Assumption 2 does not hold RP can have an arbitrarily high ratio. The instance given at the end of Section 4.1 is an example of this, which can the readers can easily verify themselves.

## 6 Computational Results

The heuristics  $G_2$  and RP were tested on problem instances that reflect the telecommunication application described in the introduction. The instances have realistic sizes and are randomly generated. Parameter values were chosen so as to generate problems that were challenging for the heuristics. The test problems have the following characteristics.

The number of services ranges from 500 to 1000 and the number of scenarios from 10 to 50, with 5 to 13 services peaking in each scenario. In each scenario, a group of related services (the subservices involved in a service in the context of the telecommunication problem) experience a peak in demand. The single resource has a constant capacity of 10,000 units. There are 72 sets of test problems generated each with different relationships between the various problem data. For each set, 30 test problems are generated with the parameters for each service chosen randomly, giving a total of 2160 test problems.

The test problems include problems with both  $\sum r \leq s$  and  $\sum r > s$ . Roughly two-thirds have  $\sum r \leq s$  and of the remainder about half have  $s < \sum r < 2s$  and half have  $4s < \sum r < 5s$ . The last two groups were generated to test the heuristic (RP) on problems for which it is not guaranteed to perform within the constant performance ratio derived.

The problems were all solved on a PC Pentium III, with 500MB of RAM, 2GB of swap space, running at over 800MHz, and running Microsoft Windows 2000. AMPL and CPLEX 7.5.0 were used to find the optimal solutions to the integer problem and the linear relaxation. The heuristics were coded in Matlab 6, using the linear relaxation solution as generated by CPLEX.

In solving the integer problem by AMPL and CPLEX 7.5.0, a limit of 30 minutes was given. About 30% of the problem instances could not be solved within that time. For the remaining 70% the average time was over 13 minutes. The linear relaxation was solved with an average time of 19 seconds and a standard deviation of 30 seconds. About 92% of the linear relaxations were solved within 60 seconds.

The running times of heuristics  $G_2$  and RP were recorded independently of the time required to solve the linear relaxation. For the heuristic  $G_2$  the mean running time was 0.08 seconds. For heuristic RP the mean running time was 0.09 seconds. The maximum running time of either heuristic was 0.6 seconds.

Quantities  $\pi^{\text{OPT}}/\pi^{G_2}$  and  $\pi^{\text{OPT}}/\pi^{\text{RP}}$  were calculated. These ratios only differed for 1% of the test problems. For heuristic RP, the average ratio over all

test problems was 1.08, with standard deviation 0.08. For 4.4% of the cases the solution found was within the default MIP gap accepted by CPLEX (of 0.0001). In addition, 34% of the cases were within 1% of the optimal solution and 67% of the cases within 10%. The worst ratio encountered for a test problem was 1.52. These results are shown in Figure 3.

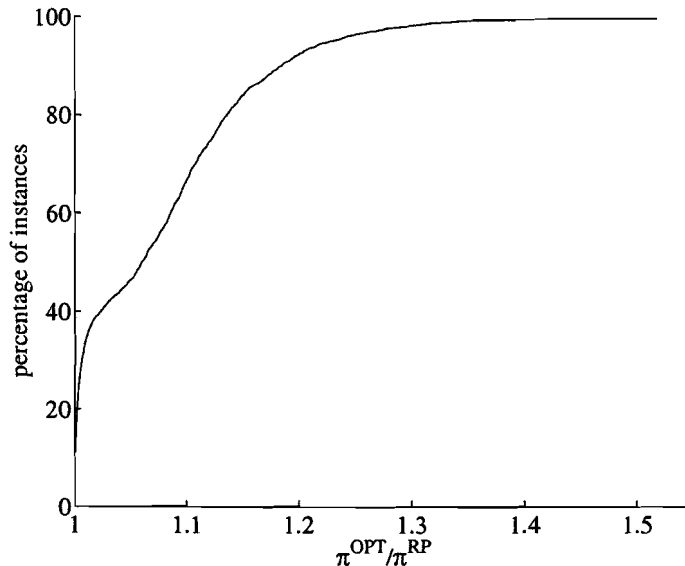


Figure 3: Percentage of instances with given performance ratio or better.

When only those test problems with  $\sum r \leq s$  are considered, the average ratio is 1.06 and the worst ratio 1.33.

From the test problems studied the performance seemed most affected by the average standard deviation of the demand (over all services),  $\frac{1}{n} \sum_{j=1}^n \text{std}(\delta_j)$ . This is shown in Figure 4. The figure includes all instances. Restricting to instances with  $\sum r \leq s$  shows a similar relationship.

We notice that the standard packages, which we applied without any performance tuning, take too much time for the typical telecommunication problem that motivated this research. The time between the signal and the occurrence of the peak is typically no more than 20 minutes, and often much less. The approximation algorithms obviously give very good solutions, quickly. Therefore, they are recommended in favor of the exact optimization methods. In other practical applications for which there is much more time between the installation of services and the actual use, one may still aim for optimality.

