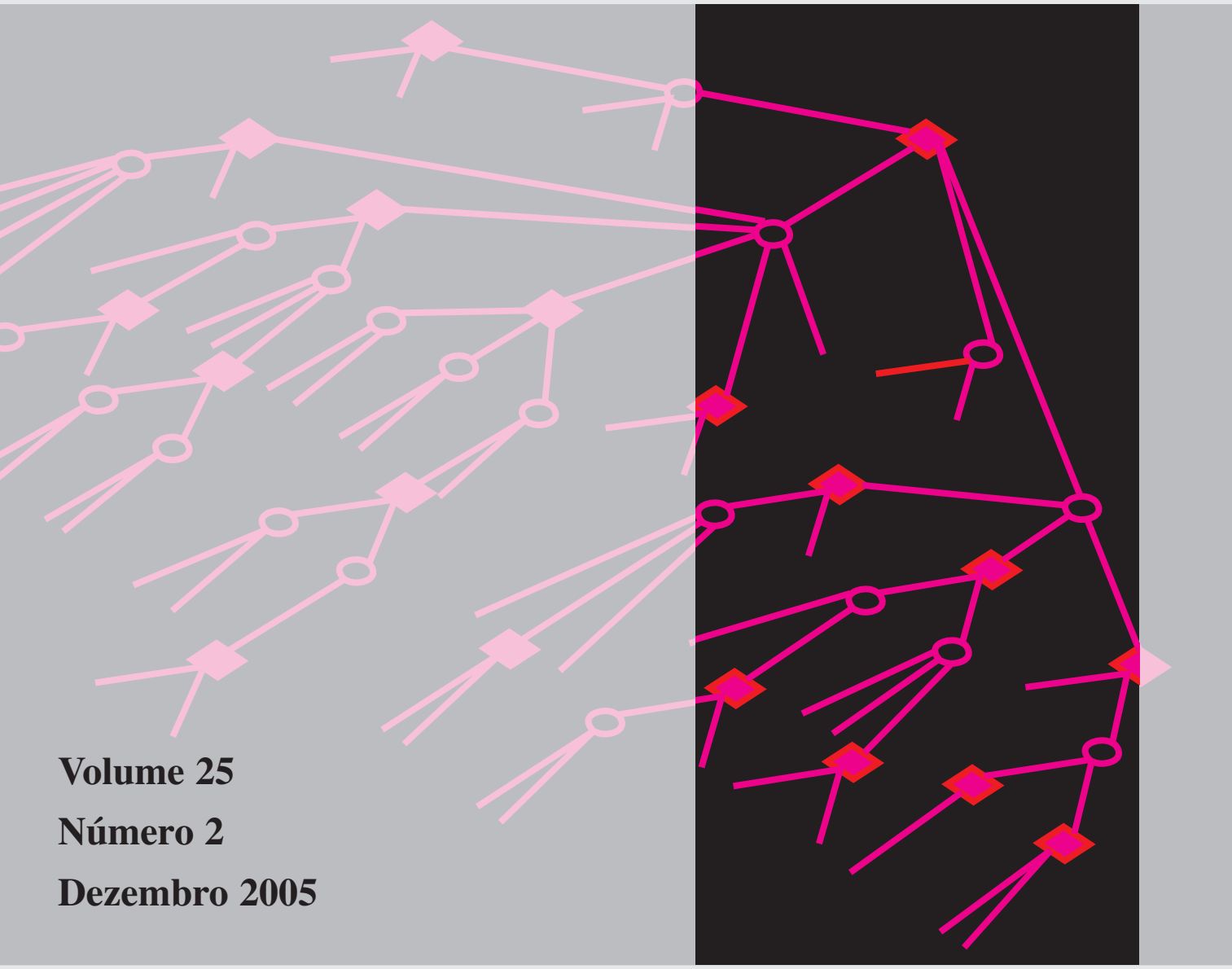
ISSN: 0874-5161

Investigação Operacional



Volume 25

Número 2

Dezembro 2005

INVESTIGAÇÃO OPERACIONAL

Propriedade:
APDIO
Associação Portuguesa
de Investigação Operacional

Patrocinadores

Fundação Calouste Gulbenkian

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA E DA TECNOLOGIA

Apoio do Programa Operacional Ciência, Tecnologia,
Inovação do Quadro Comunitário de Apoio III.

ISSN nº 0874-5161
Dep. Legal nº 130 761 / 98
Execução Gráfica: J. F. Macedo - Astrografe
700 Ex.
2005/12

INVESTIGAÇÃO OPERACIONAL

Volume 25 — nº 2 — Dezembro 2005 Publicação Semestral

Editor Principal: José F. Oliveira
Universidade do Porto

Comissão Editorial

M. Teresa Almeida
Inst. Sup. Economia e Gestão

C. Henggeler Antunes
Univ. de Coimbra

Marcos Arenales
Univ. de São Paulo

Jaime Barceló
Univ. de Barcelona

Eberhard E. Bischoff
University of Wales, Swansea

C. Bana e Costa
Inst. Superior Técnico

M. Eugénia Captivo
Univ. de Lisboa

Domingos M. Cardoso
Univ. de Aveiro

João Clímaco
Univ. de Coimbra

J. Dias Coelho
Univ. Nova de Lisboa

João P. Costa
Univ. de Coimbra

Ruy Costa
Univ. Nova de Lisboa

J. Rodrigues Dias
Univ. de Évora

Laureano Escudero
IBM, Espanha

Edite Fernandes
Univ. do Minho

J. Soeiro Ferreira
Univ. do Porto

J. Fernando Gonçalves
Univ. do Porto

Luís Gouveia
Univ. de Lisboa

Rui C. Guimarães
Univ. do Porto

Joaquim J. Júdice
Univ. de Coimbra

J. Assis Lopes
Inst. Superior Técnico

Carlos J. Luz
Inst. Polit. Setúbal

Virgílio P. Machado
Univ. Nova de Lisboa

Manuel Matos
Univ. do Porto

N. Maculan
Univ. Fed., Rio Janeiro

Rui Oliveira
Inst. Superior Técnico

J. Pinto Paixão
Univ. de Lisboa

M. Vaz Pato
Inst. Sup. Economia e Gestão

Mauricio G. Resende
AT&T Labs Research

A. Guimarães Rodrigues
Univ. do Minho

António J. L. Rodrigues
Univ. de Lisboa

J. Pinho de Sousa
Univ. do Porto

Reinaldo Sousa
Univ. Católica, Rio Janeiro

L. Valadares Tavares
Inst. Superior Técnico

B. Calafate Vasconcelos
Univ. do Porto

Luís N. Vicente
Univ. de Coimbra

Victor V. Vidal
Technical Univ. of Denmark

INVESTIGAÇÃO OPERACIONAL

Volume 25 — nº 2 — Dezembro 2005 Publicação Semestral

Editor Principal: José F. Oliveira
Universidade do Porto

Comissão Editorial

M. Teresa Almeida
Inst. Sup. Economia e Gestão

C. Henggeler Antunes
Univ. de Coimbra

Marcos Arenales
Univ. de São Paulo

Jaime Barceló
Univ. de Barcelona

Eberhard E. Bischoff
University of Wales, Swansea

C. Bana e Costa
Inst. Superior Técnico

M. Eugénia Captivo
Univ. de Lisboa

Domingos M. Cardoso
Univ. de Aveiro

João Clímaco
Univ. de Coimbra

J. Dias Coelho
Univ. Nova de Lisboa

João P. Costa
Univ. de Coimbra

Ruy Costa
Univ. Nova de Lisboa

J. Rodrigues Dias
Univ. de Évora

Laureano Escudero
IBM, Espanha

Edite Fernandes
Univ. do Minho

J. Soeiro Ferreira
Univ. do Porto

J. Fernando Gonçalves
Univ. do Porto

Luís Gouveia
Univ. de Lisboa

Rui C. Guimarães
Univ. do Porto

Joaquim J. Júdice
Univ. de Coimbra

J. Assis Lopes
Inst. Superior Técnico

Carlos J. Luz
Inst. Polit. Setúbal

Virgílio P. Machado
Univ. Nova de Lisboa

Manuel Matos
Univ. do Porto

N. Maculan
Univ. Fed., Rio Janeiro

Rui Oliveira
Inst. Superior Técnico

J. Pinto Paixão
Univ. de Lisboa

M. Vaz Pato
Inst. Sup. Economia e Gestão

Mauricio G. Resende
AT&T Labs Research

A. Guimarães Rodrigues
Univ. do Minho

António J. L. Rodrigues
Univ. de Lisboa

J. Pinho de Sousa
Univ. do Porto

Reinaldo Sousa
Univ. Católica, Rio Janeiro

L. Valadares Tavares
Inst. Superior Técnico

B. Calafate Vasconcelos
Univ. do Porto

Luís N. Vicente
Univ. de Coimbra

Victor V. Vidal
Technical Univ. of Denmark

The Art and Science of Problem Solving

René Victor Valqui Vidal †

† Informatics and Mathematical Modelling
Technical University of Denmark
2800 Lyngby
Denmark
vvv@imm.dtu.dk
www.imm.dtu.dk/∼vvv

Study the science of art
and the art of science.
- Leonardo da Vinci

Abstract

In this paper we will document that real-life problem solving in complex situations demands both rational (scientific) and intuitive (artistic) thinking. First, the concepts of art and science will be discussed; differences and similarities will be enhanced. Thereafter the concept of group problem solving facilitation both as science and art will be presented. A case study related to examination's planning will be discussed to illustrate the main concepts in practice. In addition, other cases studies will also be shortly presented.

Keywords: Problem solving, art, science, facilitation, group work

1 Introduction

In our educational institutions and in our culture in general, there is a split between art and science. It is believed that these two ways of working and thinking, the artistic attitude and the scientific attitude are two very different worlds, they are like oil and water. Although the link between art and science has historically been very close, exemplified by Leonardo da Vinci, the ideal that Leonardo represents is really not agreed upon by the art and science communities. It is the opinion of the author of this paper that this distinction between and separation of art and science is artificial and increasingly anachronistic. Fortunately things are changing; new fields arise from the synthesis of other fields. For instance, scientists are

Apdio © 2005 *Associação Portuguesa de Investigação Operacional*

relaying more and more on visual communication, and artists are working increasingly with computers. There is a common place to transfer information, ideas and knowledge. Visual problems are ultimately the same across disciplines. For example, Computer Graphics is a new field made up of art and science.

Another interesting field that demands knowledge and experience from art and science, as well as other disciplines, is Modern Design. The book of Nelson and Stolterman (2003) is a coherent meta-theoretical and holistic approach to a design theory applicable to any context. This book is valuable and useful: full of many conceptual and practical ideas; given an integrated picture of central theories and concepts about design. In addition, it brings together qualitative and creative issues from Art and Architecture with quantitative and scientific approaches of Operational Research and Engineering. The authors do this by founding their concepts and analyses on the epistemological foundations of Systems Thinking.

Another valuable contribution related to management and problem solving is the book edited by Boland and Collopy (2004). The main premise of this book is that: *managers should act not only as intelligence gatherers and decision makers, but also as creative designers.* Though decision and design are inextricably linked in management action, managers and scholars have too long emphasized the decision aspect of management over the design aspect. The main message of this book is that managing is not only decision making but also designing. This collection of papers explores the "design attitude" as opposite to the "decision attitude", that draws on examples of managing in architecture, art, and design. This volume is a critique of the predominant management education that focuses on training students to make choices among the alternatives presented to them, rather than training them to design new alternatives. In a series of brilliant and wide-ranging essays from a multitude of disciplines, the authors develop a theory of the design attitude that contrasts with the more traditionally accepted and practiced decision attitude. Their innovative view of management promises to provide a way into some of the most pressing issues facing organizational leaders, researchers and educators today. The editors have designed a thought-provoking volume that portrays management not as a science of rational problem solving but, instead as the art of generating visions and roads for achieving these visions. This collection of papers will inspire, enlight, intrigue, challenge, and teach readers about designing organisations.

In 1981, Ackoff (1981) published a paper entitled: *The art and science of mess management.* He argues that there are three kinds of thing that can done about problems:

- The *clinical approach*, where you select actions that are good enough, that *satisfices*; it relies heavily on past experience, it is qualitatively oriented, and it is rooted deeply in common sense making use of subjective judgements.
- The *research approach*, where you select actions that are the best possible outcome, which *optimizes*; it is largely based on scientific methods; it is quantitative oriented, and it makes use of mathematical models aspiring to complete objectivity. And
- *The design approach*, where you seek to change the nature, and the environment of the problem so as to remove the problem, it *dissolves* the problem; it idealizes rather than satisfice or optimize because its objective is to change the system involved or its environment in such a way as to bring it closer to an ultimately desired state, one in which the problem cannot or does not arise; it is innovative oriented and it makes use of creative and participative approaches aspiring dissolution in the containing whole.

The last approach is based on the two first approaches, practical experience and scientific approaches, but he adds design, invention, creativity, participation, and facilitation. This is the art and science of problem solving.

The main purpose of this paper is to reflect, elaborate and document about how the concept of "the art and science of problem solving" can be used in the real world to deal with important problematic situations in Society. Here, the facilitator is both the artist and scientist supporting a group to deal with a mess. As a scientist, he will be using when needed scientific approaches, experimentation, simulation and mathematical modelling in the problem solving process. As an artist, he will metaphorically speaking be like a painter who combines colours and shapes (the participants in the process) to create an art work (the problem solving process). Or, the facilitator is the director of a theatre performing a piece of art.

Art and Science are both historically and culturally laden concepts. Modern concepts of Art and Science will be discussed in Sec. 2 and 3, respectively. Sec. 4 will elaborate on the differences between Art and Science while Sec. 5 discusses their similarities. In Sec. 6 the different elements of the facilitation of problem solving will be presented as well as the art and science of problem solving. Our discussions of Sec. 6 will be more concrete by presenting a case study: the design of a decision support system for the planning of oral examinations in high schools in Denmark; this will be the subject of Sec. 7. In section 8 other case studies will be presented and finally in the last section, Section 9, the conclusions are depicted.

2 What is Art?

The answer to this question is conditioned by the fact that a definition of art has changed due to cultural and historical reasons. The boundaries of art have experienced a radical change over the last century. Previously, art was created in historically validated media and presented in a limited set of contexts for a limited set of objectives, such as search of beauty, religious glorification, or the depiction of persons and places. However, this century has produced new ways of experimentation, breaking and testing of boundaries. Artists have introduced new media, new contexts, new materials and new purposes. The art institutions have assimilated much of this experimentation, some of them depicted in the following list:

- Abstract painting (Pablo Picasso)
- Ready-mades (Marcel Duchamp)
- Interventions in non-art settings (Superflex)
- Performance art (Tracey Emin)
- Use of industrial materials, products and processes (Andy Warhol)
- Conceptual art (Joseph Beuys)
- Land art (Robert Smithson)
- Interactive art (Olafur Eliasson)
- Public art (Diego Rivera)

- Video art (Peter Land)

This expansion in art activities causes a difficulty in achieving consensus on definitions of art. The following very general definition can be easily accepted:

Making art may be depicted as the process of responding to perceptions, feelings, ideas, dreams, and other experiences by creating innovative works of art through the skillful, thoughtful, and imaginative application of tools and techniques to various media and materials. The "objects" of art that result of encounters between artists and their intentions, their interventions, their concepts and attitudes, their cultural and social realities, and the materials or media in which they choose to work.

Modern artists use unorthodox materials, tools, techniques and ideas inspired by the worlds of science, technology, humanities, economics, psychology, sociology, anthropology, etc. Some are present in non-art contexts, such as factories, laboratories, trade shows, the Internet, schools, and the street. Social interventions are manifold. The process of creating art is fill up of problems related to design and decision-making. The design attitude is related to the creative and innovative process in problem solving, while the decision attitude is related to the scientific approach to problem solving. In this sense, science can support art both providing materials and the media, and rational approaches to problem solving.

3 What is Science?

Researchers and philosophers on science suggest several defining elements. This set of core ideas, *the scientific approach*, includes the following:

- An essay to understand how and why phenomena occur
- Focus on the real (natural, social, human) world
- Focus on empirical information
- Seeking objectivity
- Use of a rational or logical approach
- Knowledge codify into laws and principles, and
- The continuous testing and refinement of hypotheses.

The crucial assumptions of the scientific approach are that the observed world is essentially orderly, and objectivity can be achieved through self-discipline and the reliance on methods such as the calibration of instruments, repeatability and multi-observed verification. There are of course variations in emphasis. That is, empiricists focus primarily on the role of observations, while rationalists emphasizes on the logical processes of theory construction and derivation. Some enhance induction built from observation; others focus on deduction drawn from theory.

Critical scientists see science as a modern delusion, challenging mainly the possibility of objectivity, noting the decisive influences of gender, social position, culture and history. Critical science is focusing in issues such as the interactions of the observer and the observed phenomena; the role of socially constructed frameworks at all stages; and the social forces and metanarratives that form the questions and paradigms used in the research process.

Several researchers have contributed to the critique of science. One describes the way dominant paradigms shape the questions that get acceptance and support. Another critiques assumptions of scientific rationality, remarking that nature gives different answers when approached differently. Others analyze the metaphoric language of science, its authoritative voice, and its unacknowledged patriarchal underlife.

In social sciences and the humanities, this kind of critique predominates. Scientists and technological innovators, however, believe in the ability to discover universal truths and assert that reform can overcome those places where scientific process falls short of its aspirations to universality and objectivity. As validity, it is usually referred to the accomplishments of the rational approach in building robust theoretical structures, and in predicting and controlling the material, organic and social world.

4 Art vs. Science: Differences

Box 1 depicts the main differences between art making and science making. Einstein (1934) has stated that the artist and the scientist each substitute a self-created world for the experiential one, with the purpose of transcendence. The main difference is that the artist is guided by an "artistic attitude" while the scientist is guided by a "scientific attitude". These attitudes are characterized in Box 1.

Artists are reflective and intuitive persons; materializing and visualising subjective experiences; and breaking the boundaries and traditions. Scientists are logical and rational persons; formulating verbally objective theories and principles; and seeking to improve and optimize. Scientists in their work are usually problem solvers, which are selecting course of actions that is believed to yield the best possible outcome. Artists are usually problem dissolvers, which are changing the nature and the environment of the system where the problem is imbedded so as to remove the problem. These differences are not exclusive; this means that sometimes the artists will be working as scientists and vice versa, during their working process and problem solving process.

5 Art vs. Science: Similarities

Box 2 describes the main similarities between art and science. Both value the careful observation of their environment to gather information through their senses. Creativity and innovation play a central role in both activities. To introduce change or improvement over the existing is of great concern to artists and scientists. Artists as well as scientists work with abstract symbols, representations for various realities and working tools. Even the language used by the two groups can be similar. Scientists working with mathematical models sometimes describe a particular good explanation or result as elegant or beautiful. The intellectual

Differences:

Art	Science
• Aesthetic, reflective	• Know, understand
• Emotion, intuition	• Reason, logic
• Idiosyncratic, personal	• Normative, principles
• Visual, sonic	• Narrative, textual
• Evocative, subjective	• Explanatory, objective
• Radical change	• Improve, optimise

Box 1: Art vs. Science: Differences

bridge of abstraction and aesthetic consideration is fundamental for both groups. Finally, both aspire to create works that have universal relevance.

These similarities provide many interest fields where art and science can support each other in their working processes. One of them is the interaction between creative and rational processes in real life problem solving. Another is artists being the facilitators of problem solving processes in scientific approaches. Modern artist are also inspired by new developments in science and technology, for instance the exhibitions of Olafur Eliasson (2004). Modern management scientists and researchers of problem solving approaches are also moving from the classical scientific approach towards a more artistic and design oriented attitude that is needed when innovation is needed in an organisation. In a modern world everything can become art, and everything can be studied scientifically.

Art vs. Science: similarities

- Observation, experimentation, sensual
- Creativity
- Change, innovation, improvement
- Models, symbols, abstraction
- Universality

Box 2: Art vs. Science: Similarities

6 The Facilitation Process

Let us see now the different elements of the facilitation process when applied to problem solving of real-life problems.

6.1 The Scene

The point of departure in our discussions is the concept of an *organization.* An organization can be a family, a community, a corporation, or an institution. What characterizes organizations is that they are purposefully designed and specialized to achieve a task. Thus an organization in a community could be a centre designed to enforce the development of the region, while firms are organizations providing some products and profits, and institutions are organizations designed to provide some services. The evolution of organizations is conditioned by external and internal factors, and sometimes organizations are experiencing *problematic situations or messes,* that are complex situations where some purposeful action is demanded to achieve some goals and visions. Problematic situations are usually related to the introduction of new technology, the re-design of the organization, the development of new strategies for the organization, the formulation of new visions for the future, or problem solving in general. In such a situation, the organization will usually appoint *a work group* to deal with the problematic situation. The *task* of this group is to analyze the mess and answer the question: What is to be done? In other words, to propose *actions plan* to be approved by the *decision-makers* of the organization. In small organizations the decision-makers (managers) are usually part of or identical to the work group. Related to these persons we have the so-called *stakeholders*, those individuals outside or inside the organization that can either affect or be affected by the actions plan. Let us see two examples to clarify the above-mentioned concepts.

Example 1: A small firm

The organization in question is a small firm in a retail business. The problematic situation is to what extent to engage in e-businesses as demanded by the bigger partners in the supply chain and what will eventually be the configuration of the technological platform to be used. The situation is also problematic because the organization has neither the technological background to identify different technological alternatives nor the experience for dealing with problematic situations. Management (the decision- makers) has appointed a work group (one person) to deal with this mess. The stakeholders are: the shareholders, the suppliers and the different type of purchasers.

Example 2: Community Work

The organization in question is a Development Centre in Odsherred (DCO), a vulnerable local region of Denmark. This is an autonomous non-profit organization which main objectives are to strengthen, develop, and inspire all type of cultural, social, environmental, and commercial activities in the region. Local innovators, in close co-operation with the relevant stakeholders of the region, carry out projects. These projects as well as the DCO itself are financed through a mix of sources: public funds, private funds, sponsors, business activities, and LEADER+, an EU-program that supports development in vulnerable regions of the countries that are members of the EU.

The problematic situation is the development of common images of ideas, projects, visions, and objectives for the region in question. These visions and objectives will be used to select the projects to be supported by the LEADER+ program. The DCO's board (the decision-makers) appointed a work group to deal with this situation. The stakeholders are: NGO's from the region, the business community, trade unions, local innovators and officials from the different municipalities.

—

To deal with messes, it is recommendable for the work group to hire *a facilitator*. A facilitator will support the group in the *problem solving process*; he or she will secure that the problem solving process ends with an action plan. The facilitator is usually the manager of the problem solving process. The facilitator could also give some expert know-how or find out if some experts have to be hired to give specialized advice. Often, the facilitators as professionals have some technical expertise for instance within information technology, so that he or she could also be the expert. To perform his job as process manager, the facilitator uses some *approaches, methods and tools* that he finds suitable for the given situation. The approaches could be quantitative (hard), qualitative (soft), participative (critical), innovative (creative) or a combination of them (multi-methodology). To facilitate groups demand the ability to both design and manage problem solving processes, creating a pro-active atmosphere and synergetic effects. Fig. 1 summarizes all the elements and concepts discussed above that will be further elaborated in the rest of this paper.

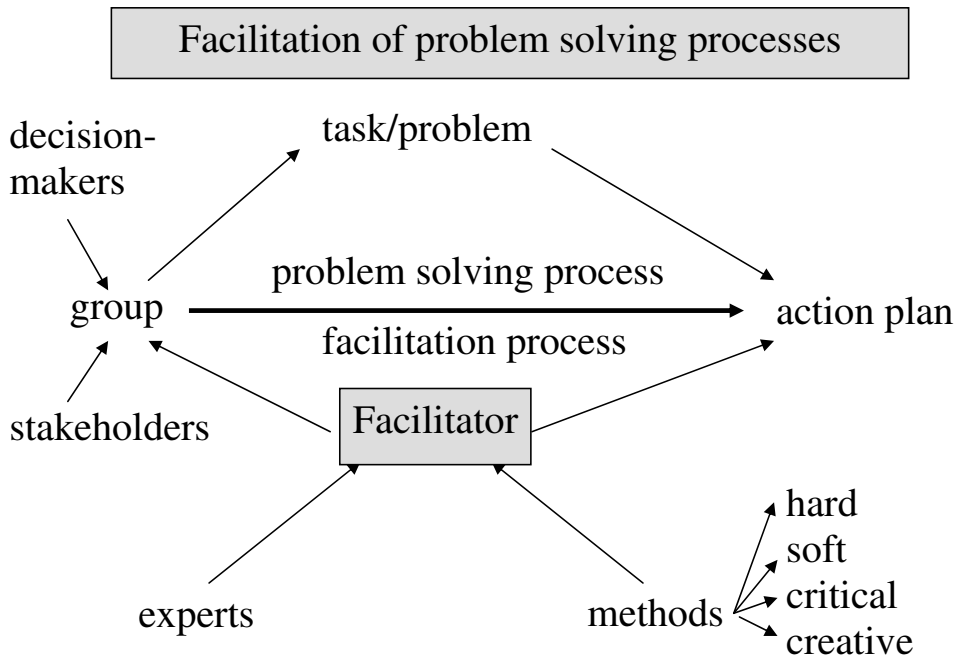


Figure 1: The facilitation process

Example 1 (continued)

In the above-mentioned example the facilitator was a student working in his MSc thesis to obtain a degree in Computer Engineering and Operational Research. The facilitator was also the technical expert. The problem solving process had a duration of around 3 months. The facilitator used several soft approaches during the problem solving process. The final product was an action plan elaborating the different realistic alternatives and a proposal for the decision-makers. The whole case study has been reported in Sørensen, Vidal, and Engström (2004).

Example 2 (continued)

The director of DCO contacted the author of this paper to support the organization and facilitation of a one-day Vision Conference. The purposes of this conference were both:

1. To generate visions and projects that would create a sustainable development of the region, and
2. To learn how to facilitate work groups, a tool that will be used during the implementation process of the LEADER+ program.

The facilitator designed and managed the Vision Conference where several creative techniques were used. The final result was a long list of potential projects. This will be used in the debates of the DCO's board while allocating funds to some selected projects. This case study has been reported in Vidal (2004a).

6.2 Social Interventions

In the two examples mentioned above, we have in principle two different kinds of social interventions. In the first one, denominated **research-driven intervention**, it is the facilitator as a researcher that takes the initiative to find a real-life case study for his MSc thesis. His objective is to test a problem solving approach and to evaluate the applicability of some methods. Obviously, the client or user will benefit by learning about the problematic situation, but there are not doubts about whose needs are ultimately driving the inquiry and helping process. This kind of intervention is quite similar to the type of interventions carried out under the name Action Research (Whyte, 1996). When Action Research was first formulated, it was clearly a case of the researchers wanting to figure out how to be more successful in implementing some changes that the researchers desired. It was found that by involving the targeted population in the research process, they became more amenable and committed to the desired change. But the initial drive came from the change agent and it was the change agent's goals that were driving the intervention. This research practice involves the client system in the researcher's agenda even though the client system might ultimately be the beneficiary. But the client did not initiate the process and it was not the client's needs that drove the process. It was the researcher's choice to involve the client.

The second example illustrates what is known as a **user-driven intervention**. The work group was composed of professionals covering different disciplines and with a lot of experience in problem solving within their own fields. In this case study, it is the client's needs that

are driving the inquiry and supporting process. During the problem solving process the work group will need support from experts, as for example the organization of a conference and the teaching of facilitation tools. In this mode of intervention, learning is a very important aspect of the problem solving process, because next time the users will organize a conference without an external facilitator and facilitation will become a tool in their future work. This form of intervention is usually found in the praxis of many consulting disciplines as for instance Management Sciences, Systems Sciences and Operational Research.

A third mode of intervention is denominated **participative intervention**, where both the work group and the facilitators co-operate and collaborate from the very beginning in the design of a problem solving process to deal with the problematic situation. This form of intervention is usually necessary when there is a need of both the practical experiences of the work group and the methodological and other expert knowledge from the facilitators. This mode of intervention can be regarded as a synthesis of the other two modes described previously.

Example 3: Public planning

It is my experience that my students start a social intervention as a research-driven form to be used as a case study for their MSc thesis work. Later, some of these studies evolve to a participative intervention seeking to develop and to implement the results achieved in the first intervention. This is for instance the case study reported in Hansen and Vidal (1995), the implementation of a computerized decision support system at The Ministry of Education in Denmark to plan the annual oral examinations at all the high schools of the country. This is a large-scale logistic and scheduling problem. The problem was solved using a whole system of hard methods developed by computer engineers under the facilitation of a systems analyst in close cooperation with a group of planners from the ministry. Since 1992, this system has been used in real-life planning. This case study will be further discussed in Sec. 7.

6.3 The Art and Science of Facilitation

The success of the problem solving process is determined by the effectiveness and creativity of the work group. Since the participants are invited or appointed, it is recommended to use some selection criteria. Some of these criteria could be: Representability, goal compatibility, process compatibility, deliberation abilities, positivism, communication abilities, and focus abilities. Obviously, the quality of performance or the piece of art created depends of the raw material you are using.

It is clear that selecting the participants is a very important task, which has to be solved seriously in order to develop effective group work and high quality results. A person, with knowledge and experience with working collaboratively with people, from the organisations involved should undertake this task.

In connection with the group work, there are two social processes to be managed by the facilitator: the problem solving process and the group process. The problem solving process is the way the group acts to solve the task supported by the facilitators and some experts. This is the rational and logical process. The group process is related to the manner in which the individuals in the group work together, how they learn, how they communicate, their social and power relationships, and how they deal with conflicts. This is the intuitive and creative

process. Obviously, these two processes interact in various degrees. In ideal group work, these two processes support each other. We talk about **group dynamics**, when energy and synergetic effects are created in the group work as a result of well-balanced processes where the task is just as important as the group trust and identity.

In addition, there is a third social process: the facilitation process. The facilitators are the managers of the social process and their main mission is to inspire, create, direct, and support group dynamics. By focusing and guiding group members' communication and decision-making processes in a creative and structured form, the facilitators can reduce the chances of engaging in faulty processes and harness the strengths of the group. The facilitator is both an artist, being the director of an artistic performance to be performed by the group, and a scientist, supporting a scientific approach to problem solving. This situation can be achieved using the following guidelines:

- Use approaches, for example creative techniques, and scientific methods;
- Specify a set of objective ground rules for the group work;
- Build on the strengths of the group and protect the group against its weakness;
- Balance members participation;
- Support the group with technological know-how;
- Support the group while dealing with conflicts;
- Plan time to close the different social processes;
- Make the group reflect and evaluate the group dynamics; and
- Empower the group.

The facilitators are constantly thinking (reflection) and listening to the deliberations in the group so they can make suitable interventions (decision making). An intervention means communicating with the group, giving information and knowledge, and encouraging the participants to think about important topics.

Let us elaborate now more theoretically about the essence of the facilitation process as opposed to its existence or its accidental qualities or, in other words, the attributes by means of which facilitation as management can be qualified or identified. As we have seen, facilitation is a purposeful process carried out by one or several persons that goes forward between two interacting processes. First, the logical/rational/legal process carried out by a purposeful group (the problem solving group) that wants to achieve some goals. This process has been called the problem solving process, and is the scene of objectivity. Secondly, the non-logical/irrational/illegal process that refers to the chaotic social process provoked by each single participant, by the participants' relations to each other, or by the participants' relations to the facilitator of the purposeful group, these bring into the participants own subjectivity, intuition, fantasy and feelings. This process can be called the problem destruction process and is the scene of subjectivity.

The facilitation process will move in the grey zone between the scene of objectivity and the scene of subjectivity. The rational and the irrational processes are fighting one another;

the one wants to impose over the other. They are in conflict with each other, but they need each other because while the problem solving process seeks to achieve realistic solutions, the irrational process will be the basis for the production of new ideas. Rationality needs chaos, and chaos needs rationality. Due to this contradiction, rationality vs. chaos, we can stipulate that facilitation is a **dialectical** process.

Let us also emphasise that facilitation is a purposeful intervention in a social process, a designed process. Facilitation is not a necessity for the evolution of the problem solving process but it is designed to support the problem solving process. The facilitation evolves very dynamically in a grey zone trying to construct a bridge between the traditional/conservative problem solving (business as usual) and the new/revolutionary power to change. The purpose of facilitation is to seek that the two above-mentioned processes do not destroy each other, but on the contrary support each other.

The facilitation process can be instructed and directed in different manners, as there are several management styles. The facilitators are the managers of this process. Note that if the group can manage itself, there is no need for a facilitator. That is, the group can learn to facilitate itself. As in any management process, it is a good idea to develop a strategy and design an action plan for the facilitation process. Managing by designing is a fundamental principle in any facilitation process (Boland and Collopy , 2004), therefore all the social processes have to be designed.

Management also involves three other central factors: Power, communication and learning. These aspects are always present in any facilitation process and should be reflected and articulated before, during and after the process. Facilitation becomes **an art** when a synergistic effect is achieved due to the constructive interaction between the rational and the irrational processes. The facilitator then becomes the director of a performance, where each participant plays a central role. By the end of the performance if synergy has been created all the participants will explode in a rush of happiness and pleasure, the pleasure of working creatively and collectively to achieve some goals. It is the same feeling that football players experience after a match where the victory has been the result of a combination of individual creativity, collective hard work and suitable facilitation (the coaching).

Summarising, we can state that the purpose of facilitation as management is not only to solve the task, but other additional goals could be:

- Each participant is a potential facilitator, therefore the importance of the learning dimension;

- Empowerment, the participants learn to be more self-confident and learn to work creatively in a group (creativity is an act of liberation from the jail of our own routines!); and

- Praxis, the facilitators should be able to learn from the experience therefore the importance of the evaluation of the intervention and the systematisation of praxis (Vidal, 2004b), in addition learning from failure is a good principle for any facilitator.

7 Case Study: Planning of High School Examinations in Denmark

In this section, the case study presented as Example 3 in Sec. 6.2 will be further discussed. This is a real-life logistic problem where a computer based support system has been developed and implemented. The system has been running at the Danish Ministry of Education since 1992.

7.1 Background

In Denmark, all planning of the official examinations at high school level is centralized at the Danish Ministry of Education. Denmark is the only country where such planning activities are centralised nationally. This cumbersome task had become increasingly difficult and time consuming due to educational reforms in 1998. In 1990, it was decided at the Ministry to develop a computer based decision support system to aid the ministerial planners in this planning process.

The Danish academic school system is divided into primary school (grade 1 through 9/10), high school (grade 10/11 through 12) and university/college, where primary school is the only compulsory school. High school, in the broad sense, has several channels, the academics as opposed to the technical or commercial high schools being the most attended ones. Approximately one half of all primary school graduates continue onto an academic high school.

The academic high school system has two major channels: The Gymnasium which is a 2 or 3 years package, 3 years being the most common, and higher preparatory school (HF), a two years package Through a system of merits, it is also possible to obtain an equivalent qualification through individual study-plans over several years (VUC). Denmark has 77 Gymnasiums, 25 HF-schools, 77 VUC-schools and 69 schools with both Gymnasium and HF curricula. This amounts to approximately 115,000 students and 12,000 teachers.

The students of the Gymnasium and HF are evaluated at the end of each school year. This evaluation includes oral and written examinations in certain courses. The planning of written examinations is much simpler since the days of examination are given before the start of the school year. This is necessary since all students answer the same examination questions and obviously they must do this at the same time. In what follows *examination* means oral examination. A *censor* is an eligible and ministerial appointed person - usually a high school teacher from another school – and an *examiner* is the person who conducts the examination – usually the teacher of the course.

An examination is carried out in the following way: A censor arrives at the school to observe the examination of each student conducted by the examiner for a fixed amount of time. After each student examination, the censor and the examiner agree on a grade for the student and then continue with the next student on the course, if any.

To encourage students to exhibit "good student behaviour", i.e. not miss classes, deliver term papers on time, etc., a bonus is granted in terms of a reduced number of examinations. Almost 95 percent of all students achieve this bonus. While a final year student could be examined in 7 subjects, "good students" will only have to attend 3 or 4 examinations. The

decision of which 3 or 4 subjects the student is to be examined in is drawn in private for each student and is not revealed until the last school day. Consequently, the student must prepare himself for all 7 subjects during the regular school year.

The examinations are gathered in a reserved 5 week period at the end of the school year from mid May to mid June. The Gymnasium only uses the last 3 weeks, except for final year students who also use the second week. First year HF-students use the last 4 weeks and VU-students and final year HF-students use all 5 weeks. Except for national holidays (which have a maximum of three whole days), the examination are placed Monday-Friday.

Previously, the examination planning was carried out by examination planners at the Ministry of Education using pencil and, especially eraser. Data was reported from each school on paper and sent by snail mail. In 1990, it was decided at the Ministry to develop an information system containing all relevant school data. The basic system is now an Oracle database with applications developed using Oracle tools and C-programming. Different systems are attached to the database, the examination system being the largest and most complex. A communication system handles the input of new data which is submitted from the schools to the ministry on floppy disks.

7.2 The problem and the approach

Summarizing, we can state that the task is to design and implement a computer based decision support system to plan and schedule the annual oral examinations for secondary education in the whole Denmark. For each student, it has to be decided:

- The number of oral examinations
- The subjects to be examined on
- The day, hour and room number for the examination
- The examinator, and
- The censor.

In practice, there are two main interrelated factors that determine the process of the solution of the above mentioned problem. The *technical approach*, i.e. the suitability of the techniques, methods, software, procedures, and so on, included in the whole decision support system, and the suitability of the *social process* related to the problem solving process itself. In Hansen and Vidal (1995), the technical approach has been described. The second factor demands close interaction and collaboration between the group work, decision makers, experts, consultants and facilitators. In this paper, we will primarily be focusing on the social processes though some aspect of the first factor will be shortly mentioned.

The planning problem described above is a complex and quite difficult combinatorial problem. It contains many decision variables; it has a variety of objectives and many feasible and satisfying solutions. We shall now elaborate on these observations.

Real life planning situations are usually complex. The examination planner has to comply with national laws and habits and must assist schools with their specific problems, making the

examination period as smooth as possible. Obviously, a computer system should support him in this task, rather than introduce additional limitations.

The examination timetabling problem is well known for its mathematical difficulty (Eiselt and Laporte, 1987). This is also true for the assignment problems related to our planning problem. Since a student will normally take more than one examination, a school may have as many as 1500 student examinations. Each student examination is to be scheduled on a specific day, which produces very many decision variables. This assignment problem will contain more than 100 million binary decision variables if formulated as a traditional 0-1 optimization problem.

Having multiple objectives is an ingrained feature of real life problems. These criteria involve a good spread of student examinations so as to provide good premises for each student, minimising the costs for the schools, the counties, and the Ministry, and sharing pedagogical benefits equally among the schools, subjects and geographical areas.

After experimenting with prototypes containing preliminary algorithms, it was concluded that finding feasible solutions did not present major difficulties. Finding satisfying solutions was more difficult but was still consider being attainable within reasonable amount of algorithm construction, system implementation effort and computational time. No demands for achieving optimal solutions were given whereas robustness and consistency were considered to be more important. This is in line with the following heuristic principle: Managerial decisions might be improved more by making them more consistent from one time to another than by approaches seeking optimality to explicit cost models; especially for situations where intangibles must otherwise be estimated or assumed.

These observations led to the conclusion that the final planning system should provide the examination planner with suitable information and optimising tools based in heuristic methods, which could be used interactively and that could be stopped at the users command yielding satisfying solutions.

To cope with the complexity of the problem at hand, it will be decomposed into four interrelated phases, each dealing with separate tasks and having well-defined goals following well-known heuristic principles (Silver et al, 1980). This decomposition approach follows to a certain extent the traditional approach (pencil and eraser) at the Ministry; this makes easier the final implementation process. This traditional approach was very time consuming for two planners with a lot of helpers. These four phases are:

- Subject draft
- Examination Chain
- Examination Scheduling
- Assignment of Censorships.

7.3 The work group and the stakeholders

The decision maker was the chief of the Examination Department at the Ministry. He is responsible that all the processes run smoothly. He played no major role in the development

of the decision support system. He gave his full support to the group work.

The work group was composed of three planners from the Examination Department at the Ministry. There experiences from many years of work at the Department were extremely useful while testing the different programmes solving each sub-problem. The leader of this group has a central position in the development of the decision support system because as a previous teacher in informatics, he has sufficient background to understand also the technical aspects of the problem and to contribute to its solution. He was at the same time the leader and a user and developer.

Stakeholders were of course the directors and teachers from the different schools that were involved in the discussions about the purpose of the new system, the first tests and the final implementation. The feedbacks from the stakeholders were important during the tuning of the whole system.

The facilitator was my previous student who had developed the technical approach in his MSc thesis, afterwards he was hired as a consultant for the Ministry. He was the facilitator of the whole development and implementation processes. As we will see below other experts were involved. He will seek for the collaboration of the users, the stakeholders, and the experts at the different stages of the development and implementation of the system.

Other experts were: One system's designer from a consulting firm and three programmers hired at the Ministry.

7.4 The facilitation process

In this case study the facilitator has two main tasks:

- First, to design, develop and implement a computerized decision support system in close cooperation with the users and other experts. As described above a satisfying system was developed by decomposing the complex problem in a series of interrelated optimization sub-problems each of them being solved using simple, fast, and reliable heuristic methods. Here the facilitator is working as a scientist using rational approaches, mathematical modelling and algorithms to find satisfying solutions and using the scientific approach to manage the problem solving process.

- Secondly, the facilitation of the group work and the work of the experts in the development and implementation stages of the problem solving process. This was a long process, it started in 1991, was used for the first time in 1992, and has been running every year since 1993. The task of the facilitator was to develop an efficient and innovative form of work, a common culture, a positive way of solving conflicts and a creative manner of finding new ideas. Here, the facilitator is working as an artist, he is instructing, directing, and coaching people to be participative, collaborative and creative in the problem solving process. He is like an instructor of a play in a theatre, supporting the different artists to perform their best to create synergetic processes. Or, more metaphorically, he is like a painter were all the participants are his colours to be combined in shapes, shadows and forms to be able to create a master piece.

The technical approaches needed to deal with the above described complex situation are relatively easy to develop. Similar complex logistic problems have been previously solved using mathematical models and heuristics and special dedicated computerized systems.

The real complexity of the problematic situation in question is the social complexity related to the development and implementation of the system by the actors in a participative and collaborative way. It is very complex the management of these social processes. Here the manager, that is the facilitator, is not only a rational and intelligent decision-maker, but also a creative and artistic designer. This managing attitude, managing as designing, is found in architecture, art and design professions.

Of course as with any practical project there have been conflicts, delays, and other problems related to negativity of some of the users or programmers leaving the Ministry; but in spite of the facilitator's lack of practical experience, he and the leader of the working group believed that it could be done and were highly motivated to do the task. The system has now been used for 14 years in practice. This has been a great success. For the Ministry, the examination system is the most prestigious system since the examinations have intensive attention from the schools, the public and the politicians; if things go wrong, from the press! Fortunately most people, including many students and teachers, are not aware of the existence of such a decision support system.

8 Other Case Studies

The facilitation processes do not need to take so long time and not all problems have to be solved using mathematical models, computers and experts. The case studies shortly described below show real life problem solving were the facilitator was working as an educator, a scientist and an artist.

8.1 Community Facilitation (Whyte, 1996)

Noorvick is an arctic Eskimo community located somewhere in Alaska. The community is organised following the old traditions and some new ideas coming from the outside, the so-called developed world. The community has to deal with the authorities; located far away in a big city, in what concerns development, planning and transfer of resources. Many problems, conflicts and disputes had not been suitable solved. The result was misunderstandings and poor cooperation. These two cultures, Eskimos and public servants, were not able to communicate to each other. Meanwhile lack of work, lack of education centres, poverty, and resignation was the actual situation of the community.

One day a stranger, an artist, arrived to the community. He was a young man, who just finished his education in communication, filmmaking and video techniques. He did not want to make military service and applied to do some social work somewhere in the poor world. He expected to be assigned to some Caribbean island, to an on going project where he could use his artistic qualifications. The Eskimo community did not know about Tim's speciality and Tim did not know about the needs of the community. Tim was supposed to stay for two years.

In reality the artist knew nothing about Eskimo culture. Coming from a warm region, he

had to adapt to cold, snow and ice. He started learning some Eskimo language. He was a fast learner and his natural open-mindness to something new helped a lot.

After this first period, that probably took a couple of months, the artist became an explorer. He formulated the following questions: How can I contribute in the best way with my qualifications to the development of this community? Can I?

After some talks with the leaders of the community and some of the authorities he was able to identify the central problem: The lack of a constructive and cooperative communication between these two cultures. What ought to be done? Both parts were willing to do their best, but they did not know what to do. The lack of trust to each other was deteriorating the situation.

After some incubation time, one night the artist woke up after a fantastic dream. In his dream the conflict was dissolved and a fantastic cooperation between the community and the authorities had begun. He knew exactly what to do. It was 3 o'clock in the morning. He went to his working table and started outlining his ideas. By 6 o'clock he went to bed again, but he could not sleep. He was anxious to meet the leaders of the community to discuss his proposal.

His proposal was: To use the art of filmmaking and the video technology to solve the communication problem. The wishes, problems and ideas of the community would be filmed and sended as videotape to the authorities. The authorities would look at the videos and start a dialogue process to achieve consensus about some proposals.

The leaders of the community were positive about the idea, they wanted to give it a trial but a central question was formulated: Who was going to make the videotapes? The leaders concluded: it had to be the artist; he had the expertise. He did not agree he did not like the idea that the community would be depending on his expertise. What would happen when he leaves the community? The other problem was the language, the films had to be made in the local language and thereafter professionally translated to English, and he was not able to do it in the local language.

Some persons from the community were selected; they formed the group who was going to learn about facilitation, the art of filmmaking and video technology. To select the persons was not an easy task, all the people wanted to learn the art of filmmaking! The artist elaborated a textbook and a series of lessons and workshops. His education program was following two principles: learning by doing and dialectical thinking. The students were very fast learners and highly motivated.

To make the story short, here is the final evaluation: The artist's ideas and actions were a complete success. His students learned to make films to communicate the needs and problems of the community. In his final report he wrote: *The villagers had established ownership of the technology and learned how to facilitate discussions with other villagers and with the authorities.* He was happy but a lot still has to be done. He resided in Noorvick for 11 years.

8.2 The Vision Conference (Vidal, 2004a)

Nykøbing is a region in Zealand, part of the kingdom of Denmark. This region has a beautiful nature, nice beaches and many historical sites. During summertime there are too many people. During wintertime there are not so many people, only the old and the very young. There are

not so many possibilities for working or studying in this region. The administrators of EU consider this region vulnerable. There are many vulnerable regions in the whole Europe. The administrators want to support economically the development of these types of regions in the whole Europe.

Economical support was conditioned by the demand that the local people organise themselves in a NGO, including all the relevant stakeholders of the region. The wise people of the region created an organisation: DCO. Many wise and experienced people were involved. They found out that the region had very solid human resource capabilities and multidisciplinary know-how. The problem was how to develop visionary strategies, projects, and actions for the future that would push forward a sustainable development. Creative ideas and innovative projects were badly needed.

The director of DCO invited a multi-disciplinary visual artist for a meeting. The artist had developed a concept called Vision Conference, a combination of creative workshops and facilitation processes, arranged for a large group of people. The main purpose of such a conference was the creative production of visions, ideas and projects to innovate and develop a region or an organisation.

At the meeting the artist explained his work in elementary terms. He emphasised that he could be regarded as the instructor of a performance at an improvisation theatre or a leader of a jazz orchestra. He pointed out that his goal was to get the best of each participant in creative and collaborative terms. Finally, by the end of the meeting it was agreed to organise a one-day Vision Conference for the region. A plan for the conference day was outlined. Four of the artist's students were hired as facilitators and four persons from DCO were appointed as helpers to the facilitators. These last four persons should be the organisers of future conferences for the region.

The Vision Conference was an extraordinary experience for everybody. There were forty people; haft men halt women. It was like being part of a jam session, everybody playing a different instrument. It was at the same time chaotic and well planned. The four facilitators made a fantastic job guiding each of the four groups. A lot of energy was released; you could feel that at the plenary sessions.

In the morning, all the groups worked freely and collaboratively creating ideas for projects that could develop the region according to the guidelines elaborated by DCO. Quantity was demanded. During the afternoon, the groups selected some of the most promising ideas for further elaboration and design of an action plan. Quality was demanded. At the end, the artist made a short evaluation; he was very pleased with the performance of everybody. The participants were tired but very happy. Most participants had not before experienced situations were you were working seriously and having fun at the same time. Somebody remarked that this conference was a big exercise in creative communication.

Some days latter, the artist delivered a report with the achieved results, an evaluation of the whole conference and a list of tasks to be carried out by the leaders of DCO in their future strategic work. He ended his report with his favourite quotation*: Our task is to make the familiar strange and the strange familiar.*

8.3 Facilitation in a Firm

Rational, a modern consulting firm was located in Copenhagen. Three engineers started the firm twenty years ago. Now the firm employs more than thirty people with very different specialities. But they have something in common: They are highly qualified and super rational consultants. Rational is an international firm; they sell knowledge, know-how and analyses. Knowledge about the applications of IT and economic and technological analyses are Rational's specialities

At a meeting of the Board of Directors, where future strategies were discussed some concern was expressed about the fact that the firm and the employees were not very creative. They were very good solving problems in a rational way, but they were not good if creativity and innovation was demanded. One of the directors mentioned that in the future many of the tasks they would get would demand such qualifications. In addition, it was remarked that the employees were not very good at looking for new opportunities, new tasks, new costumers, etc. Moreover, the employees were not good working together in projects in a collaborative way; they had to learn to cooperate.

At the next meeting one of the directors informed that searching at the Internet; he had found a multi-disciplinary artist that was specialist in the facilitation of creative processes. Some firms had used him as an adviser with very positive results. Moreover, he was also giving courses about creative problem solving. It was agreed to invite him for a meeting.

During that meeting, the artist registered that the director was a very rational person, but he had a serious interest in visual art and music. The director was a collector, but he dared not to practice and develop his artistic tendencies. They agreed about starting slowly a kind of cooperation. Several meetings were planned where the artist would be a kind of adviser about how to introduce more creativity and innovation in the firm. The artist emphasised that he could teach the employees about artistic thinking to complement their rational thinking, and dialectical thinking in order to synthesise these two thinking modes.

Later, the artist participated in some projects as a mentor for some of the experts from Rational. A course was given for the employees of the firm focusing on management and the facilitation of creative processes. Both the artist and the director were very positive about this form of cooperation. Some of the employees were using creative tools in some projects. Moreover, the director and some of the employees discovered that the artist was also a painter, and some paintings were commissioned.

The director wanted to go further in the cooperation; he wanted to hire the artist full-time. The artist refused, he still wanted to have some time to paint, to make installations and sculptures. Then the artist proposed to the director to adopt his concept of Vision Conference as a way of working. He was interested in the idea that his concepts would be diffused to the whole world. Two of the artist's students, two engineers, were employed in the firm.

One Saturday in a nice hotel, a one-day Vision Conference was run for all the employees of Rational. The purpose of the conference was twofold:

- To learn how to organise such a workshop for a large group, and
- To discuss ideas for a knowledge policy for Rational.

The Vision Conference was both a failure and a success. The participants learned to plan, organise and carry out a Vision Conference. But they were not able to develop creative ideas and actions for a knowledge policy. They had too many mental locks.

Some days later, the artist was writing his final report. He outlined some suggestions of activities for the future specially focusing on knowledge management. He concluded*: The main condition for creative development is motivation of the participants.*

9 Conclusions

Everything can de approached scientifically and everything can become art. Our main message is that in what concerns problem solving in complex situations, it is advisable to use both the scientific and the artistic attitudes. More satisfying results will be achieved, the risk of failures will be minimized, all the participants will be empowered, and everybody will learn from the experience, even the facilitator.

In the case of the planning problem, the Ministry could have ordered the decision support system from a firm instead of in-house development. But in such situation the consequences of failure were too serious and could easily become a political issue. In Denmark, there are too many bad experiences with implemented computerized decision support systems that were extremely expensive to develop and implement and that did not solve the problem, on the contrary caused more problematic situations.

In the case study related to the planning of the examinations the facilitator was educated as an engineer, but in the social process he was managing he was an artist although he was not aware of that. He used his intuition to solve conflicts, supervised the experts and used time to dialogue with the users. He was able to create a common language, a common culture and motivate all participants. He was managing by designing.

In the other case studies, the facilitators were both professional artists and scientists, being able to use these complementary approaches when it was needed.

10 References

Ackoff, R. L. (1981) The Art and Science of Problem Solving, Interfaces, Vol. 11, No. 1.

Boland, J. Richard Jr. and Collopy, Fred (eds.) (2004) Managing as Designing, Stanford Business Books, Stanford University Press, California, USA.

Eiselt, H. A. and Laporte, G. (1987) Combinatorial Optimization with Soft and Hard Requirements, Journal of Operations Research, Vol 38, No 9, pp. 785-795.

Eistein, A. (1934) Essays in Science, Philosophical Library, NY.

Eliasson, O. (2004) Minding the World, ARoS Aarhus Kunstmuseum, Aarhus, Denmark

Hansen, M. P. and Vidal, R. V. V. (1995) Planning of High School Examinations in Denmark, European Journal of Operational Research, Vol. 87, pp. 519-534.

Nelson, H. G. and Stolterman, E. (2003) The Design Way: Intentional Change in an Unpredictable World – Foundations and Fundamentals of Design Competence, Educational Technology Publications, Englewood Cliffs, New Jersey, USA.

Silver, E. A., Vidal R. V. V., and De Werra, D (1980) A Tutorial on Heuristic Methods, European Journal of Operational Research, Vol 5, pp. 153-162.

Sørensen, L., Vidal, R.V.V., and Engström, E. (2004) Using Soft OR in a Small Company – The case of Kirby, European Journal of Operational Research, Vol. 152 pp. 555-570.

Vidal, R.V.V. (2004a) The Vision Conference: Facilitating Creative Processes, Systems Practice and Action Research, Vol. 17, No 3, pp. 385-405

Vidal, R.V.V. (2004b) From Action to Learning: The systematisation of alternative consulting experiences, AI&Society, Vol. 18, No 2, pp. 134-148.

Whyte, F.W. (1996) Emancipatory practice Through the Sky River Project, Systems practice, Vol. 9, nr.2, pp. 151-157.

A Gestão Logística dos Resíduos em Portugal

Ana Paula Barroso * Virgínia Helena Machado *

* Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa
{apb, vhm}@fct.unl.pt

Abstract

Following the communitarian politics, Portugal has to implement wasted management politics namely the recovery logistic management systems. This paper is focused in the analysis of the effects of the introduction of the discarded products at the end of their service life recovery in the traditional logistic chain. This will be analyzed in that it concerns to the reverse logistics system configuration, being characterized the associated critical factors of some wastes. The implementation degree of directives of the European Parliament and of the European Union in Portugal will be analyzed specifically for the Waste Electrical and Electronic Equipment and End-of-Life Vehicles. The critical factors that constrain these politics implementation will be evidenced.

Resumo

Para dar sequência à política nacional de gestão de resíduos e, consequentemente, à política comunitária, têm vindo a ser desenvolvidos e implementados em Portugal sistemas de gestão logística de resíduos. Neste artigo, serão analisados os efeitos da introdução da recuperação de resíduos na cadeia logística tradicional no que concerne à configuração do sistema de logística inversa, sendo caracterizados os factores críticos que lhe estão associados. Será analisado também o grau de implementação das directivas do parlamento Europeu e do Conselho da União Europeia em Portugal, para alguns resíduos, especificamente os Resíduos de Equipamentos Eléctricos e Electrónicos e de Veículos em Fim de Vida, evidenciando factores críticos que condicionam a sua implementação.

Keywords: Supply Chain Management, Reverse Logistics, Waste Management

Title: Recovery Logistics Management in Portugal

Apdio © 2005 *Associação Portuguesa de Investigação Operacional*

1 Introdução

A última revisão do Plano Estratégico dos Resíduos Industriais, realizada em 2001 pelo Instituto dos Resíduos sob a alçada do Ministério do Ambiente e do Ordenamento do Território, fixa os principais objectivos do plano em conformidade com as quatro linhas mestras dos objectivos estabelecidos para a política do ambiente, constantes do "Plano Nacional de Desenvolvimento Económico e Social 2000-2006": (i) gestão sustentável dos recursos naturais, (ii) protecção e valorização ambiental do território, (iii) conservação da natureza, protecção da biodiversidade e da paisagem e, por último, (iv) integração do ambiente nas políticas sectoriais e de desenvolvimento local e regional.

Os objectivos discriminados para cada uma das linhas mestras, no âmbito da gestão logística dos resíduos, passam (i) pelo aumento das taxas de reutilização e reciclagem, (ii) programação da fase de infra estruturação básica que integra a recolha, transporte, tratamento e destino final em aterro, e (iii) desenvolvimento de novos modelos de gestão dos resíduos.

O Decreto-Lei n° 237/97, da Assembleia da República, define resíduos, em conformidade com o Catálogo Europeu de Resíduos (CER), como sendo "quaisquer substâncias ou objectos de que o detentor se desfaz ou tem intenção ou obrigação de se desfazer".

De facto, nesta definição de tão vasta abrangência, podem estar incluídos, para além dos produtos em fim de vida útil, os produtos danificados, avariados e em fim de utilização. Como um produto pode ficar impróprio para utilização/consumo durante o seu abastecimento entre o fabricante e o cliente final (por exemplo ter chegado ao termo de validade), em qualquer nível da cadeia de abastecimento pode haver necessidade do detentor proceder à sua devolução. O fabricante pode ter necessidade de devolver matérias-primas que não satisfazem a qualidade especificada, o distribuidor pode proceder à devolução do produto devido à gestão interna do seu nível de *stock* ou a acordos comerciais pré estabelecidos e, por último, o cliente pode querer devolver produtos que se avariaram e estão dentro da garantia, produtos em fim de vida útil ou, mesmo, em fim de utilização.

Com a melhoria do nível de vida, sobretudo nos países industrializados, tem-se verificado um aumento cada vez maior dos resíduos, em número e em quantidade. Até um passado muito recente, os resíduos eram eliminados, normalmente, através da deposição em aterro. Com os aterros sobrelotados e a escassez de incineradoras (em número e capacidade), têm sido envidados esforços no sentido de reintegrar os resíduos nos processos produtivos originais tendo em vista a minimização das substâncias a depor em aterro bem como a redução do consumo de recursos naturais. A reintegração dos resíduos nos processos produtivos permite um desenvolvimento mais sustentável, não colocando em risco as gerações futuras.

A preocupação quanto ao estado actual do ambiente e à tendência de degradação está patente em Directivas Comunitárias e Legislação Portuguesa, através de princípios e normas aplicáveis aos sistemas de gestão de resíduos. O cumprimento dos objectivos estabelecidos nas Directivas Comunitárias e Decretos-Lei, que visam principalmente a redução do consumo de recursos através da reintegração/aproveitamento dos resíduos, pode passar por alterações ao nível das filosofias de produção e de gestão dos fluxos de material bem como de informação.

A reintegração dos resíduos recuperados na cadeia de abastecimento, implicará um fluxo de material e de informação adicional, em sentido inverso ao tradicional, o que permite fechar o circuito. Assim, a cadeia de abastecimento em circuito fechado terá de englobar não só

as actividades logísticas tradicionais, abastecimento, produção, distribuição e consumo como, também, as actividades associadas à logística inversa, recolha, inspecção/separação, reprocessamento, deposição e redistribuição de resíduos recuperados [9].

Neste artigo, serão analisados os efeitos da introdução da recuperação de resíduos na cadeia logística tradicional no que concerne à configuração do sistema de logística inversa, sendo caracterizados os factores críticos que lhe estão associados. Será analisado também o grau de implementação das Directivas do parlamento Europeu e do Conselho da União Europeia em Portugal, para alguns resíduos, especificamente os Resíduos de Equipamentos Eléctricos e Electrónicos (REEE) e de Veículos em Fim de Vida (VFV), evidenciando os factores críticos, ao nível estratégico e operacional, que condicionam a sua implementação.

O artigo está estruturado do seguinte modo. Na Secção 2 faz-se uma breve alusão aos princípios e normas aplicáveis aos sistemas de gestão logística de resíduos. Na Secção 3 é abordada a gestão da cadeia de abastecimento na qual se evidência a importância da integração da logística inversa. Na secção 4 são identificados os factores críticos que afectam as actividades da logística inversa. Na Secção 5 é realizada a descrição, de um modo sucinto, dos objectivos fundamentais a alcançar com a implementação das Directivas da União Europeia (UE), 2002/96/CE e 2003/108/CE, que enquadram legalmente a gestão de REEE e é analisado o grau de implementação das directivas no nosso país. Na Secção 6 é realizada a descrição dos objectivos fundamentais a alcançar com a implementação da Directiva da UE, 2000/53/CE, que enquadra legalmente a gestão de VFV sendo, também, analisado o grau de implementação da Directiva em Portugal. Finalmente, na secção 7, são apresentadas as conclusões.

2 Gestão Logística de Resíduos. Princípios e Normas

O processo de recuperação de produtos é uma peça fundamental na cadeia de abastecimento tal como, actualmente, é definida. A gestão integrada da recuperação de produtos e da cadeia de abastecimento tradicional pode contribuir para a redução dos custos globais e do impacte ambiental conduzindo, assim, a uma gestão mais eficiente e eficaz de todo o sistema. Embora muitas empresas negligenciem a recuperação de produtos, normalmente por não valorizarem a sua importância, este processo pode ajudar a empresa a alcançar uma vantagem competitiva sustentável.

Assim, tanto ao nível da UE como de Portugal, já existem princípios e normas aplicáveis a sistemas de gestão de resíduos que estão patentes em Directivas e em Decretos-Lei e que, fundamentalmente, imputam aos produtores a responsabilidade pelo ciclo de vida total dos seus produtos.

Qualquer Directiva da UE, que enquadre legalmente a gestão de resíduos a nível comunitário, visa, em primeiro lugar, a prevenção dos resíduos, em segundo lugar, a sua reutilização, reciclagem e outras formas de valorização e, em terceiro lugar, a minimização dos riscos e impactes ambientais decorrentes do seu tratamento e eliminação. Deve contribuir também para a harmonização das medidas nacionais em matéria de gestão de resíduos, por forma a assegurar o bom funcionamento do mercado interno. Os objectivos devem ser alcançados por meio de uma vasta gama de medidas, nomeadamente em matéria de recolha separada de resíduos, seu tratamento e valorização.

A política comunitária no domínio do ambiente baseia-se nos princípios da acção preventiva de correcção dos danos causados ao ambiente, prioritariamente na fonte, e do poluidor-pagador. A ideia subjacente ao princípio da responsabilidade do produtor é a de responsabilizar, pela poluição do ambiente, as entidades que têm possibilidade de melhorar esta situação. Por conseguinte, são os produtores/fabricantes pois concebem o produto, determinam as suas especificações e seleccionam os seus materiais. A criação de um elo entre produtores/fabricantes e a gestão dos resíduos contribui para uma melhor concepção dos produtos, tendo em vista facilitar a sua recuperação, reciclagem e eliminação quando passam a ser designados por resíduos.

No entanto, em matéria de gestão de resíduos, as diferentes políticas nacionais de cada país da UE prejudicam a eficácia das políticas de reciclagem, dado haver a possibilidade de se verificarem movimentos transfronteiriços de resíduos para sistemas de gestão de resíduos menos dispendiosos e, por outro lado, diferentes aplicações nacionais do princípio da responsabilidade do produtor levam a disparidades substanciais nos encargos financeiros impostos aos operadores económicos.

A fim de reduzir os custos para os produtores resultantes da gestão de resíduos provenientes de produtos colocados no mercado antes da entrada em vigor de qualquer Directiva, normalmente, está previsto um período de transição após a sua entrada em vigor.

Em Portugal existe legislação que já estabelece um conjunto de regras de gestão, para determinados tipos de resíduos que visa, principalmente, a criação de circuitos de recolha selectiva, o seu correcto armazenamento e pré-tratamento, nomeadamente no que diz respeito à separação das substâncias perigosas neles contidas, e o posterior envio para reutilização e reciclagem, desencorajando, por conseguinte, a sua eliminação por via da simples deposição em aterro.

3 Gestão da Cadeia de Abastecimento em Circuito Fechado

Na última década, foram diversos os factores, de índole económica, legislativa e de cidadania, que contribuíram para o desenvolvimento de teorias, casos de estudo e modelação quantitativa relativamente à gestão da cadeia de abastecimento no sentido inverso ao tradicional, ou seja, da logística inversa. De facto, há quem invista na área da logística inversa porque se sente socialmente motivado para o fazer, para tentar antecipar-se à legislação, para aproveitar oportunidades financeiras oferecidas pelo mercado ou, mesmo, para obter uma imagem "verde". Especialmente na UE, tem havido um aumento da legislação ambiental que obriga as organizações a investir na logística inversa, nomeadamente no que diz respeito ao estabelecimento de quotas mínimas de reciclagem e valorização, regras relativas à embalagem de produtos e de responsabilização do produtor/fabricante quanto à retoma dos resíduos provenientes dos produtos que produz/fabrica.

3.1 Logística Inversa

Associado à crescente consciencialização das pessoas e instituições para os problemas de índole ambiental e económica, à preocupante escassez de matérias-primas a nível mundial e ao aumento da procura de produtos "amigos" do ambiente, surgiu a necessidade de gerir um fluxo

logístico inverso ao tradicional (entre o ponto de consumo final e o ponto de origem), que deu origem ao conceito de logística inversa.

A missão da logística inversa é, por conseguinte, planear, implementar e controlar de um modo eficiente e eficaz a recuperação de resíduos preconizando a redução do consumo de matérias-primas, recorrendo a meios como a reciclagem, a substituição e a reutilização de materiais, a deposição e a reparação e refabrico de produtos, fechando o circuito da cadeia de abastecimento. Alguns autores atribuem a designação de inverso apenas quando os sentidos dos fluxos inverso e tradicional partilham o mesmo canal de distribuição ([2], [7] e [10]). Outros autores são menos restritivos admitindo na fase de recuperação dos resíduos a utilização de diferentes canais de distribuição ([9], [17] e *Reverse Logistics Executive Council*), sendo esta a definição adoptada pelas autoras.

3.2 Actividades de Logística Inversa

Os produtos a recuperar podem ser introduzidos no fluxo inverso por diversas razões: devoluções ao nível do fabricante, devoluções comerciais (B2B e B2C), "chamadas" de produto, retornos do produto dentro da garantia, produtos em fim de utilização e produtos em fim de vida útil.

Dependendo do estado do produto, das obrigações contratuais com o vendedor e da procura, as empresas podem ter várias opções quanto ao modo de recuperar os produtos [14].

Em qualquer fluxo inverso, um dos objectivos cruciais é recuperar o maior valor possível dos produtos. Para atingir este objectivo, devem ser realizadas, sequencialmente, as seguintes etapas [9]: recolha, inspecção/separação, reprocessamento, deposição e/ou redistribuição.

A recolha compreende todas as tarefas que envolvem a recolha propriamente dita dos resíduos e a sua movimentação física para tratamento (aquisição, transporte e armazenagem). A inspecção/separação inclui as operações que dividem o fluxo de resíduos de acordo com as diferentes opções de recuperação e deposição (desmontagem/desmantelamento, prensagem, teste e armazenagem, ...). A transformação de um resíduo noutro produto, componente ou material novamente utilizável designa-se por reprocessamento. Nesta actividade está incluída a reciclagem, a reparação e o refabrico podendo estar também envolvidas as operações de limpeza, substituição e remontagem. A deposição é a actividade adequada para os resíduos que por razões técnicas ou económicas não são recuperados (pode incluir transporte, aterro e incineração). Por último, a redistribuição refere-se à colocação no mercado de produtos recuperados podendo incluir actividades de venda, transporte e armazenagem [9].

Assim, quando em presença de um fluxo inverso, deve ser decidido o que fazer com cada produto. Deve começar por se identificar o produto, avaliar o seu estado, decidir qual o modo de recuperação mais adequado e, após a recuperação, reintroduzi-lo na cadeia de abastecimento. Os produtos, peças ou materiais recuperados, não têm necessariamente de entrar na mesma cadeia de abastecimento de onde foram originários. Na figura 1, encontram-se representadas as etapas intervenientes na cadeia de recuperação de resíduos.

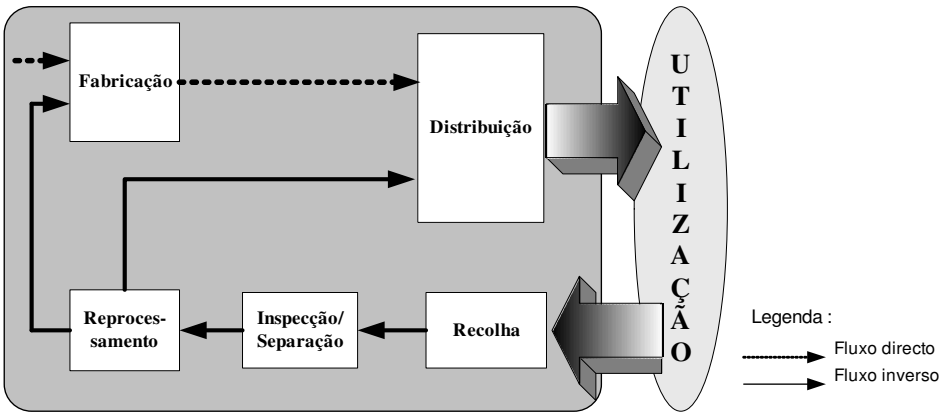


Figura 1: Cadeia de recuperação de resíduos

3.3 Sistema de Logística Inversa

A implementação de um processo logístico inverso (cliente-fornecedor) requer, tal como num directo (fornecedor-cliente), a definição de uma estrutura adequada para lidar com os fluxos de entrada (resíduos) e os fluxos de saída (produtos, componentes e/ou materiais recuperados). Assim, torna-se necessário desenvolver instalações de processamento e armazenagem de resíduos bem como sistemas de transporte que liguem, de um modo eficaz, os pontos onde os resíduos são recolhidos às instalações onde serão sujeitos a operações de tratamento, reprocessamento e/ou deposição.

Existe um conjunto de paradigmas que pode ser identificado num sistema de logística inversa. O mais relevante está associado à centralização e descentralização do sistema de gestão. Num sistema descentralizado, todas as decisões relativas ao destino a dar ao resíduo são tomadas ao nível do retalhista. Embora sejam evitados alguns custos de transporte, uma vez que os resíduos não são todos encaminhados para um centro de processamento central, o custo total de transporte dos produtos aumenta, uma vez que todos os produtos provenientes de mercados secundários estão dispersos na rede de pontos de venda da organização e, directa ou indirectamente, o retalhista tem de pagar o custo de recolha dos produtos.

A centralização da recolha dos resíduos é mais eficiente do que se for realizada por centros de distribuição [15]. Centralizar o fluxo inverso cria volumes maiores, o que promove não só a constituição de massa crítica para a aquisição de equipamento especializado, como também focaliza as actividades na logística inversa.

Na Europa, a responsabilidade de recolha dos produtos, tais como consumíveis electrónicos e automóveis, está atribuída aos fabricantes. Em colaboração ou sob imposição dos governos, têm sido criados sistemas de recolha e recuperação. Estes sistemas podem ser organizados a nível nacional quando envolvem a recuperação de volumes elevados de resíduos de baixo valor acrescentado (reciclagem de materiais) ou organizados a nível das empresas no caso da recuperação envolver volumes reduzidos de resíduos mas com elevado valor acrescentado (refabrico).

3.4 Integração da Logística Inversa na Cadeia de Abastecimento

O conceito de gestão da cadeia de abastecimento, apesar de relativamente recente, não é mais do que um complemento à lógica associada à cadeia logística. A gestão logística diz respeito, principalmente, à optimização dos fluxos dentro da organização, enquanto que a gestão da cadeia de abastecimento reconhece que a integração interna não é suficiente, sendo necessário estendê-la a entidades externas que se relacionam com a organização [3].

Em 1998, o *Global Supply Chain Forum* definiu a gestão da cadeia de abastecimento como a integração dos processos de negócio chave desde os fornecedores directos e indirectos, de produtos, serviços e informação, que adicionam valor aos clientes e outros actores da cadeia de abastecimento, até aos utilizadores finais. Deste modo, a gestão da cadeia de abastecimento focaliza-se na integração dos processos que ocorrem no sentido tradicional na cadeia de abastecimento. Por outro lado, a análise integrada de toda a cadeia de abastecimento é o melhor ponto de partida para a obtenção de soluções dado que procura a optimização global em detrimento da local [13].

A logística inversa, de um modo geral, tem sido considerada um módulo independente da cadeia de abastecimento tradicional. Focaliza-se essencialmente na gestão dos processos que ocorrem no sentido inverso na cadeia de abastecimento. A optimização dos processos que lhe estão afectos é realizada localmente. Se for integrada na cadeia de abastecimento pode contribuir para uma maior redução tanto dos custos globais de toda a cadeia de abastecimento como do impacte ambiental, apesar da sua complexidade aumentar.

Com a integração da logística inversa na cadeia de abastecimento, verifica-se que [19]:

- O planeamento da rede de transportes é mais complexo;
- Os prazos de entrega/aprovisionamento são maiores, uma vez que passa a ser necessário incluir o tempo associado à recolha dos produtos.
- É necessário maior espaço de armazenagem, uma vez que os produtos a recuperar têm que ser armazenados;
- São necessários sistemas de informação mais complexos pois têm de incluir o "seguimento/acompanhamento" dos produtos em recuperação; e
- É necessário mais formação dado que é preciso formar os trabalhadores para a realização das actividades associadas à gestão dos resíduos.

No entanto, as melhores soluções são feitas à luz da optimização global. Um programa de logística inversa adequado pode conduzir a economias significativas no custo de aquisição de materiais, deposição em aterro, manutenção de *stock* e transporte [13].

Na Figura 2 estão representados, esquematicamente, os principais processos associados à gestão de resíduos e as entidades constituintes da cadeia de abastecimento tradicional. A fase designada de inspecção/separação deve ajudar a determinar qual o processo de recuperação mais adequado a cada produto/componente/material a reintroduzir na cadeia de abastecimento tradicional, bem como indicar qual o tratamento prévio a realizar quando se procede à sua eliminação por deposição em aterro ou incineração.

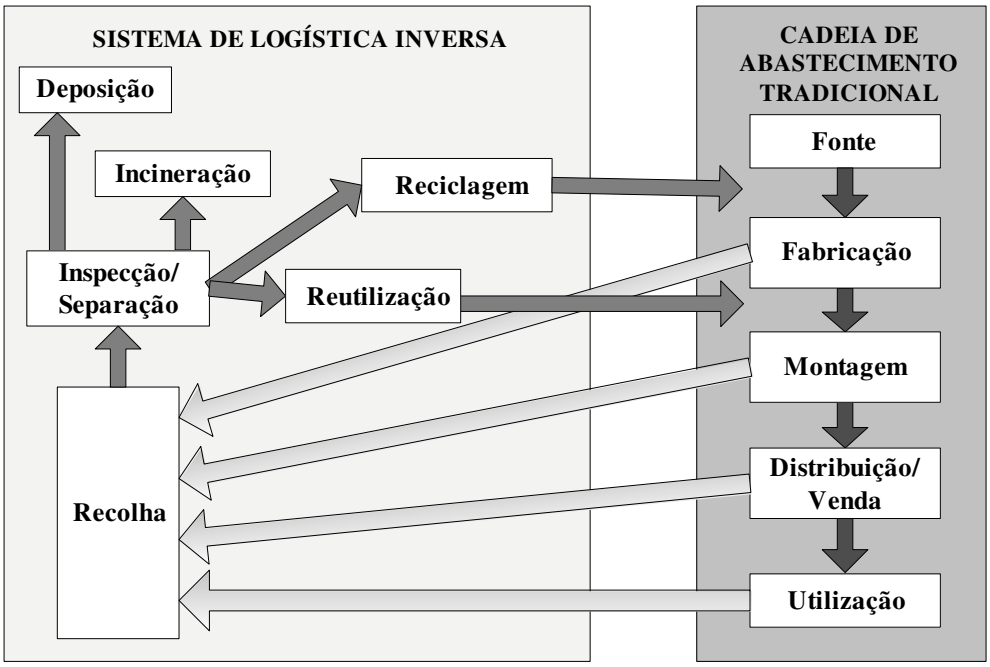


Figura 2: Cadeia de abastecimento em circuito fechado

4 Factores Críticos de um Sistema de Gestão de Resíduos

A logística inversa é importante como estratégia de negócio sustentável e lucrativa [7]. No entanto, existem diversos factores que afectam um sistema deste tipo.

Carter e Ellram desenvolveram um modelo de factores (internos e externos) que afectam a logística inversa [2] tendo por base a integração dos conhecimentos obtidos a partir de uma extensa revisão bibliográfica com uma ferramenta de análise de marketing [1]. Um outro autor, identificou mecanismos internos estratégicos e operacionais que afectam os sistemas de logística inversa [7]. São estes mecanismos combinados com o modelo de [2] que servem de base ao modelo conceptual dos factores críticos, que afectam os sistemas de logística inversa, e que se encontra representado na figura 3.

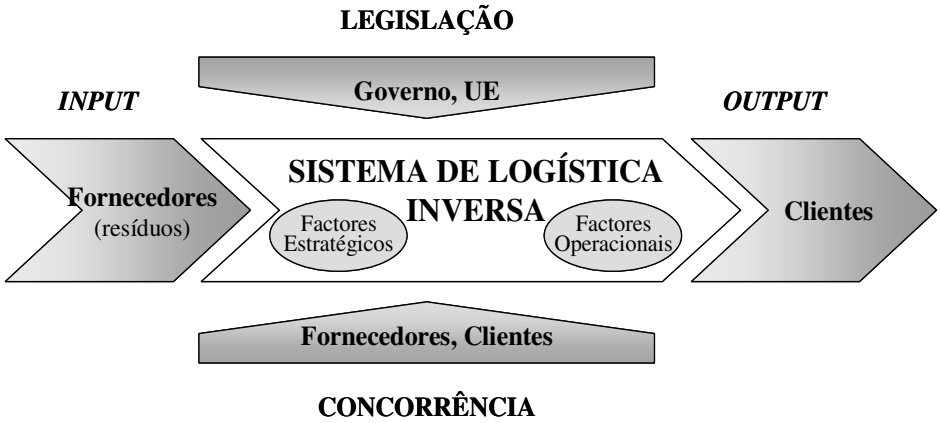


Figura 3: Modelo conceptual dos factores críticos que afectam o sistema de logística inversa

O modelo conceptual identifica os factores críticos externos que afectam as actividades inerentes à logística inversa de uma organização, bem como os factores críticos internos, agrupados em estratégicos e operacionais, a considerar quando da implementação de um sistema de logística inversa ([2], [7], [14], [15] e [16]). Como o modelo sugere, os factores externos sofrem a influência do ambiente em que o sistema está inserido, nomeadamente, *Input*, *Output*, Legislação e Concorrência.

i) Factores Externos

Para estes factores contribuem diversas organizações chave tais como fornecedores, concorrentes, organizações governamentais, organizações ambientais e consumidores.

O *Input* diz respeito aos potenciais fornecedores de resíduos. Qualquer sistema de logística inversa para ser bem sucedido precisa de um volume considerável de resíduos para recuperar [15] logo, será necessário fazer, numa primeira fase, a sua recolha. No entanto, é muito difícil suportar um sistema de recolha de resíduos economicamente viável atendendo à dispersão dos potenciais fornecedores. Relativamente à Legislação, há a referir que em matéria de legislação ambiental no âmbito da gestão de resíduos, existe uma dinâmica a todos os níveis governamentais. O factor externo designado por *Output*, compreende a procura de produtos recuperados, na forma de material reciclado e/ou produtos/componentes reutilizáveis. Por último, o factor Concorrência diz respeito ao nível de competitividade dos fornecedores de resíduos, clientes de produtos/componentes reutilizáveis e clientes de material reciclado.

ii) Factores Internos

Internamente, a empresa deve analisar os factores estratégicos e operacionais sendo os primeiros críticos. Nos factores estratégicos podem ser incluídos os custos estratégicos, a qualidade total, o serviço ao cliente e os aspectos ambientais e legislativos.

O sucesso de um sistema de logística inversa depende dos custos estratégicos que lhe estão associados. Estes custos podem incluir o custo do equipamento para desmontar/desmantelar os resíduos, o custo de qualificação dos recursos humanos para operar o sistema de logística inversa e o custo associado à necessidade de armazenagem adicional. Os custos associados à mão-de-obra, transporte e armazenagem, estarão dependentes do fluxo de *input* e de *output* do sistema. A custos estratégicos reduzidos corresponde, normalmente, uma utilização eficiente dos recursos, métodos e tecnologias ([7], [16] e [17]). Especificamente, a decisão estratégica de colocar um maior ênfase na reutilização do produto em detrimento da reciclagem deverá considerar custos superiores associados ao transporte e armazenagem.

Para qualquer sistema logístico ser eficiente é essencial a identificação e satisfação dos requisitos do serviço ao cliente ([11], [12], [16] e [18]). O principal requisito, de importância estratégica, do serviço ao cliente é a obtenção de um fluxo de produtos recuperados (*output* do sistema) consistente, standard e de elevada qualidade.

Os aspectos ambientais e legislativos considerados na definição da estratégia da empresa podem conduzir à redução de custos e a melhorias ambientais pelo facto de promoverem a utilização de recursos recuperados ([2] e [7]).

Genericamente, os factores operacionais influenciam os sistemas de logística inversa no que diz respeito à análise custo benefício, transporte, armazenagem, gestão do abastecimento, embalagem, bem como reprocessamento e desmantelamento. Estes factores nem sempre têm

igual importância devendo as empresas considerar cada um deles em termos da sua importância relativa [7]. A gestão do abastecimento, o reprocessamento e o transporte parecem ser os factores operacionais mais importantes.

O transporte é normalmente o maior custo associado à logística inversa, contribuindo em 25% ou mais para o custo total [15]. Pode ser identificado como um custo variável chave.

Tal como para o factor transporte, o tipo de *input* do sistema pode ter um forte impacto no espaço necessário para o armazenamento e manuseamento dos resíduos a recuperar.

A gestão do abastecimento para sistemas de logística inversa focados na reutilização de componentes e material reduz a utilização e os custos das matérias-primas [16], sendo necessário estabelecer relações de cooperação com fornecedores capazes de fornecer um *input* consistente e de elevada qualidade.

A embalagem para um sistema de logística inversa de qualquer tipo de resíduo tem, prioritariamente, o objectivo de protecção, principalmente quando se pretende seguir a via da reutilização em detrimento da reciclagem.

5 Os REEE em Portugal

5.1 Introdução

A produção de equipamentos eléctricos e electrónicos (EEE) é um dos domínios da indústria transformadora com um crescimento mais rápido no mundo ocidental. Tanto a inovação tecnológica como a expansão do mercado continuam a acelerar o processo de substituição. Esta evolução leva a um importante aumento dos resíduos de equipamentos eléctricos e electrónicos (REEE) que é imperativo gerir de um modo sustentado.

Em termos gerais, todos os equipamentos que necessitam de electricidade para funcionarem adequadamente são eléctricos ou electrónicos. Os REEE são constituídos por uma mistura complexa de materiais e componentes e representam cerca de 4% dos resíduos urbanos. O seu fluxo difere do fluxo dos outros resíduos urbanos por várias razões:

- O crescimento dos REEE é cerca de três vezes superior ao crescimento dos resíduos urbanos normais, prevendo-se que o volume aumente, anualmente, pelo menos 3-5%.
- Devido ao seu teor em matérias perigosas, os EEE provocam problemas ambientais importantes durante a fase de gestão dos resíduos.
- A problemática ambiental associada aos REEE, devido sobretudo à presença de metais pesados, é maior que a associada aos outros constituintes dos resíduos urbanos.

Geralmente, os pequenos REEE, que podem ser eliminados com os resíduos urbanos normais, são directamente incinerados ou depositados em aterros. A percentagem destas opções de gestão de resíduos difere substancialmente consoante o Estado membro da UE. Actualmente, mais de 90% dos REEE são depositados em aterros, incinerados ou valorizados, sem qualquer

tratamento prévio [8]. Este facto origina emissões consideráveis de substâncias perigosas para o meio ambiente.

O fluxo de REEE contribui significativamente para as substâncias halogenadas e metais pesados contidos no fluxo de resíduos urbanos, podendo verificar-se efeitos nocivos específicos durante a incineração. Assim, a recolha separada e o tratamento do fluxo de REEE, contribui para um fluxo de resíduos urbanos mais limpo e, também, para uma redução das emissões causadas pela incineração ou fusão de REEE contendo metais pesados e substâncias halogenadas.

Durante a deposição dos REEE em aterro verificam-se efeitos ambientais negativos que são consideravelmente maiores nos aterros não controlados. Estima-se que cerca de 70% dos aterros são não controlados em alguns Estados membros e na maior parte dos países que aderiram recentemente à UE [5]. Em Portugal, o número de aterros não controlados é aproximadamente de 300 [6].