## 7 Conclusions

This paper considered a service-provision problem. The main motivation for studying the problem was an application on a distributed processing telecommunication network. In that problem there is uncertain demand for the services. It was shown that the natural stochastic programming model is strongly NP-hard. It is worthwhile to stress this as the deterministic counterpart, having

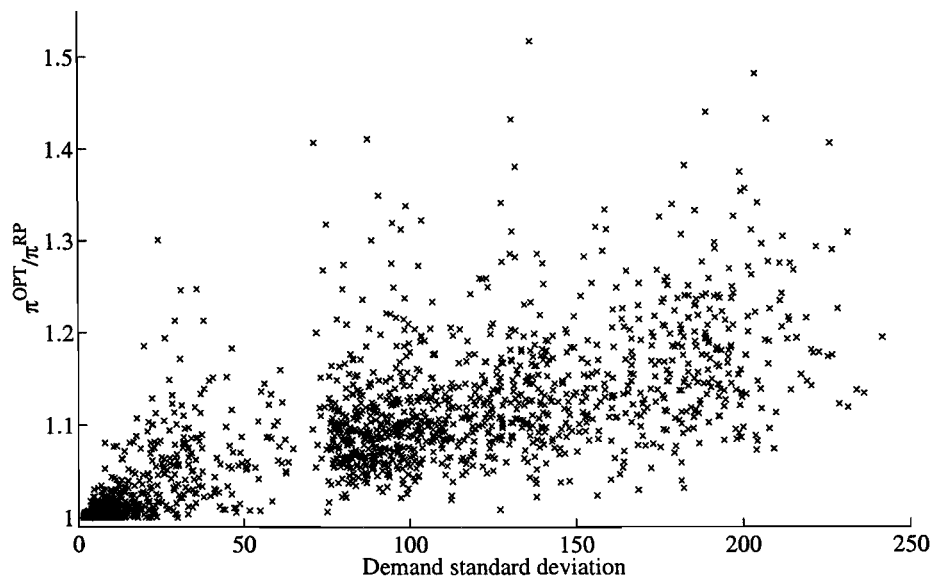


Figure 4: Heuristic performance compared to the demand standard deviation.

the same number of binary decision variables, is weakly NP-hard. Thus, the complexity of the problem increases by introducing stochasticity, even if it only means adding continuous decision variables for each scenario of the problem.

We studied approximation algorithms for the stochastic service provision problem. A first algorithm based on the LP relaxation of the deterministic equivalent stochastic problem has worst-case performance ratio equal to the minimum of the number of services and the number of (demand) scenarios plus one. The second algorithm has a constant worst-case performance ratio for a restricted class of problems. The assumption defining this subclass is, however, not unnatural and indeed satisfied for the telecommunication problem.

Results of this type are standard in combinatorial optimization, but, to the best knowledge of the authors, the results in this paper are the first worst-case performance analysis results in the field of stochastic programming. They give guarantees on the quality of the solutions provided by the algorithms.

The numerical testing shows that these guarantees are quite pessimistic in practice, for the random instances tested. The approximation algorithms typically performed very well, while exact optimization methods required impractical running times.

## References

- [1] J.R. Birge, F.V. Louveaux. *Introduction to stochastic programming*. Springer Verlag, New York, 1997.
- [2] S. Dye, L. Stougie, and A. Tomasgard. Approximation algorithms and relaxations for a service provision problem on a telecommunication network. Working paper #2-98, Department of Industrial Economics and Technol-

ogy Management, Norwegian University of Science and Technology, N-7034 Trondheim, Norway, 1998, to appear in *Discrete Applied Mathematics*.

- [3] S. Dye, A. Tomasgard, and S.W. Wallace. Feasibility in transportation networks with supply eating arcs. *Networks*, 31:165–176, 1998.
- [4] M.R Garey and D.S. Johnson. *Computers and Intractability, a guide to the theory of NP-completeness*. Freeman, New York, 1979.
- [5] P. Kall and S.W. Wallace. *Stochastic Programming*. Wiley, Chichester, 1994.
- [6] S. Martello and P. Toth. *Knapsack Problems: Algorithms and Computer Implementations*. John Wiley & sons, 1990.
- [7] A. Tomasgard, S. Dye, S.W. Wallace, J.A. Audestad, L. Stougie, and M.H. van der Vlerk. Stochastic optimization models for distributed communication networks. Working paper #3-97, Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, N-7034 Trondheim, Norway, 1997.
- [8] A. Tomasgard, J.A. Audestad, S. Dye, L. Stougie, M.H. van der Vlerk, and S.W. Wallace. Modelling aspects of distributed processing in telecommunication networks. *Annals of Operations Research*, 82:161–184, 1998.