Vários problemas ambientais e de saúde decorrentes da exposição a substâncias perigosas contidas nos EEE e ligados à actual gestão de REEE podem ser reduzidos através da criação de sistemas de recolha separada, tratamento e valorização dos REEE e da substituição das substâncias que são mais problemáticas. Esta última etapa constitui a forma mais eficaz de garantir uma redução significativa dos riscos para a saúde e o ambiente. Razões de ordem técnica (qualidade do produto, normas, requisitos de ensaio, entre outras) e económica (maiores custos) impedem, actualmente, a sua substituição generalizada.

5.2 Legislação

A Directiva 2002/96/CE do Parlamento Europeu e do Conselho, de 27 de Janeiro de 2003, enquadra legalmente a gestão de REEE (Jornal Oficial da União Europeia L 037, de 13/02/2003, p. 24-39). Muito recentemente foi promulgada a Directiva 2003/108/CE (Jornal Oficial da União Europeia L 345, de 08/12/2003, p. 106-107) que faz pequenas alterações à anterior. Estas directivas pretendem alargar o papel tradicional dos produtores, tornando-os responsáveis pela gestão dos produtos eléctricos e electrónicos em fim de vida útil. A criação de uma ligação entre os produtores e a gestão dos resíduos contribui para uma melhor concepção dos produtos, tendo em vista facilitar a sua reciclagem e eliminação quando passam a resíduos.

Um dos principais objectivos destas Directivas é promover o aumento da reciclagem de REEE. Em geral, o aumento da reciclagem permite a redução de recursos e de capacidades de eliminação, em especial no que se refere à deposição em aterro. No entanto, existem algumas limitações. A título de exemplo, vários estudos sugerem que o risco de geração de dioxinas constitui uma razão para a total ausência de reciclagem dos plásticos que contêm retardadores de chama bromados.

Em Portugal, o estabelecimento de um conjunto de regras e objectivos de gestão que visa a criação de circuitos de recolha selectiva de REEE, o seu correcto armazenamento e pré tratamento, e o posterior envio para reutilização ou reciclagem, está legislado no Decreto-Lei (DL) n° 20/2002, de 30 de Janeiro. É de realçar que esses objectivos devem estar, actualmente, alcançados pelos produtores, uma vez que a data limite de entrada em vigor já foi ultrapassada (2003/12/31). No que concerne aos objectivos de gestão, o artigo 4° do DL refere que os produtores devem adoptar as medidas necessárias para que, a partir do início de 2004, seja obrigatoriamente garantida a recolha selectiva de REEE, numa proporção de, pelo menos,

2Kg *per capita* por ano. Relativamente à percentagem de reutilização e reciclagem dos REEE recolhidos, esta varia entre os 50 e os 75%, dependendo da categoria em que é classificado o equipamento.

5.3 Análise Comparativa

A análise comparativa da legislação portuguesa e da UE será realizada em relação ao número de produtos em que é aplicável, aos objectivos e à responsabilidade de gestão dos REEE.

As Directivas da UE estabelecem um conjunto de regras de gestão para um número muito abrangente de EEE, 101 produtos classificados em 10 categorias: grandes electrodomésticos, pequenos electrodomésticos, equipamentos informáticos e de telecomunicações, equipamentos de consumo, equipamentos de iluminação, ferramentas eléctricas e electrónicas, brinquedos e equipamento de desporto e lazer, aparelhos médicos (com excepção de todos os produtos implantados e infectados), instrumentos de monitorização e controlo e distribuidores automáticos. O DL n° 20/2002, de 30 de Janeiro, visa apenas 18 dos produtos identificados nas Directivas, em que 11 estão incluídos na categoria grandes electrodomésticos (classe 1), 5 na categoria equipamentos informáticos e de telecomunicações, 1 na categoria equipamentos de consumo (classe 2), e 1 na categoria equipamentos de iluminação (classe 3).

No âmbito dos objectivos de gestão, o DL estabelece que, actualmente, seja garantida uma recolha selectiva de REEE numa proporção de, pelo menos, 2 kg/habitante/ano e as percentagens de reutilização e reciclagem dos REEE recolhidos de, pelo menos, 75, 65 e 50%, em peso, por equipamento, para as classes 1, 2 e 3, respectivamente, tabela 1.

Tabela 1. Percentagem de reutilização e reciclagem dos REEE recolhidos, DL

Classe	**% Mínima de Reutilização e Reciclagem**
1	75
2	65
3	50

A Directiva da UE estabelece uma taxa mínima de recolha separada de 4 kg/habitante/ano provenientes de particulares, a partir de 2006/12/31. No que respeita aos REEE enviados para tratamento, até 2006/12/31, os produtores devem atingir os objectivos apresentados na tabela 2, percentagens de valorização, reutilização e reciclagem dos REEE recolhidos, do peso médio por aparelho.

Tabela 2. Percentagem mínima de valorização, reutilização e reciclagem, do peso médio, por aparelho, dos REEE recolhidos, UE.

Categoria	**1 e 10**	**3 e 4**	**2, 5, 6, 7 e 9**	**Lâmpadas (*)**
Valorização (%)	80	75	70	
Reutilização e Reciclagem (%)	75	65	50	80

(*) de descarga de gás

A Directiva acrescenta que é imperativo fixar novos objectivos (mais ambiciosos), tanto para a recolha selectiva como para a valorização, até 2008/12/31.

No que concerne à responsabilidade de gestão, o DL considera que é da responsabilidade dos municípios a recolha de REEE considerados resíduos urbanos e da responsabilidade dos produtores a gestão dos REEE não abrangidos pela definição de resíduo urbano. No entanto, a Directiva imputa aos produtores essa responsabilidade para todos os REEE.

6 Os VFV em Portugal

6.1 Introdução

Os veículos em fim de vida útil (VFV) são resíduos, cujo tratamento é uma prioridade no âmbito da gestão dos resíduos.

O parque automóvel tem vindo a aumentar, sendo de esperar que o fluxo dos VFV também aumente. Em 1999, a *European Environment Agency* estimou um aumento de 21% no número de VFV na UE entre os anos 1995 e 2010.

Actualmente, os VFV na UE dão origem a cerca de 9 milhões de toneladas de resíduos por ano. Na generalidade da Europa dos quinze, mais de 70% destes materiais são valorizados por agentes económicos que se dedicam a actividades de reciclagem ou reutilização de componentes. No entanto, aproximadamente 25% da massa dos veículos (considerada como resíduo de fragmentação) representa até 10% da quantidade total de resíduos perigosos gerados actualmente na UE e que irão ser depositados em aterro. Esta solução deve ser evitada tanto quanto possível, devido às quantidades que comporta, à área de ocupação de terrenos a que obriga, bem como ao valor potencial de alguns dos componentes depositados.

Neste contexto, a UE enquadrou legalmente o problema dos VFV através da Directiva 2000/53/CE do Parlamento e do Conselho.

6.2 Legislação

A Directiva 2000/53/CE do Parlamento e do Conselho, de 18 de Setembro (Jornal Oficial da União Europeia L 269, de 21/10/2000, p. 34-43), tem em vista, principalmente, a prevenção da produção de resíduos provenientes de veículos e a promoção de reutilização, de reciclagem e de outras formas de valorização de VFV. Como objectivos consequentes e indispensáveis para um desenvolvimento sustentável, esta Directiva estabelece a redução da quantidade de resíduos a eliminar e a melhoria do desempenho ambiental de todos os operadores intervenientes durante o ciclo de vida dos veículos, sobretudo aqueles directamente envolvidos em operações de tratamento de VFV. Assim, fixa metas de reutilização, valorização e reciclagem a cumprir em determinadas datas. Até 2003/07/01, os veículos colocados no mercado não podem conter chumbo, mercúrio, cádmio ou crómio hexavalente, com excepção de alguns casos referidos no anexo II da Directiva. Até 2006/01/01, para os VFV produzidos antes de 1980, a taxa mínima de reutilização e valorização deve ser de 75% e a taxa mínima de reutilização e reciclagem deve ser de 70% em massa, em média, por veículo e por ano. Para os VFV produzidos depois de 1980, a taxa mínima de reutilização e valorização deve ser de 85% e a taxa mínima de

reutilização e reciclagem deve ser de 80% em massa, em média, por veículo e por ano. Até 2015/01/01, para todos os VFV, a taxa mínima de reutilização e valorização deve ser de 95% e a taxa mínima de reutilização e reciclagem deve ser de 85% em massa, em média, por veículo e por ano.

A Directiva estabelece, ainda, a necessidade dos VFV serem entregues em operadores autorizados, sem custos para o último proprietário, e submetidos a operações de despoluição.

O DL n° 196/2003, de 23 de Agosto, visa transpor para a ordem jurídica interna a Directiva n° 2000/53/CE e estabelecer um conjunto de normas de gestão que visa a criação de circuitos de recepção de VFV, o seu correcto transporte, armazenamento e tratamento, desencorajando, sempre que possível, o recurso a formas de eliminação tais como a sua deposição em aterros. Este diploma prevê a constituição de um sistema integrado de gestão. Os fabricantes ou importadores de veículos ficam obrigados a submeter a gestão dos VFV a um sistema integrado ou a um sistema individual. As operações de gestão de VFV encontram-se legisladas no DL, nomeadamente no que concerne a transporte, centros de recepção e operadores de desmantelamento e de fragmentação.

Tem-se vindo a verificar uma redução percentual (em massa) de metais (ferro e aço) na composição dos veículos, acompanhada por um crescente uso de plásticos e de outros materiais menos densos (como o alumínio e o magnésio). Esta crescente substituição de componentes metálicos por poliméricos tem provocado uma diminuição das taxas de reciclagem e de reutilização de materiais provenientes de VFV, porque os plásticos são materiais menos densos representando, por isso, um índice de reciclagem inferior, o que contraria os objectivos da Directiva da UE.

6.3 Análise Comparativa

A quantificação do número de VFV produzido em Portugal é incerto, visto que, apesar do abate de matrículas dos veículos ser previsto por lei, este não é efectuado na maior parte dos casos. Torna-se, assim, difícil fazer uma análise realista do panorama nacional, dificultando qualquer iniciativa que tenha em vista a realização de um plano nacional de resolução do problema de VFV. Este facto, aliado à necessidade de se proceder a uma mais intensa recuperação de peças e componentes de VFV, visto a actual taxa de reutilização e reciclagem dos VFV gerados a nível nacional (em massa, aproximadamente 76% [4]) não atingir os objectivos estabelecidos na Directiva (80%), impõe a necessidade de uma reestruturação das infra-estruturas nacionais e implementação de um sistema de processamento de VFV capaz de cumprir os objectivos comunitários.

As metas estabelecidas na Directiva 2000/53/CE só poderão ser atingidas com uma intensificação da recuperação dos constituintes não metálicos dos VFV mais importantes e através da criação de opções para esses materiais recuperados. A questão mais importante que se coloca é de índole económica, sendo de extrema importância o *eco-design* e a existência de infra estruturas de processamento de VFV eficientes.

O *eco-design* dos veículos, e em particular, o *design* para reciclagem, são factores importantes a ter em consideração nos novos veículos a produzir, implicando um significativo esforço a curto prazo da indústria automóvel na concepção de futuros modelos e permitindo uma redução dos custos inerentes à remoção de componentes.

7 Conclusões

Actualmente, a gestão e redução de resíduos tornou-se uma prioridade dos países industrializados pelo que representa quer para o crescimento/desenvolvimento sustentável quer para o impacte ambiental. Uma política integrada de gestão de resíduos traduz-se, prioritariamente, na prevenção da quantidade e perigosidade dos resíduos e na maximização das quantidades recuperadas para valorização, tendo em vista a minimização de resíduos enviados para eliminação. Com os aterros sobrelotados e a insuficiente capacidade das incineradoras, têm sido desenvolvidos esforços no sentido de reintegrar os resíduos nos processos produtivos originais. Por conseguinte, tanto ao nível da UE como de Portugal, já existem princípios e normas aplicáveis aos sistemas de gestão de resíduos que estão patentes em Directivas e em Decretos-Lei.

Para cumprir os objectivos estabelecidos nas Directivas, poderá ser necessário alterar as filosofias de produção e de gestão, e os fluxos, de materiais e de informação, de modo a reduzir quer o consumo de recursos quer a formação de resíduos. O fluxo de materiais, que passa a fazer-se nos dois sentidos, terá de englobar não só as actividades logísticas tradicionais mas também as inerentes à logística inversa.

Assim, frequentemente, é necessário adaptar o desenho da cadeia de abastecimento tradicional à passagem do fluxo dos materiais em dois sentidos. Adicionalmente é fundamental a percepção dos factores críticos, estratégicos e operacionais, tanto de índole técnica como económica e de marketing que condicionam o redesenho da cadeia de abastecimento.

Relativamente às Directivas da UE, no caso dos REEE a data limite de transposição ainda não foi atingida (2004/08/13), enquanto que para os VFV foi transposta recentemente para a legislação portuguesa.

A implementação destas Directivas a nível nacional está, fundamentalmente, condicionada por factores de índole técnica e económica. As Directivas são cada vez mais exigentes em termos de valorização dos resíduos, pelo que a concepção de novos produtos ou o redesenho de produtos existentes é um factor crítico para os fabricantes/produtores, atendendo à responsabilização que lhes é imputada nas Directivas através do princípio do poluidor pagador.

A existência de infra estruturas eficientes de processamento de resíduos, nomeadamente REEE e VFV, é de igual modo um factor crítico para os sistemas de logística inversa, atendendo quer aos investimentos quer ao *know-how* que exige.

8 Referências

[1] Achrol, R., Reve, T. e Stern, L., "The Environment of Marketing Channel Dyads: a Framework for Comparative Analysis", *Journal of Marketing*, Vol. 47, Fall, 1983, 55 – 67.

[2] Carter, C. e Ellram, L., "Reverse Logistics: A Review of Literature and Framework for Future Investigation", *Journal of Business Logistics*, Vol. 19, 1, 1998, 85 – 102.

[3] Christopher, M., "Logistics and Supply Chain Management", Richard D. Irwin, Inc., New York, 1994.

[4] Coelho, M. C., "Que Fazer dos Veículos em Fim de Vida?", *Ingenium*, 2ª série, N° 65, 2002, 76 – 79.

[5] Conferência para Planeamento da Gestão de Resíduos, Grécia, 16-17 de Janeiro de 1997.

[6] Conferência para Planeamento da Gestão de Resíduos, Portugal, 23-24 de Janeiro de 1997.

[7] Dowlatshahi, S., "Developing a Theory of Reverse Logistics", *Interfaces*, Vol. 30, 3, 2000, 143 – 154.

[8] Environmental Consequences of Incineration and Landfilling of Waste from Electronic Equipment (Copenhaga 1995), Conselho Nórdico de Ministros. Segundo o estudo "Pilotsammlung von Elektroaltgeräten in Bregenz", 95% dos REEE produzidos na Áustria são simplesmente deitados fora com os resíduos urbanos ou introduzidos na cadeia de reciclagem sem qualquer pré-tratamento.

[9] Fleischmann, M., H. R. Krikke, R. Dekker e S. D. P. Flapper, "A Characterisation of Logistics Networks for product Recovery", OMEGA, *The International Journal of Management Science*, 28-6, 2000, 653-666.

[10] Guide Jr, V.D.R. e Van Wassenhove, "Full Cycle Supply Chains", Carnegie Mellon Press, USA, 2003.

[11] Guintini, R. e Andel, T., "Advance with Reverse Logistics", *Transportation & Distribution*, Vol. 36, 2, 1995a, 73.

[12] Guintini, R. e Andel, T., "Reverse Logistics Role Models", *Transportation & Distribution*, Vol. 36, 4, 1995b, 97.

[13] Krikke, H., Bloemhf-Ruwaard, J. e Van Vassenhove, L.N., "Design of Closed Loop Supply Chains: A Production and Return Network for Refrigerators", ERIM Report Series Research in Management, ERS-2001-45-LIS, 2001, 1 – 33.

[14] Rogers, D.S. e Tibben-Lembke, R.S., Going Backwards: Reverse Logistics Practice, Reverse Logistics Executive Council, 1998.

[15] Stock, J., Development and Implementation of Reverse Logistics Programs, Council of Logistics Management, Oak Brook, IL, 1998.

[16] Stock, J., "Reverse Logistics", Council of Logistics Management, Oak Brook, IL, 1992.

[17] Thierry, M., Salomon, M., Van Nunen, J. e Van Wassenhove, L., "Strategic Issues in Product Recovery Management", *California Management Review*, Vol. 37, 2, 1995, 114 – 135.

[18] Witt, C., "Distribution: a differentiator in 2000", *Material Handling Engineering*, Vol. 50, 11, 1995, 57-77.

[19] Wu, H. e Dunn, S.C., "Environmentally Responsible Logistics Systems", *International Journal of Physical Distribution & Logistics Management*, Vol.25, 2, 1994, 20 – 38.

9 Abreviaturas

CE - Comissão Europeia

CER - Catálogo Europeu de Resíduos

DL - Decreto-Lei

EEE - Equipamentos Eléctricos e Electrónicos

REEE - Resíduos de Equipamentos Eléctricos e Electrónicos

UE - União Europeia

VFV - Veículos em Fim de Vida

A model and a heuristic for the primary health care coverage planning problem in Portugal – application to "Cova da Beira" case

Rogério A . F. Monteiro * António J. Pascoal †

* Instituto Superior de Matemática e Gestão
6230 Fundão

† Departamento de Matemática
Universidade Portucalense - Infante D. Henrique
4000 Porto

Abstract

This communication presents an attempt in order to help on the Planning of a Primary Health Care Delivery System in Portugal. The research was conducted in order to get the optimal location and "districting" of hierarchical facilities in a network system. A linear integer multiobjective model is presented, considering location-allocation theory and also demographic evolution, as well as coherent structure. A metaheuristic solution method is provided. The model and the metaheuristic was applied to a set of instances and "Cova da Beira" Case.

Resumo

Esta comunicação apresenta uma tentativa no sentido de ajudar no Planeamento do Sistema de Cobertura de Cuidados de Saúde Primários em Portugal. A investigação foi conduzida no sentido de obter a partição de zonas de procura e localização óptima de instalações hierárquicas num sistema de rede. É apresentado um modelo multiobjectivo linear inteiro, considerando a teoria de localização-afectação e também a evolução demográfica, bem como a estrutura coerente. É também apresentada uma metaheurística. O modelo e a metaheurística foram aplicados a um conjunto de instâncias e ao Caso da "Cova da Beira".

Keywords: Integer Programming, demographic models, hierarchical location-allocation models, hierarchical location-delivery models, multi-heuristic methods.

Apdio © 2005 *Associação Portuguesa de Investigação Operacional*

1 Introdução

Tornou-se relevante o desenvolvimento da investigação em saúde pela necessidade do planeamento deste sector e, em particular, nos Cuidados de Saúde Primários.

Assim, este artigo pretende constituir-se como um contributo para o Planeamento da Cobertura de Cuidados de Saúde Primários em Portugal. Para tal, procedeu-se:

- na secção 2 à definição do Problema de Planeamento de Cuidados de Saúde Primários em Portugal;
- na secção 3 à indicação da Complexidade Computacional do Problema do Planeamento da Cobertura de Cuidados de Saúde Primários em Portugal (PPCCSPP) e ao desenvolvimento de um modelo para o PPCCSPP;
- na secção 4 ao desenvolvimento de uma heurística para o PPCCSPP;
- na secção 5 á apresentação de uma abordagem de resolução do PPCCSPP em sistemas de instalações existentes;
- na secção 6 à determinação da complexidade computacional do método de resolução do PPCCSPP;
- na secção 7 à aplicação do modelo e do método a um conjunto de instâncias;
- na secção 8 à aplicação da heurística e do método exacto do Sistema CPLEX às instâncias do Concelho de Belmonte (IB), dos Concelhos da Covilhã (ICC) e Fundão (ICF) e da Cova da Beira (ICB). Nesta secção são ainda apresentados outros testes de validação;

Finalmente nas últimas secções são apresentadas as conclusões e sugestões de trabalho futuro.

2 Definição do PPCCSPP

Matematicamente, a rede de Cuidados de Saúde Primários (CSP) define-se como um processo pelo qual uma dada zona de procura é particionada em zonas de procura pequenas e, numa mesma zona, todas as localidades (freguesias) afectadas a uma instalação de nível $m-1$(Extensões do Centro de Saúde – Instalações tipo A) são afectadas a uma e uma só instalação de nível superior m(Centro de Saúde – Instalações tipo B), onde se considerarmos I^k a instalação mais inclusiva (Centro de Saúde) com o serviço mais inclusivo e I^1 a instalação menos inclusiva (Extensões do Centro de Saúde) com o serviço menos inclusivo, se tem:$I^k \supset I^{k-1} \supset \ldots \supset I^1$.

Considerando, como se entende em Serrão et al.(1998)[15], os valores da Universalidade e Generalidade na cobertura através do planeamento e da distribuição dos recursos e que, segundo Santana (1993)[12], os incentivos para trazer os utilizadores aos Cuidados de Saúde Primários (CSP) públicos podem implicar o estabelecimento de locais de oferta mais acessíveis à população e maior facilidade em termos de transportes públicos e que, sempre que seja necessário, nos locais de menor acessibilidade geográfica, os CSP possam ser fornecidos por

equipas móveis que com facilidade se possam deslocar a esses locais, bem como provocar a diminuição do tempo de espera; o Planeamento da cobertura dos CSP caracteriza-se por um processo de análise demográfica através de Cenários de Perspectivas de Evolução Demográfica de médio prazo (10 a 15 anos), por processos hierárquicos de localização-afectação e localização-distribuição coerentes que permitam incorporar:

- a incerteza demográfica;
- restrições de capacidade/disponibilidade;
- restrições de zonalidade;
- a coerência hierárquica de instalações e serviços;
- e outros critérios de localização estabelecidos pelo Ministério da Saúde;
- a localização-distribuição de Extensões Móveis de Cuidados de Saúde Primários, com restrições de capacidade/disponibilidade e de extensão total.

Especificamente, os Centros de Saúde e Extensões de Saúde têm as fichas de caracterização constantes de (DGOTDU,1997)[6].

3 Um Modelo para o PPCCSPP

Nesta secção, é desenvolvido um modelo para o PPCCSPP. Em particular, o modelo tem em consideração os critérios de modelação estabelecidos em (DGOTDU, 1997)[6].

Assim o modelo é composto por três submodelos. Considerem-se as seguintes abreviaturas:

- **SPED** - Submodelo de perspectivas de Evolução Demográfica;
- **SLAI** - Submodelo de Localização-Afectação de Instalações Coerentes de Cuidados de Saúde Primários;
- **SLD** - Submodelo de Localização-Distribuição de Extensões Móveis de Cuidados de Saúde Primários.

3.1 Formulação global

O **SPED** permite a construção de perspectivas de Evolução Demográfica, por forma a estimar a procura por localidade (freguesia) a médio prazo (10-15 anos). Assim, o submodelo, partindo de hipóteses de evolução, estima a população, tendo em conta a informação demo-socio-económica da localidade, nomeadamente, as taxas de fecundidade, de mortalidade e migratórias e a probabilidade de sobrevivência dos grupos etários considerados.

O **SLAI** é uma extensão do problema da pq-mediana de Serra e ReVelle (1993)[14] que permite a localização-afectação de instalações coerentes de cuidados de saúde primários de acordo com:

- Os critérios descritos em (DGOTDU, 1997)[6];
- O princípio de equidade de acesso aos cuidados de saúde, referido por Santana (1993)[12] como um factor determinante no acesso para todos à saúde;
- Os dois objectivos estratégicos do Sistema de Saúde Português referidos em Serrão et al.(1998)[15];

considerando explicitamente a minimização dos efeitos das distâncias, das fracas redes de transportes e da ausência de transportes compatíveis que separam os utentes dos locais de oferta de cuidados de saúde. O modelo SLAI pode ser resolvido para vários cenários conforme se descreve em 4.

Atendendo a que um dos objectivos deste trabalho é debelar, de alguma forma, as barreiras geográficas no tocante ao acesso aos cuidados de saúde primários, o **SLD** permite o acesso a estes através de Extensões Móveis de Cuidados de Saúde Primários (veículos), aos utentes de todas as localidades (freguesias) com população inferior a 1500 habitantes cujo tempo de percurso no meio de transporte habitual para as unidades de cuidados de saúde primários já existentes seja superior a 30 minutos, ou não existam, para o efeito, pelo menos dois transportes diários. O **SLD** é uma formulação baseada na estrutura do problema da p-mediana hamiltoniana com restrições de capacidade de Branco e Coelho (1984)[2], em que os **p** circuitos (**rotas**) podem não ser na prática real circuitos hamiltonianos, mas estrelas (entende-se por estrela uma árvore com $n$vértices em que um destes tem grau $n-1$ e designa-se por centro). A formulação integra restrições de extensão total de uma qualquer rota e permite que o número de veículos seja uma variável de decisão. Assim, o submodelo pressupõe alguns requisitos específicos a considerar explicitamente:

1. Todas as rotas têm de começar e terminar no depósito (Instalação do tipo B);
2. Toda a localidade (freguesia) excepto a localidade (freguesia onde está instalado o depósito ou instalação do tipo B) é servida uma única vez por rota;
3. A cada instalação do tipo A estão associados:
   - um número de atendimentos $q_i \quad (q_0 = 0)$;
   - um tempo de serviço $t_i \quad (t_0 = 0)$.
4. A cada rota estão associados:
   - um número máximo de atendimentos Q;
   - uma extensão total (tempo de percurso mais tempo de serviço) máximo T.

Matematicamente, o **modelo global** tem a seguinte estrutura:

SPED:

$$\mathbf{Pop}(t+10)=\mathbf{Pop}(t)\times p_{\text{sobrevivência}}+\mathbf{Pop}(t)\times p_{\text{fecundidade}}+\mathbf{Pop}(t)\times p_{\text{migrar}}$$

onde,

$\text{Pop}(t)$	População no instante t
$\text{Pop}(t+10)$	População no instante $t+10$, sendo o ano a unidade de tempo
$p_{\text{sobrevivência}}$	Probabilidade de sobrevivência dos indivíduos ao fim de dez anos
$p_{\text{fecundidade}}$	Taxa de fecundidade prevista para o período $t+10$
p_{migrar}	Probabilidade de migrar

SLAI:

$$MinA = \sum_{i\in I}\sum_{j\in J} p_i d_{ij} e_{ij} \tag{3.1}$$

$$MinB = \sum_{i\in I}\sum_{k\in K} p_i d_{ik} s_{ik} \tag{3.2}$$

Sujeito a:

$$\sum_{j\in J} e_{ij} = 1 \qquad \forall i\in I \tag{3.3}$$

$$\sum_{k\in K} s_{ik} = 1 \qquad \forall i\in I \tag{3.4}$$

$$\sum_{k\in K} es_{jk} = 1 \qquad \forall j\in J \tag{3.5}$$

$$e_{ij} \le e_j \qquad \forall i\in I\ \ \forall j\in J_A \tag{3.6}$$

$$e_{ij} \le e_j + se_j \qquad \forall i\in I\ \ \forall j\in J_{AB} \tag{3.7}$$

$$e_{ij} \le se_j \qquad \forall i\in I\ \ \forall j\in K_B\backslash J_A \tag{3.8}$$

$$s_{ik} \le s_k \qquad \forall i\in I\ \ \forall k\in K \tag{3.9}$$

$$e_{ij} + es_{jk} \le s_{ik} + 1 \qquad \forall i\ \ \forall j\in J\ \ \forall k\in K \tag{3.10}$$

$$P_{\min} e_j \le \sum_{i\in I} p_i e_{ij} \le P_{\max} e_j \qquad \forall j\in J \tag{3.11}$$

$$P'_{\min} s_k \le \sum_{i\in I} p_i s_{ik} \le P'_{\max} s_k \qquad \forall k\in K \tag{3.12}$$

$$e_z^{\min} \le \sum_{j\in Z_z} e_j \le e_z^{\max} \qquad \forall j\in J_A\ \ \forall e_z^{max}, e_z^{min}\in \mathrm{Z} \tag{3.13}$$

$$s_z^{\min} \le \sum_{k\in Z_z} s_k \le s_z^{\max} \qquad \forall k\in K_B\ \ \forall s_z^{max}, s_z^{min}\in \mathrm{Z} \tag{3.14}$$

$$\left.\begin{array}{l} e_{ij}, s_{ik}, es_{jk}, e_j, s_k, r_{ij}, r'_k = (0,1) \\ d_{ij} = \begin{cases} \text{distância entre } i \text{ e } j \text{ se } r_{ij}=0 \\ M \text{ se } r_{ij}=1 \end{cases} \\ d_{ik} = \begin{cases} \text{distância entre } i \text{ e } k \text{ se } r'_k=1 \\ M \text{ se } r'_k=0 \end{cases} \\ M = \text{ Número Suficientemente Grande.} \end{array}\right\} \qquad \forall i\in I\ \ \forall j\in J\ \ \forall k\in K$$

onde,

i, I	Índice e conjunto de localidades de procura
j, J	Índice e conjunto de potenciais localizações de instalações do tipo A
k, K	Índice e conjunto de potenciais localizações de instalações do tipo B
J_A	Conjunto de potenciais localizações somente de instalações do tipo A; i.é. $J \setminus (J \cap K)$
J_{AB}	Conjunto de potenciais localizações para instalações do tipo A e B; i.é. $J \cap K$
K_B	Conjunto de potenciais localizações somente de instalações do tipo B; i.é. $K \setminus (J \cap K)$
p_i	População da localidade de procura i
r_{ij}	1, se o tempo de percurso no meio de transporte habitual entre i e j for superior a 30 minutos, ou não existam, para o efeito, pelo menos dois transportes diários; 0, caso contrário
$r'k$	1, se a localidade k integra ou está próxima de outros equipamentos sociais e estabelecimentos de ensino; 0, caso contrário
P_{min}	População mínima necessária para a localização de uma instalação do tipo A em j
P_{max}	População máxima necessária para a localização de uma instalação do tipo A em j
P'_{min}	População mínima necessária para a localização de uma instalação do tipo B em k
P'_{max}	População máxima necessária para a localização de uma instalação do tipo B em k
z	Índice de zona
Z_z	Conjunto de localidades compreendidas na zona z
e^z_{min}	Número mínimo de instalações do tipo A que podem ser localizadas na zona z
e^z_{max}	Número máximo de instalações do tipo A que podem ser localizadas na zona z
s^z_{min}	Número mínimo de instalações do tipo B que podem ser localizadas na zona z
s^z_{max}	Número máximo de instalações do tipo B que podem ser localizadas na zona z
es_{jk}	1, se a instalação do tipo A localizada em j está afectada na hierarquia a uma instalação do tipo B localizada em k; 0, caso contrário
e_{ij}	1, se a procura da localidade i está afectada a uma instalação do tipo A localizada em j; 0, caso contrário
s_{ik}	1, se a procura da localidade i está afectada a uma instalação do tipo B localizada em k; 0, caso contrário
e_j	1, se existe uma instalação do tipo A localizada em j; 0, caso contrário
se_j	1, se existe uma instalação do tipo B localizada em j; 0, caso contrário
s_k	1, se existe uma instalação do tipo B localizada em k; 0, caso contrário

SLD:

$$Min \sum_{k=1}^{p} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} \sum_{l=1}^{m} x^l_{ijk} \tag{3.15}$$

Sujeito a:

$$\sum_{l=1}^{m}\sum_{k=1}^{p} y_{ik}^{l} = 1 \qquad i = 1,2,...,n \tag{3.16}$$

$$\sum_{i=1}^{n} x_{ijk}^{l} = y_{jk}^{l} \qquad j = 1,2,...,n \;\; k = 1,2,...,p \tag{3.17}$$

$$\sum_{j=1}^{n} x_{ijk}^{l} = y_{ik}^{l} \qquad i = 1,2,...,n \;\; k = 1,2,...,p \;\; l = 1,2,...,m \tag{3.18}$$

$$\sum_{i=1}^{n} v_{ik} = 1 \qquad k = 1,2,...,p \tag{3.19}$$

$$\sum_{l=1}^{m} v_{ik}^{l} = m \qquad i = 1,2,...,n \;\; k = 1,2,...,p \tag{3.20}$$

$$y_{ik} \geq v_{ik} \qquad i = 1,2,...,n \;\; k = 1,2,...,p \tag{3.21}$$

$$\sum_{k=1}^{p}\sum_{i,j\in S}^{n} x_{ijk}^{l} \leq |S| - 1 + \sum_{k=1}^{p}\sum_{i\in S} v_{ik} \qquad \forall S : |S| > 1 \text{ e } S \subset N \quad l = 1,2,...,m \tag{3.22}$$

$$\sum_{i=1}^{n} q_i \left(\sum_{j=1}^{n} x_{ijk}^{l}\right) \leq Q \qquad k = 1,2,...,p \;\; l = 1,2,...,m \tag{3.23}$$

$$\sum_{i=1}^{n} t_i \sum_{j=1}^{n} x_{ijk}^{l} + \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij} x_{ijk}^{l} \leq T \qquad k = 1,2,...,p \;\; l = 1,2,...,m \tag{3.24}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ijk}^{l} = 1 \qquad k = 1,2,...,p \;\; l = 1,2,...,m \tag{3.25}$$

$$x_{ijk} \in \{0,1\} \qquad i,j = 1,2,...,n \;\; k = 1,2,...,p$$

$$y_{ik} \in \{0,1\} \qquad i = 1,2,...,n \;\; k = 1,2,...,p$$

$$v_{ik} \in \{0,1\} \qquad i = 1,2,...,n \;\; k = 1,2,...,p$$

onde,

i, j	Índices de localidades a percorrer ou de procura
k	Índice do circuito
l	Índice do veículo ou Móvel
S	Conjunto das localidades já integradas em algum dos p subconjuntos
$\|S\|$	Número de localidades integradas em S
d_{ij}	Distância ou tempo de percurso entre as localidades i e j
$\|S\|$	Número de localidades integradas em S
d_{ij}	Distância ou tempo de percurso entre as localidades i e j
q_i	Número de atendimentos por localidade
Q	Número máximo de atendimentos por rota
t_i	Tempo de serviço em cada localidade
T	Extensão total (tempo de percurso mais tempo de serviço)
l	Número de veículos ou móveis, onde $l \in [1, m]$ e $m \in [m', m'']$ com $m' \geq 1$ e $m'' \leq n$
y_{ik}	1, se a localidade i pertence ao circuito k; 0, caso contrário
y_{ik}^{l}	1, se a localidade i pertence ao circuito k e é servida pelo veículo l; 0, caso contrário
v_{ik}	1, se um depósito é localizado no vértice i do circuito k; 0, caso contrário
y_{ik}^{l}	1, se o depósito localizado em i no circuito k é servido pelo veículo l; 0, caso contrário
x_{ijk}^{l}	1, se o veículo l vai de i para j directamente no circuito k; 0, caso contrário

Em **SLAI**:

A função objectivo (3.1) contem explicita e implicitamente três e quatro parâmetros, respectivamente. Esta função minimiza a soma das distâncias ponderadas entre cada localidade e a instalação do tipo A mais próxima. A função objectivo (3.2) contem explicita e implicitamente o mesmo número de parâmetros de (3.1). Esta função minimiza a soma das distâncias ponderadas entre cada localidade e a instalação do tipo B mais próxima. As áreas de procura são afectadas a uma e uma só instalação do tipo A e a uma e uma só instalação do tipo B, através das restrições (3.3) e (3.4), respectivamente. Com (3.5), uma instalação do tipo A é afectada a uma e uma só instalação do tipo B. As restrições (3.6) asseguram que se a área de procura i está afectada para serviços de nível A ao nodo j ($j \in J_A$), então este nodo deve ter uma instalação do tipo A . Um nodo i pode ser afectado a uma instalação do tipo A ou a uma instalação do tipo B para receber serviços do tipo A, através de (3.7). Se o nodo i está livre para se afectar ao nodo j para serviços tipo A e o nodo j só pode ter instalações do tipo B então deve existir em j uma instalação tipo B, o que é estabelecido por (3.8). O conjunto de restrições (3.9) define que se a área de procura i está afectada ao nodo k para serviços do tipo B, então deve existir uma instalação do tipo B no nodo k. O reforço da coerência para a localização, isto é, todos as áreas afectadas à mesma instalação do tipo A estão afectadas à mesma instalação do tipo B, é assegurado por (3.10). A capacidade/disponibilidade mínima e máxima por tipo de instalação, incluindo o número de instalações de cada tipo a serem localizadas, é conseguida pelas desigualdades (3.11) e (3.12), respectivamente, consoante se trate de instalações do tipo A ou B. Finalmente, os números mínimo e máximo de cada tipo de instalações por zona são conseguidos pelas desigualdades (3.13) e (3.14) consoante se trate de instalações do tipo A ou B, respectivamente.

Em **SLD**:

A função objectivo (3.15) consiste em minimizar o custo total de transporte. A garantia que cada localidade i fica afectada a um e um só circuito e um só veículo é dada pelos conjuntos de condições (3.16), (3.17) e (3.18). A criação de p instalações (depósitos), sendo localizada uma em cada circuito é assegurada pelas condições (3.19). O depósito do circuito k é servido por todos os veículos do circuito k através das condições (3.20). A garantia de que a instalação (depósito) a abrir em cada circuito k só pode ser localizada numa das localidades pertencentes a esse circuito é dada pelas condições (3.21). A formação de circuitos em subconjuntos de N que contenham pelo menos uma instalação (depósito) é traduzida pelas restrições (3.22). A capacidade máxima (a quantidade máxima transportada por cada veículo) por rota é assegurada pelas restrições (3.23). O custo (tempo de percurso mais tempo de serviço) máximo por rota é assegurado pelas restrições (3.24). Finalmente, para que cada veículo seja usado uma só vez, consideram-se as restrições (3.25).

3.1.1 Formulação do SLAI para Hierarquia de Serviços Sucessivamente Exclusivos

A formulação apresentada refere-se a um sistema de instalações de serviços sucessivamente inclusivos, no entanto se a hierarquia se caracterizar por serviços sucessivamente exclusivos, ou seja, se as instalações do tipo B não prestarem os serviços prestados pelas instalações do tipo A, as restrições (3.6), (3.7) e (3.8) podem ser substituídas pelo seguinte conjunto de restrições:

$$e_{ij} \leq e_j + se_j \qquad \forall i \in I \;\; \forall j \in J \tag{3.26}$$

3.1.2 Formulação do SLAI para Hierarquia de Serviços Localmente Inclusivos

Se a hierarquia se caracterizar por serviços localmente inclusivos, isto é, uma instalação do tipo B presta serviços do tipo A apenas na localidade onde se encontra localizada, as restrições (3.6), (3.7) e (3.8) são substituídas pelas restrições:

$$e_{ij} \leq e_j + se_j \qquad \forall i \in I \cap J_{AB} \;\; \forall j \in J_{AB} \tag{3.27}$$
$$e_{ij} \leq se_j \qquad \forall i \in I \cap K_B \;\; \forall j \in K_B \tag{3.28}$$

3.2 Reformulações do PPCCSPP para Sistemas de Instalações Existentes

O PPCCSPP existe sempre que se pretenda constituir um novo sistema de instalações de cuidados de saúde primários ou sempre que se pretenda melhorar e expandir um sistema de instalações de cuidados de saúde primários existente. No caso em que se pretende constituir um novo sistema de instalações, o modelo do PPCCSPP apresentado em 3, pode ser directamente aplicado. Todavia, na maioria das situações reais trata-se de melhorar e/ou expandir um sistema já existente. Atendendo a que o Problema da Cobertura de Cuidados de Saúde Primários em Portugal se enquadra em critérios bem definidos onde não cabe teoricamente o efeito **ad-hoc**, é necessário introduzir no modelo a possibilidade de actualização do sistema existente. Neste caso e para que o modelo possa ser aplicado, determinados parâmetros de decisão devem ser fixados e devem ser efectuadas pequenas alterações para alguns parâmetros de sistema ou input nas situações seguintes:

1. A DIMINUIÇÃO DA POPULAÇÃO NA(S) LOCALIDADE(S) DA(S) ZONA(S).
2. A EXTINÇÃO DE LOCALIDADE(S) DE UMA OU MAIS ZONAS.
3. A EXTINÇÃO DE UMA OU MAIS ZONAS.
4. O AUMENTO DA POPULAÇÃO NA(S) LOCALIDADE(S) DA ZONA(S).
5. A CRIAÇÃO DE ZONAS.

3.3 Formulação do SLAI com Restrições de Tempo de Espera - (SLAITE)

O **SLAI** pode também ser formulado no sentido de minimizar os efeitos do tempo de espera no acesso aos cuidados de saúde primários que, tal como concluiu Santana (1993)[12], é um dos factores desincentivantes que mais influencia o custo do referido acesso. Acontece, porém, que esta formulação envolve uma componente estocástica, pois pressupõe o conhecimento da procura e da sua distribuição dentro do período de serviço ou atendimento. Assim, a formulação admite a aplicação do modelo de Poisson como modelo de comportamento aleatório e então admite-se que:

1. O modelo de Poisson traduz a realidade;

2. As chegadas processam-se de modo homogéneo dentro do período de serviço ou atendimento: sem concentrações numa parte do período e escassez de procura noutra parte do período e com intensidade f_i por parte de cada localidade.

A formulação supõe, ainda, um sistema de fila de espera M/M/m (com m servidores) por instalação.

Desta forma, constitui-se uma extensão do SLAI designada por SLAITE, que integra as restrições do SLAI e as seguintes restrições estocásticas (restrições de congestionamento) do tipo das usadas por Marianov e Serra (1998)[10]:

$$\sum_{i \in I} f_i e_{ij} \leq \mu_j \rho_{\alpha j} \qquad \forall i \in I \;\; \forall j \in J_A \tag{3.29}$$

$$\sum_{i \in I} f_i s_{ik} \leq \mu_k \rho_{\alpha k} \qquad \forall i \in I \;\; \forall k \in K \tag{3.30}$$

Onde $\forall\, i \in I \;\; \forall\, j \in J \;\; \forall\, k \in K$, supomos na função objectivo respectiva:

$$d_{ij} = \begin{cases} \text{distância entre } i \text{ e } j, \text{ se } r_{ij} = 0 \text{ ou se existem } m_j \text{ servidores em cada uma das instalações do tipo A localizadas em } j \\ M(\text{suficientemente grande}), \text{ se } r_{ij} = 1 \text{ ou se não existem } m_j \text{ servidores em cada uma das instalações do tipo A localizadas em } j \end{cases}$$

$$d_{ij} = \begin{cases} \text{distância entre } i \text{ e } k, \text{ se } r'_k = 1 \text{ ou se existem } m_k \text{ servidores em cada uma das instalações do tipo B localizadas em } k \\ M(\text{suficientemente grande}), \text{ se } r'_k = 0 \text{ ou se não existem } m_k \text{ servidores em cada uma das instalações do tipo B localizadas em } k \end{cases}$$

f_i	designa a intensidade de procura na localidade i,
μ_j, μ_k	Taxa média de procura da localidade j ou k e $\mu_j \geq m_k \lambda_k$, em que $\lambda_j = \sum_{f_i e_{ij}}$ ou $\lambda_k = \sum_{f_i s_{ik}}$
$\rho_{\alpha j},\ \rho_{\alpha k}$	É a taxa $\frac{\lambda}{\mu}$ da localidade j ou k para pelo menos uma probabilidade α de um utente encontrar uma fila com não mais de b utentes, isto é, $p_0 + p_1 + ... + p_{m+b} \geq \alpha$. Em que p_w representa a probabilidade do sistema se encontrar no estado constante w: o estado w corresponde a w utentes na instalação (no sistema) a serem atendidos, isto é, $w = m + b$ representa m utentes a serem atendidos e b na fila de espera.

Os conjuntos de restrições (3.29) e (3.30) asseguram que os utentes de cada localidade i quando afectados a serviços do tipo A ou B, consoante se trate de instalações do tipo A ou B, localizados em j ou k, têm uma probabilidade α de encontrar uma fila com não mais de b utentes, isto é, $p_0 + p_1 + ... + p_{m+b} \geq \alpha$.

4 Uma Heurística para o PPCCSPP

É possível demonstrar que o PPCCSPP pertence à classe dos problemas $NP-$*difíceis*, pelo que não existe nenhum algoritmo polinomial em tempo para resolver o problema. Desta

forma, apenas pequenas instâncias podem ser resolvidas por algoritmos de solução óptima. A aplicação de qualquer algoritmo de solução óptima, como o que faz parte do sistema CPLEX 6.01, à resolução do submodelo de localização-afectação de instalações (Centros de Saúde e Extensões dos Centros de Saúde) envolve $n^3 + 4n^2 + 6n + 4$ restrições e $3(n^2 + n)$ variáveis.

Atendendo à estrutura matemática do problema, este caracteriza-se por uma dimensão significativa na prática, pois como é sabido os problemas da vida real caracterizam-se por dimensões consideráveis que, devido à sua complexidade, raramente requerem a obtenção da solução óptima, considerando-se satisfatória a obtenção de uma solução "razoavelmente" óptima, desenvolve-se neste artigo uma abordagem de resolução heurística para o problema. A heurística pode ser aplicada à resolução do modelo para vários cenários. A heurística é aplicada à resolução do SLAI para cada um dos cenários. A partir das soluções obtidas para cada um dos cenários, o decisor obtém a solução final a utilizar com base no cenário sustentado na hipótese de evolução demográfica que lhe parece mais realista ou adequada à capacidade e aos meios que dispõe ou à resposta ao desafio que se lhe coloca. A abordagem que desenvolvemos é uma multi-metaheurística, incluída nos algoritmos de melhoramento, desenvolvida para o modelo apresentado. Concretamente, o método de resolução aqui proposto é constituído por três fases:

Fase 1 - Método de Resolução do Submodelo de Perspectivas de Evolução Demográfica (SPED), que designamos por MSPED;
Fase 2 - Método de Resolução do Submodelo de Localização-Afectação de Instalações (SLAI), que designamos por MSLAI;
Fase 3 - Método de Resolução do Submodelo de Localização-Distribuição (SLD), que designamos por MSLD.

4.1 Um método de Resolução do SPED - MSPED

O MSPED é baseado na metodologia do "Cohort-Survival" com integração de taxas migratórias. Assim, tem-se:

Passo 1 - Construir a probabilidade de sobrevivência dos grupos etários considerados na população de partida a partir de elementos de cálculo de uma tábua abreviada de mortalidade. Esta tábua é deduzida de taxas de mortalidade médias para o período em análise. Calcular $\mathbf{Pop}(i,t) \times p_{\text{sobrevivência}}$.

Passo 2 - Calcular $\mathbf{Pop}(i,t) \times t_{\text{fecundidade}}$ tendo em conta a taxa de fecundidade média verificada segundo os nados-vivos por grupo etário das mães num determinado período, ou seja, considerando cenários de evolução da taxa de fecundidade.

Passo 3 - Calcular $\mathbf{Pop}(i,t) \times p_{\text{migrar}}$ tendo em conta as taxas migratórias observadas noutros períodos, isto é, considerando cenários de evolução das taxas específicas de saldos migratórios.

Passo 4 - Calcular as $\text{Pop}(i+10, t+10)$ tendo em conta os resultados, respectivos, obtidos nos Passos anteriores.

4.2 Um Método de Resolução do SLAI - MSLAI

Como foi referido anteriormente o **MSLAI** é uma meta-heurística baseada na heurística de Teitz e Bart (1967)[16], melhorada por Desham e Rushton (1992)[5], e na meta-heurística de

Benati e Laporte (1994)[1].

Desta forma o método que desenvolvemos tem a seguinte articulação de Fases e Passos respectivos:

Fase 1: **Localização e afectação de instalações do tipo B**

Subfase 1: **Pré-processamento de dados**

Passo 1 - Construção da matriz distância D ou matriz tempo T:

$$D = d_{lc} = \begin{bmatrix} d_{11} \cdots d_{1n} \\ \vdots \quad \ddots \quad \vdots \\ d_{n1} \cdots d_{nn} \end{bmatrix} \text{ ou } T = t_{lc} = \begin{bmatrix} t_{11} \cdots t_{1n} \\ \vdots \quad \ddots \quad \vdots \\ t_{n1} \cdots t_{nn} \end{bmatrix}$$

onde $l, c = 1, \cdots, n$.

Passo 2 - Construir o vector população por localidade, P:
$P = P_l = [P_1, \cdots, P_n]$, $l = 1, \cdots, n$.

Passo 3 - Actualizar o vector P.
Fazer $\frac{P_l}{v\max} = NIB$, $l = 1, \cdots, n$, onde
$v\max$ =Valor máximo de população afectada a um serviço do tipo B
e
NIB =Número de instalações do tipo B.
Se $NIB \in$ N e $NIB > 1$ então $B_c = NIB - 1$ (B_c é o número de serviços tipo B localizados em c) e $P_l = v\max$. Se NIB é fraccionário maior que 2 e se o algarismo decimal for menor a 5 ou igual ou superior a 5, arredondar NIB à unidade por defeito ou por excesso, respectivamente, diminuído de 1 e $P_l = v\max$; Se não $P_l = P_l$. Actualizar P em conformidade.

Passo 4 - Construção da matriz ponderada MP:

$$MP_{lc} = \begin{bmatrix} P_1 d_{11} \cdots P_1 d_{1n} \\ \vdots \quad \ddots \quad \vdots \\ P_n d_{n1} \cdots P_n d_{nn} \end{bmatrix} \text{ ou } MP_{lc} = \begin{bmatrix} P_1 t_{11} \cdots P_1 t_{1n} \\ \vdots \quad \ddots \quad \vdots \\ P_n t_{n1} \cdots P_n t_{nn} \end{bmatrix} \text{ onde}$$

$l, c = 1, \cdots, n$.

Passo 5 - Restrições de capacidade (População necessária para abrir uma instalação tipo B):

$$v\min < \sum_{l \in LOCB_c} P_l < v\max, \qquad c \in SB$$

onde:
$v\min$ =Valor mínimo de população afectada a cada instalação tipo B.
$v\max$ =Valor máximo de população afectada a cada instalação tipo B.
SB =Conjunto de localidades com instalações tipo B localizadas.
$LOCB$ =Conjunto de localidades afectadas ao serviço tipo B.
Introdução dos parâmetros $v\min$ e $v\max$.

Passo 6 - Restrições de zonalidade:
Se $l \notin Z_z$ e $c \in Z_z$, isto é $l \not\subset z$ e $c \subset z$, ou $l \in Z_z$ e $c \notin Z_z$, isto é $l \subset z$ e $c \not\subset z$, fazer:
MP_{lc}Número suficientemente grande. Caso contrário $MP_{lc} = MP_{lc}$;
onde z e Z são definidos como índice da zona e conjunto e conjunto de localidades de z, respectivamente.

Subfase 2: Algoritmo Tipo Ganancioso

Passo 1 - Seja $S_0 = \emptyset$ e $p = 1$

Passo 2 - Seja $S_p = S_{p-1} \cup V_c$, onde V_c representa o índice do vértice (localidade) que permite a maior redução da distância média:
$\Delta\left(S_{p-1}, V_c\right) = \min\limits_{V_c \in V} \left[Z\left(S_{p-1} \cup V_c\right) - Z\left(S_{p-1}\right)\right]$ onde $V = \{\text{Vértices}\}$.
Enquanto $S_{p-1} \cup V_c$ não é admissível, $p = p+1$ e repetir o Passo 2. Caso contrário ir para a Subfase 3.

Comentário: Sendo $V = \{V_1, V_2, \ldots, V_n\}$, o conjunto dos vértices ou localidades, S_p é um subconjunto de V com p elementos, P_l é a população ou peso da localidade l e, para qualquer subconjunto S_p de V, tem-se:
$d\left(V_l, S_p\right) = \min\limits_{V_c \in S_p} d\left(X_l, X_c\right)$. Donde $S_{p-1} \cup V_c = \sum\limits_{l \in V} P_l d\left(S_{p-1}, V_c\right)$ e $Z\left(\, S_{p-1} \cup V\,\right)_c = \sum\limits_{l \in V} \min\left[\, P_l d\left(S_{p-1}, V_c\right)\,\right]$.

Subfase 3: Heurística Tipo Teitz e Bart

Passo 1 - Seja $S^* = S_p$ e $Z_s^* = Z\left(S_p\right)$

Passo 2 - Seja $I = 0$, $Z_S^0 = Z_S^*$

Passo 3 - Fazer $S_I = S_{I-1} - V_S + V_A$, onde $V_S \in S_{I-1}$ e $V_A \in (V \backslash S_{I-1})$. Verificar a admissibilidade. Caso seja admissível guardar a solução.

Passo 4 - Se $Z\left(S_I\right) < Z_S^*$ e S_I é admissível, $S^* = S_I$ e $Z_S^* = Z\left(S_I\right)$. Repetir o Passo 2 até que todos os vértices tenham sido permutados.

Passo 5 - Se $Z_S^* < Z\left(S_0\right)$, seja $Z_S^0 = Z_S^*$ e voltar ao Passo 2. Caso contrário, ir para a Subfase 4.

Subfase 4: Meta-heurística (Pesquisa Tabu)

Passo 1 - Seja $I = 0$

Passo 2 - Seja $Z_S^0 = Z_S^*$. Nenhum vértice é tabu.

Passo 3 - Considerem-se todas as soluções adjacentes Z_I^i de Z_I, obtidas permutando uma instalação do vértice $V_i' \in S_I$ com o vértice $V_i'' \notin S_I$, guardar as soluções admissíveis. Ordenar as soluções Z_I^i por ordem crescente em função de $Z(S_I^i)$. Ordenar os vértices da mesma maneira.

Passo 4 - Se $Z\left(S_I^i\right) < Z_S^*$ ou se V_i'' não é tabu, fazer $S_{I+1} = S_I$, $Z_S^* = Z\left(S_I\right)$, declarar V_i' tabu até à iteração $I + \theta$, com $\theta \in \left[\overline{\theta}, \underline{\theta}\right]$, e ir para o Passo 5. Caso contrário, fazer $i = i + 1$. Se todos os vértices foram visitados, escolher o vértice V_i'' com a menor marca tabu (o menos tabu) $I + \theta$ e levantar o estado tabu de V_i''. Repetir o Passo 4.

Passo 5 - Fazer $I = I + 1$. Se $I < I_{\max}$, voltar ao Passo 2 da Subfase 4. Caso contrário, fazer $p = p + 1$ e voltar ao Passo 2 da Subfase 2 sempre que a última solução encontrada seja admissível. Se assim não acontecer, guardar a última solução admissível como solução final.

Onde:

$\overline{\theta} = 5$, $\underline{\theta} = 10$ e $I_{\max} = 20$, seguindo a sugestão de Glover e Laguna (1993)[8].

Fase 2: **Localização e afectação de instalações do tipo A**

Subfase 1: **Pré-processamento de dados para instalações do tipo A**

Passo 1 - Restrições para a localização de uma instalação do tipo A:

R_1 : População necessária para abrir uma instalação do tipo A, traduz-se por: $P_{\min} \leq \sum_{l \in LOCA_c} P_l \leq P_{\max}$, $c \notin SB$, onde:
$LOCA$ =Conjunto das localidades afectadas à instalação do tipo A.
$P_{\max}$ =População máxima afectada a cada instalação do tipo A .
$P_{\max}$ =População mínima afectada a cada instalação do tipo A .

R_2 : Distância máxima ou tempo máximo de percurso no meio de transporte habitual, entre as unidades existentes, necessário para ser localizada uma instalação do tipo A, traduz-se:
$d_{lc} > D_{\max}$ ou $t_{lc} > T_{\max}$, com $c \in SB$ ou $c \in SA$ e $l \in LSI$, onde:
SA =Conjunto das localidades com instalações do tipo A
LSI =Conjunto das localidades sem instalação.

R_3 : Número de transportes públicos diário necessário para criar uma instalação do tipo A, traduz-se:
$t'_{lc} = 0$ com $c \in SB$ ou $c \in SA$ e $l \in LSI$.

Passo 2 - Actualizar o vector P. Seja $P' = P$.
-Fazer $\frac{P'_l}{P_{\max}} = NIA$, $l \notin SB$. Se $NIA \in$ N e $NIA > 1$ então $E_c = NIA - 1$ e $P'_l = P_{\max}$. Se NIA for fraccionário maior que 2 e se o algarismo decimal for menor que 5 ou maior ou igual a 5, arredondar NIA à unidade por defeito ou excesso, respectivamente, diminuído de 1 e $P'_l = P_{\min}$. Caso contrário $P'_l = P'_l$. Actualizar P'.
- Senão, fazer $\frac{P'_l}{P_{\min}} = NIA$, $l \notin SB$. Se $NIA \in$ N e $NIA > 1$ então $E_c = NIA-1$ e $P'_l = P_{\min}$. Se NIA for fraccionário maior que 2 e se o algarismo decimal for menor que 5 ou maior ou igual a 5, arredondar NIA à unidade por defeito ou excesso, respectivamente, diminuído de 1 e $P'_l = P_{\min}$.Caso contrário $P'_l = P'_l$. Actualizar P'.

Passo 3 - Reconstruir, para $l, c \notin SB$, a matriz ponderada MPR:

1. Fazer $P'_l d_{lc}$ ou $P'_l t_{lc}$ = Número suficientemente grande , para $t'_{lc} = 0$ e MPR_{lc} com $c \notin LOCB$;

   (a) Fazer $MPR_{lc} = P'_l d_{lc}$ ou $MPR_{lc} = P'_l t_{lc}$ para $c \in LOCB$.

Passo 4 - Reiniciar o MSLAI na Subfase 2 da Fase 1. Ir para o Passo 5.

Passo 5 - Calcular $Z^*(Novo) = \beta Z^*_S + \alpha Z^*_{S'}$, com $\alpha = \frac{Z^*_{S'}}{Z^*_S + Z^*_{S'}}$ e $\beta = 1 - \alpha$.
Guardar $Z^*(Novo)$.

Subfase 2: **Repetir a Subfase 1 da Fase 2 para cada solução admissível encontrada na fase anterior. Se for encontrado um $Z^*_N(Novo)$ inferior, fazer $Z^*(Novo) = Z^*_N(Novo)$.**

4.3 Um Método de Resolução do SLD - MSLD

O MSLD é uma bi-metaheurística baseada na metaheurística do MSLAI e na metaheurística desenvolvida por Rego (1994)[11] para o problema da determinação de Rotas de Veículos com

restrições de capacidade e extensão (duração) de rota[1].

O MSLD consiste:

1. Em localizar, tendo em conta as restrições de zonalidade, as localidades (depósitos), de entre as localidades não afectadas a quaisquer instalações, de onde saem as Extensões Móveis de Cuidados de Saúde Primários, e afectar as restantes localidades às localidades consideradas como depósitos;
2. Determinar para cada localidade (depósito) e respectivas localidades afectadas o número de Extensões Móveis de Cuidados de Saúde Primários e as rotas de custo mínimo tendo em conta as restrições do problema apresentadas.

Assim, o algoritmo tem a seguinte estrutura:

Fase 1: **Localização e afectação de localidades (depósitos)**

Subfase 1: **Pré-processamento de dados para Extensões Móveis de Cuidados de Saúde Primários a prestarem serviços do tipo A a localidades não afectadas a quaisquer instalações do tipo A ou B.**

Passo 1 - Construção da matriz distância D ou matriz tempo de percurso T:

$$D = d_{lc} = \begin{bmatrix} d_{11} \cdots d_{1n} \\ \vdots \quad \ddots \quad \vdots \\ d_{n1} \cdots d_{nn} \end{bmatrix} \text{ ou } T = t_{lc} = \begin{bmatrix} t_{11} \cdots t_{1n} \\ \vdots \quad \ddots \quad \vdots \\ t_{n1} \cdots t_{nn} \end{bmatrix}$$

Passo 2 - Construir o vector população por localidade, P:

$P = P_l = [P_1, \cdots, P_n], \quad l = 1, \cdots, n.$

Passo 3 - Construção da matriz ponderada MP:

$$MP_{lc} = \begin{bmatrix} P_1 d_{11} \cdots P_1 d_{1n} \\ \vdots \quad \ddots \quad \vdots \\ P_n d_{n1} \cdots P_n d_{nn} \end{bmatrix} \text{ ou } MP_{lc} = \begin{bmatrix} P_1 t_{11} \cdots P_1 t_{1n} \\ \vdots \quad \ddots \quad \vdots \\ P_n t_{n1} \cdots P_n t_{nn} \end{bmatrix}$$

onde $l, c = 1, \cdots, n$.

Passo 4 - Restrições de zonalidade:

Se $l \notin Z_z$ e $c \in Z_z$, isto é $l \not\subset z$ e $c \subset z$, ou $l \in Z_z$ e $c \notin Z_z$, isto é $l \subset z$ e $c \not\subset z$, fazer:

MP_{lc} =Número suficientemente grande. Caso contrário $MP_{lc} = MP_{lc}$;

onde Z e z são definidos como em 3.2.1.

Subfase 2: **Subfase 2 da Fase 1 do MSLAI**

Subfase 3: **Subfase 3 da Fase 1 do MSLAI**

Subfase 4: **Subfase 4 da Fase 1 do MSLAI**

Fase 2: **Determina o número de Extensões Móveis de Cuidados de Saúde Primários e as rotas de custo mínimo para cada uma das zonas (estrelas) determinadas na fase anterior. Assim, é aplicada a metaheurística de Rego (1994) [11] ao modelo inteiro seguinte:**

$$Min \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=1}^{m} t_{ij} x_{ijk} \tag{4.1}$$

[1]Em substituição da metaheurística de Rego (1994)[11], pode ser aplicada à resolução do MSLD a heurística baseada em algoritmos genéticos desenvolvida por Schutz et al. (1998) apresentada no IO'98.

Sujeito a:

$$\sum_{k=1}^{m} y_{ik} = \begin{cases} 1 \\ m \end{cases} \qquad i = 1, 2, ..., n \;\; i = 0 \tag{4.2}$$

$$\sum_{i=1}^{n} x_{ijk} = y_{jk} \qquad j = 1, 2, ..., n \;\; k = 1, 2, ..., n \tag{4.3}$$

$$\sum_{j=1}^{n} x_{ijk}^{=} y_{ik} \qquad i = 0, 1, ..., n \;\; k = 1, 2, ..., m \tag{4.4}$$

$$\sum_{i=1}^{n} q_i y_{ik} \leq Q_k \qquad k = 1, 2, ..., m \tag{4.5}$$

$$\sum_{i=1}^{n} t_i y_{ik} + \sum_{i=0}^{n} \sum_{j=1}^{n} t_{ij} x_{ijk} \leq T_k \qquad k = 1, 2, ..., m \tag{4.6}$$

$$\sum_{i,j \in S \times S = 1} x_{ijk} \leq |S| - 1 \qquad \forall S : 1 < |S| < n \;\; S \subseteq \{1, \cdots, n\} \; k = 1, 2, ..., m \tag{4.7}$$

$$x_{n0k} = 1 \qquad k = 1, 2, ..., m \tag{4.8}$$

$$x_{ijk} \in \{0, 1\} \qquad i = 0, 1, ..., n \;\; j = 1, 2, ..., n \;\; k = 1, 2, ..., m \tag{4.9}$$

$$y_{ik} \in \{0, 1\} \qquad i = 1, 2, ..., n \;\; k = 1, 2, ..., m \tag{4.10}$$

onde,

n	Número de localidades, o depósito corresponde à localidade 0
m	Número de veículos
t_{ij}	Duração do percurso entre i e j, $i, j = 0, \cdots, n$
q_i	Procura da localidade i, $i = 0, \cdots, n$
t_i	Tempo para servir a localidade i, $i = 0, \cdots, n$
Q_k	Capacidade do veículo k, $k = 0, \cdots, m$
T_k	Limite temporal da rota k, $k = 0, \cdots, m$

$$x_{ijk} = \begin{cases} 1, & \text{se o veículo } k \text{ vai de } i \text{ para } j \text{ directamente} \\ 0, & \text{caso contrário} \end{cases}$$

$$y_{ik} = \begin{cases} 1, & \text{se a localidade } i \text{ é servida pelo veículo } k \\ 0, & \text{caso contrário} \end{cases}$$

A função objectivo (4.1) minimiza o tempo total de percurso. As restrições (4.2) garantem que as localidades estão afectadas a um veículo, o depósito é servido por todos os veículos. As equações (4.3) e (4.4) garantem que um veículo que chega a uma localidade para a servir também parte dessa localidade servida. As desigualdades (4.5) e (4.6) não permitem que a capacidade dos veículos e o limite da duração das rotas sejam excedidos. As desigualdades (4.7) eliminam a formação de subcircuitos. As igualdades (4.8) obrigam a que os veículos sigam imediatamente da localidade n servida para o depósito. As igualdades (4.9) e (4.10) definem as variáveis como binárias.

As restrições (4.7) e (4.8) podem ser substituídas pelas restrições de Rego (1994) [11]: onde para $i \neq j = 1, 2, ..., n$ e y_i arbitrário

$$y_i - y_j + n \sum_{k=1}^{m} x_{ijk} \leq n - 1 \tag{4.11}$$

as quais eliminam a formação de subcircuitos e impõem que cada rota passe pelo depósito.

5 Uma Abordagem de Resolução do PPCCSPP em Sistemas de Instalações Existentes

Como foi referido, na maioria das situações reais trata-se de melhorar e/ou expandir um sistema de instalações existente. Tendo em conta esta realidade é necessário introduzir no método a possibilidade de actualização do sistema existente de cuidados de saúde primários. Assim, constroem-se para as zonas (estrelas) com alterações, resultantes de movimentos demográficos, excepto os casos de extinção de zona(s), o vector P, a matriz de distância D (ou a matriz de tempo T) e a matriz de adjacência A que passam a constituir os parâmetros de sistema para aplicação do MSLAI. Tendo em conta os resultados do MSLAI constroem-se, de acordo com as restrições de zonalidade, para as localidades não afectadas a qualquer instalação do tipo A ou B, a matriz distância D (ou a matriz de tempo T) e o vector P para aplicação do MSLD.

6 Complexidade Computacional do Método de Resolução do PPCCSPP

Consideremos os subalgoritmos mais significativos MSLAI e MSLD e, de entre todos os procedimentos que os compõem, o Passo 3 da Subfase 4 da Fase 1 do MSLAI como o Passo mais significativo. Deste resultam $(n-p)p$ trocas que implicam $(n-p)p$ cálculos do valor da solução correspondente e a reordenação das soluções. Como o cálculo do valor de cada solução considera, para além das $n-1$ adições, np comparações. E a reordenação das soluções no pior-caso é um procedimento $O(np-p)^2 = O(n^2p^2 - 2np^3 + p^4)$, tem-se, usando o conceito de notação assimptótica do tipo "BIG-O", $O(PPCCSPP) = O(n^2p^2)$ no pior-caso.

7 Experiência Computacional

O SLAI e o MSLAI são avaliados para um conjunto de instâncias geradas aleatoriamente. Inicialmente, para cada instância foram gerados $n \in [o, 63]^2$ nodos seguindo uma distribuição uniforme. Os arcos de conexão obtiveram-se também aleatoriamente para assegurar a conectividade de cada rede. Uma vez conhecidos os arcos, calculou-se a distância euclideana entre todos os nodos conectados entre si, aplicando-se posteriormente um algoritmo de caminho mínimo para obter as matrizes distância de cada rede. Cada matriz distância foi modificada por forma a originar uma matriz distância composta por duas zonas e uma outra matriz distância composta por três zonas. Para que as localidades se agrupassem por zonas foram seleccionados aleatoriamente o número e as localidades a afectar a cada zona. As procuras nos nodos foram geradas aleatoriamente no intervalo $[0, 40000]$. Para cada n foram geradas 100 instâncias. As soluções óptimas foram obtidas usando o algoritmo exacto que faz parte do

sistema CPLEX 6.01. O método foi codificado em C++ e executado num Pentium III a 500 MHZ com 384 MB de RAM. Os resultados são os constantes da Tabela 1.

O uso do método de ponderação para um problema de programação inteira implica não ter a certeza de que todos os pontos-solução são encontrados (Cohon, 1978) [4]. Contudo, no sentido de identificar um conjunto de pontos-solução no espaço objectivo, foram realizadas várias execuções com um conjunto diferente de ponderações em α e β do SLAI[2]. Assim, foram realizadas, para cada instância e para cada composição (número de zonas) correspondente, 12 execuções do CPLEX e da Heurística com diferentes ponderações.

Tabela 1: Comparação das soluções heurísticas com as soluções óptimas

n	NZ	% de Soluções Óptimas	% de Soluções Óptimas ATG	% de Soluções Óptimas T T B	% de Soluções Óptimas Tabu	% Média de Desvio do Óptimo	Tempo Médio Consumido (segundos)	Tempo Médio CPLEX (segundos)
5	1	100	92	98	100	0.00	0.00	0,05
	2	100	52	97	100	0.00	0.00	0,05
	3	100	37	98	100	0.00	0.00	0,05
31	1	100	17	97	100	0.00	0.73	39,11
	2	91.7	9	82	91.7	0.01	1.11	308,34
	3	99. 3	9	79	99.3	0.02	1.79	380,01
62	1	100	15	98	100	0.00	6,03	18211,03
	2	92.4	7	81	85	0.02	6.74	30888,27
	3	93.4	2	82	93.4	0.01	7,01	34446,22

Na Tabela 1: a coluna 1 representa o número de nodos de cada instância; a coluna 2 indica o número de zonas (NZ) que compõem a instância; a coluna 3 apresenta a percentagem de vezes que se obteve a solução óptima para cada composição; a coluna 4 indica a percentagem de vezes que a solução óptima foi obtida pelo ATG - Algoritmo Tipo Ganancioso; a coluna 5 apresenta a percentagem de vezes que a solução óptima foi obtida pela heurística TTB - Tipo Teitz e Bart; a coluna 6 exibe a percentagem de vezes que a solução óptima foi obtida pela Pesquisa Tabu e a coluna 7 mostra o desvio médio do óptimo para os casos em que não foi possível obtê-lo. Da tabela pode observar-se que em todas as instâncias com apenas uma zona, sem decomposição, a heurística obteve a solução óptima. Nos casos de instâncias com duas ou três zonas obteve o valor óptimo em mais de 90% das instâncias onde o efeito da percentagem média de desvio do óptimo sobre a solução é praticamente insignificante. É de salientar que a heurística não possui qualquer sofisticação numérica ou outra como acontece com qualquer versão do CPLEX, constatando-se apenas que nos casos em que não foi possível obter o valor óptimo se verificou um aumento de instalações do tipo B (Centros de Saúde), permitindo uma diminuição das distâncias ponderadas para as instalações do tipo B e um aumento das distâncias ponderadas para as instalações do tipo A (Extensões dos Centros de Saúde), por serem instalações de serviços sucessivamente inclusivos. Nas instâncias sem decomposição foram obtidas melhores

[2] $OBJ_{\text{Novo}} = \alpha \times OBJ_A + \beta \times OBJ_B$, onde OBJ_{Novo} é o novo valor objectivo ponderado, OBJ_A é o valor objectivo que minimiza a distância ponderada às instalações do tipo A e OBJ_B o valor objectivo que minimiza a distância ponderada às instalações do tipo B.

soluções do que nas correspondentes instâncias com composição. Em todas as execuções o tempo de computação da heurística é consideravelmente inferior ao tempo de computação do algoritmo exacto. O aumento da dimensão do problema não parece, pelo menos nesta experiência, influenciar significativamente a qualidade das soluções, apenas é acompanhado por um aumento do Tempo Médio consumido. Em suma, a aplicação do modelo e do método revelou-se bastante eficiente em relação ao método exacto, pelo menos na nossa experiência, e capaz de resolver problemas de grande dimensão eficientemente.

8 Um Exemplo de Aplicação: O Caso da "COVA DA BEIRA"

Considerando a importância de uma abordagem analítica e prática na aplicação do modelo e do método a uma situação real, ilustramos a nossa abordagem com o exemplo da "Cova da Beira" permitindo uma comparação com a situação real e o estabelecimento de parâmetros admissíveis de capacidade/disponibilidade para uma aplicação integrada. Na aplicação consideraram-se os dados da Tabela A1.

Assim, antes de apresentarmos quaisquer resultados da aplicação da Fase 3 do método às localidades não afectadas a Extensões de Saúde e a Serviços tipo A em Centros de Saúde resultantes da aplicação das Fases 1 e 2 do método (ver Tabela 2), importa fazer uma análise dos resultados da aplicação das Fases 1 e 2 do método (ver Tabela 3), considerando a cobertura e a disponibilidade média por habitante do sistema de cuidados de saúde primários da Cova da Beira. Comecemos por analisar, tendo em conta os dados de 2001 cedidos pela Sub-Região de Saúde de Castelo Branco para: o Distrito de Castelo Branco; os Concelhos de Belmonte, Covilhã e Fundão e para a Cova da Beira, a cobertura e a disponibilidade média de consulta por habitante. Na Tabela A2 podem observar-se, para os referidos Concelhos e Cova da Beira, a população residente constante dos resultados preliminares dos Censos 2001 (P), o número de primeiras consultas (NPC), o número de primeiras consultas e seguintes (NPCS), a cobertura (C) e a disponibilidade média de consultas por habitante (DMCH) em 2001. Na mesma tabela pode, também, ser observada a mesma informação para o Continente, tendo por base os dados provisórios relativos a 2001 cedidos pela Direcção-Geral da Saúde. Na tabela $C = NPC/P$ e $DMCH = NPCS/P$. Da observação da coluna DMCH da Tabela 5 e atendendo a que em relação a 1999 em 2001 o NPCS baixou em média 22% (em 1999 a DMCH para o Distrito de Castelo Branco foi de 2,9), parece-nos possível considerar como disponibilidade média de consultas por habitante a meta de 3 consultas por ano. Atendendo aos resultados da Tabela 2 e ao constante em Tien et al. (1984)[17], que cada consulta tem a duração de 20 minutos e que entre cada consulta é necessário um tempo de preparação de 5 minutos, obtemos a Tabela 4 que permite observar o número de horas/dia-útil (NHD) requeridos por cada localidade num intervalo temporal de 11 meses.

Tendo em conta a disponibilidade média de 30 horas/semana de cada médico de família, obtemos uma disponibilidade média de 6 horas/dia-útil por médico de família. Como normalmente por razões de boa gestão e de eficiência os turnos de intervenção do médico de família se desenvolvem por períodos de 3 horas no máximo e a afectação de Extensões Móveis não deverá ser de dedicação exclusiva a uma só localidade, então a afectação de Extensões Móveis de Cuidados de Saúde Primários só poderá ser feita na presença de pelo menos duas localidades cujo comprimento total (tempo de percurso mais tempo de serviço) de cada rota seja não superior a 3 horas. Assim, neste caso particular, apenas as localidades Z2(3) e Z2(18),

Tabela 2: Localidades (freguesias) não afectadas a Extensões de Saúde nem a serviços tipo A em Centros de Saúde.

	LOCALIDADES (FREGUESIAS) NÃO AFECTADAS
Belmonte (Zona 1)	Z1(3,4)
Covilhã (Zona 2)	Z2(1,3,4,15,18,22,28,30)
Fundão (Zona 3)	Z3(21,22)
Cova da Beira (Zona 1, Zona 2 e Zona 3)	Z1(3,4);Z2(1,3,4,15,18,22,28,30);Z3(21,22)

para o Concelho da Covilhã, e Z3(21), para o Concelho do Fundão, poderão ser consideradas à partida para uma possível aplicação da Fase 3 do método desenvolvido. Na Tabela 5 constam os tempos de percurso (TP), o tempo de serviço (TS), o tempo de preparação entre consultas (TE) e o tempo total (TT), para servir cada localidade. Da tabela conclui-se que apenas as localidades Z2(3) e Z3(21) verificam Tempos Totais não superiores a 3 horas. Como Z2(3) e Z3(21) são localidades de Concelhos distintos (Centros de Saúde distintos) e no sistema actual cada localidade e cada Extensão Móvel só pode estar afectada a um só Centro de Saúde (a uma única zona) e não é, como foi referido, aconselhável a dedicação exclusiva de Extensões Móveis a uma localidade apenas, deverão as localidades Z2(3), Z3(21) e todas as localidades não afectadas a Extensões de Saúde terem Extensões de Saúde próprias ou serem afectadas às localidades mais próximas que poderão oferecer os serviços tipo A que recebiam nos Centros de Saúde a que estavam afectadas. Para isso, refazem-se as matrizes de Distâncias Ponderadas e de Transporte, colocando um número suficientemente grande e 0(zero) na matriz de distâncias ponderadas e na matriz de transportes, respectivamente, entre as localidades mais próximas (localidades que distam de Z2(3), Z3(21) e das restantes localidades não afectadas a Extensões de Saúde menos de trinta minutos) e as localidades onde se localizam os Centros de Saúde. No caso da gestão dos recursos se processar numa lógica de inter-zonas (Sistema Local de Saúde da Cova da Beira) e não de zona, parece-nos que poderia ser rentável adquirir uma Extensão Móvel para servir em dois turnos/dia-útil as localidades Z2(3) e Z3(21) com depósito no Centro de Saúde mais equidistante localizado em Z3(18).

Na Tabela A3 apresentam-se para as IB, ICC, ICF e ICB os valores dos parâmetros $P_{\min}$,$P_{\max}$, $P'_{\min}$, $P'_{\max}$,$v\min$ e $v\max$ que verificam a admissibilidade para um cenário de cobertura de localidades que não satisfazendo o critério de População-Base, verificam a conjunção de critérios de Localização de Extensões de Saúde. Na Tabela 6 está representada a cobertura para a Cova da Beira numa lógica inter-zonas.

8.1 Observações Significativas Resultantes da Aplicação da Heurística e do CPLEX 6.01 às IB, ICC, ICF e ICB.

A Tabela 7 mostra, para as IB, ICC, ICF e ICB: as soluções e tempos de CPU (segundos) obtidos pelo CPLEX; as soluções obtidas pelo Algoritmo Tipo Ganancioso (A.T.G.) e as

Tabela 3: Localização de Centros de Saúde e Afectação de Localidades, Localização de Extensões de Saúde e Afectação de Localidades, Valores Objectivos e Tempos de CPU para IB, ICC, ICF e ICB, com a aplicação do CPLEX e da Metaheurística, considerando a população residente dos Censos 2001.

INSTÂNCIA	**ZONA[CS(LA)]**	**ZONA[ES(LA)]**	**VALOR OBJECTIVO**	**TEMPO DE CPU HEURISTICA (segundos)**	**TEMPO DE CPU CPLEX (segundos)**
IB	1(1,2,3,4,5)	2(2,5)	78048	0,00	0,05
ICC	11(2,3,5,6,10,11,13,16,17,20,23,25,28,30,31) 12(1,4,7,8,9,12,14,15,18,19,21,22,24,26,27,29)	1;3;4;7(7,24) 15;18;19;22;28 30	1388173	1,23	309,85
ICF	18(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17) 18(18,19,20,21,22,23,24,25,26,27,28,29,30,31)	19(9,10,19);21;22 24(7,24);25(16,25) 26(8,26);27(13,27)	895870	0,76	39,27
ICB	Z1[1(1,2,3,4,5)] Z2[11(2,3,5,6,10,11,13,16,17,20,23,25,28,30,31) Z2[12(1,4,7,8,9,12,14,15,18,19,21,22,24,26,27,29) Z3[18(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17)] Z3[18(18,19,20,21,22,23,24,25,26,27,28,29,30,31)]	Z1[2(2,5)] Z2[1;3;4;7(7,24)] Z2[15;18;19;22] Z2[28;30] Z3[19(9,10,19)] Z3[21;22;24(7,24)] Z3[25(16,25)] Z3[26(8,26)] Z3[27(13,27)]	2362091	7,3	37574,48

Tabela 4: Número horas/dia-útil requeridos por cada localidade de procura não afectada a Extensões de Saúde.

Localidades Não Afectadas	Belmonte	Covilhã	Fundão	Cova da Beira
Z1(3)	4,7 h/d	-	-	4,7 h/d
Z1(4)	4,4 h/d	-	-	4,4 h/d
Z2(1)	-	3,6 h/d	-	3,6 h/d
Z2(3)	-	1,4 h/d	-	1,4 h/d
Z2(4)	-	3 h/d	-	3 h/d
Z2(15)	-	4,4 h/d	-	4,4 h/d
Z2(18)	-	2,2 h/d	-	2,2 h/d
Z2(22)	-	3,6 h/d	-	3,6 h/d
Z2(28)	-	3,3 h/d	-	3,3 h/d
Z2(30)	-	4,6 h/d	-	4,6 h/d
Z3(21)	-	-	1,1 h/d	1,1 h/d
Z3(22)	-	-	4,2 h/d	4,2 h/d

Tabela 5: Tempo de Percurso, tempo de serviço, tempo de preparação entre consultas e tempo total, para servir as localidades Z2(3), Z2(18) e Z3(21) não afectadas a Extensões de Saúde.

	TP	TS	TE	TT
Z2(3)	0,83 h	1,4 h	0,35 h	2,58 h
Z2(18)	1 h	2,2 h	0,55 h	3,75 h
Z3(21)	0,83 h	1,1 h	0,27 h	2,2 h

Tabela 6: Localização de Centros de Saúde e Afectação de Localidades, Localização de Extensões de Saúde e Afectação de Localidades para IB, ICC, ICF e ICB, considerando que a Extensão Móvel de Saúde tem depósito no Centro de Saúde do Fundão Z3(18) e as rotas verificam-se entre Z3(18) e as Freguesias Z2(3) e Z3(21).

INSTÂNCIA	**ZONA[CS(LA)]**	**ZONA[ES(LA)]**
IB	1(1,2,3,4,5)	2(2,5)
ICC	11(2,3,5,6,10,11,13,16,17,20,23,25,28,30,31) 12(1,4,7,8,9,12,14,15,18,19,21,22,24,26,27,29)	1;4;7(7,24);15;18;19;22;28;30
ICF	18(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17) 18(18,19,20,21,22,23,24,25,26,27,28,29,30,31)	19(9,10,19);22;24(7,24);25(16,25);26(8,26);27(13,27)
ICB	Z1[1(1,2,3,4,5)] Z2[11(2,3,5,6,10,11,13,16,17,20,23,25,28,30,31) Z2[12(1,4,7,8,9,12,14,15,18,19,21,22,24,26,27,29) Z3[18(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17)] Z3[18(18,19,20,21,22,23,24,25,26,27,28,29,30,31)]	Z1[2(2,5)];Z2[1;4;7(7,24);15;18;19;22;28;30] Z3[19(9,10,19);22;24(7,24);25(16,25);26(8,26);27(13,27)]

soluções e tempos de CPU (segundos) obtidos pela aplicação da heurística Tipo Teitz & Bart e da pesquisa Tabu, algoritmos que compõem a Heurística desenvolvida. Como se pode ler na Tabela 7, da aplicação dos dois métodos observa-se que:

- a solução obtida é a mesma para as quatro instâncias;
- para todos as instâncias, os tempos de computação consumidos pelo CPLEX são superiores aos consumidos pela Metaheurística;
- com a aplicação do Algoritmo Tipo Ganancioso aos dados pré-processados se obteve o valor objectivo do CPLEX 6.01 para 75% das instâncias;
- a Heurística Tipo Teitz e Bart obteve para 75% das instâncias o valor objectivo do CPLEX 6.01;
- a Heurística Tipo Teitz e Bart obteve para 75% das instâncias o valor objectivo da pesquisa Tabu;
- em relação ao Algoritmo Tipo Ganancioso e à Heurística Tipo Teitz e Bart, a Pesquisa Tabu obteve melhoramento do valor objectivo em 25% das instâncias;
- a Pesquisa Tabu obteve o valor objectivo do CPLEX 6.01 para as quatro instâncias.

Da aplicação da heurística e do CPLEX a IB, ICC, ICF e ICB foi possível observar a robustez da heurística e a grande sensibilidade do CPLEX a variações dos valores dos parâmetros, principalmente aquando do aumento do valor de $P'_{\max}$ e/ou diminuição do valor de $s_z^{\max}$, produzindo um número muito significativo de nodos. Semelhante tendência se verifica, embora menos significativamente, com o estreitamento dos intervalos populacionais e com a variação dos parâmetros $P_{\max}$e $e_z^{\max}$.

Tabela 7: Soluções e tempos de computação obtidos por CPLEX 6.01 e pela Metaheurística para IB, ICC, ICF e ICB.

INSTÂNCIA	CPLEX 6.01		HEURÍSTICA				
	SOLUÇÃO	TEMPO CPU (SEGUNDOS)	A.T.G.	TIPO TEITZ & BART		TABU	
			SOLUÇÃO	SOLUÇÃO	TEMPO CPU (SEGUNDOS)	SOLUÇÃO	TEMPO CPU (SEGUNDOS)
IB	78048	0,05	78048	78048	0,00	78048	0,00
ICC	1388173	309,85	1388173	1388173	0,6	1388173	0,63
ICF	895870	39,27	895870	895870	0,2	895870	0,56
ICB	2362091	37574,48	17893643	17893643	2,1	2362091	5,2

8.2 Aplicação da Metaheurística e do CPLEX 6.01 às IB, ICC, ICF e ICB Considerando as Perspectivas de Evolução Demográfica para 2011

Como era de esperar, por se tratar de evoluções populacionais proporcionais em relação a 2001 e tendo em conta a Tabela A4, da observação dos resultados conclui-se que em termos de localização-afectação se obtêm os mesmos resultados de 2001.

8.3 Algumas Observações Gerais Significativas

Na Tabela 8 pode observar-se, para cada Concelho da Cova da Beira e para a Cova da Beira, os números de Centros de Saúde, Extensões de Saúde e Extensões Móveis de Saúde determinados pela aplicação da Metaheurística e, para a configuração real. A primeira coluna designa as Instâncias, a segunda coluna representa a configuração real (CONFIGURAÇÃO 1), a terceira coluna representa a configuração sem prever a existência de Extensões Móveis de Saúde (CONFIGURAÇÃO 2) e a quarta coluna descreve a configuração prevendo a existência de Extensões Móveis de Saúde (CONFIGURAÇÃO 3). Da tabela podem observar-se algumas diferenças que se supõe deverem-se à falta de rigor na aplicação dos critérios estabelecidos na construção do sistema de cuidados de saúde primários existente. Com efeito, dos dados da Sub-Região de Saúde de Castelo Branco, pode observar-se que:

- nalguns casos o critério de irradiação para as Extensões de Saúde não são freguesias mas sim localidades anexas das freguesias, como são os casos de: Carvalhal Formoso, Gaia e Montes do Bispo no Concelho de Belmonte, Quintãs, Quintas da Torre e S. Martinho no Concelho do Fundão e Barroca Grande no Concelho da Covilhã;

- no Concelho do Fundão, Escarigo (atendendo aos critérios esta localidade não deve receber serviços do tipo A no Centro de Saúde) não tem Extensão de Saúde e não tem Extensões de Saúde na sua adjacência. No entanto Lavacolhos (atendendo aos critérios deve receber serviços tipo A no Centro de Saúde), que tem menos população que Escarigo, tem uma Extensão de Saúde e tem Extensões de Saúde na sua adjacência.

Legenda:

NCS – Número de Centros de Saúde

NES – Número de Extensões de Saúde

NEMS – Número de Extensões Móveis de Saúde

Tabela 8: Número de Centros de Saúde, Extensões de Saúde e Extensões Móveis de Saúde para as configurações real, sem Extensões Móveis e com Extensões Móveis.

INSTÂNCIA	CONFIGURAÇÃO 1			CONFIGURAÇÃO 2			CONFIGURAÇÃO 3		
	NCS	NES	NEMS	NCS	NES	NEMS	NCS	NES	NEMS
IB	1	7	0	1	3	0	1	3	0
ICC	1	25	0	2	10	0	2	9	1
ICF	1	23	0	1	7	0	1	6	1
ICB	3	55	0	4	20	0	4	18	2

Em suma, verifica-se que a configuração real não corresponde a nenhuma das duas outras configurações pois os critérios aplicados são obviamente diferentes e, por isso, não nos parece cientificamente correcto fazermos qualquer comparação para além da que consta da Tabela 8. A propósito convém salientar que os valores da função objectivo da configuração real são menores por consequência do efeito ad-hoc não se traduzindo contudo sempre em menos custo para os utentes e para o estado. No entanto importa salientar que o modelo e o método desenvolvidos permitem, por um lado, a resolução do Problema do Planeamento da Cobertura de Cuidados de Saúde Primários para sistemas existentes e/ou com critérios diferentes e, por outro lado, o estabelecimento de critérios de Planeamento da Cobertura de Cuidados de Saúde Primários consoante os objectivos.

8.4 Aplicação da Heurística e do Método Exacto às Instâncias IB, ICC, ICF e ICB com Diferentes Ponderações para as Funções Objectivos em SLAI

No sentido de identificar um conjunto de pontos-solução no espaço objectivo, foram realizadas várias execuções com um conjunto diferente de ponderações em α e β do SLAI[3]. Assim, foram realizadas, para cada instância e para cada cenário populacional (2001 e 2011), 12 execuções do CPLEX e da Heurística com diferentes ponderações na tentativa de se obter um conjunto solução. A Tabela 9 mostra que das execuções nenhuma solução produziu diferença objectiva relativamente significativa. Contudo outras execuções foram realizadas tendo-se registado diferenças relativamente significativas nas soluções resultantes do aumento do número de instalações do tipo B (Centros de Saúde), na medida em que permite uma diminuição das distâncias ponderadas para as instalações do tipo B e um aumento das distâncias ponderadas para as instalações do tipo A (Extensões dos Centros de Saúde), por serem instalações de serviços sucessivamente inclusivos.

[3] $OBJ_{Novo} = \alpha \times \ OBJ_A + \beta \times OBJ_B$, onde OBJ_{Novo} é o novo valor objectivo ponderado, OBJ_A é o valor objectivo que minimiza a distância ponderada às instalações do tipo A e OBJ_B o valor objectivo que minimiza a distância ponderada às instalações do tipo B.

Tabela 9: Comparação entre a Heurística e o CPLEX

INSTÂNCIA	% DE SOLUÇÕES ÓPTIMAS	% DE DESVIO MÉDIO DO ÓPTIMO	TEMPO MÉDIO CONSUMIDO (segundos)
IB	100%	0%	0
ICC	83%	1%	1,17
ICF	100%	0%	0,73
ICB	83%	1%	7,16

8.5 Outras aplicações da heurística

Um dos objectivos do trabalho é aplicar o modelo e a heurística ao Caso concreto da Cova da Beira. Não sendo praticável o levantamento de um conjunto mais significativo de instâncias de dimensão real e útil, propomo-nos neste trabalho obter uma base indicadora sobre os limites e a qualidade das soluções da heurística aplicada a instâncias de dimensão real e útil. Neste sentido, na subsecção 8.5.1 descreve-se a aplicação do CPLEX às instâncias IB, ICC, ICF e ICB a partir da solução produzida pela heurística e na subsecção 8.5.2 descreve-se a aplicação da heurística a instâncias de dimensão real e útil com solução conhecida.

8.5.1 Aplicação do CPLEX 6.01, usando "Priority Orders", às instâncias IB, ICC, ICF, ICB

Consideraram-se as soluções encontradas pela heurística para IB, ICC, ICF e ICB como soluções iniciais para o CPLEX, procurando, desta forma, melhorar as soluções (obter limites inferiores) para as instâncias. Para cada instância construiu-se um ficheiro ***.ord*** onde constam as variáveis inteiras obtidas pela heurística e o respectivo valor e a ordem de fixação de cada uma das variáveis. Em apêndice encontra-se apenas, por razões de espaço, parte do ficheiro "***icb.ord***" para a instância da Cova da Beira com a indicação apenas das variáveis com valor 1. Em todo o caso, atendendo a que nas variáveis: $i, j, k = 0, ..., 4$, correspondem a localidades de IB; $i, j, k = 5, ..., 35$, correspondem a localidades de ICC e $i, j, k = 36, ..., 66$, correspondem a localidades de ICF, é possível extraírem-se a partir da porção do ficheiro "***icb.ord***" de ICB em apêndice os ficheiros ***.ord*** para IB, ICC e ICF, respectivamente. Dos ficheiros ***.log*** do CPLEX em apêndice, podem observar-se todos os procedimentos realizados, todas as parametrizações estabelecidas e os resultados obtidos nesta aplicação. A Tabela 10 apresenta os resultados principais retirados dos ficheiros ***.log***. A coluna 1 designa as instâncias, a coluna 2 apresenta as soluções obtidas pela heurística para cada instância, a coluna 3 apresenta as soluções obtidas pelo CPLEX a partir das soluções produzidas pela heurística, a coluna 4 apresenta o número de iterações do CPLEX, a coluna 5 apresenta o tempo de CPU consumido pela heurística para cada instância e a coluna 6 apresenta o tempo de CPU consumido pelo CPLEX para cada instância. Da Tabela 10 e dos ficheiros ***.log*** pode observar-se que o CPLEX, apesar de tomar como solução inicial o limite inferior da heurística, não obteve um limite inferior melhor que o da heurística (a distância das soluções da heurística ao óptimo do CPLEX é zero) e que apresenta um consumo de tempo de CPU muito superior ao da heurística. Dos ficheiros ***.log***

constata-se que nestas condições o CPLEX não produziu nodos adicionais[4].

Tabela 10: Resultados principais da aplicação do CPLEX, usando "Priority Orders", a IB, ICC, ICF e ICB.

INSTÂNCIA	SOL. HEURÍSTICA	SOL. CPLEX	NÚMERO ITERAÇÕES CPLEX	TEMPO CPU CPLEX (segundos)	TEMPO CPU HEURÍSTICA (segundos)
IB	78048	78048	55	0,05	0,00
ICC	1388173	1388173	4253	310,73	1,23
ICF	895870	895870	2691	39,24	0,76
ICB	2362091	2362091	29130	37575,03	7,3

8.5.2 Aplicação do CPLEX 6.01 e da heurística a três problemas adaptados com solução conhecida

Foram aplicados o CPLEX e a heurística a um problema (PC) adaptado de Christofides (1975)[3], a um problema (PS) real e útil adaptado de Serra (1990)[13] e a um problema (PE) real e útil adaptado de El-Tell (1983)[7]. Os resultados são os constantes da Tabela 11, onde se pode observar que as soluções obtidas tanto pelo CPLEX como pela heurística são as conhecidas.

Tabela 11: Soluções conhecidas e obtidas pelo CPLEX e pela Heurística para PC, PS e PE.

PROBLEMA	N.º DE NODOS	SOL. CONHECIDA	SOL. CPLEX	SOL. HEURÍSTICA
PC	12	72	72	72
PS	79	631[5]	631	631
PE	31	21433	21433	21433

9 Conclusões Significativas

Sumariamente, propomos um modelo e um método de resolução para tratar o Problema de Planeamento da Cobertura de Cuidados de Saúde Primários em Portugal. O método heurístico caracteriza-se por ser multi-heurístico e revelou-se bastante eficiente e robusto em relação ao método exacto, pelo menos para as instâncias tratadas, e capaz de resolver problemas de grande dimensão eficientemente. A aplicação do método desenvolvido ao modelo, por um lado, revelou um desajustamento (na aplicação dos critérios estabelecidos para a localização e afectação de instalações de cuidados de saúde primários em Portugal) na constituição do sistema de cuidados de saúde primários da Cova da Beira e, por outro lado, permitiu observar que o método é em si uma ferramenta de simulação e de apoio à decisão (uma metodologia de acção) no Planeamento

[4]Não produziu "Branch-and-Bound Tree".

[5]Para 6 instalações do tipo A e 3 instalações do tipo B, de acordo com a solução de Serra [13].

da Cobertura de Recursos de Cuidados de Saúde Primários em Portugal, ao estabelecer um quadro referencial de parâmetros admissíveis de capacidade. No caso concreto, que pode ser generalizado, registou a necessidade de existência de duas ou mais localidades, não afectadas a qualquer Extensão de Saúde, com uma capacidade máxima de 270 habitantes/localidade, para a constituição de um sistema de Extensões Móveis de Saúde por zona ou inter-zonas. O método desenvolvido aplicado ao modelo permite a aplicação potencial a outros sistemas hierárquicos como o Sistema Educativo, o Sistema Bancário, etc.. A aplicação simples do Algoritmo Tipo Ganancioso à solução resultante da aplicação do modelo e do método permite, tendo em conta eventuais prioridades e/ou escassez de recursos, um faseamento (tendo em conta os critérios população e distância) da implementação da solução.

10 Sugestões de Trabalho Futuro

Ao longo deste estudo, entre outras, algumas dificuldades foram identificadas que merecem ser estudadas no futuro. Assim, o método poderá, com as necessárias adaptações, ser aplicado ao SLAITE (Submodelo de Localização-Afectação de Instalações com Restrições de Tempo de Espera) apresentado, por forma a permitir tirar vantagem da sua estrutura. Nomeadamente a inclusão de boas soluções estocásticas para os parâmetros de sistema não-determinísticos que integram as restrições estocásticas. Para qualquer dos submodelos SLAI e SLAITE apresentados é importante que o método aqui desenvolvido adaptado permita considerar um esquema de Custo-afectação com base na Teoria dos Jogos, baseado no esquema proposto por Granot (1983)[9].

Finalmente, para qualquer modelo que constitua uma combinação dos submodelos desenvolvidos importa que o método desenvolvido considere:

- uma Análise Demográfica no sentido de se obterem Perspectivas de Evolução Demográficas mais rigorosas, por forma a compensar o mais possível a estocasticidade do processo;
- a integração num Sistema Visual Interactivo de Apoio à Decisão como uma metodologia de acção.

11 Agradecimentos

Os autores são muito gratos a todas, e foram muitas, as instituições e ilustres personalidades que possibilitaram a realização deste trabalho. Os autores expressam a sua gratidão a um revisor anónimo pelo contributo dos seus comentários e sugestões.

12 Referências Bibliográficas

[1] Benati, S. e Laporte, G. (1994).Tabu Search Algorithms for the (r—Xp)-Medianoid and (r—p)-Centroid Problems, Location Science, Vol. 2, n.º4, pp.193-204.

[2] Branco, I. M. e Coelho, J. D. (1984). Formulações Matemáticas da P-Mediana Hamiltoniana, Nota n.º 10, Departamento de Estatística, Investigação Operacional e Computação, Faculdade de Ciências de Lisboa.

[3] Christofides, N. (1975). Graph Theory: An Algorithmic Approach, Academic Press, London.

[4] Cohon, J. L. (1978). Multiobjective Programming and Planning, Academic Press, Inc.

[5] Densham, P. e Rushton, G. (1992). Strategies for solving large location-allocation problems by heuristic methods, Environment and Planning A, Vol. 24, pp. 289-304.

[6] DGOTDU (1997). Equipamentos Colectivos e Normas para a Programação e Caracterização de Redes, 9 Colecção Divulgação.

[7] El-Tell, K. (1983). An Analytic Approach to Primary Health Care Planning in Developing Countries with an Application to Jordan, Ph.D., Rensselaer Polytechnic Institute.

[8] Glover, F. e Laguna, M. (1993). Tabu Search. In C. Reeves, editor, Modern Heuristics Techniques for Combinatorial Problems, Blackwell, Oxford.

[9] Granot, D. (1987). The role of cost allocation in locational models, Operations Research, Vol. 35, p. 234.

[10] Marianov, V. e Serra, D. (1998). Probabilistic, Maximal Covering Location-Allocation Models for Congested Systems, Journal of Regional Science, Vol. 38, N.º3, pp. 401-424.

[11] Rego, C. (1994). Uma Heurística Tabu para a Determinação de Rotas de Veículos, Investigação Operacional, Vol. 14, pp. 207-232.

[12] Santana, P. (1993). Acessibilidade e Utilização dos Serviços de Saúde. Ensaio Metodológico em Geografia da Saúde, C.C.R.C e A . R. S. C..

[13] Serra, D. (1990). Location and districting of hierarchical facilities, Ph.D., The Johns Hopkins University.

[14] Serra, D. e ReVelle, C. (1993).The PQ-Median Problem: Location and Districting of Hierarchical Facilities, Location Science, 1(4), pp. 299-312.

[15] Serrão, D. et al (1998). Recomendações para uma reforma estrutural, Conselho de Reflexão sobre a Saúde.

[16] Teitz, M. B. e Bart, P. (1968). Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph, Operations Research 16, pp. 955-965.

[17] Tien, J. M. e El-Tell, K. (1984). A Quasihierarchical Location-Allocation Model for Primary Health Care Planning, IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-14, N.º3.

APÊNDICE :

Legenda: NF (Número da Freguesia)

Tabela A.1: Resultados preliminares dos Censos 2001 da população residente para o mesmo ano, e Perspectivas de Evolução Demográfica da população residente em 2011 nos Concelhos de Belmonte, Covilhã e Fundão.

BELMONTE (Z1)			
Freguesias	**NF**	**2001**	**2011**
Belmonte	1	3227	2840
Maçainhas	5	384	338
Inguias	4	846	744
Col.Torre	3	894	787
Caria	2	2240	1971
Totais		7591	6680

COVILHÃ (Z2)			
Freguesias	**NF**	**2001**	**2011**
Cantar-Galo	6	2403	2115
Boidobra	5	2842	2501
Barco	4	573	504
Ald. Souto	3	264	232
Ald.S.F.Assis	1	686	604
Ald.Carvalho	2	2074	1825
Casegas	7	701	617
Conceição	10	7684	6762
Cortes Meio	8	980	862
Coutada	9	475	418
Dominguiso	14	1188	1045
Erada	15	845	744
Ferro	16	1846	1624
Orjais	17	859	756
Ourondo	18	418	368
Paul	619	1816	1598
Peraboa	620	1074	945
Peso	21	780	686
Santa Maria	611	2477	2180
S.J. da Beira	22	697	613
S. Martinho	12	4937	4345
S. Pedro	13	2732	2404
Sarzedo	23	174	153
S.S.Miguel	24	694	611
Teixoso	25	4428	3897
Tortosendo	26	5405	4756
U. da Serra	27	1369	1205
Vale Formoso	28	638	561
Vales do Rio	29	839	738
Verdelhos	30	876	771
Canhoso	31	1733	1525
Totais		54507	47965

FUNDÃO (Z3)			
Freguesias	**NF**	**2001**	**2011**
Alpedrinha	6	1179	1038
A .N.Cabo	5	689	606
Ald.Joanes	4	969	853
Alcongosta	3	575	506
Alcaria	2	1270	1118
Alcaide	1	767	675
At.Campo	7	650	572
Barroca	8	607	534
B.de Baixo	9	275	242
B.de Cima	10	466	410
Capinha	11	590	519
Castelejo	12	821	722
Cast.Novo	13	362	319
Donas	14	931	819
Enxames	15	605	532
Escarigo	16	314	276
Fatela	17	485	427
Fundão	18	8934	7862
J.de Cima	19	352	310
Lavacolhos	20	245	216
M.Rainha	21	214	188
Orca	22	816	718
Peroviseu	23	830	730
P.Atalaia	24	789	694
Salgueiro	25	826	727
Silvares	26	1089	958
Soalheira	27	1140	1003
S.da Casa	28	984	866
Telhado	29	623	548
V.Prazeres	30	1507	1326
Valverde	31	1452	1278
Totais		31356	27592

Tabela A.2: População residente, número de primeiras consultas, número de primeiras consultas e seguintes, cobertura e a disponibilidade média de consultas/habitante, por Concelho, Cova da Beira, Distrito e Continente.

	P	**NPC**	**NPCS**	**C**	**DMCH**
Belmonte	7591	3695	13906	49%	1,8
Covilhã	54507	35185	109760	65%	2
Fundão	31356	17187	58341	55%	1,9
Cova da Beira	93454	56067	182007	60%	1,9
Distrito de Castelo Branco	216511	142397	458535	66%	2,1
Continente	10355824	6725911	20420714	65%	2

Tabela A.3: Valores mínimos e máximos de população necessários para a localização de Extensões de Saúde e Centros de Saúde em IB, ICC, ICF e ICB, com a aplicação do CPLEX 6.01 e da Metaheurística.

INSTÂNCIA	$P_{\min}$	$P_{\max}$	$P'_{\min} = v \min$	$P'_{\max} = v \max$
IB	846	2624	6000	7591
ICC	264	1816	6000	32104
ICF	214	1696	6000	31356
ICB	214	2624	6000	39695

Tabela A.4: Valores mínimos e máximos de população necessários para a localização de Extensões de Saúde e Centros de Saúde em IB, ICC, ICF e ICB, com a aplicação da Metaheurística.

INSTÂNCIA	$P_{\min}$	$P_{\max}$	$P'_{\min} = v \min$	$P'_{\max} = v \max$
IB	744	2309	6000	6680
ICC	232	1598	6000	28251
ICF	188	1492	6000	27592
ICB	188	2309	6000	28251

Porção do ficheiro "icb.ord"

sp0	186	1	pp3	148	1	s0625	110	1	s5448	72	1	p5048	34	1
sp23	185	1	pp6	147	1	s0825	109	1	s6048	71	1	p5648	33	1
sp25	184	1	pp7	146	1	s1125	108	1	s6548	70	1	p4248	32	1
sp48	183	1	pp8	145	1	s1325	107	1	s4048	69	1	p5548	31	1
ps00	182	1	pp11	144	1	s1425	106	1	s5248	68	1	p6248	30	1
ps2323	181	1	pp16	143	1	s1525	105	1	s4448	67	1	p5748	29	1
ps2525	180	1	pp19	142	1	s1625	104	1	s4348	66	1	p4148	28	1
ps4848	179	1	pp20	141	1	s1925	103	1	s5948	65	1	p5448	27	1
p00	178	1	pp24	140	1	s2025	102	1	s5348	64	1	p4448	26	1
p2323	177	1	pp32	139	1	s2225	101	1	p22	63	1	p4348	25	1
p2525	176	1	pp34	138	1	s2525	100	1	p33	62	1	p5948	24	1
p4848	175	1	pp36	137	1	s2425	99	1	p11	61	1	p5348	23	1
ps10	174	1	pp40	136	1	s2825	98	1	p41	60	1	p3736	22	1
ps20	173	1	pp52	135	1	s3025	97	1	p0523	59	1	p3836	21	1
ps30	172	1	pp60	134	1	s3125	96	1	p0923	58	1	p3636	20	1
ps0625	171	1	pp63	133	1	s3325	95	1	p1023	57	1	p6160	19	1
ps0723	170	1	pp64	132	1	s5848	94	1	p1223	56	1	p6060	18	1
ps0825	169	1	pp65	131	1	s4648	93	1	p2623	55	1	p6665	17	1
ps1125	168	1	s00	130	1	s4948	92	1	p1723	54	1	p6565	16	1
ps1625	167	1	s10	129	1	s4748	91	1	p1823	53	1	p3940	15	1
ps1925	166	1	s20	128	1	s4548	90	1	p2123	52	1	p4040	14	1
ps2025	165	1	s30	127	1	s5048	89	1	p2723	51	1	p5152	13	1
ps2425	164	1	s40	126	1	s6148	88	1	p2923	50	1	p5252	12	1
ps3223	163	1	s0523	125	1	s3948	87	1	p3523	49	1	p66	11	1
ps3423	162	1	s0723	124	1	s3748	86	1	p1325	48	1	p77	10	1
ps3648	161	1	s0923	123	1	s3848	85	1	p1425	47	1	p88	9	1
ps4048	160	1	s1023	122	1	s5648	84	1	p1525	46	1	p1616	8	1
ps5248	159	1	s1223	121	1	s4248	83	1	p2225	45	1	p1919	7	1
ps6048	158	1	s2323	120	1	s5148	82	1	p3025	44	1	p2020	6	1
ps6348	157	1	s2623	119	1	s5548	81	1	p3125	43	1	p2424	5	1
ps6448	156	1	s1723	118	1	s6248	80	1	p3325	42	1	p3232	4	1
ps6548	155	1	s1823	117	1	s6648	79	1	p1111	41	1	p3434	3	1
ss0	154	1	s2123	116	1	s5748	78	1	p2811	40	1	p6363	2	1
ss23	153	1	s2723	115	1	s4848	77	1	p5848	39	1	p6464	1	1
ss25	152	1	s2923	114	1	s3648	76	1	p4648	38	1			
ss48	151	1	s3223	113	1	s4148	75	1	p4948	37	1			
pp1	150	1	s3423	112	1	s6348	74	1	p4748	36	1			
pp2	149	1	s3523	111	1	s6448	73	1	p4548	35	1			

Legenda : Em relação ao modelo SLAI $sp_j = se_j$;$ps_{jk} = es_{jk}$; $pp_j = e_j$; $p_{ij} = e_{ij}$; $ss_k = s_k$

ib.log

```
Log started (V6.0.1) Thu Mar 10 18:23:14 2005

New value for time limit in seconds: 1.296e+06
New value for use of starting variable values: 1
New value for interval for printing mixed integer node display: 1
New value for mixed integer optimality gap tolerance: 0

Problem 'ib.lp' read.
Read time =    0.01 sec.
Priority order 'ib.ord' read.
Tried aggregator 1 time.
MIP Presolve eliminated 47 rows and 5 columns.
MIP Presolve modified 15 coefficients.
Reduced MIP has 212 rows, 85 columns, and 680 nonzeros.
Presolve time =   -0.00 sec.
Pseudo-cost variable selection.
MIP start values provide initial solution.

Integer optimal solution:  Objective =    7.8048000000e+04
Solution time =    0.05 sec.  Iterations = 55  Nodes = 0
```

icf.log

```
Log started (V6.0.1) Thu Mar 10 19:21:54 2005

New value for time limit in seconds: 1.296e+06
New value for interval for printing mixed integer node display: 1
```

```
New value for use of starting variable values: 1
New value for mixed integer optimality gap tolerance: 0

Problem 'icf.lp' read.
Read time =    0.91 sec.
Priority order 'icf.ord' read.
Tried aggregator 1 time.
MIP Presolve eliminated 1893 rows and 31 columns.
MIP Presolve modified 961 coefficients.
Reduced MIP has 31932 rows, 2945 columns, and 100130 nonzeros.
Presolve time =    1.18 sec.
Pseudo-cost variable selection.
MIP start values provide initial solution.

Integer optimal solution:  Objective =    8.9587000000e+05
Solution time =   39.24 sec.  Iterations = 2691  Nodes = 0
```

icc.log

```
Log started (V6.0.1) Thu Mar 10 18:58:43 2005

New value for time limit in seconds: 1.296e+06
New value for use of starting variable values: 1
New value for interval for printing mixed integer node display: 1
New value for mixed integer optimality gap tolerance: 0

Problem 'icc.lp' read.
Read time =    0.93 sec.
Priority order 'icc.ord' read.
Tried aggregator 1 time.
MIP Presolve eliminated 1893 rows and 31 columns.
MIP Presolve modified 930 coefficients.
Reduced MIP has 31932 rows, 2945 columns, and 100130 nonzeros.
Presolve time =    1.19 sec.
Pseudo-cost variable selection.
MIP start values provide initial solution.
Objective is integral.
Generating cliques on row 17664
Clique cuts generated:  1922

        Nodes                                        Cuts/
   Node  Left     Objective  Iinf  Best Integer     Best Node    ItCnt         Variable B Parent  Depth

      0     0  1386595.0000     4  1388173.0000  1386595.0000     3938
                     cutoff        1388173.0000    Covers:  1     4190

Clique cuts were discarded.
Cover cuts applied:  1

Integer optimal solution:  Objective =    1.3881730000e+06
Solution time =  310.73 sec.  Iterations = 4253  Nodes = 0
```

icb.log

```
Log started (V6.0.1) Tue Mar 22 13:40:46 2005

New value for time limit in seconds: 1.296e+06
New value for interval for printing mixed integer node display: 1
New value for use of starting variable values: 1
New value for mixed integer optimality gap tolerance: 0

Problem 'icb.lp' read.
Read time =   11.41 sec.
Priority order 'icb.ord' read.
```

```
Presolve has eliminated 8917 rows and 67 columns...
Tried aggregator 1 time.
Presolve has eliminated 8917 rows and 67 columns...
MIP Presolve eliminated 8917 rows and 67 columns.
MIP Presolve modified 4489 coefficients.
Reduced MIP has 310216 rows, 13601 columns, and 952070 nonzeros.
Presolve time =   12.05 sec.
Pseudo-cost variable selection.
MIP start values provide initial solution.
Objective is integral.

Clique cuts generated:  8980

        Nodes                                          Cuts/
   Node  Left     Objective  IInf  Best Integer     Best Node    ItCnt  Variable B Parent  Depth

      0     0  2360513.0000     8  2362091.0000  2360513.0000    28810
                    cutoff        2362091.0000    Covers:  1    29096

Clique cuts were discarded.
Cover cuts applied:  1

Integer optimal solution:  Objective =    2.3620910000e+06
Solution time = 37575.03 sec.  Iterations = 29130  Nodes = 0
```

Métodos de Melhora de Ordenação em DEA Aplicados à Avaliação Estática de Tornos Mecânicos

Fabiana Rodrigues Leta [♮] João Carlos C. B. Soares de Mello [*†]
Eliane Gonçalves Gomes [‡] Lidia Angulo Meza [§]

[♮] Programa de Pós-Graduação em Engenharia Mecânica - Universidade Federal Fluminense
Rua Passo da Pátria 156, São Domingos, 24210-240, Niterói, RJ, Brasil
fabiana@ic.uff.br

[†] Departamento de Engenharia de Produção - Universidade Federal Fluminense
Av. Dr. Júlio Soares de Arruda 803, Parque São Quirino, 13088-300, Campinas, SP, Brasil
jcsmello@producao.uff.br

[‡] Embrapa – Parque Estação Biológica
Av. W3 Norte final, Asa Norte, 70770-901, Brasília, DF, Brasil
eliane.gomes@embrapa.br

[§] Departamento de Ciência dos Materiais - Universidade Federal Fluminense
Av. dos Trabalhadores 420, 27255-125, Volta Redonda, RJ, Brasil
lidia_a_meza@yahoo.com

Abstract

In this paper we present a turning machine study with the use of Data Envelopment Analysis (DEA). The main goal is to obtain a unique performance measure aggregating all the partial measures without personal opinions. We use DEA CCR classic model together with certain techniques for ranking improvement. We include a brief description of those techniques as well as some new developments in order to achieve the goal of our problem. The models are used to evaluate four turning machines belonging to Machine-Tools Laboratory of Universidade Federal Fluminense.

Resumo

Este artigo faz um estudo preliminar da qualidade de máquinas-ferramenta com Análise Envoltória de Dados (DEA). O objectivo é agregar várias medidas de erro estáticas em uma única medida, mas sem incorporar opiniões subjectivas. Para este fim, além do modelo DEA CCR são usadas técnicas adicionais para melhorar a discriminação. Para adequar-se

[*] Autor correpondente.

Apdio © 2005 *Associação Portuguesa de Investigação Operacional*

à especificidade do problema, foi necessário introduzir algumas alterações à forma clássica de aplicação dos métodos mencionados. Os modelos são aplicados a um conjunto de quatro tornos do Laboratório de Usinagem da Universidade Federal Fluminense.

Keywords: DEA; Ranking improvement; Machine tools.

Title: Turning Machine Statica Evaluation using Methods for Ranking Improvement of DEA Models

---

1 Introdução

Na fabricação de uma peça, suas dimensões, seu acabamento superficial e seus desvios geométricos devem satisfazer ao exigido no projecto, o que envolve, em muitos casos, tolerâncias estreitas. Alcançar tal objectivo depende sobremaneira da qualidade de fabricação da máquina-ferramenta utilizada. Desse modo, o aumento da demanda de componentes fabricados com maior qualidade tem conduzido a consideráveis pesquisas sobre os meios de melhorar e preservar a qualidade de fabricação das máquinas-ferramenta (Branco e Leta, 1996).

Para melhorar a qualidade é necessário ter meios de medi-la. Ora, a qualidade de uma máquina-ferramenta depende de vários factores e, portanto, a medida de qualidade considera múltiplos critérios, com a necessária subjectividade da atribuição de maior ou menor importância a alguns factores, dependendo da opinião de um decisor.

Para conseguir uma medida mais confiável é recomendável usar um método que permita considerar vários factores com um mínimo de subjectividade. A Análise Envoltória de Dados (*Data Envelopment Analysis* – DEA) permite esse tipo de avaliação, já que os diversos factores em análise são ponderados por meio de problemas de programação linear, sem a interferência dos decisores.

Neste artigo são avaliados quatro tornos do Laboratório de Usinagem da Universidade Federal Fluminense com um modelo DEA CCR (Charnes et al., 1978) de *output* constante (Lovell e Pastor, 1999) e orientado a *inputs*. Devido à relação entre o número de variáveis e o número de DMUs não seguir a recomendação empírica de ao menos três DMUs para cada variável, surgem empates entre as unidades. Assim, são necessários modelos adicionais que promovam uma melhor discriminação.

Para promover o desempate entre as unidades eficientes são usadas quatro abordagens, com variações teóricas em relação aos modelos clássicos, de forma a melhor adaptarem-se ao problema abordado. As abordagens em questão, cujos aspectos teóricos são revistos na secção 5, são: supereficiência (Andersen e Petersen, 1993; Seiford e Zhu, 1999), restrições directas aos pesos (Dyson e Thanassoulis, 1988), avaliação cruzada (Sexton, 1986) e fronteira invertida (Yamada et al., 1994; Entani et al., 2002; Angulo Meza et al., 2003; Soares de Mello et al., 2005). Os resultados obtidos pelas diversas abordagens são então comparados com o objectivo de justificar a escolha da melhor máquina de acordo com o perfil de cada decisor.

2 Verificação de Máquinas-Ferramenta

2.1 Principais aspectos

Ao planear o procedimento de fabricação de uma peça, o engenheiro de fabricação deve conhecer o estado das máquinas em relação à qualidade de fabricação, com o objectivo de reduzir o seu custo operacional. Em conjunto com outras actividades para prevenção de paradas e falhas, o controle de qualidade sobre a máquina é parte importante de qualquer sistema organizado de produção. Este controle pode ser chamado de Avaliação da Integridade Funcional da Máquina-Ferramenta, e seu objectivo é determinar que variações ocorrem na máquina e qual a sua influência nas tolerâncias de fabricação. Essa determinação deve ser objectiva e exacta.

Uma vantagem deste tipo de avaliação é a garantia da rastreabilidade da máquina aos padrões de qualidade, permitindo, deste modo, a uniformização da produção. Se os padrões de qualidade forem estabelecidos por normas internacionais, pode-se vislumbrar esta uniformização segundo parâmetros mundiais de qualidade, ou seja, um produto manufaturado em qualquer país seguindo tais normas possuirá aceitação internacional. Entretanto, ainda não existem normas que tratem especificamente da avaliação funcional de máquinas-ferramenta durante seu uso. O objectivo fundamental desta análise é minimizar o erro geométrico da máquina-ferramenta, que gera peças não conformes.

Os fabricantes de máquinas-ferramenta são cada vez mais solicitados a reduzir o tempo de produção de componentes nas máquinas, e melhorar a sua a qualidade. Em outras palavras, as máquinas-ferramenta devem ser capazes de produzir peças respeitando-se as tolerâncias de projeto cada vez mais estreitas, em condições econômicas de usinagem, quando manipuladas por um operador qualificado.

É necessário que a qualidade de fabricação da peça dependa apenas dos seguintes factores (Agostinho et al., 1977): a) fixação na base, rigidez da máquina, das partes componentes e dispositivos de fixação; b) alinhamento das várias partes da máquina em relação umas às outras, já que a execução de uma geometria de formas variadas é baseada no movimento relativo entre várias partes da máquina e, por isso, no alinhamento destas partes; c) qualidade e exactidão dos dispositivos de controle e dos mecanismos motores.

Cada máquina é submetida a testes de aceitação para avaliar sua qualidade construtiva. O conjunto destes testes é denominado de Verificação de Máquinas-Ferramenta e é dividido em verificações geométricas e provas práticas. As verificações geométricas avaliam o alinhamento das várias partes de uma máquina-ferramenta e são realizadas sob condições estáticas. Nas provas práticas os testes de alinhamento são realizados com a máquina submetida a carregamentos dinâmicos e excitações na freqüência de trabalho. Consideran-se a influência de vibrações e deflexões das várias partes.

Vários parâmetros são definidos para avaliar cada máquina. Mesmo para um único parâmetro, a medida pode ser efetuada em diversos pontos, fornecendo assim uma série de critérios de avaliação que, no presente estudo, serão agregados considerando-se o modelo de Análise Envoltória de Dados de *output* constante. No caso de estudo aqui apresentado será considerada a avaliação de tornos mecânicos, destacando-se em especial a verificação geométrica referente ao paralelismo entre as guias da contra ponta e as guias do carro.

2.2 Verificações geométricas

Entende-se por verificação geométrica a confirmação das dimensões, formas e posições dos componentes da máquina, bem como dos seus movimentos relativos no campo de trabalho. Inclui todas as operações de medida que afectam esses componentes e fazem referência apenas aos elementos que possam prejudicar a qualidade de trabalho da máquina. São verificações ou testes de precisão e não têm relação com os testes ou com as verificações de funcionamento, como vibrações, ruídos anormais, movimento de prende solta de componentes, verificações de torque residual, etc., nem com as verificações das características das máquinas, tais como velocidades e avanços.

A apreciação das definições de dimensões, forma, posições e movimentos relativos conduz a uma necessária distinção entre definições geométricas e definições metrológicas. "As definições geométricas são abstractas e relatam somente linhas e planos imaginários" (Jain, 1981). Na prática, as definições geométricas geralmente não são aplicáveis, pois não consideram os meios de execução nem as possibilidades de verificação.

As verificações geométricas geralmente feitas em máquina-ferramenta são (Mateus, 1974): nivelamento de barramento e mesa; movimento rectilíneo ou perpendicular de linha recta e plano; paralelismo, equidistância e coincidência; retilineidade; rotações.

2.3 Provas práticas

As provas práticas constituem, na sua essência, em pôr a máquina em funcionamento e verificar se trabalha sob as condições especificadas. Dessa forma, as verdadeiras provas práticas devem ser estabelecidas pelo utilizador da máquina. Apesar disso, e como muitas vezes as máquinas não são adquiridas para executar um único tipo de trabalho, é necessário estabelecer uma série de trabalhos-tipo, com tolerâncias tais que, executados por um operário neutro, possam definir a qualidade da máquina.

As provas práticas incluem a usinagem de peças de teste, apropriadas para os propósitos fundamentais a que a máquina tenha de ser designada, tendo já pré-determinado limites e tolerâncias (Jain, 1981). Para determinar a exactidão da máquina-ferramenta, devem ser executados ensaios práticos com operações de acabamento condizentes com o projeto da máquina.

Caso seja dispendioso ou difícil conduzir ambos os tipos de testes, a máquina pode ser avaliada somente por verificações geométricas ou somente por provas práticas (Branco e Leta, 1996). Quando ambos os métodos não fornecerem as mesmas indicações, "os resultados obtidos por testes práticos devem ser aceitos como válidos" (Jain, 1981).

As provas práticas representam com maior realidade as condições de operação da máquina e por isso os seus resultados são de maior valor do que as verificações geométricas. Entretanto, um ensaio não exclui o outro, sendo ambos de grande importância para avaliar a qualidade de fabricação da máquina-ferramenta.

2.4 Paralelismo entre as guias da contra ponta e as guias do carro

O procedimento de verificação deste paralelismo verifica se, ao usinar cilindros de comprimentos diferentes, estes terão o mesmo desvio de cilindricidade. Para máquinas em uso permite descrever o desgaste do barramento em relação às guias da contra ponta. Pode-se esperar que o desgaste no barramento concentre-se na região próxima ao cabeçote, enquanto o desgaste das guias da contra ponta terá maior intensidade na região próxima ao outro extremo do barramento.

No recebimento de uma máquina nova, os seguintes erros são aceitáveis: a) no plano vertical, 40 mm; b) no plano horizontal, 30 mm.

A metodologia para obtenção das medidas consiste na execução das seguintes etapas: fixa-se o relógio comparador no suporte e este no adaptador para o porta ferramentas; o sensor do relógio é posicionado sobre a guia da contra ponta, tomando-se cuidado para que fique na região de deslizamento da base da contra ponta, que é a região de atrito e desgaste; desloca-se o carro e a cada 100 mm realiza-se a leitura das indicações do relógio. Na Figura 1 tem-se um desenho esquemático da montagem dos instrumentos para verificação (em perspectiva e em vista lateral).

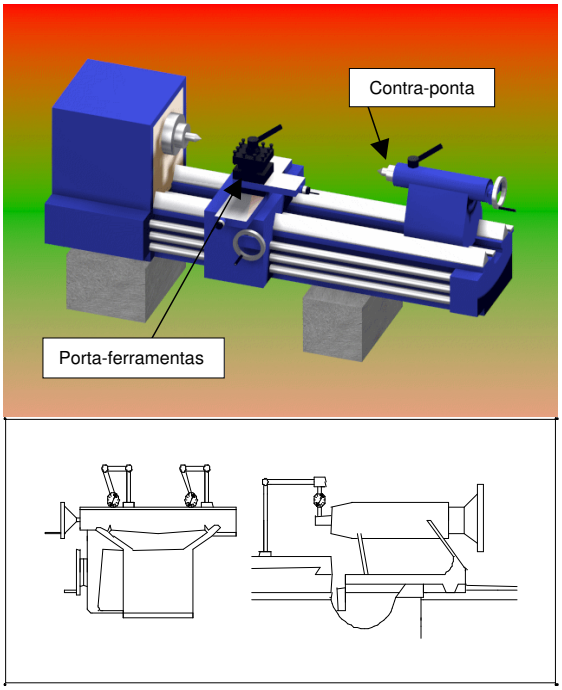


Figura 1: Verificação do paralelismo entre as guias da contra ponta e as guias do carro.

3 Análise Envolvatória de DadosN

A Análise Envoltória de Dados (DEA) tem como objectivo medir a eficiência de unidades tomadoras de decisão, designadas por DMUs (*Decision Making Units*), na presença de múltiplos *inputs* (entradas, recursos ou factores de produção) e múltiplos *outputs* (saídas ou produtos).

Neste artigo, o modelo DEA a ser usado é o CCR (também conhecido por CRS ou *constant returns to scale*), que trabalha com retornos constantes de escala (Charnes et al., 1978). Em sua formulação matemática considera-se que cada DMU k, $k = 1, ..., n$, é uma unidade de produção que utiliza m *inputs* x_{ik}, $i = 1, \ldots, m$, para produzir s *outputs* y_{jk}, $j = 1, \ldots, s$. Esse modelo maximiza o quociente entre a combinação linear dos *outputs* e a combinação linear dos *inputs*, com a restrição de que para qualquer DMU esse quociente não pode ser maior que 1.

Esse modelo de programação fracionária pode ser linearizado, transformando-se em um Problema de Programação Linear (PPL) apresentado em (2), onde h_o é a eficiência da DMU $o$em análise; x_{io} e y_{jo} são os *inputs* e *outputs* da DMU o; v_i e $u_j$são os pesos calculados pelo modelo para *inputs* e *outputs*, respectivamente.

$$\begin{array}{l} \max h_o = \sum_{j=1}^{s} u_j y_{jo} \\ \text{sujeito a} \\ \sum_{i=1}^{m} v_i x_{io} = 1 \\ \sum_{j=1}^{s} u_j y_{jk} - \sum_{i=1}^{m} v_i x_{ik} \leq 0 \ , \ k = 1, ..., n \\ v_i, u_j \geq 0 \ \forall i, j \end{array} \tag{1}$$

Apesar de os modelos DEA apresentarem a vantagem de permitir fazer ordenações sem depender de opiniões de decisores, são extremamente benevolentes com as unidades avaliadas. Estas podem ser eficientes desconsiderando vários dos critérios de avaliação. Assim, é comum haver um grande número de unidades com eficiência 100%. Tradicionalmente, considera-se que o número de unidades deve ser, pelo menos, o dobro ou o triplo do número de variáveis para que seja obtida uma boa classificação. No presente artigo, o número de variáveis é maior que o número de unidades. Tal situação não inviabiliza o uso de DEA, mas recomenda a utilização de técnicas adicionais para aumentar a discriminação.

4 Modelamento

Para a avaliação de quatro tornos do Laboratório de Usinagem da Universidade Federal Fluminense, considerando as medidas efectuadas de paralelismo entre guias, foi usado o modelo DEA CCR. Ao usar DEA é permitido que cada torno busque sua eficiência ao especializar-se na sua melhor medida, ou seja, escolher sua melhor faixa de operação.

Os desvios de paralelismo são interpretados como um preço a pagar para o torno funcionar, sendo, portanto, considerados *inputs*. Como trata-se de uma avaliação estática, não há nenhum *output*, já que nada foi produzido. Essa é uma situação que conduz a paradoxos que podem ser contornados assumindo-se que o *output* é constante e unitário para todos os tornos (Lovell e Pastor, 1999). O *output* representaria, assim, a própria existência do torno, numa abordagem semelhante à usada por Soares de Mello e Gomes (2004). Esses modelos são equivalentes a modelos multicritério aditivo, com a particularidade de que as próprias alternativas atribuem pesos a cada critério, ignorando qualquer opinião de um eventual decisor. Ou seja, DEA é usado apenas como ferramenta multicritério, e não como uma medida de eficiência clássica. Estudos sobre este tipo de uso para DEA podem ser encontrados em Farinaccio e Ostanello (1999), Belton e Stewart (1997), Soares de Mello et al. (2002b, 2004b) e Senra (2004).

A Tabela 1 resume os dados utilizados. No modelamento do problema foram ignoradas as medidas entre 0 e 100 mm, já que apresentam erros pequenos e semelhantes, não contribuindo para distinguir as máquinas entre si.

Tabela 1. *Inputs* e *Output* para o modelo DEA CCR.

DMU	***Output***	***Input***			
	Existência do torno	**d200**	**d300**	**d400**	**d500**
T1	1	35	58	52	35
T2	1	2	5	40	50
T3	1	10	30	48	49
T4	1	25	40	35	22

A Tabela 2 apresenta os resultados do modelo DEA CCR clássico. Verifica-se um empate para duas DMU, T2 e T4. Para que haja uma distinção entre essas duas DMUs eficientes, devem-se empregar modelos adicionais.

Tabela 2. Resultados do modelo DEA CCR clássico.

DMU	**Eficiência (%)**
T1	67,8
T2	100,0
T3	85,7
T4	100,0

5 Desempate entre Unidades Eficientes

5.1 Formulação teórica

Angulo Meza e Estellita Lins (2002) apresentam uma revisão dos modelos para aumento de discriminação em DEA. Os autores dividem esses modelos em dois grandes grupos: modelos que incorporam a informação *a priori* do decisor, e modelos que não usam aquela informação para seus cálculos. Dentro do primeiro grupo temos os modelos de restrições aos pesos e um modelo DEA surgido de uma fusão com o apoio multicritério à decisão, chamado de Análise de Eficiência do Valor. Já o segundo grupo compõe-se de três modelos: supereficiência, avaliação cruzada e um modelo multiobjectivo.O método da supereficiência, não estudado pelos autores referidos, enquadra-se no segundo grupo.

Neste artigo optou-se por usar quatro modelos de aumento de discriminação em DEA, cujos resultados são comparados. Os modelos usados, supereficiência, restrições aos pesos, avaliação cruzada e fronteira invertida, são descritos sucintamente a seguir.

5.1.1 Supereficiência

A ideia básica desse modelo é comparar a unidade que é avaliada com uma combinação linear de todas as outras unidades da amostra, sendo que a unidade de referência é excluída. Assim, dado que a unidade que em avaliação é retirada do conjunto das unidades de comparação, a eficiência obtida pode ser maior que 100%. Como este modelo permite que as DMUs obtenham eficiências superiores a 100%, ele consegue desempatar as unidades eficientes.

O método tem como vantagem o facto de fornecer uma discriminação entre as unidades eficientes, sem alterar a ordenação das ineficientes. A desvantagem é que a ordenação obtida depende apenas de condições locais da fronteira, não de propriedades gerais das DMUs ou da fronteira. A Figura 2 mostra a avaliação da unidade C utilizando o modelo da supereficiência.

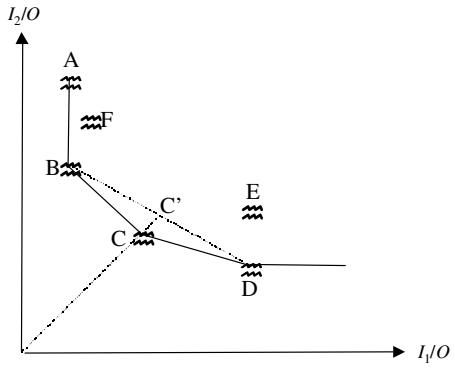


Figura 2: Exemplo da supereficiência.

Em (2) apresenta-se o modelo (formulação do envelope) para o cálculo da supereficiência (Charnes et al., 1994). Nesse modelo θ é a eficiência; s^+ e s^- são as folgas; e é um vector unitário; X_k e Y_k representam, respectivamente, o conjunto dos *inputs* e dos *outputs*; λ_k representa a contribuição da DMU k na formação do alvo da DMU em avaliação.

$$
\begin{aligned}
&\text{Min } \theta - e's^- - e's^+ \\
&\text{sujeito a} \\
&\theta X_j = \sum_{\substack{k=1 \\ k \neq j}}^{n} \lambda_k X_k + s^- \\
&Y_j = \sum_{\substack{k=1 \\ k \neq j}}^{n} \lambda_k Y_k - s^+ \\
&\lambda_k, s^+, s^- \geq 0
\end{aligned}
\tag{2}
$$

5.1.2 Restrições aos Pesos

Quando há julgamentos de valor sobre a importância relativa entre os *inputs* e/ou *outputs*, estes podem ser incorporados aos modelos DEA através de restrições aos pesos associados aos *inputs* e/ou aos *outputs* das unidades avaliadas. Allen et al. (1997) apresentam uma completa revisão da evolução da incorporação de julgamentos de valor através de restrições aos pesos.

A incorporação de julgamentos de valor através de restrições aos pesos pode ser dividida em três grupos de métodos (Estellita Lins e Angulo Meza, 2000): restrições directas sobre os multiplicadores; ajuste dos níveis de *input-output* observados para a captura de julgamentos de valor; restrição a *inputs* e *outputs* virtuais.

Neste artigo foram usadas restrições directas aos pesos. Nesse enfoque, desenvolvido por Dyson e Thanassoulis (1988) e generalizado por Roll et al. (1991), são impostos limites numéricos aos multiplicadores com o objectivo de não superestimar ou ignorar *inputs* e *outputs* na análise.

Seja $I_o = \sum_i v_i x_{io}$ o numerador da função objectivo na formulação original, no qual I_o é o *input* virtual consumido pela DMU o. Os limites impostos aos multiplicadores de *inputs*, v_i, e de *outputs*, u_j, são dados pelas relações apresentadas em (3), onde II, SI, IO, SO são os limites inferior e superior para *inputs* e *outputs*, respectivamente.

$$\begin{array}{l} II_i \leq v_i \leq SI_i \\ IO_i \leq u_i \leq SO_i \end{array} \tag{3}$$

Esse tipo de restrição pode levar à inviabilidade do PPL, já que estabelecer um limite superior ao peso de um *input* implica em um limite inferior no *input* virtual do restante das variáveis. Estellita Lins e Silva (2002) discutem em que condições as restrições aos pesos não tornam o PPL inviável. Neste artigo buscam-se, por tentativas sucessivas, os valores máximos que não inviabilizam os PPLs.

5.1.3 Avaliação Cruzada

Esta abordagem, introduzida por Sexton (1986) e ampliada por Doyle e Green (1994), faz uma avaliação de conjunto, isto é, na avaliação cruzada as DMUs têm uma avaliação própria (DEA clássico) e também são avaliados pelas outras DMUs utilizando os pesos ótimos dados pelo modelo. Dessa forma, cada DMU determina os pesos para o cálculo de seu índice de eficiência e utiliza esses pesos para determinar os índices de eficiência das outras DMUs.

Pode-se dizer que enquanto em DEA clássico cada DMU é avaliada segundo seu próprio ponto de vista, na avaliação cruzada ela também é avaliada segundo os pontos de vista das outras DMUs. Chama-se eficiência cruzada à média dos pontos de vista de todas as DMUs.

Por outro lado, os PPLs que determinam a eficiência de cada DMU podem ter múltiplas soluções ótimas para determinar o índice de eficiência, ou seja, os pesos (ou multiplicadores) podem não ser únicos. Para escolher entre os vários possíveis valores para os pesos ótimos de cada DMU, arbitra-se que estes quando aplicados às outras DMUs devem minimizar a sua eficiência (formulação agressiva) ou, ao contrário, maximizá-la (formulação benevolente). Doyle e Green (1994) estabeleceram o PPL (4) para o cálculo dos pesos na formulação agressiva, onde a eficiência da DMU s usando os pesos da DMU k é dada em (5).

$$
\begin{aligned}
&\text{Min} \sum_i \left(u_{ki} \sum_{s \neq k} y_{si} \right) - \sum_j \left(v_{kj} \sum_{s \neq k} x_{sj} \right) \\
&\text{sujeito a} \\
&\sum_j v_{kj} x_{kj} = 1 \\
&u_{ki} y_{ki} - E_{kk} \sum_j v_{kj} x_{kj} = 0 \\
&\sum_i u_{ki} y_{si} - \sum_j v_{kj} x_{sj} \leq 0 \quad , \forall s \neq k \\
&u_{ki}, v_{kj} \geq 0
\end{aligned}
\tag{4}
$$

$$
E_{ks} = \frac{\sum_i u_{ki} y_{si}}{\sum_j v_{kj} x_{sj}} \tag{5}
$$

5.1.4 Fronteira Invertida

O quarto modelo proposto para aumento da discriminação entre as unidades em avaliação é a fronteira invertida (Yamada et al., 1994; Entani et al., 2002). O uso de fronteira invertida para contornar o problema da baixa discriminação foi proposto inicialmente por Angulo Meza et al. (2003).

A fronteira invertida é uma avaliação pessimista das DMUs. Para tanto é feita uma troca dos *inputs* com os *outputs* do modelo original. Esta fronteira invertida é composta pelas DMUs com as piores práticas gerenciais (e pode ser chamada de fronteira ineficiente). Pode-se igualmente afirmar que as DMUs pertencentes à fronteira invertida têm as melhores práticas sob uma óptica oposta.

Para ordenar as DMUs é, então, calculado um índice de eficiência agregado, que é a média aritmética entre a eficiência em relação à fronteira original e a ineficiência (1 menos eficiência) em relação à fronteira invertida. Este índice pode ser apresentado de forma normalizado, para o que basta dividir todos os valores pelo maior índice calculado.

5.2 Adaptações e resultados

O primeiro modelo adicional usado para desempatar as DMUs T2 e T4 foi o modelo de superreficiência. Na Tabela 3, que apresenta os resultados, verifica-se que o torno 2 é melhor que o torno 4.

São, em seguida, apresentados os outros três modelos: restrições aos pesos, avaliação cruzada e fronteira invertida. Nos dois primeiros casos a avaliação foi conduzida de maneira diferente da que é usual na literatura, com alterações metodológicas propostas pelos autores para este caso particular.

Tabela 3. Resultados do modelo de supereficiência DEA CCR.

DMU	Eficiência (%)
T1	67,8
T2	600,0
T3	85,7
T4	159,0

No caso do modelo de restrições aos pesos foram usadas restrições directas aos pesos dos *inputs* em vez de usar regiões de segurança. Considerou-se como a melhor DMU aquela que permitiu restrições mais rígidas aos pesos, sem inviabilizar o PPL. Com este procedimento não foi necessário que o decisor arbitrasse relações entre pesos, o que permitiu que a avaliação continuasse independente de qualquer opinião subjectiva.

Para este caso, o torno T2 deixou de ser eficiente quando foi imposto que o valor mínimo do peso de cada *input* fosse, pelo menos, 0,01. Já para o torno T4, este parâmetro foi 0,005 e, dessa forma, o torno T2 é o melhor, já que permitiu restrições mais rígidas aos pesos, sem perder a eficiência. Registre-se que, neste caso, perder a eficiência significa tornar inviável o PPL que a calcula.

Tal como foi descrito, no modelo de avaliação cruzada todas as DMUs avaliam-se mutuamente obtendo-se uma média final das avaliações. Isso torna a classificação extremamente sensível à inclusão ou exclusão de qualquer alternativa, mesmo que ineficiente. Para reduzir essa desvantagem, foi usada a abordagem proposta por Soares de Mello et al. (2002a) de só usar as DMUs eficientes como unidades avaliadoras. Ao fazer com que a DMU T2 avalie a DMU T4, a eficiência de T4 foi igual a 8%. No caso contrário, ao T4 avaliar T2, a eficiência de T2 foi de 44%. Como há apenas duas DMUs avaliadoras, não é necessário calcular eficiências médias para ratificar-se a conclusão anterior de superioridade do torno T2.

Finalmente foi usado o método da fronteira invertida. Para os cálculos deste método foi usado o programa SIAD (Angulo Meza et al., 2004), que fornece o índice agregado, já normalizado. A Tabela 4 resume os resultados obtidos:

Tabela 4. Resultados obtidos com o uso da fronteira invertida.

DMU	Eficiência invertida (%)	Índice agregado normalizado (%)
T1	100,0	52,8
T2	100,0	77,8
T3	100,0	66,7
T4	71,4	100,0

Nota-se que este método inverteu a ordenação relativa de T2 e T4, quando comparado com os métodos anteriores. Este resultado é explicável pelo facto de que os outros métodos usaram informação relativa a todos os *outputs*. Já o índice agregado usa apenas as informações mais otimistas (eficiência DEA clássica) e pessimistas (fronteira invertida). No caso presente, observa-se que os tornos T1, T2 e T3 apresentam eficiência de 100% em relação à fronteira invertida, o que significa que cada um deles é o que apresenta pior resultado em relação a, pelo

menos, uma medida. Como o torno T4 nunca é o pior e, tal como T2, é o melhor em relação a alguma medida, teve o melhor resultado neste método.

5.3 Comparação de resultados

Os resultados obtidos por três do quatro métodos usados foram idênticos. Isto decorre do facto de que esses métodos (supereficiência, avaliação cruzada e restrições aos pesos) mantêm o caracter optimista das avaliações feitas por DEA. Decisores que queiram que a melhor peça fabricada por um dos tornos seja a melhor possível, devem usar esses métodos.

Já o método da fronteira invertida considera, com igual importância, uma avaliação optimista e outra pessimista. Ou seja, além de se exigir que as melhores peças sejam muito boas (embora não sejam as melhores possíveis), exige-se ainda que as piores não sejam de qualidade muito baixa

As alterações introduzidas no método da avaliação cruzada resultaram em uma simplificação dos cálculos. O uso apenas das DMUs eficientes pode ser estendido a outros casos, desde que se lida rigorosamente com o problema da multiplicidade de soluções óptimas do modelo dos multiplicadores. Em Soares de Mello et al. (2004a) é mostrado um caso em que a técnica aqui proposta não pode ser usada. A outra alteração proposta, de não fazer médias, só foi possível porque apenas duas DMUs eram eficientes.

Já as alterações ao método de restrição aos pesos são válidas em qualquer problema em que não se deseja levar em conta opiniões de decisores ou de especialistas. No entanto, a sua aplicação a casos em que haja um grande número de DMUs eficientes, pode tornar-se extremamente morosa, ou exigir o uso de técnicas de análise de sensibilidade.

O método da supereficiência tem a desvantagem de ser baseado em propriedades locais da fronteira e não tem interpretação prática. Embora sempre possa ser usado, os resultados não são muito fiáveis.

Por outro lado, o método da fronteira invertida pode ter aplicação quase universal. A excepção é quando se deseja permitir que as unidades em avaliação especializem-se em algumas tarefas. O método tem encontrado várias aplicações práticas (Angulo Meza et al., 2003; Pimenta et al., 2004; Gomes et al., 2004) e teóricas (Soares de Mello et al., 2005).

6 Conclusões

A qualidade de uma máquina-ferramenta depende de vários factores. Esses factores podem ser avaliados conjuntamente utilizando-se a abordagem DEA, que realiza uma avaliação sem informação adicional dos decisores com uso de Problemas de Programação Linear.

No caso em estudo, devido ao grande número de variáveis em relação às unidades avaliadas, duas DMUs foram consideradas eficientes, sem que o modelo básico distinguisse entre as duas. Para desempata-las foram utilizadas abordagens adicionais, que conduziram, em sua maioria, ao resultado de superioridade do torno T2 em relação ao T4.

A questão da escolha entre T2 e T4 fica condicionada ao objectivo do decisor. Caso ele

deseje garantir que no pior dos casos a máquina não terá o funcionamento muito deficiente (atitude pessimista) deve preferir o T4. Qualquer outra posição do decisor conduz à escolha de T2.

Os resultados aqui apresentados constituem um estudo preliminar, já que outros parâmetros devem ser considerados numa avaliação completa de máquinas-ferramenta. Em especial, a avaliação de provas práticas deve ser objecto de estudo posterior, onde DEA será usado como uma verdadeira medida de eficiência, já que em vez de *output* unitário poderá ser levado em conta, por exemplo, a velocidade de usinagem.

7 Agradecimentos

Ao CNPq pelo apoio financeiro, processo 301095/2003-5.

8 Referências

Agostinho, O.L., Rodrigues, A.C.S., Lirani, J. Tolerâncias, Ajustes, Desvios e Análise de Dimensões. Edgard Blücher, São Paulo, 1977.

Allen, R., Athanassopoulos, A., Dyson, R.G., Thanassoulis, E. Weights restrictions and value judgements in data envelopment analysis: evolution, development and future directions. Annals of Operations Research, v. 73, p. 13-34, 1997.

Andersen P., Petersen N.C. A Procedure for Ranking Efficient Units in Data Envelopment Analysis. Management Science, v. 39, n. 10, p. 1261-1264, 1993.

Angulo Meza, L., Estellita Lins, M.P. Review of Methods for Increasing Discrimination in Data Envelopment Analysis. Annals of Operations Research, v. 116, p. 225-242, 2002.

Angulo Meza, L., Biondi Neto, L., Soares de Mello, J. C. C. B., Gomes, E. G., Coelho, P. H. G. FSDA - Free Software for Decision Analysis (SLAD - Software Livre de Apoio à Decisão): A software package for Data Envelopment Analysis models. Memórias del XII Congreso Latino-Iberoamericano de Investigación de Operaciones y Sistemas - CLAIO, 2004.

Angulo Meza, L., Gomes, E.G., Soares de Mello, J.C.C.B., Biondi Neto, L. Fronteira DEA de dupla envoltória no estudo da evolução da ponte aérea Rio-São Paulo. Panorama Nacional da Pesquisa em Transportes 2003, v. 2, p.1158-1166, 2003.

Belton, V., Stewart, T. J. DEA and MCDA: competing or complementary approaches?. Working Paper, University of Strathclyde, Scotland, 1997.

Branco, D.F., Leta, F.R. Avaliação de Tornos Mecânicos em Uso. Anais do 2° Congresso Internacional de Engenharia Industrial e XVI ENEGEP, Piracicaba, SP,1996.

Charnes, A., Cooper, W.W., Rhodes, E. Measuring the efficiency of decision-making units. European Journal of Operational Research, v. 2, p. 429-444, 1978.

Charnes, A., Cooper, W.W.,Lewin, A.Y., Seiford, L.M. Data Emvelopment Analysis: Theory, methodology and applications. Kluwer Academic Publishers, USA, 1994.

Doyle, J., Green, R. Efficiency and cross-efficiency in DEA derivations, meanings and uses. Journal of the Operational Research Society, v. 45, p. 567-578, 1994.

Dyson, R.G., Thanassoulis, E. Reducing Weight Flexibility in Data Envelopment Analysis. Journal of the Operational Research Society, v. 39, n. 6, p. 563-576, 1988.

Entani, T., Maeda, Y., Tanaka, H. Dual Models of Interval DEA and its extensions to interval data. European Journal of Operational Research, v. 136, p. 32-45, 2002.

Estellita Lins, M.P., Angulo Meza, L. Análise Envoltória de Dados e perspectivas de integração no ambiente de Apoio à Decisão. Editora da COPPE/UFRJ, Rio de Janeiro, 2000.

Estellita Lins, M.P., Silva, A.C.M. Evitando a inviabilidade em modelos DEA com restrições aos pesos. Relatório Técnico EP03/01-PO, Programa de Engenharia de Produção/COPPE/UFRJ, Rio de Janeiro, 2001.

Farinaccio, F., Ostanello, A. Evaluation of DEA Validity as a MCDA/M Tool: Some Problems And Issues. Technical Report TR-99-06, Dipartimento de Informatica, Università di Pisa, Itália, 1999.

Gomes, E.G., Mangabeira, J.A.C., Soares de Mello, J.C.C.B., Angulo Meza, L The use of Data Envelopment Analysis in agriculture in the presence of interval data. Memorias del XII Congreso Latino-Iberoamericano de Investigación Operativa – CLAIO, 2004.

Jain, R. K. Egineering Metrology. Khanna Publisher, New Delhi, 1981.

Lovell, C.A.K., Pastor, J.T. Radial DEA models without inputs or without outputs. European Journal of Operational Research, v. 118, n. 1, p. 46-51, 1999.

Mateus, A G. Tolerâncias e Ajustes. Editora Polígono,1974.

Pimenta, H.L.N, Macedo, M.A., Soares de Mello, J.C.C.B. Decisão de realização de investimentos em tecnologia da informação com análise envoltória de dados. Revista Produção On Line, v. 4, n. 2, p. 1-16, 2004.

Roll, Y., Cook, W.D., Golany, B. Controlling factor weights in DEA. IIE Transactions, v. 23, n. 1, p. 2-9, 1991.

Seiford, L.M., Zhu, J. Infeasibility of super efficiency data envelopment analysis models. Information Systems and Operational Research, v. 27, p. 174-187, 1999.

Senra, L.F.A.C. Métodos de seleção de variáveis em DEA: Estudo de Caso no Setor Elétrico. Dissertação de Mestrado, Programa de Engenharia de Produção - Universidade Federal Fluminense, Niterói, 2004.

Sexton, T.R. Measuring Efficiency: An assessment of Data Envelopment Analysis. New Directions For Program Ev*aluation*, n. 32, San Francisco, Jossey-Bass, 1986.

Soares de Mello, J.C.C.B., Gomes, E.G., Angulo Meza, L., Biondi Neto, L., Coelho, P.H.G. A modified DEA model for Olympic evaluation. Memorias del XII Congreso Latino-Iberoamericano de Investigación Operativa – CLAIO, 2004a.

Soares de Mello, J.C.C.B., Gomes, E.G., Angulo Meza, L., Estellita Lins, M.P. Selección de Variables Para el Incremento del Poder de Discriminacion de los Modelos DEA. Revista de La Escuela de Perfeccionamiento En Investigación Operativa, n. 24, p. 40-52, 2004b.

Soares de Mello, J.C.C.B., Gomes, E.G., Angulo Meza, L., Biondi Neto, L., Sant'Anna, A.P. Fronteiras DEA difusas. Investigação Operacional, v. 25, n. 1, p. 85-103, 2005.

Soares de Mello, J.C.C.B, Estellita Lins, M.P., Gomes, E.G. Construction of a smoothed DEA frontier. Pesquisa Operacional, v. 22, n. 2, p. 183-201, 2002a.

Soares de Mello, J.C.C.B., Gomes, E.G., Soares de Mello, M.H.C., Estellita Lins, M.P. Método Multicritério para Seleção de Variáveis em modelos DEA. Pesquisa Naval, v. 15, p.55-66, 2002b.

Soares de Mello, J.C.C.B., Gomes, E.G. Eficiências Aeroportuárias: Uma Abordagem Comparativa com Análise de Envoltória de Dados. Revista de Economia e Administração, v. 3, n. 1, p. 15-23, 2004.

Yamada, Y, Matui, T., Sugiyama, M. New analysis of efficiency based on DEA. Journal of the Operations Research Society of Japan, v. 37, n. 2, p. 158-167, 1994.

Limites para Eficiências não Radiais em DEA: uma Abordagem Vectorial

João Carlos Correia Baptista Soares de Mello * Lidia Angulo Meza †
Eliane Gonçalves Gomes ‡ Luiz Biondi Neto §

* Departamento de Engenharia de Produção - Universidade Federal Fluminense
Rua Passo da Pátria 156, 24210-240, São Domingos, Niterói, RJ, Brasil
jcsmello@producao.uff.br

† Departamento de Ciência dos Materiais - Universidade Federal Fluminense
Av. dos Trabalhadores 420, 27255-125, Volta Redonda, RJ, Brasil
lidia@metal.eeimvr.uff.br

‡ Empresa Brasileira de Pesquisa Agropecuária (Embrapa) - SGE
Parque Estação Biológica, W3 Norte final, 70770-901, Asa Norte,Brasília, DF
eliane.gomes@embrapa.br

§ Departamento de Eletrônica e Telecomunicações - Universidade do Estado do Rio de Janeiro
Rua São Francisco Xavier, 524, Bl. E, Sala 5025, 20550-900,Maracanã, Rio de Janeiro, RJ, Brasil
lbiondi@uerj.br

Abstract

This paper presents a new approach to evaluate non-radial DEA efficiencies. This approach uses vector formulation to non-radial efficiencies. We propose the construction of an efficiency vector that leads to an index, which is an upper bound for the non-radial efficiency module. We also perform the evaluation of a lower limit for the same index. The developments here proposed are for DEA BCC model, as for the CCR model the efficiencies are the same independently from the projection region. The proposed score is applied in a numeric example.

Resumo

Este artigo apresenta uma nova abordagem para o estudo de eficiências não radiais em DEA, qual seja, uma formulação vectorial para o seu cálculo. Propõe-se a construção de vectores de eficiência que fornecem um majorante para o módulo da eficiência não radial e um limite inferior para a mesma grandeza. Os desenvolvimentos referem-se ao modelo DEA BCC, já que no modelo CCR as eficiências são iguais independentemente da região de projecção. O índice proposto é aplicado a um exemplo numérico.

Apdio © 2005 *Associação Portuguesa de Investigação Operacional*

Keywords: DEA; Directional efficiency; Vectors projection.

Title: Upper and Lower Bounds for Non- Radial Efficiences in DEA: a Vector Approach

---

1 Introdução

As medidas de eficiência em DEA têm sido abordadas considerando dois enfoques clássicos. O primeiro considera índices de eficiência radiais, isto é, que consideram a redução ou aumento equiproporcional dos *inputs* ou dos *outputs* (Debreu, 1951; Farrel, 1957) e que podem ser obtidos pelos modelos clássicos de DEA, CCR (Charnes et al., 1978) e BCC (Banker et al., 1984). Nos modelos radiais tem-se optado pela redução dos *inputs* ou aumento dos *outputs*, não ocorrendo alterações simultâneas.

O segundo enfoque considera índices de eficiência não radiais, ou seja, variações não proporcionais dos *inputs* e *outputs* e permissão de alteração simultânea dos *inputs* e *outputs*. Alguns dos índices existentes que adoptam este enfoque são Ruggiero e Bretschneider (1998), que introduziram a medida ponderada de Russell; Pastor et al. (1999), que propõem medidas de eficiência globais – *Global Efficiency Measures*, GEMs; Briec (2000), que apresenta uma modificação e extensão das medidas de Russell; Tone (2002), que propõe uma medida baseada nas folgas utilizando o conceito de super eficiência. Ruggiero (2000) faz um estudo das múltiplas medidas de eficiência formuladas até a data do estudo e propõe uma nova medida para corrigir os problemas com as medidas existentes. Em Tone (2001) encontra-se uma revisão bibliográfica dos índices de eficiência não radiais até a data, baseados nas folgas. Nestes dois últimos artigos é enfatizado o tratamento escalar da eficiência.

Estas abordagens acarretam várias inconsistências aos índices obtidos. Entre elas citam-se os casos de eficiências negativas e maiores do que um (Gonzalez Araya, 2003) e a combinação matematicamente inadequada de índices. A solução para esse problema está no tratamento vectorial da eficiência seguindo uma analogia com a derivada direccional.

De facto, uma vez que a eficiência depende do ponto de projecção, ela é caracterizada por um número e por uma direcção de projecção. Grandezas caracterizadas desta forma são vectoriais. Assim, considerar a eficiência como escalar, como feito usualmente na literatura (Pastor, 1995; Tone, 2001), não tem sentido. Seria como postular que a velocidade de um corpo devesse ser escalar.

Este artigo propõe a construção de um vector de eficiências que projectado na direcção do vector fornece um majorante para o módulo da eficiência não radial. É ainda construído um segundo vector que fornece um limite inferior para a mesma grandeza e são verificadas as condições em que os dois limites são iguais. Destaca-se que os desenvolvimentos que se seguem referem-se ao modelo DEA BCC, já que no modelo DEA CCR as eficiências são iguais independentemente da região de projecção.

2 Abordagem Proposta

2.1 Formulação Bidimensional

2.1.1 Modelo baseado em analogia com derivada direccional

A Figura 1 ilustra os conceitos que serão a seguir definidos. O vector $\overrightarrow{ef}$, aqui chamado vector eficiência, tem como coordenadas aparentes o simétrico da eficiência orientada a *inputs* e a eficiência orientada a *outputs*. O simétrico da eficiência orientada a *inputs* é usado no lugar da própria eficiência, visto que a projecção é feita na direcção decrescente do eixo horizontal.

Quanto à eficiência orientada a *outputs*, existe um problema de nomenclatura. A teoria clássica de DEA considera como eficiência orientada a outputs o resultado de um PPL que minimiza um quociente não inferior à unidade. Por esta definição, uma DMU será tanto mais eficiente quanto menor o índice de eficiência. O uso deste número na formulação vectorial de eficiência levaria a um vector com uma coordenada no intervalo [0,1} e a outra no intervalo [1, ∞[. Na hora de compor um índice único com estas coordenadas haveria a predominância de uma coordenada sobre outra, pelo simples facto de serem definidas de forma diferente.

Há ainda que considerar que não é intuitivo, principalmente para leigos, falar em eficiências superiores à unidade, e que indicam melhores DMUs quanto menor for o índice. Uma solução clássica para este problema de interpretação é trabalhar com o inverso da eficiência orientada a outputs, que passa a ser um número no intervalo [0,1], tal como a orientada a inputs. Pode-se ainda dizer que o índice obtido do PPL é, na verdade, uma ineficiência e a verdadeira eficiência é o inverso desse índice.Usando esta interpretação, doravante o inverso da eficiência orientada a *outputs* será, denominada apenas por eficiência orientada a *outputs*. Matematicamente, o vector eficiência $\overrightarrow{ef}$ é dado por (1), onde e_i e e_o são as eficiências nas orientações a *inputs* e *outputs*.

$$\overrightarrow{ef} = -e_i\vec{i} + e_o\vec{j} \tag{1}$$

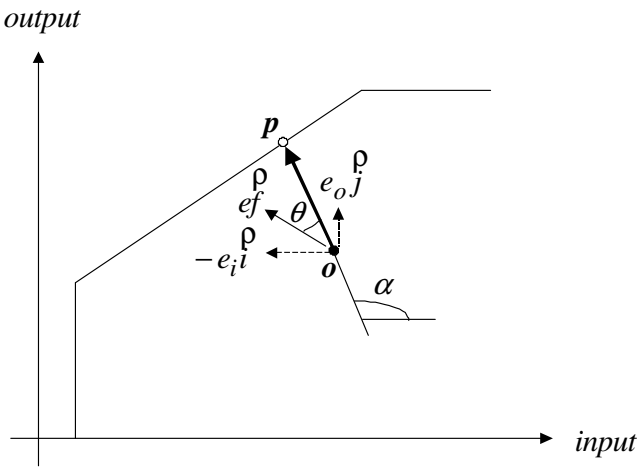


Figura 1: Representação bidimensional da fronteira DEA BCC.

Dado um ponto p na fronteira, define-se o vector $\vec{v}$ como aquele que tem origem na DMU

o e extremidade no ponto p. O vector $\vec{v}$ determina a direcção da projecção da eficiência, sendo que esta direcção faz um ângulo θ com o vector eficiência $\overrightarrow{ef}$.

Da definição de produto interno tem-se que o módulo da projecção do vector eficiência no vector $\vec{v}$, ou o módulo da eficiência no vector $\vec{v}$ (doravante chamado eficiência direccional) é dado por (2), onde α é o ângulo que o vector $\vec{v}$ faz com o semi-eixo positivo dos *inputs* (havendo decréscimo de *inputs*, como ocorre em projecções dominantes, α é um ângulo maior que 90°). Note-se que em (2), $\cos\alpha$ e $sen\alpha$ são as coordenadas do vector unitário $\frac{\vec{v}}{\|\vec{v}\|}$, ou seja, são os cosenos directores do vector $\vec{v}$.

$$ef_{\vec{v}} = \overrightarrow{ef}\frac{\vec{v}}{\|\vec{v}\|} = -e_i\cos\alpha + e_o sen\alpha \tag{2}$$

Tem-se, por outro lado, que o produto interno também pode ser calculado em função dos módulos dos dois vectores e do ângulo formado por eles, como apresentado em (3), ou ainda em (4). De (4) deduz-se que a eficiência direccional é máxima quando $\theta = 0°$.

$$\overrightarrow{ef}\frac{\vec{v}}{\|\vec{v}\|} = \overrightarrow{ef}\frac{\vec{v}}{\|\vec{v}\|}\cos\theta \tag{3}$$

$$\overrightarrow{ef}\frac{\vec{v}}{\|\vec{v}\|} = \overrightarrow{ef}\cos\theta \tag{4}$$

Segundo o modelo até agora adoptado, conclui-se que a eficiência direccional máxima é obtida na direcção do vector eficiência. Tendo em vista essa propriedade, é fácil observar que a formulação proposta apresenta inconsistências. Com efeito, no caso de $e_o = e_i = e$, tem-se $ef_{\vec{v}\ \max} = e\sqrt{2} > 1$, o que é uma clara contradição. Repare-se, ainda, que uma aplicação deste índice ao modelo CCR conduziria a valores da eficiência direccional $ef_{\vec{v}}$ variáveis com a direcção, o que contradiz a propriedade deste modelo de ter a eficiência técnica invariante com a orientação.

2.1.2 Modelo com factor de correcção

As inconsistências mencionadas no item anterior são particularmente evidentes no caso da DMU o já estar na fronteira. Neste caso, o vector $\vec{v}$ não pode ser determinado pela diferença de pontos (uma vez que as duas são coincidentes), mas pelos seus cosenos directores. Para tal DMU, $\overrightarrow{ef} = (-1, 1)$ e, portanto, a eficiência direccional é dada pela expressão (5).

$$ef_{\vec{v}} = -\cos\ \alpha + sen\ \alpha\ \neq\ 1 \tag{5}$$

Para uma DMU na fronteira, a eficiência direccional deveria ser unitária qualquer que seja a direcção. Para obter este resultado basta dividir o valor obtido por $-\cos\ \alpha + sen\ \alpha$. A proposta para a eficiência direccional de qualquer outra DMU é corrigir o índice (4), dividindo-o pela eficiência direccional do seu alvo, ou seja, pelo resultado da expressão (5). Obtém-se, assim, um número entre 0 e 1. De acordo com esta proposta, a eficiência direccional da DMU o na direcção do vector $\vec{v}$ é dada pela equação (6).

$$ef_{\vec{v}} = \frac{-e_i \cos\alpha + e_o sen\alpha}{-\cos\alpha + sen\alpha} \tag{6}$$

Evidentemente, a expressão (6) somente é válida para o caso em que o denominador não seja zero. Como na formulação em análise só se consideram projecções dominantes, $\cos\alpha$e $sen\alpha$nunca assumiram o mesmo valor e, como consequência, o denominador será sempre diferente de zero.

Para investigar as propriedades do índice proposto é conveniente calcular a sua derivada em relação ao angulo α, conforme apresentado em (7).

$$\frac{def_{\vec{v}}}{d\alpha} = \frac{-e_o + e_i}{(sen\alpha - \cos\alpha)^2} \tag{7}$$

Três situações podem ocorrer:

- $e_i > e_o$: Neste caso, a derivada é positiva e a eficiência direccional é estritamente crescente com o ângulo α. Consequentemente, assume o maior valor na orientação a *input* e o menor na orientação a *output*.
- $e_i < e_o$: Analogamente, a derivada é sempre negativa e, portanto, a orientação a *output* fornece o valor máximo e a orientação a *input* o valor mínimo.
- $e_i = e_o$: Nesta situação, a derivada é nula e, portanto, a eficiência direccional é constante.

Como caso particular deste último tem-se o modelo CCR. Comprova-se, assim, que o índice corrigido mantém coerência com as propriedades do modelo DEA CCR.

É importante destacar que todos os desenvolvimentos anteriores consideraram que as projecções orientadas a *input*, a *output* ou não radiais teriam os alvos sempre na mesma face. Como pode ser observado na Figura 2 nem sempre isso acontece. Neste caso, os índices propostos supõem que os alvos encontram-se no segmento $\overline{AC}$, quando na verdade estão na face $\overline{BC}$. Portanto, a eficiência direccional calculada é maior ou igual que a eficiência direccional real. Dessa forma, a equação (VI) transforma-se na inequação (8).

$$\overrightarrow{ef} \leq \frac{-e_i \cos\alpha + e_o sen\alpha}{-\cos\alpha + sen\alpha} \tag{8}$$

Para obter uma melhor avaliação da eficiência direccional é necessário, além do limite superior, estimar um limite inferior. Para tal, estende-se a face onde o alvo é projectado e calculam-se as eficiências orientadas a *input* e a *output* em relação a uma fronteira fictícia, composta apenas por essa face. Como a DMU em observação está mais afastada desta fronteira fictícia que da fronteira real, os cálculos efectuados fornecem um valor inferior à eficiência direccional real. Se as duas estimativas forem iguais, ou seja, todas as projecções forem na mesma face, tem-se o valor exacto da eficiência direccional. Caso contrário, são válidas as desigualdades apresentadas em (9), onde x_o e y_o são o *input* e o *output* da DMU o em análise; a e b são, respectivamente, os coeficientes angular e linear da face de projecção.

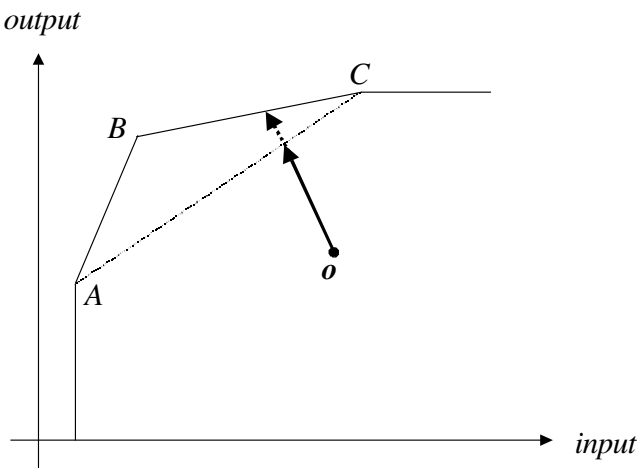


Figura 2: Projecção em diferentes faces.

$$\frac{(ax_o + b)(y_o - b)\cos\ \alpha\ + ax_o y_o sen\ \alpha}{ax_o(ax_o + b)(sen\ \alpha\ + \cos\ \alpha)} \leq \overrightarrow{ef} \leq \frac{-e_i \cos\alpha + e_o sen\alpha}{-\cos\alpha + sen\alpha} \tag{9}$$

Cabe observar que a determinação do limite inferior exige o cálculo da equação da face de projecção, o que pode acarretar problemas computacionais de ordem de dimensão superior (Dulá, 2002; Gonzalez Araya, 2003).

Vale ainda ressaltar que caso o ponto de projecção seja uma DMU extremo-eficiente não existirá uma única face de suporte. Neste caso, pode-se usar a recta tangente à fronteira suavizada (Soares de Mello et al., 2002).

2.2 Formulação Multidimensional

A formulação bidimensional do problema determina a direcção de projecção através de um único ângulo α. Para dimensões superiores vários ângulos devem ser considerados. Estes ângulos são determinados através dos cosenos directores do vector que indica a direcção de projecção, ou seja, pelas coordenadas do unitário de tal vector (o vector dividido pela sua norma), conforme apresentado em (10), onde $\cos\alpha_i$é a i-ésima coordenada do unitário do vector de projecção.

$$\vec{ef_v} = \frac{\vec{ef}\frac{\vec{v}}{\|\vec{v}\|}}{\sum_i cos\alpha_i} \tag{10}$$

Deve-se considerar que o vector eficiência tem número de coordenadas igual à dimensão do problema, isto é, a soma do número de *inputs* e *outputs*. Portanto, o simples cálculo das eficiências radiais orientadas a *input* e a *output* não é suficiente para caracterizar tal vector.

Observe-se que cada coordenada do vector está associado, *ceoteris paribus*, ao índice de contracção (expansão) do i-ésimo *input* (*output*) para atingir a fronteira. No caso de *input* este índice é o simétrico do factor de contracção, e no caso do *output* é o inverso do factor de

contracção. O factor de contracção é calculado segundo o PPL (11) e o de expansão segundo o PPL (12).

$$
\begin{array}{l}
\text{Min } h_o \\
\text{sujeito a} \\
h_o x_1 \geq \sum_{j=1}^{n} \lambda_j x_{j1} \\
x_2 \geq \sum_{j=1}^{n} \lambda_j x_{j2} \\
\ldots \\
y_1 \leq \sum_{j=1}^{n} \lambda_j y_{j1} \\
y_2 \leq \sum_{j=1}^{n} \lambda_j y_{\ j2} \\
\\
\ldots. \\
\sum_{j=1}^{n} \lambda_j = 1
\end{array}
\tag{11}
$$

$$
\begin{array}{l}
\text{Max} h_o \\
\text{sujeito a} \\
x_1 \geq \sum_{j=1}^{n} \lambda_j x_{j1} \\
x_2 \geq \sum_{j=1}^{n} \lambda_j x_{j2} \\
\ldots \\
h_o y_1 \leq \sum_{j=1}^{n} \lambda_j y_{j1} \\
y_2 \leq \sum_{j=1}^{n} \lambda_j y_{j2} \\
\ldots. \\
\sum_{j=1}^{n} \lambda_j = 1
\end{array}
\tag{12}
$$

A Tabela 1 apresenta um exemplo numérico para o caso de dois *inputs* e um *output*. A DMU D é BCC ineficiente e em qualquer direcção é projectada com dominância na face composta pelas DMUs A, B e C.

A Tabela 2 apresenta os resultados do cálculo dos factores de contracção e de expansão para a DMU D. Os resultados não radiais foram obtidos pelos PPLs já apresentados e os resultados radiais obtidos pelo programa SIAD (Angulo Meza et al., 2004).

Para a DMU D o vector eficiência é dado por $(-0,80, \ -0,80, \ 0,89)$ (Tabela 2). Observa-se que tanto as eficiências radiais quanto as componentes do vector eficiência são relativas a alvos da mesma face. Portanto, para esta DMU, os valores obtidos por qualquer eficiência direccional são valores exactos, já que os majorantes são iguais aos minorantes.

Embora os valores apresentados sejam meramente um exemplo numérico, sem conexão com problemas práticos, podem ser encontrados na literatura vários exemplo de modelos com

Tabela 1: Dados para o exemplo numérico multidimensional.

DMU	***Input* 1**	***Input* 2**	***Output***
A	1	1	1
B	11	1	11
C	1	11	11
D	5	5	8

Tabela 2: Resultados da DMU D para os dados da Tabela 1.

	$\boldsymbol{Ef}_{Input1}$	$\boldsymbol{Ef}_{Input2}$	$\boldsymbol{Ef}_{Output}$	$\boldsymbol{Ef}_{Inputradial}$
Eficiência	0,80	0,80	0,89	0,90
λ_A	0,30	0,30	0,20	0,30
λ_B	0,30	0,40	0,40	0,35
λ_C	0,40	0,30	0,40	0,35

2 inputs e 1 output. Por exemplo, em Soares de Mello et al (2003) apresenta-se um modelo de avaliação de companhias aéreas em que os inputs são pessoal e assentos.km oferecidos, e o output é assentos.km vendidos. No contexto desse exemplo, pode-se interpretar a Ef_{Input1} como o percentual de redução de pessoal, se apenas fosse possível mexer nessa variável, para a DMU tornar-se eficiente. Já Ef_{Input2} representaria o percentual de redução de assentos.km oferecidos, caso apenas essa variável fosse alterada, para a DMU ficar eficiente. Já Ef_{Output} representaria o quanto é necessário aumentar percentualmente a quantidade de passageiros.km vendidos para atingir a fronteira eficiente. E, por fim, $\mathrm{Ef}_{Inputradial}$ representa o quanto os dois *inputs* devem ser reduzidos (equiproporcionalmente) para a DMU atingir a fronteira, sem alterar o *output*.

Dentre as várias direcções possíveis, pode-se calcular eficiência na direcção do próprio vector eficiência. Para este caso, é possível obter uma forma geral do majorante, que é dada pela expressão (13).

$$ef_{\vec{ef}} \leq \frac{\overrightarrow{ef}.\overrightarrow{ef}}{\left\|\overrightarrow{ef}\right\| . \sum\limits_i \cos\alpha_i} = \frac{\left\|\overrightarrow{ef}\right\| . \left\|\overrightarrow{ef}\right\|}{\left\|\overrightarrow{ef}\right\| \sum\limits_i \cos\alpha_i} . \cos\theta = \frac{\left\|\overrightarrow{ef}\right\|}{\sum\limits_i \cos\alpha_i} = \frac{\overrightarrow{ef}.\overrightarrow{ef}}{\sum\limits_n ef_o + \sum\limits_r ef_i} \tag{13}$$

Para o exemplo das Tabelas 1 e 2, a desigualdade transforma-se em igualdade e, assim, tem-se que $ef_{\vec{ef}} = \frac{\overrightarrow{ef}.\overrightarrow{ef}}{\sum\limits_n ef_o + \sum\limits_r ef_i} = \frac{(0{,}8)^2+(0{,}8)^2+(0{,}89)^2}{0{,}8+0{,}8+0{,}89} \cong 0,83$.

No exemplo apresentado, destaca-se que devido a pequena dimensão o esforço computacional envolvido não foi alto, mas ao incrementar-se o número de variáveis a situação muda, já que não existe algoritmo eficiente para calcular a equação do todas as faces eficientes (Dulá, 2002). Para resolver este problema a formulação multiobjectivo de DEA de Lins et al (2004) pode ser usada para determinar a intersecção de qualquer recta, que contenha uma dada DMU, com a fronteira eficiente, como mostrado em Soares de Mello et al (2005).

3 Conclusões

Este artigo apresenta uma nova abordagem para o cálculo de eficiências não radiais. O índice proposto não apresenta inconvenientes caso o alvo esteja situado na região de dominância. De facto, o sinal negativo nas eficiências orientadas a *input* impede o surgimento das eficiências negativas.

Por outro lado, o factor de correcção impede que a eficiência seja maior que um. Este factor de correcção poderia acarretar outro problema, qual seja o da divisão por zero. No entanto, mais uma vez, observa-se que tal facto não acontece na região de dominância.

A abordagem apresentada apenas leva ao cálculo de um índice em casos particulares. Em casos mais gerais, fornece limites inferiores e superiores da eficiência direccional. Estes limites são baseados na geometria do problema. São, portanto, mais consistentes que os índices tradicionais baseados em somas (sem sentido para o cálculo da eficiência, já que esta é baseada em distâncias euclidianas) ou em ponderações, que sempre exigem arbitrariedade de um decisor (situação contrária aos paradigmas de DEA). Os limites aqui apresentados podem servir como teste de consistência para os numerosos índices não radiais já existentes.

4 Agradecimentos

Ao CNPq pelo apoio financeiro, processo 470423/2004-8.

5 Referências

Angulo-Meza, L., Biondi, L. N., Soares de Mello, J.C.C.B., Gomes, E.G., Coelho, P.H.G. (2004). FSDA – Free Software for Decision Analysis (SLAD –Software Livre de Apoio a Decisão): A Software Package for Data Envelopment Analysis Models. *Congreso Latino Iberoamericano de Investigación de Operaciones y Sistemas*, La Habana – Cuba.

Banker, R.D., Charnes, A., Cooper, W.W. (1984). Some Models For Estimating Technical Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30 (9), 1078-1092.

Briec, W. (2000). An extended Färe-Lovell technical efficiency measure. *International Journal of Production Economics*, v. 65, p. 191-195.

Charnes, A., Cooper, W.W., Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, v. 2, p. 429-454.

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica*, v. 19, p. 273-292.

Dulá, J.H. (2002). Computations in DEA. *Pesquisa Operacional*, v. 22, n. 2, p. 165-182.

Lins, M.P.E., Angulo Meza, L., Silva, A.C.M. (2004). A multi-objective approach to determine alternative targets in data envelopment analysis. *Journal of the Operational Research Society*, v. 55, p. 1090-110.

Farrell, M.J. (1957). The measurement of productive efficiency. *Journal of the Royal Static Society*. Ser. A. 120, p. 253-281.

Gonzalez-Araya, M.C. (2003). *Projeções Não Radiais em Regiões Fortemente Eficientes da Fronteira DEA - Algoritmos e Aplicações*. Tese de Doutorado. Programa de Engenharia de Produção. COPPE/UFRJ.

Pastor, J.T. (1995). Improving the new DEA-efficiency measure of Tone. *Working paper*. Departamento de Estadística e Investigación Operativa. Universidad de Alicante.

Pastor, J.T., Ruiz, J.L., Sirvent, I. (1999). An enhanced DEA Russell graph efficiency measure. *European Journal of Operational Research*, v. 115. p. 596-607.

Ruggiero, J., (2000). Measuring technical efficiency. *European Journal of Operational Research*, v. 121, p. 138-150.

Ruggiero, J., Bretschneider, S. (1998). The weighted Russell measure of technical efficiency. *European Journal of Operational Research*, v. 108, p. 438-451.

Soares de Mello, J.C.C.B., Estellita-Lins, M.P., Gomes, E.G. (2002). Construction of a Smoothed DEA Frontier. *Pesquisa Operacional*, v. 22, n. 2, p. 183-201.

Soares de Mello, J.C.C.B., Angulo-Meza, L., Gomes, E. G.; Serapião, B.P.; Lins, M.P.E. (2003). Análise de Envoltória de Dados no estudo da eficiência e dos benchmarks para companhias aéreas brasileiras. *Pesquisa Operacional*, v. 23, n. 2, p. 325-345.

Soares de Mello, J.C.C.B., Gomes, E.G., Angulo-Meza, L., Biondi, L.N., Sant'Anna, A. P. (2005). Fronteiras DEA difusas. *Investigação Operacional*, v. 25, n. 1, p. 85-103.

Tone, K. (2002). A slacks-based measure of super-efficiency in data envelopment analysis. *European Journal of Operational Research*, v. 143, p. 32-41.

Tone, K. (2001). A slacks based measure of efficiency in Data Envelopment Analysis. *European Journal of Operational Research*, v. 130, p. 498-509.

An Extension of a Variant of a Predictor-Corrector Primal-Dual Method from Linear Programming to Semidefinite Programming

F. Bastos * ‡ A. Teixeira * †

‡ Departamento de Estatística e Investigação Operacional
Universidade de Lisboa
Lisboa - Portugal
fbastos@fc.ul.pt

† Departamento de Matemática
Universidade de Trás-os-Montes e Alto Douro
Vila Real - Portugal
ateixeir@utad.pt

* CIO - Centro de Investigação Operacional

Abstract

We extend a variant of a predictor-corrector primal-dual method for Linear Programming to Semidefinite Programming. Two versions are proposed. One of the versions uses the HKM direction and the other the NT direction. We present the algorithms associated with these versions and the computational experience using the SDPLIB 1.2 collection of Semidefinite Programming test problems. We show that, in general, the algorithm using the HKM direction is the best and is also better than the one relative to the classical method.

Keywords: Semidefinite Programming, predictor-corrector interior point variant, HKM direction, NT direction.

1 Introduction

Semidefinite Programming, SDP, is a recent, but important, field of Mathematical Programming and, although its roots are older [3, 12], most of its remarkable advances were made in the 90's [1, 14]. With this work, we intend to obtain a predictor-corrector primal-dual interior

Apdio © 2005 *Associação Portuguesa de Investigação Operacional*

point algorithm with better performance and more precise than the other algorithms of the same type already known.

Many interior-point methods for SDP are extensions of the methods that already exist for Linear Programming, LP [1, 14]. In [2] a variant of Mehrotra's predictor-corrector method for LP was presented and it was shown to be more efficient than the original for the class of problems studied in that work. We propose a new variant for SDP that is an extension of that one for LP.

We present two versions of the new variant. One of them uses a direction, that was independently proposed by Helmberg, Rendl, Vanderbei and Wolkowicz in [8] and Kojima, Shindoh and Hara [11] and later rediscovered by Monteiro [13], which is usually refered as *HKM direction*. The other version uses the *NT direction* introduced by Nesterov and Todd in [15]. We also compare the results of these versions and the ones obtained with the classical method.

This paper is organized as follows. In Section 2, we present some useful notions and results about Semidefinite Programming and the solution of the the *Lyapunov equation*

$$AX + XA^T = H, \tag{1}$$

with $A \in \mathcal{S}_n^{++}$ and $H \in \mathcal{M}_n$. In Section 3, the primal-dual and the predictor-corrector algorithms, bases of the new variant, are described. In Section 4, the version of the variant with the HKM direction is presented. In Section 5, the NT direction for the new variant is computed and the corresponding version of the variant is presented. In Section 6, the computational experience is described. A brief description of the used tools is given and the obtained results with the new versions and the classical method are compared.

2 Preliminaries

The set of real $m \times n$ matrices, $\mathcal{M}_{m,n}$, is a real vector space isomorphic to $\mathbb{R}^{mn}$. In this vector space we define the inner product between $A, B \in \mathcal{M}_{m,n}$ by

$$\langle A, B\rangle = \mathrm{Tr}\,(B^T A) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} b_{ij},$$

where $\mathrm{Tr}\,(\cdot)$ represents the trace, which is the sum of the main diagonal entries of a matrix. The associated norm, $\langle A, A\rangle^{1/2}$, is called *Frobenius norm* and is represented by $\|A\|_F$. The particular case of the square matrices of order n will be represented by $\mathcal{M}_n$. The set of real symmetric matrices of order n, $\mathcal{S}_n = \{A \in \mathcal{M}_n : A^T = A\}$, is a subspace of $\mathcal{M}_n$ isomorphic to $\mathbb{R}^{n(n+1)/2}$. For $A, B \in \mathcal{S}_n$ $\langle A, B\rangle = \mathrm{Tr}\,(AB)$. The set of positive semidefinite matrices, $\mathcal{S}_n^+$, is a full dimensional, non polyhedral convex cone in $\mathcal{S}_n$ and induces a partial order on the set of symmetric matrices. We write $A \succeq B$ if $A - B \in \mathcal{S}_n^+$. The interior of the cone $\mathcal{S}_n^+$ is the set of positive definite matrices and is denoted by $\mathcal{S}_n^{++}$. Similarly, $A \succ B$ means that $A - B \in \mathcal{S}_n^{++}$.

Consider the matrices $A_i \in \mathcal{S}_n$, $i = 1, ..., m$ and the linear operator $\mathcal{A} : \mathcal{S}_n \longrightarrow \mathbb{R}^m$ defined

by

$$\mathcal{A}(X) = \begin{bmatrix} \langle A_1, X\rangle \\ \vdots \\ \langle A_m, X\rangle \end{bmatrix}.$$

The adjoint operator of $\mathcal{A}$, the operator $\mathcal{A}^T : \mathbb{R}^m \to \mathcal{S}_n$ defined by $\langle \mathcal{A}(X), y\rangle = \langle X, \mathcal{A}^T(y)\rangle$, for all $X \in \mathcal{S}_n$ and $y \in \mathbb{R}^m$, is $\mathcal{A}^T(y) = \sum_{i=1}^m y_i A_i$.

The *standard* primal-dual pair of Semidefinite Programming problems is

$$\begin{array}{llll} \text{(P)} & \min\limits_{X} & \langle C, X\rangle & \\ & \text{s.t} & \mathcal{A}(X) = b & \\ & & X \succeq 0 & \end{array} \qquad \begin{array}{llll} \text{(D)} & \max\limits_{y,Z} & b^T y & \\ & \text{s.t} & \mathcal{A}^T(y) + Z = C & \\ & & Z \succeq 0, & \end{array}$$

where $C \in \mathcal{S}_n$ and $b \in \mathbb{R}^m$.

The primal problem is (P) while (D) is the Lagrangian dual of (P). Any X $((y, Z))$ that satisfy the constraints of (P) ((D)) is called a *feasible solution* of (P) ((D)). If X $((y, Z))$ is a feasible solution of (P) ((D)) and $X \succ 0$ $(Z \succ 0)$ then X $((y, Z))$ is called a *strictly feasible solution* of (P) $((D))$.

Any primal-dual pair of feasible solutions verifies the relation $\langle C, X\rangle - b^T y = \langle X, Z\rangle \geqslant 0$, called *weak duality* property. If $\langle X, Z\rangle = 0$, then the primal-dual pair (X, Z) is optimal but the reciprocal is not true, that is, the *strong duality* property is not always true. A sufficient condition for the optimal values of (P) and (D) be the same is the existence of strictly feasible solutions for both the primal and the dual problems.

We end this section with a result that gives the solution of the *Lyapunov equation*

$$AX + XA^T = H, \tag{2}$$

with $A \in \mathcal{S}_n^{++}$ and $H \in \mathcal{M}_n$ known [10].

Let us first introduce the definition of Hadamard product [9], necessary to the next theorem.

Definition 2.1. *The Hadamard product of $A = [a_{ij}], B = [b_{ij}] \in \mathcal{M}_{m,n}$ is $A \circ B = [a_{ij} b_{ij}]$.*

Theorem 2.1. *Let $A \in \mathcal{S}_n^{++}$ be a matrix satisfying $A = S\Lambda S^T$, $S, \Lambda \in \mathcal{M}_n$, S orthogonal and $\Lambda = \mathrm{Diag}(\lambda_1(A), \ldots, \lambda_n(A))$, where $\lambda_i(A)$, $i = 1, \ldots, n$ are the eigenvalues of A. Consider that $H \in \mathcal{M}_n$, then the solution of (2) is*

$$X = S[L \circ (S^T H S)]S^T,$$

with $L = [l_{ij}]$, where $l_{ij} = (\lambda_i(A) + \lambda_j(A))^{-1}$.

3 A primal-dual method and a classical predictor-corrector variant

The SDP algorithms presented in this paper correspond to different versions of a new predictor-corrector variant of the primal-dual method developed by Christopher Helmberg, Franz Rendl,

Robert J. Vanderbei and Henry Wolkowicz [8]. In this section we will briefly describe the infeasible versions of this primal-dual method and of the classical predictor- -corrector variant.

Although it is not needed a feasible solution, the method requires the existence of an initial primal-dual pair (X^0, Z^0) satisfying $X^0, Z^0 \in \mathcal{S}_n^{++}$.

The primal barrier problem is

$$\begin{array}{rll} \mathrm{PB}(\mu): & \min\limits_{X} & \langle C, X\rangle - \mu \ln(\det(X)) \\ & \text{s.a} & \mathcal{A}(X) = b, \end{array}$$

where $\mu > 0$ is the *barrier parameter.*

For each μ, the corresponding Lagrangean is

$$\mathcal{L}_\mu(X, y) = \langle C, X\rangle - \mu \ \ln(\det(X)) + \langle y, b - \mathcal{A}(X)\rangle.$$

The first order optimality conditions for $\mathrm{PB}(\mu)$, necessary and sufficient in this case, are

$$\begin{array}{rcl} \nabla_X \mathcal{L}_\mu & = & C - \mu X^{-1} - \mathcal{A}^T(y) = 0 \\ \nabla_y \mathcal{L}_\mu & = & b - \mathcal{A}(X) = 0. \end{array}$$

Taking $Z = \mu X^{-1}$ we can rewrite these conditions in the following form

$$\mathrm{CPS}(\mu) \quad \left\{ \begin{array}{c} \mathcal{A}(X) = b \\ \mathcal{A}^T(y) + Z = C \\ XZ = \mu I \\ X \succ 0 \\ Z \succ 0. \end{array} \right. \tag{3}$$

The set of solutions of $\mathrm{CPS}(\mu)$ for $\mu > 0$,

$$\mathcal{C} = \{(X_\mu, y_\mu, Z_\mu) : \mu > 0\},$$

is called *central path.* Note that, for a point (X, y, Z) in the central path we obtain $\mu = \langle X, Z\rangle/n$.

Given $(X, y, Z) \in \mathcal{C}$ we want to find a search direction $(\Delta X, \Delta y, \Delta Z)$ so that $(X + \Delta X, y + \Delta y, Z + \Delta Z)$ is still in the central path. Let

$$F_\mu(p) = \left[\begin{array}{c} \mathcal{A}(X) - b \\ \mathcal{A}^T(y) + Z - C \\ XZ - \mu I \end{array} \right], \tag{4}$$

with $p = (X, y, Z)$. Using Newton's method we can obtain a search direction $\Delta p = (\Delta X, \Delta y, \Delta Z)$ by solving

$$F_\mu(p) + \nabla F_\mu(p) \Delta p^T = 0.$$

This system can be written as

$$\text{NFS}(\mu)\left\{\begin{array}{lcl}\mathcal{A}(\Delta X) & = & F_p\\ \mathcal{A}^T(\Delta y)+\Delta Z & = & F_d\\ \Delta XZ+X\Delta Z & = & \mu I-XZ,\end{array}\right. \tag{5}$$

where $F_p = b - \mathcal{A}(X)$ and $F_d = C - Z - \mathcal{A}^T(y)$.

Rewriting NFS(μ) we obtain

$$\left\{\begin{array}{lcl}\mathcal{A}(X\mathcal{A}^T(\triangle y)Z^{-1}) & = & -\mu\mathcal{A}(Z^{-1})+b+\mathcal{A}(XF_dZ^{-1})\\ \Delta Z & = & F_d-\mathcal{A}^T(\Delta y)\\ \Delta X & = & \mu Z^{-1}-X-XF_dZ^{-1}+X\mathcal{A}^T(\Delta y)Z^{-1}.\end{array}\right. \tag{6}$$

From the first equation we get Δy and, replacing it in the remaining equations, we get ΔX and ΔZ. The obtained ΔZ is symmetric, consequence of the second equation, but, as X and Z do not commute, this is not the case for ΔX. As the first and the third components of the next iterate must belong to $\mathcal{S}_n^{++}$ we have to obtain a symmetric ΔX. There are several ways of doing this, one of them is just using the symmetric part of ΔX, that is, supposing that $(\widehat{\Delta X}, \Delta y, \Delta Z)$ is the solution of (6) then, our search direction will be $(\Delta X, \Delta y, \Delta Z)$ with $\Delta X = (\widehat{\Delta X} + (\widehat{\Delta X})^T)/2$. This search direction is known as *HRVW/KSH/M direction.*

In each iteration, two step lengths are computed, one associated with the primal direction, α_p, and the other with the dual, α_d. Both values are obtained using the backtracking method. After computing α_p and α_d the algorithm moves from the current iterate (X, y, Z) to the next iterate

$$(X+\alpha_p\Delta X, y+\alpha_d\Delta y, Z+\alpha_d\Delta Z).$$

Brian Borchers [4] developed an infeasible predictor-corrector variant of the primal-dual method presented in [8]. We will call it *classical predictor-corrector*, PCC. In this variant, to obtain the search direction we assume that is possible to take a step of length one in the direction $(\Delta X, \Delta y, \Delta Z)$ and replace in the first three equations of CPS(μ), in (3), the current iterate (X, y, Z) by the new iterate $(X + \Delta X, y + \Delta y, Z + \Delta Z)$. The result is the system

$$\left\{\begin{array}{lcl}\mathcal{A}(\Delta X) & = & F_p\\ \mathcal{A}^T(\Delta y)+\Delta Z & = & F_d\\ X\Delta Z+\Delta XZ & = & \mu I-XZ-\Delta X\Delta Z.\end{array}\right. \tag{7}$$

To obtain the predictor direction, in (7) we take $\mu = 0$ and neglect the non-linear term $\Delta X\Delta Z$. Solving this system we get the triple $(\widehat{\Delta X^\circ}, \Delta y^\circ, \Delta Z^\circ)$ satisfying

$$\left\{\begin{array}{lcl}\mathcal{A}(X\mathcal{A}^T(\Delta y^\circ)Z^{-1}) & = & b+\mathcal{A}(XF_dZ^{-1})\\ \Delta Z^\circ & = & F_d-\mathcal{A}^T(\Delta y^\circ)\\ \widehat{\Delta X^\circ} & = & -X-XF_dZ^{-1}+X\mathcal{A}^T(\Delta y^\circ)Z^{-1},\end{array}\right.$$

with $F_d = C - Z - \mathcal{A}^T(y)$. As in the primal-dual case, the predictor direction $(\Delta X^\circ, \Delta y^\circ, \Delta Z^\circ)$ is obtained symmetrizing $\widehat{\Delta X^\circ}$.

The corrector direction is also obtained from (7). We use $(\Delta X^\circ, \Delta y^\circ, \Delta Z^\circ)$ to compute μ and to approximate the non-linear term $\Delta X\Delta Z$ by $\Delta X^\circ\Delta Z^\circ$. The final direction is the

sum of the predictor and the corrector directions, that is, we take $(\Delta X, \Delta y, \Delta Z) = (\widehat{\Delta X^\circ} + \widehat{\Delta X^c}, \Delta y^\circ + \Delta y^c, \Delta Z^\circ + \Delta Z^c)$. Now, considering the conditions satisfied by the predictor direction we obtain the following system

$$\left\{\begin{array}{rcl} \mathcal{A}(\widehat{\Delta X^c}) & = & 0 \\ \mathcal{A}^T(\Delta y^c) + \Delta Z^c & = & 0 \\ X\Delta Z^c + \widehat{\Delta X^c} Z & = & \mu I - \Delta X^\circ \Delta Z^\circ. \end{array}\right. \tag{8}$$

From (8) we get the triple $(\widehat{\Delta X^c}, \Delta y^c, \Delta Z^c)$ satisfying

$$\left\{\begin{array}{rcl} \mathcal{A}(X\mathcal{A}^T(\Delta y^c)Z^{-1}) & = & -\mathcal{A}(\mu Z^{-1} - \Delta X^\circ \Delta Z^\circ Z^{-1}) \\ \Delta Z^c & = & -\mathcal{A}^T(\Delta y^c) \\ \widehat{\Delta X^c} & = & \mu Z^{-1} - \Delta X^\circ \Delta Z^\circ Z^{-1} + X\mathcal{A}^T(\Delta y^c)Z^{-1}. \end{array}\right.$$

We obtain the corrector direction $(\Delta X^c, \Delta y^c, \Delta Z^c)$ after symmetrizing $\widehat{\Delta X^c}$. The final direction is

$$(\Delta X, \Delta y, \Delta Z) = (\Delta X^\circ + \Delta X^c, \Delta y^\circ + \Delta y^c, \Delta Z^\circ + \Delta Z^c).$$

In the classical predictor-corrector algorithm, like in the primal-dual, different choices for the primal and dual step lengths are made, using the backtracking method. In each iteration,

$$\begin{array}{rcl} \alpha_p^k & = & 0.99 \sup\limits_{\alpha}\{\alpha \in [0,1] : \ X^k + \alpha \Delta X^k \succ 0\} \\ \alpha_d^k & = & 0.99 \sup\limits_{\alpha}\{\alpha \in [0,1] : \ Z^k + \alpha \Delta Z^k \succ 0\}. \end{array} \tag{9}$$

To obtain the supreme, we begin with $\alpha = 1$ and α is actualized to be 90% of the previous α, this is, $\alpha_{\text{new}} = 0.9\alpha_{\text{old}}$.

The computation of the duality measure, μ, is not only based on the quantities $\langle X, Z\rangle/n$ and $\langle X, Z\rangle/2n$ but it also uses two line searches associated with the predictor direction.

The initial solution used is

$$X^0 = 100\alpha I, \qquad y = 0, \qquad Z^0 = 100\beta I, \tag{10}$$

with

$$\begin{array}{rcl} \alpha & = & n \max\limits_{1 \leqslant i \leqslant m} \dfrac{1 + |b_i|}{1 + \|A_i\|_F}, \\ \beta & = & \dfrac{1 + \max\left(\max\limits_{1 \leqslant i \leqslant m} \|A_i\|_F, \|C\|_F\right)}{\sqrt{n}}. \end{array} \tag{11}$$

Finally, the termination criteria is

$$\begin{array}{ll} \dfrac{\|b - \mathcal{A}(X^k)\|_2}{1 + \|b\|_2} \leqslant 1.0 \times 10^{-06}; & \dfrac{\|C - \mathcal{A}^T(y^k) - Z^k\|_F}{1 + \|C\|_F} \leqslant 1.0 \times 10^{-06} \\ \dfrac{|\langle C, X^k\rangle - b^T y^k|}{1 + |b^T y^k|} \leqslant 1.0 \times 10^{-07}; & X^k \succ 0; \quad Z^k \succ 0. \end{array} \tag{12}$$

4 A new predictor-corrector variant

From the previous section, we know that the classical predictor-corrector algorithm performs four line searches by iteration, two of them after obtaining the predictor direction to estimate the duality measure and the other two after computing the final direction. The algorithm we extend from LP [2] just makes two line searches per iteration and so it has the advantage of decreasing the running time of each iteration.

The major differences between this new variant and the classical predictor-corrector for LP are:

- the predictor direction is computed as in the primal-dual method;
- it uses the same μ for both the predictor and the corrector directions;
- to obtain μ no line search is needed.

We first tried to extend this variant to SDP literally. In particular, we used $\mu = \langle X, Z\rangle/2n$, indicated in [8] as a good heuristic for the primal-dual algorithm, in the predictor and corrector steps. We verified that this option did not work in this context. We also tried some other heuristics mentioned in [2] such as

$$\mu = \frac{\langle X, Z\rangle}{\theta(n)} \qquad \text{with} \qquad \theta(n) = \begin{cases} n^2 & \text{if} \quad n \leqslant 5000 \\ n\sqrt{n} & \text{if} \quad n > 5000, \end{cases}$$

always using the same μ in both steps.

After some computational experience, we verified that we could get substantial improvements if we consider different values of μ in the predictor and corrector steps. From that computational experience we concluded that the best heuristics for μ are:

$$\begin{aligned} \mu = \mu_p &= \frac{\langle X, Z\rangle}{2.2n^2}, && \text{in the predictor step} \\ \mu = \mu_c &= \frac{\langle X, Z\rangle}{2n^{5/4}}, && \text{in the corrector step.} \end{aligned} \tag{13}$$

These expressions for the duality measure, although basically empirical, have some theoretical foundation. Let us assume this is a feasible method, that is, $F_p = 0$ and $F_d = 0$ in NFS(μ), (5). Then, at k-th iteration

$$\begin{aligned} \langle C, X^k\rangle - \langle b, y^k\rangle &= \langle \mathcal{A}^T(y^k) + Z^k, X^k\rangle - \langle \mathcal{A}(X^k), y^k\rangle \\ &= \langle \mathcal{A}^T(y^k) + Z^k, X^k\rangle - \langle \mathcal{A}^T(y^k), X^k\rangle \\ &= \langle Z^k, X^k\rangle. \end{aligned}$$

Suppose now that, instead of considering α_p^k and α_d^k, we would consider $\alpha^k = \min\{\alpha_p^k, \alpha_d^k\}$. Then, we would get

$$\begin{aligned} \langle C, X^{k+1}\rangle - \langle b, y^{k+1}\rangle &= \langle Z^{k+1}, X^{k+1}\rangle = \langle Z^k + \alpha^k \Delta Z^k, X^k + \alpha^k \Delta X^k\rangle \\ &= \langle C, X^k\rangle - \langle b, y^k\rangle + \alpha^k \left(\langle Z^k, \Delta X^k\rangle + \langle \Delta Z^k, X^k\rangle\right) + \\ &\quad + (\alpha^k)^2 \langle \Delta Z^k, \Delta X^k\rangle. \end{aligned}$$

On the other side, from the third equation of NFS(μ), (5), we have $\mathrm{Tr}\,(\Delta X^k Z^k + X^k \Delta Z^k) = \mathrm{Tr}\,(\mu^k I - X^k Z^k)$ that is equivalent to

$$\langle Z^k, \Delta X^k\rangle + \langle \Delta Z^k, X^k\rangle = n\mu^k - \langle Z^k, X^k\rangle.$$

Furthermore, using the definition of adjoint operator and the assumption of the method being feasible, we obtain

$$\begin{aligned}\langle \Delta Z^k, \Delta X^k\rangle &= \langle -\mathcal{A}^T(\Delta y^k), \Delta X^k\rangle = -\langle \mathcal{A}(\Delta X^k), \Delta y^k\rangle \\ &= \langle 0, \Delta y^k\rangle = 0.\end{aligned}$$

Then, $\langle C, X^{k+1}\rangle - \langle b, y^{k+1}\rangle = \langle C, X^k\rangle - \langle b, y^k\rangle + \alpha^k \left(n\mu^k - \langle Z^k, X^k\rangle\right)$. This allow us to conclude that

$$\langle C, X^{k+1}\rangle - \langle b, y^{k+1}\rangle < \langle C, X^k\rangle - \langle b, y^k\rangle$$

if

$$\mu^k < \frac{\langle Z^k, X^k\rangle}{n}. \tag{14}$$

Both options for μ^k in (13) satisfy (14). But, as the method is not feasible and we do not use $\alpha^k = \min\{\alpha_p^k, \alpha_d^k\}$, we can not guarantee a decrease of $\langle C, X\rangle - \langle b, y\rangle$. However, in practice, we verified that these options are much better than any other already proposed.

To implement this new variant we first used the HKM direction and the backtracking method to compute the step length, as in the classical predictor-corrector. For the backtracking method the proceeding is similar to the one described in (9), but the new α is chosen as 85% of the previous one and not 90% as it happened before. This version of the new variant can be formalized in the following algorithm:

Algorithm 1 - BHKM

Require: (X^0, y^0, Z^0) satisfying (10) and (11);
Take $k = 0$;
while the termination criteria (12) is not satisfied **do**
(a) Take $\mu_p^k = \sigma^k \dfrac{\langle X^k, Z^k\rangle}{2.2n^2}$, with $\sigma^k \in [0, 1]$;
(b) Solve system (6) for $(\widehat{\Delta X^{\circ^k}}, \Delta y^{\circ^k}, \Delta Z^{\circ^k})$, considering $\mu = \mu_p^k$;
Symmetrize $\widehat{\Delta X^{\circ^k}}$ to get $(\Delta X^{\circ^k}, \Delta y^{\circ^k}, \Delta Z^{\circ^k})$;
(c) Take $\mu_c^k = \sigma^k \dfrac{\langle X^k, Z^k\rangle}{2n^{5/4}}$, with $\sigma^k \in [0, 1]$;
(d) Solve system (8) for $(\widehat{\Delta X^{c^k}}, \Delta y^{c^k}, \Delta Z^{c^k})$, considering $\mu = \mu_c^k$;
Symmetrize $\widehat{\Delta X^{c^k}}$ to get $(\Delta X^{c^k}, \Delta y^{c^k}, \Delta Z^{c^k})$;
(e) Take $(\Delta X^k, \Delta y^k, \Delta Z^k) = (\Delta X^{\circ^k}, \Delta y^{\circ^k}, \Delta Z^{\circ^k}) + (\Delta X^{c^k}, \Delta y^{c^k}, \Delta Z^{c^k})$;
(f) Compute α_p^k and α_d^k using (9) with $\alpha_{\text{new}} = 0.85\alpha_{\text{old}}$;
(g) Take $X^{k+1} = X^k + \alpha_p^k \Delta X^k$ and $(y^{k+1}, Z^{k+1}) = (y^k, Z^k) + \alpha_d^k(\Delta y^k, \Delta Z^k)$;
(h) Take $k = k + 1$;
end while

5 The new predictor-corrector variant with NT direction

Another direction that is often used in practice and also theoretically studied is due to Nesterov and Todd and is called *NT direction* [15, 16, 17]. Thus, we decided to implement this direction

in our predictor-corrector variant.

As this direction is not directly obtained, as it happens with HKM, we will present its computation. Once more, we will assume that is possible to take a step of length one in the direction $(\Delta X, \Delta y, \Delta Z)$. This way we get the system

$$\left\{\begin{array}{ccc} \mathcal{A}(\Delta X) & = & F_p \\ \mathcal{A}(\Delta y) + \Delta Z & = & F_d \\ (X+\Delta X)(Z+\Delta Z) & = & \mu I, \end{array}\right. \tag{15}$$

with $F_p = b - \mathcal{A}(X)$ and $F_d = C - Z - \mathcal{A}^T(y)$.

The NT direction is obtained using a scaling matrix to reformulate the third equation of (15), so that the solution of the new system is symmetric. Let us consider the Cholesky factorizations of X and Z, $X = LL^T$ and $Z = RR^T$, the singular value decomposition of R^TL, $UDV^T = R^TL$, and the scaling matrix $W = X^{1/2}(X^{1/2}ZX^{1/2})^{-1/2}X^{1/2}$. It can be shown that W satisfies the condition $W^{-1}X = ZW$ and that it can be written as $W = GG^T$, with $G = LVD^{-1/2}$ [17]. Then, $G^{-1}XG^{-T} = G^TZG$.

Consider now $D = G^{-1}XG^{-T} = G^TZG$, $D_X = G^{-1}\Delta XG^{-T}$ and $D_Z = G^T\Delta ZG$. Thus, the third equation of (15) is equivalent to $(D+D_X)(D+D_Z) = \mu I$. Weakening this equation, by replacing the left member by its symmetric part, we get

$$\frac{1}{2}\left[(D_X+D_Z)D + D(D_X+D_Z)\right] = \mu I - D^2 - \frac{1}{2}\big(D_XD_Z + D_ZD_X\big). \tag{16}$$

To obtain the predictor direction we take $\mu = \mu_p$ and neglect the non-linear terms in (16). This way we get a Lyapunov equation in $D_X^\circ + D_Z^\circ$ which has the symmetric solution $\mu_pD^{-1} - D$. Multiplying both members of

$$D_X^\circ + D_Z^\circ = \mu_pD^{-1} - D$$

by G on the left and G^T on the right we get

$$\Delta X^\circ = \mu_pZ^{-1} - X - W\Delta Z^\circ W.$$

This equation, jointly with (15), allow us to obtain

$$\left\{\begin{array}{ccc} \mathcal{A}(\Delta X^\circ) & = & F_p \\ \mathcal{A}^T(\Delta y^\circ) + \Delta Z^\circ & = & F_d \\ \Delta X^\circ & = & \mu_pZ^{-1} - X - W\Delta Z^\circ W. \end{array}\right. \tag{17}$$

Solving (17) we get the triple $(\Delta X^\circ, \Delta y^\circ, \Delta Z^\circ)$ satisfying

$$\left\{\begin{array}{ccc} \mathcal{A}(W\mathcal{A}^T(\Delta y^\circ)W) & = & \mathcal{A}(WF_dW + X - \mu_pZ^{-1}) + F_p \\ \Delta Z^\circ & = & F_d - \mathcal{A}^T(\Delta y^\circ) \\ \Delta X^\circ & = & \mu_pZ^{-1} - X - WF_dW + W\mathcal{A}^T(\Delta y^\circ)W. \end{array}\right. \tag{18}$$

The corrector direction is obtained taking in (16) $D_X = D_X^\circ + D_X^c$, $D_Z = D_Z^\circ + D_Z^c$ and $\mu = \mu_p + \mu_c$. The non-linear terms D_XD_Z and D_ZD_X are also approximated by $D_X^\circ D_Z^\circ$ and

$D_Z^\circ D_X^\circ$, respectively. Then, considering the conditions satisfied by the predictor direction, we can write

$$(D_X^c + D_Z^c)D + D(D_X^c + D_Z^c) = H_c,$$

with $H_c = 2\mu_c I - (D_X^\circ D_Z^\circ + D_Z^\circ D_X^\circ)$. By Theorem 2.1, the solution of this Lyapunov equation is $L \circ H_c$, with $L = [l_{ij}]$ where $l_{ij} = (d_i + d_j)^{-1}$ and d_i and d_j represent the i-th and the j-th entries of the diagonal matrix D, respectively. Multiplying both members of

$$D_X^c + D_Z^c = L \circ H_c$$

by G on the left and by G^T on the right we get

$$\Delta X^c = G(L \circ H_c)G^T - W\Delta Z^c W. \tag{19}$$

Now, consider $D_X^\circ = G^{-1}\Delta X^\circ G^{-T}$, $D_X^c = G^{-1}\Delta X^c G^{-T}$, $D_Z^\circ = G^T\Delta Z^\circ G$ and $D_Z^c = G^T\Delta Z^c G$. Then, it is possible to show that the equalities $D_X = D_X^\circ + D_X^c$ and $D_Z = D_Z^\circ + D_Z^c$ are equivalent to $\Delta X = \Delta X^\circ + \Delta X^c$ and $\Delta Z = \Delta Z^\circ + \Delta Z^c$, respectively. Furthermore, $D_X^\circ D_Z^\circ + D_Z^\circ D_X^\circ = G^{-1}\Delta X^\circ \Delta Z^\circ G + (G^{-1}\Delta X^\circ \Delta Z^\circ G)^T$. Thus, taking $(\Delta X, \Delta y, \Delta Z) = (\Delta X^\circ + \Delta X^c, \Delta y^\circ + \Delta y^c, \Delta Z^\circ + \Delta Z^c)$ and considering (19) and the conditions satisfied by the predictor direction we get the system

$$\left\{\begin{array}{rcl} \mathcal{A}(\Delta X^c) & = & 0 \\ \Delta Z^c + \mathcal{A}^T(\Delta y^c) & = & 0 \\ \Delta X^c & = & G(L \circ H_c)G^T - W\Delta Z^c W, \end{array}\right. \tag{20}$$

with $H_c = 2\mu_c I - [(G^{-1}\Delta X^\circ \Delta Z^\circ G) + (G^{-1}\Delta X^\circ \Delta Z^\circ G)^T]$. The solution of (20) is the corrector direction $(\Delta X^c, \Delta y^c, \Delta Z^c)$ that satisfies

$$\left\{\begin{array}{rcl} \mathcal{A}(W\mathcal{A}^T(\Delta y^c)W) & = & -\mathcal{A}(G(L \circ H_c)G^T) \\ \Delta Z^c & = & -\mathcal{A}^T(\Delta y^c) \\ \Delta X^c & = & G(L \circ H_c)G^T + W\mathcal{A}^T(\Delta y^c)W. \end{array}\right.$$

The final direction, $(\Delta X, \Delta y, \Delta Z)$, is obtained adding the predictor and the corrector directions.

From the computational experience with this version we verified that we got better results when we used a different heuristic for the duality measure in the predictor direction,

$$\mu_p = \frac{\langle X, Z\rangle}{2.3n^2}.$$

The algorithm corresponding to the version of our predictor-corrector variant that uses NT direction will be called *BNT algorithm* and differs from the BHKM algorithm in the following points:

(a) Take $\mu_p^k = \sigma^k \dfrac{\langle X^k, Z^k \rangle}{2.3n^2}$, with $\sigma^k \in [0,1]$;
(b) Solve system (17) for $(\Delta X^{\circ^k}, \Delta y^{\circ^k}, \Delta Z^{\circ^k})$, considering $\mu_p = \mu_p^k$;
(d) Solve system (20) for $(\Delta X^{c^k}, \Delta y^{c^k}, \Delta Z^{c^k})$, considering $\mu_c = \mu_c^k$;

6 Computational experience

Now, we will describe the computational experience that we have done to compare the two versions of our predictor-corrector variant and the classical predictor-corrector method, described in the previous sections.

6.1 Brief description of the used tools

The computational tests were performed in a Pentium III, 1100 Mhz, 245232 KB of RAM, with the 7.3 version of the Red Hat Linux operating system and the 2.96 version of the gcc compiler. To implement the new predictor-corrector variant we used the 3.2 version of the source code of the package CSDP by Brian Borchers [4, 6, 7]. The code was modified to achieve two main purposes: it was adapted to be possible to implement the different versions of the predictor-corrector variant and it was optimized to become faster and more robust.

To compare the performance of the algorithms we used the SDPLIB 1.2 collection of SDP test problems [5]. It has 92 problems in a wide range of sizes and drawn from many different applications, including truss topology design, control systems engineering and combinatorial optimization. It also includes primal and dual infeasible instances.

6.2 Results

We will present the results corresponding to both versions of the predictor-corrector variant described earlier and compare those results with the ones obtained with the classical predictor-corrector algorithm (PCC algorithm). We will use tables with information about the number of iterations, It, the total CPU time (in seconds), T, and the relative duality gap, RDgap. We will also include information about the relative admissibility, that specifies if the algorithm stops with a solution that is far from satisfying any of the admissibility criteria or not. All the solutions obtained satisfy the relative dual admissibility termination criteria but the same does not happen with the primal one. Thus, in RDgap column will appear * if the relative primal admissibility belongs to $]1.0 \times 10^{-06}, 1.0 \times 10^{-03}[$ and ** if it is non inferior to 1.0×10^{-03}. In all the experiences we used (12) as termination criteria and (10), satisfying (11), as initial solution.

Due to memory problems it was not possible to obtain any results for the problems *maxG55* and *maxG60*, so they were not included in the tables. For the same reason, it was also not possible to obtain some results with the BNT algorithm. In this case, the fields of the table will contain $\underline{\text{m}}$.

In the next section we will analyse the results presented.

Name	It	T	RDgap
arch0	30	31	3.51e-08
arch2	30	34	6.28e-08
arch4	29	33	9.99e-09
arch8	30	34	8.66e-08
control1	26	< 1	4.07e-08
control2	30	1	1.33e-08
control3	37	9	3.49e-07
control4	31	29	7.52e-08
control5	38	134	6.85e-07
control6	39	349	6.68e-08 *
control7	39	792	5.45e-08
control8	44	1680	1.69e-06
control9	39	2629	5.84e-07
control10	39	4476	2.87e-06
control11	41	7466	3.35e-06 *
equalG11	21	3378	1.06e-08
equalG51	24	7727	1.35e-08
gpp100	20	7	3.11e-08
gpp124-1	29	18	7.38e-07 *
gpp124-2	26	14	3.70e-08
gpp124-3	21	12	9.55e-08
gpp124-4	24	13	8.34e-08
gpp250-1	32	157	2.53e-05
gpp250-2	28	125	4.77e-07
gpp250-3	21	83	5.86e-08
gpp250-4	27	113	4.04e-08
gpp500-1	26	1070	7.56e-08
gpp500-2	34	1715	4.83e-07 *
gpp500-3	39	2099	5.49e-07
gpp500-4	32	1618	3.56e-08
hinf1	40	< 1	7.46e-06
hinf2	40	< 1	9.89e-09
hinf3	40	< 1	6.25e-06
hinf4	19	< 1	3.04e-08
hinf5	40	1	3.23e-05
hinf6	40	< 1	1.55e-05
hinf7	33	< 1	1.25e-08
hinf8	40	< 1	2.13e-06
hinf9	23	< 1	1.89e-08
hinf10	40	< 1	1.36e-04
hinf11	40	< 1	7.13e-05
hinf12	49	< 1	5.62e-04
hinf13	40	< 1	1.92e-05 *
hinf14	40	1	1.96e+00 **
hinf15	45	1	3.89e-06 *

Name	It	T	RDgap
infd1	14	< 1	2.69e+00
infd2	15	< 1	3.03e-01
infp1	40	2	2.09e+00
infp2	40	3	1.93e+00
maxG11	19	1691	5.22e-08
maxG32	20	51059	7.84e-08
maxG51	21	3962	1.00e-08
mcp100	16	4	1.33e-08
mcp124-1	16	5	9.59e-08
mcp124-2	17	5	1.00e-08
mcp124-3	18	6	1.00e-08
mcp124-4	17	7	1.48e-08
mcp250-1	18	35	1.00e-08
mcp250-2	17	34	8.03e-08
mcp250-3	18	40	1.00e-08
mcp250-4	17	42	6.99e-08
mcp500-1	20	430	9.50e-09
mcp500-2	20	459	2.62e-08
mcp500-3	20	483	1.00e-08
mcp500-4	20	526	1.00e-08
qap5	19	1	4.95e-06
qap6	18	2	6.07e-08
qap7	41	21	1.04e-07
qap8	26	49	3.92e-07
qap9	19	83	3.51e-08
qap10	22	234	4.42e-08
qpG11	20	15225	9.50e-09
qpG51	22	44005	1.00e-08
ss30	30	466	7.28e-08
theta1	18	< 1	9.50e-09
theta2	19	26	6.19e-08
theta3	19	247	4.14e-08
theta4	21	1387	3.14e-08
theta5	20	4727	6.84e-08
theta6	20	13951	6.73e-08
thetaG11	27	6221	2.19e-08
thetaG51	40	193855	2.19e-07
truss1	16	< 1	1.00e-09
truss2	18	< 1	2.42e-08
truss3	18	< 1	3.13e-08
truss4	16	< 1	1.00e-08
truss5	22	2	6.05e-08
truss6	24	1	5.11e-08
truss7	21	1	8.29e-08
truss8	27	41	4.71e-08

Table 1: PCC algorithm

Name	It	T	RDgap
arch0	31	28	2.18e-08
arch2	31	31	3.49e-08
arch4	30	31	1.42e-08
arch8	31	32	3.10e-08
control1	35	< 1	6.93e-08
control2	34	1	7.34e-08
control3	39	9	3.20e-07
control4	32	30	7.10e-08
control5	39	136	3.38e-07
control6	37	329	9.50e-07 *
control7	40	766	1.66e-06
control8	42	1515	1.61e-06
control9	38	2538	7.18e-07
control10	39	4457	1.12e-05 *
control11	41	7441	4.72e-06 *
equalG11	22	2888	6.60e-08
equalG51	24	6220	9.09e-08
gpp100	25	6	7.17e-08
gpp124-1	25	11	1.54e-08
gpp124-2	24	10	5.48e-08
gpp124-3	27	12	1.59e-08
gpp124-4	24	11	4.83e-08
gpp250-1	25	86	5.62e-08
gpp250-2	25	86	5.71e-08
gpp250-3	25	86	7.37e-08
gpp250-4	23	79	8.90e-09
gpp500-1	26	895	1.33e-08
gpp500-2	25	865	4.49e-08
gpp500-3	24	831	4.03e-08
gpp500-4	24	832	2.56e-08
hinf1	24	< 1	8.23e-08
hinf2	34	< 1	1.00e-07
hinf3	23	< 1	8.38e-08
hinf4	25	< 1	2.45e-08
hinf5	40	< 1	5.70e-06 *
hinf6	31	< 1	2.22e-06
hinf7	43	< 1	5.93e-08
hinf8	34	< 1	3.99e-07
hinf9	29	< 1	8.73e-09
hinf10	30	< 1	1.94e-05
hinf11	30	< 1	1.24e-05
hinf12	40	< 1	7.89e-02
hinf13	32	< 1	1.08e-05 *
hinf14	37	< 1	5.51e-06 *
hinf15	35	2	2.83e-05 *

Name	It	T	RDgap
infd1	14	< 1	2.58e+00
infd2	13	< 1	8.76e-01
infp1	30	1	1.87e+00
infp2	30	1	1.93e+00
maxG11	20	1364	3.49e-08
maxG32	22	52270	6.08e-08
maxG51	21	3047	6.61e-08
mcp100	17	2	5.37e-08
mcp124-1	18	4	6.91e-08
mcp124-2	19	5	3.67e-08
mcp124-3	18	5	3.76e-08
mcp124-4	19	5	3.02e-08
mcp250-1	20	31	1.55e-08
mcp250-2	20	32	3.44e-08
mcp250-3	19	34	3.51e-08
mcp250-4	19	39	5.91e-08
mcp500-1	20	322	2.04e-08
mcp500-2	20	351	2.11e-08
mcp500-3	19	353	4.92e-08
mcp500-4	20	403	1.83e-08
qap5	28	< 1	2.06e-08
qap6	26	3	6.25e-08
qap7	37	18	1.09e-07
qap8	25	39	8.69e-08
qap9	24	101	5.81e-08
qap10	27	280	9.31e-08
qpG11	21	12223	7.80e-08
qpG51	25	43256	8.52e-08
ss30	27	369	8.14e-08
theta1	22	< 1	3.20e-08
theta2	22	28	6.82e-08
theta3	22	276	3.65e-08
theta4	22	1426	3.57e-08
theta5	22	5115	5.05e-08
theta6	22	15187	1.86e-08
thetaG11	28	5233	2.85e-08
thetaG51	38	142052	1.68e-07
truss1	24	< 1	5.04e-08
truss2	20	< 1	2.19e-08
truss3	22	< 1	3.79e-08
truss4	23	< 1	3.88e-08
truss5	27	3	2.25e-08
truss6	24	1	1.47e-08
truss7	26	< 1	1.76e-08
truss8	27	38	2.82e-08

Table 2: BHKM algorithm

Name	It	T	RDgap
arch0	32	65	3.13e-08
arch2	30	65	8.68e-08
arch4	31	67	2.53e-08
arch8	31	67	8.54e-08
control1	34	1	4.98e-08
control2	33	1	4.75e-08
control3	39	10	2.43e-07
control4	32	37	3.99e-08
control5	44	160	1.58e-06 *
control6	38	354	2.61e-07 *
control7	39	786	5.64e-07
control8	38	1463	1.41e-06
control9	38	2708	7.68e-07
control10	41	4932	9.54e-06 *
control11	39	7211	3.33e-06 *
equalG11	21	6474	4.39e-08
equalG51	23	15046	3.19e-08
gpp100	25	12	9.99e-08
gpp124-1	26	23	1.33e-08
gpp124-2	29	26	2.36e-08
gpp124-3	23	20	9.13e-08
gpp124-4	22	19	4.62e-08
gpp250-1	25	207	1.55e-08
gpp250-2	25	210	4.53e-08
gpp250-3	23	193	7.33e-09
gpp250-4	23	194	4.96e-08
gpp500-1	27	2097	2.28e-08
gpp500-2	26	1995	6.01e-08
gpp500-3	24	1847	2.04e-08
gpp500-4	22	1678	3.84e-08
hinf1	33	< 1	1.91e-07
hinf2	33	< 1	1.11e-07
hinf3	32	< 1	1.45e-06
hinf4	25	< 1	8.64e-08
hinf5	30	< 1	2.02e-05 *
hinf6	20	< 1	7.88e-08
hinf7	39	< 1	1.83e-07 **
hinf8	32	< 1	1.18e-07
hinf9	28	< 1	8.52e-08
hinf10	30	< 1	4.06e-05
hinf11	30	< 1	4.41e-05
hinf12	40	< 1	7.49e-02
hinf13	30	2	7.23e-05 *
hinf14	38	< 1	9.57e-08 *
hinf15	33	1	2.15e-04 *

Name	It	T	RDgap
infd1	30	1	2.92e+00
infd2	30	2	4.36e-01
infp1	9	1	1.95e+00
infp2	10	< 1	1.98e+00
maxG11	19	5325	2.79e-08
maxG32	m	m	m
maxG51	22	12100	1.96e-08
mcp100	18	7	8.00e-08
mcp124-1	18	13	2.83e-08
mcp124-2	17	13	1.87e-08
mcp124-3	17	13	2.36e-08
mcp124-4	17	14	1.86e-08
mcp250-1	19	143	5.31e-08
mcp250-2	18	138	4.09e-08
mcp250-3	18	138	6.58e-08
mcp250-4	18	140	5.59e-08
mcp500-1	21	1426	2.83e-08
mcp500-2	20	1410	2.71e-08
mcp500-3	19	1347	4.02e-08
mcp500-4	19	1333	1.25e-08
qap5	19	< 1	1.60e-06
qap6	26	3	5.62e-08
qap7	28	14	7.95e-08
qap8	27	48	9.04e-08
qap9	28	128	7.89e-09
qap10	22	246	4.61e-08
qpG11	m	m	m
qpG51	m	m	m
ss30	27	655	9.30e-08
theta1	22	2	8.78e-08
theta2	22	33	2.01e-08
theta3	20	279	6.08e-08
theta4	22	1523	3.10e-08
theta5	21	4975	2.22e-08
theta6	24	16751	2.20e-8
thetaG11	28	11021	4.89e-08
thetaG51	m	m	m
truss1	24	< 1	4.07e-08
truss2	19	1	7.57e-08
truss3	23	< 1	3.83e-08
truss4	23	< 1	8.57e-08
truss5	22	3	3.38e-08
truss6	25	3	6.44e-08
truss7	24	< 1	3.72e-09
truss8	25	39	1.16e-08

Table 3: BNT algorithm

6.3 Analysis of the results

To help the analysis of the previous tables, for each class of problems, we will construct graphics with information about the time and the number of iterations. As the running times of the problems in *hinf*, *inf* and *truss* classes are very small, we do not present the corresponding graphics with the time information. Furthermore, when, for a certain class, the amplitude of the interval of time is too big we will present two graphics to facilitate the reading.

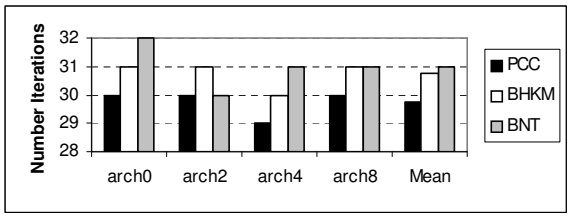


Figure 1: Number of iterations for the *arch* class.

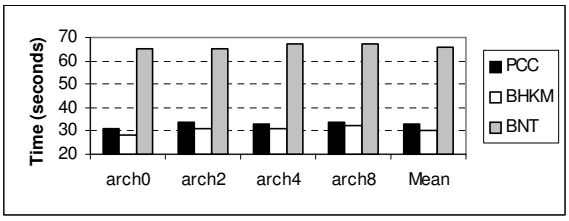


Figure 2: Time, in seconds, for the *arch* class.

As we can observe in Figure 1, for the *arch* class the algorithms that performed less and more iterations were PCC and BNT, respectively. From Figure 2 we can conclude that BHKM was the fastest while BNT was the most time consuming.

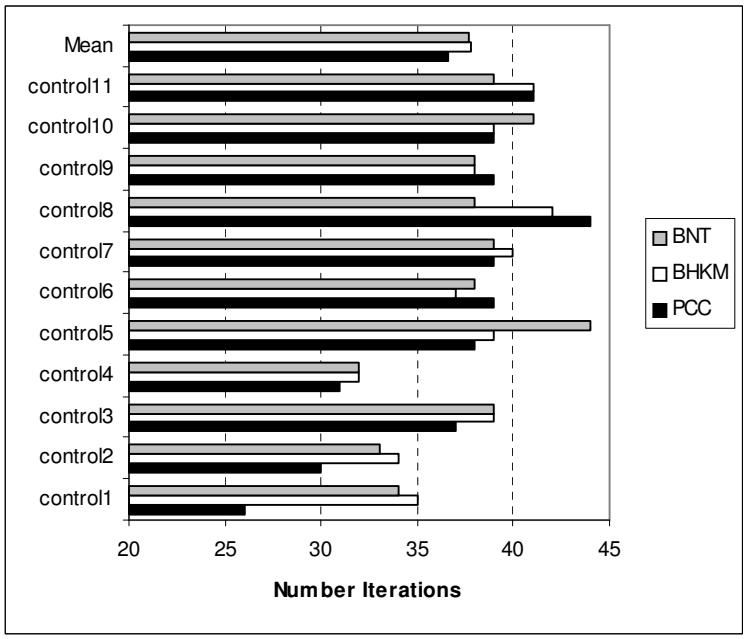


Figure 3: Number of iterations for the *control* class.

Consider now the *control* class. In Figure 3 is possible to observe that PCC was the algorithm with less iterations while the other two presented similar means for the number of iterations. For the time, Figures 4 and 5, all the algorithms presented identical means.

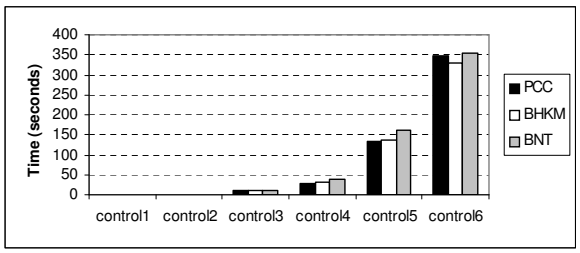


Figure 4: Time, in seconds, for the problems *control1* to *control6*.

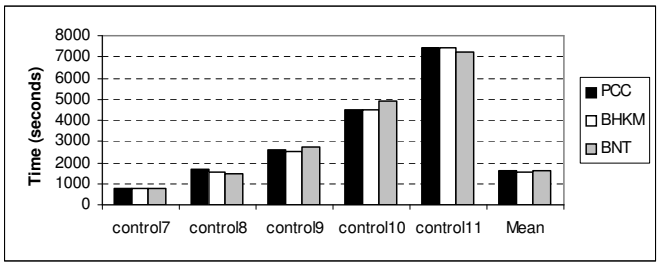


Figure 5: Time, in seconds, for the problems *control7* to *control11*.

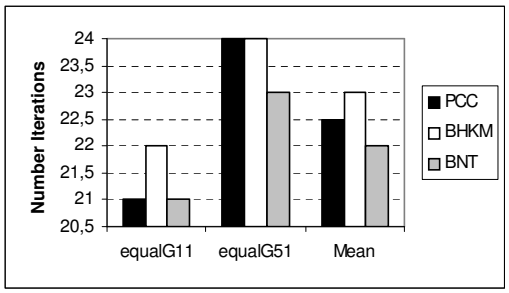


Figure 6: Number of iterations for the *equalG* class.

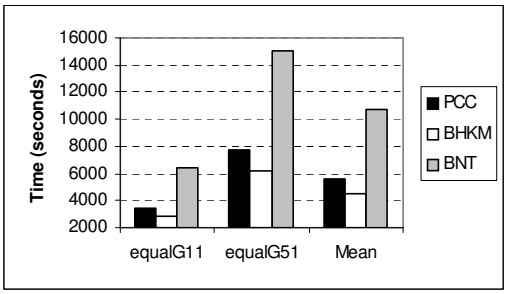


Figure 7: Time, in seconds, for the *equalG* class.

For the *equalG* class, the algorithm with less iterations was BNT while the one with more was BHKM, Figure 6. These positions were inverted for the time, Figure 7.

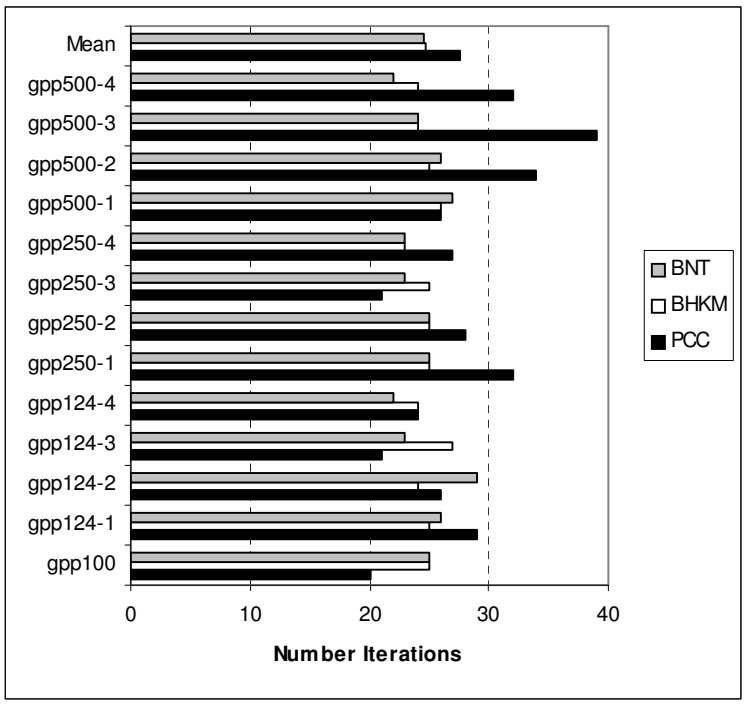


Figure 8: Number of iterations for the *gpp* class.

Relatively to the *gpp* class, we can observe in Figure 8 that the algorithms with less iterations were BHKM and BNT, being the first one the fastest and the second the slowest, Figures 9 and 10.

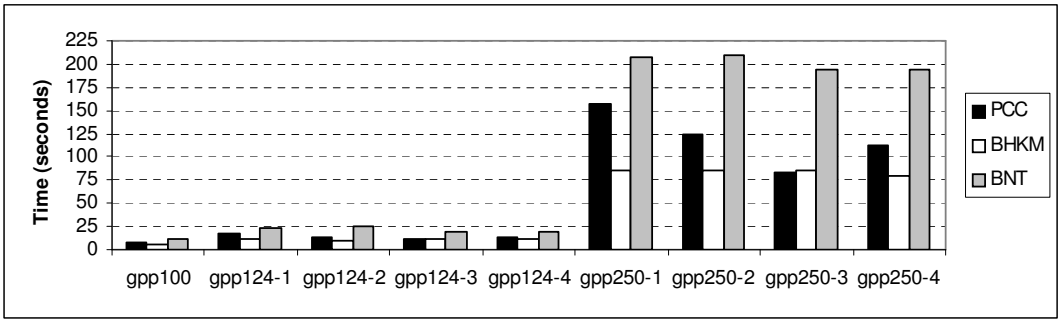


Figure 9: Time, in seconds, for the problems *gpp100* to *gpp250-4*.

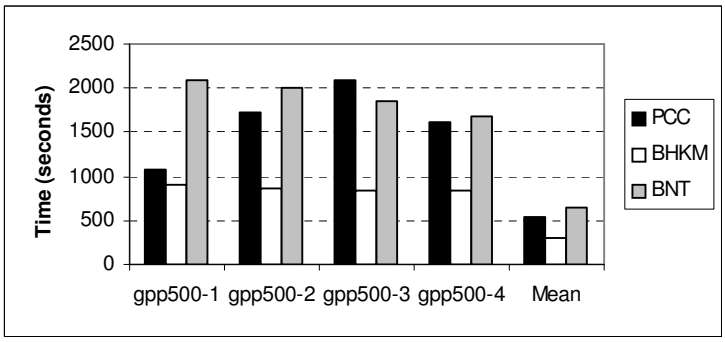


Figure 10: Time, in seconds, for the problems *gpp500*.

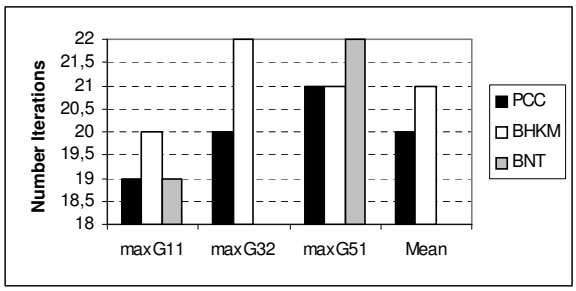


Figure 11: Number of iterations for the *maxG* class.

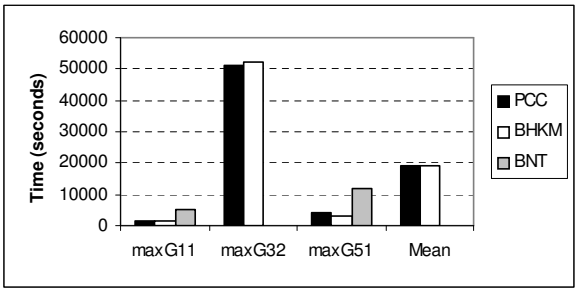


Figure 12: Time, in seconds, for the *maxG* class.

Consider now the *maxG* class. Due to memory problems, it was not possible to obtain any results for *maxG32* with BNT. But, in general, this algorithm was the worst one for the remaining problems. PCC was the algorithm that presented less iterations, Figure 11. PCC and BHKM spent similar amounts of time, Figure 12.

For the *mcp* class, in Figure 13 is possible to see that PCC was the algorithm with less iterations while BHKM was the one with more. But, this last one was the fastest, being BNT the most expensive, Figures 14 and 15.

For the *qap* class, the best and the worst algorithms were PCC and BHKM, respectively, Figures 16, 17 and 18.

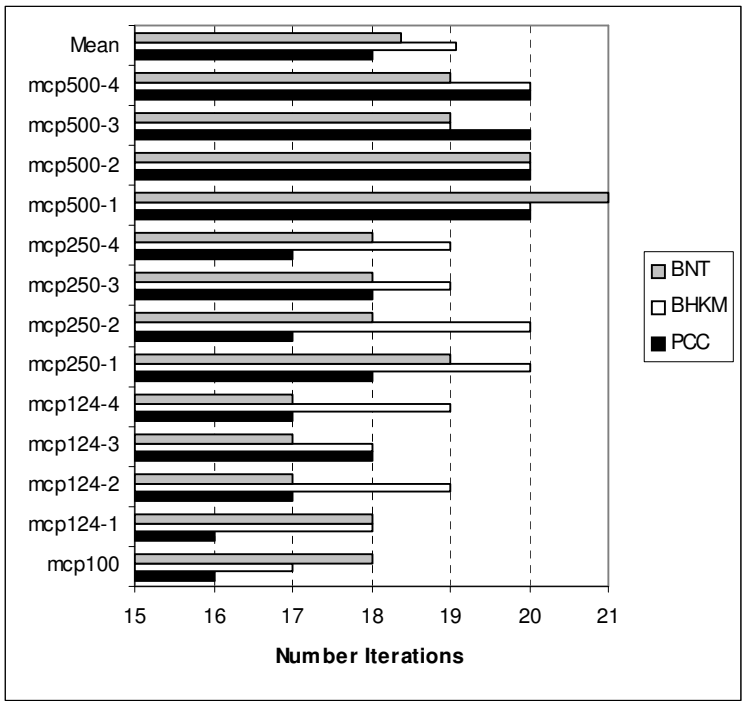


Figure 13: Number of iterations for the *mcp* class.

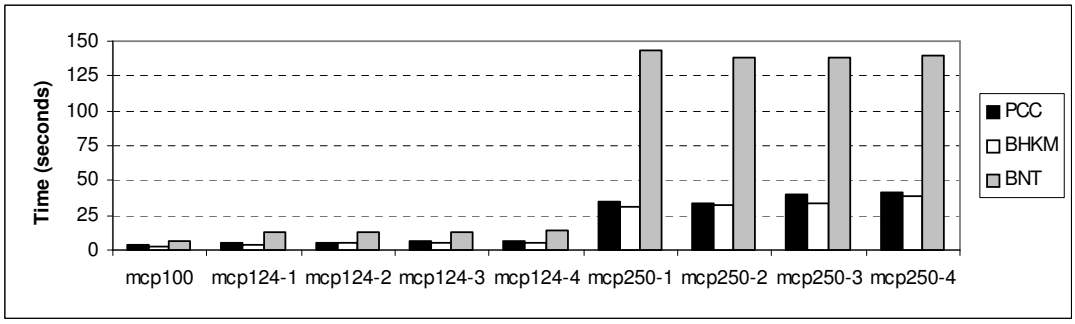


Figure 14: Time, in seconds, for the problems *mcp100* to *mcp250-4*.

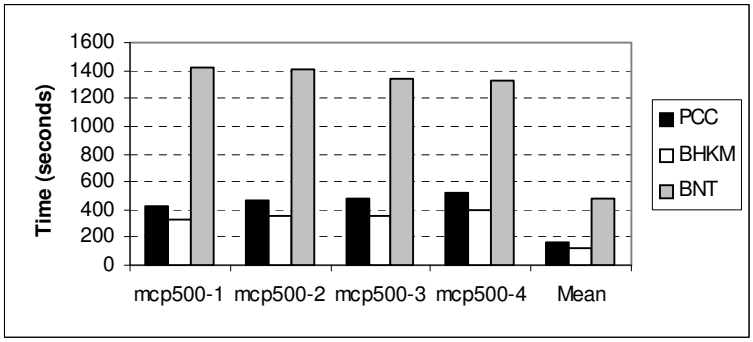


Figure 15: Time, in seconds, for the problems *mcp500*.

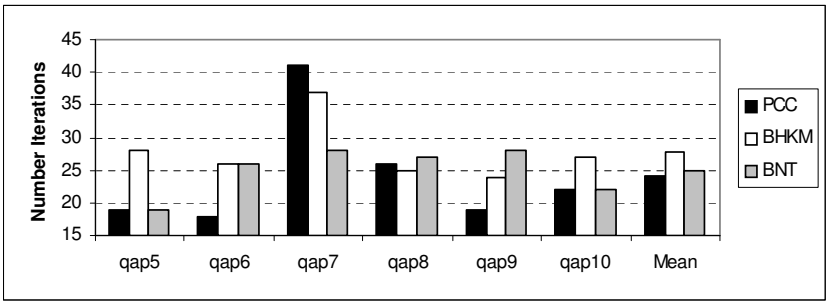


Figure 16: Number of iterations for the *qap* class.

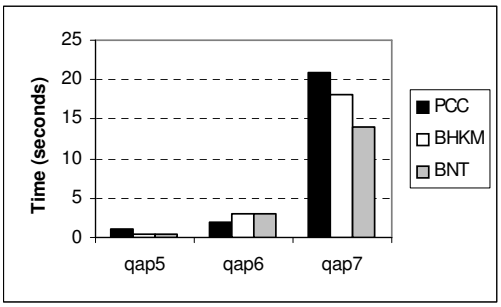


Figure 17: Time, in seconds, for the problems *qap5* to *qap7*.

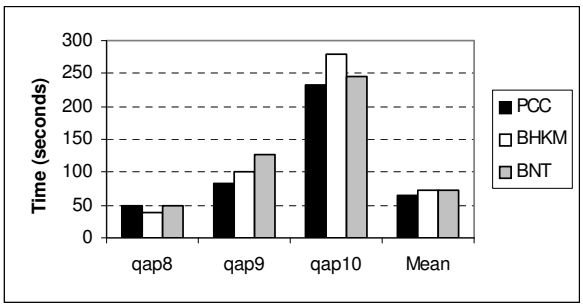


Figure 18: Time, in seconds, for the problems *qap8* to *qap10*.

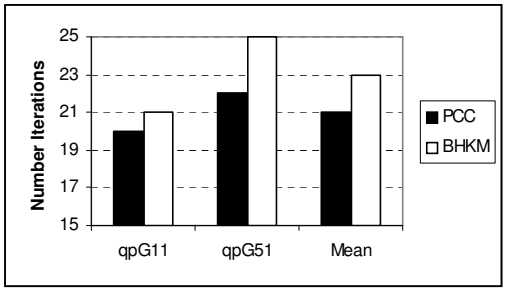


Figure 19: Number of iterations for the *qpG* class.

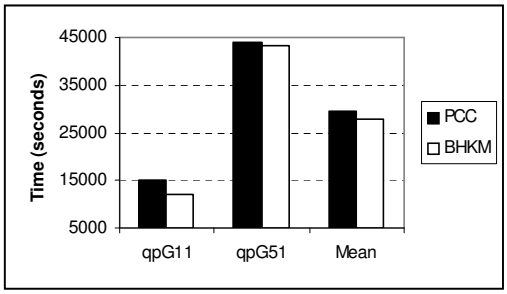


Figure 20: Time, in seconds, for the *qpG* class.

Consider, now, the *qpG* class. Due to memory problems, it was not possible to obtain any results with BNT. Once more, BHKM was the algorithm with more iterations, Figure 19, but also the fastest, Figure 20.

From Figures 21 and 22 is possible to conclude that, for the *ss30* problem, BHKM and BNT were the algorithms that presented less iterations being the first one the fastest and the second one the slowest.

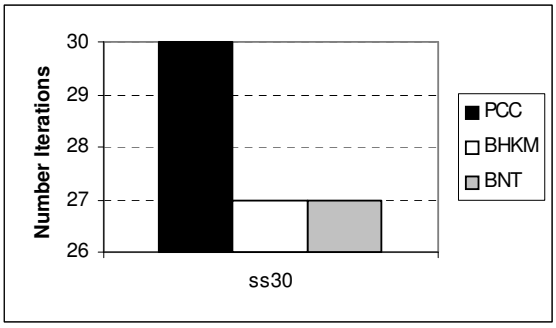


Figure 21: Number of iterations for the *ss30* problem.

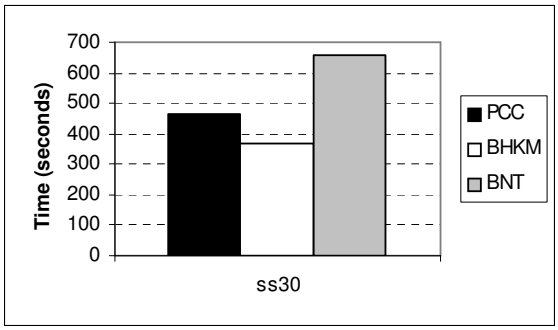


Figure 22: Time, in seconds, for the *ss30* problem.

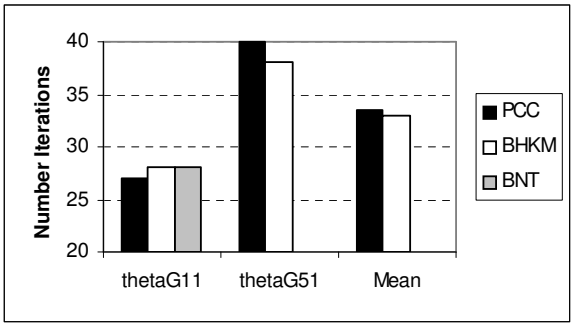


Figure 23: Number of iterations for the *thetaG* class.

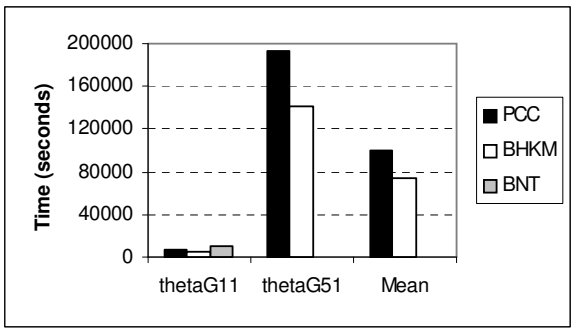


Figure 24: Time, in seconds, for the *thetaG* class.

As it happened before, due to memory problems, it was not possible to obtain any results for *thetaG51* with BNT. For *thetaG11* this algorithm was the most expensive. As we can see in Figures 23 and 24, the BHKM algorithm was the best one for the *thetaG* class.

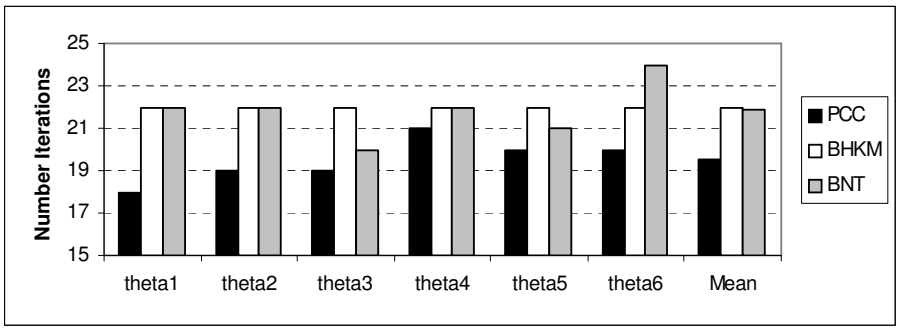


Figure 25: Number of iterations for the *theta* class.

From Figures 25, 26 and 27 we can conclude that PCC was the best algorithm for the *theta* class, while BHKM was the one with more iterations and BNT was the slowest.

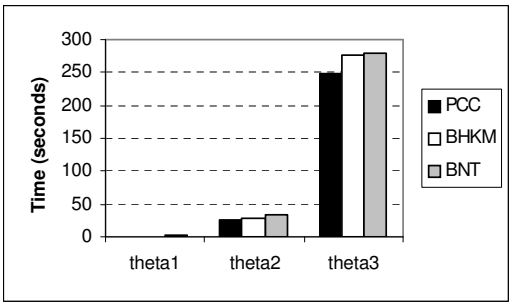


Figure 26: Time, in seconds, for the problems *theta1* to *theta3*.

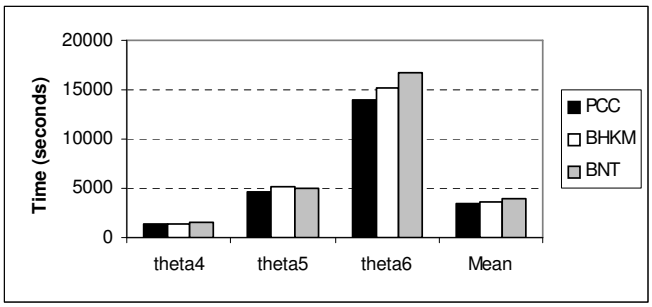


Figure 27: Time, in seconds, for the problems *theta4* to *theta6*.

It remains to analyse the results relating to *truss*, *hinf* and *inf* classes. As these problems are of quick resolution we decided to study only the results about the number of iterations.

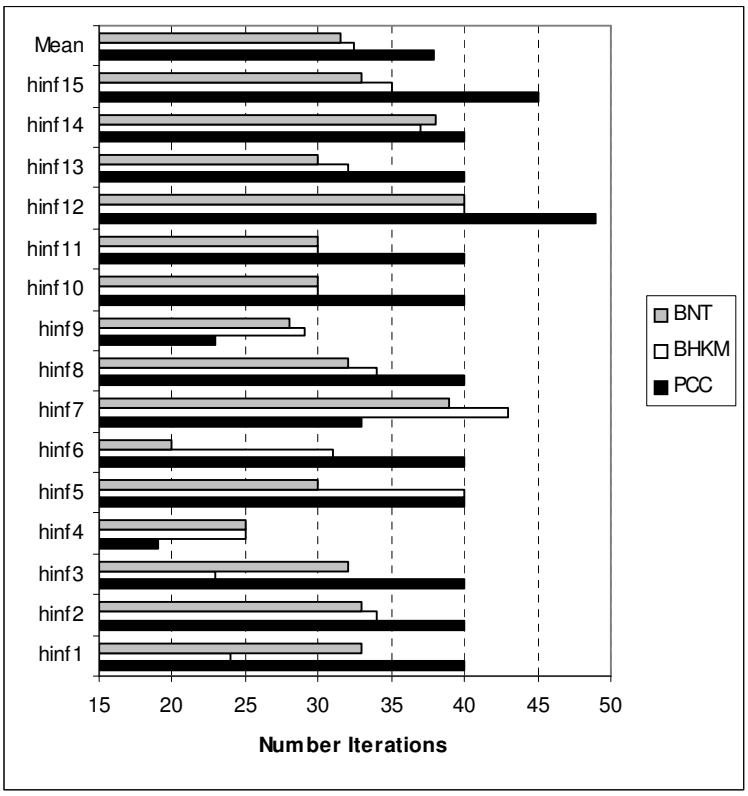


Figure 28: Number of iterations for the *hinf* class.

As we can see in Figure 28, for the *hinf* class the algorithm that presented less iterations was BNT while the one with more was PCC.

For the *truss* class, Figure 29, the algorithm with less iterations was PCC and the one with more was BHKM.

For the *infd* problems, Figure 30, BHKM was the algorithm that took less iterations while BNT was the one with more. Finally, for the *infp* problems, Figure 31, BNT was the best

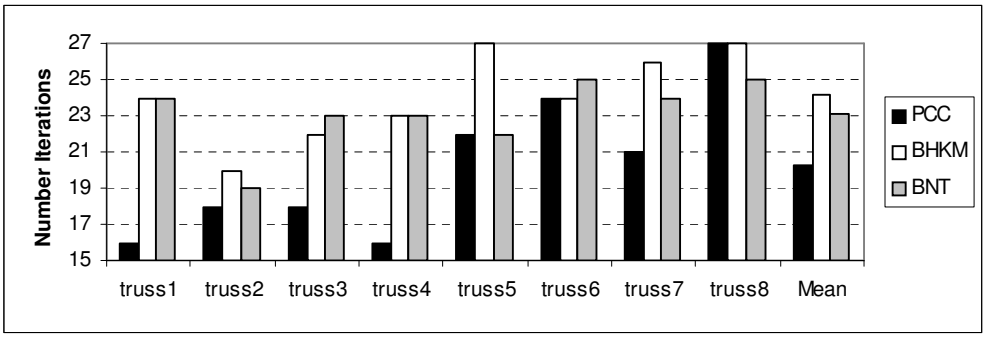


Figure 29: Number of iterations for the *truss* class.

algorithm while PCC was the worst.

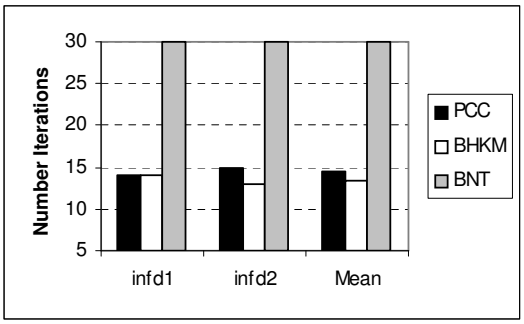


Figure 30: Number of iterations for the *infd* problems.

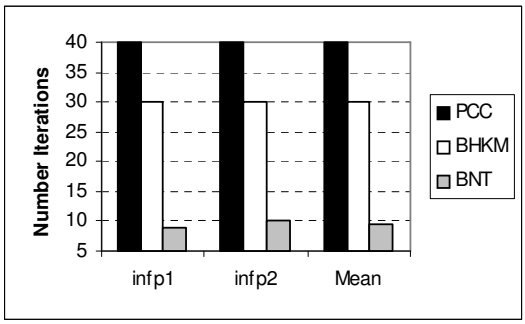


Figure 31: Number of iterations for the *infp* problems.

To summarize we can say that:

- although BHKM was the algorithm that, in general, presented more iterations it was the fastest one;
- PCC was the algorithm that presented, for most of the cases, less iterations;
- BNT was the algorithm that spent more time.

For the classes of problems with results that satisfy all the termination criteria, the efficiency of the algorithms can be measured just in terms of the two previously analysed characteristics. For the other classes we considered that the efficiency should be measured in terms of how far the solution is from satisfying all those criteria. The *control* is one of these classes. For this class, PCC was the algorithm that performed better while BHKM and BNT presented similar behaviour. For the *gpp* class the only algorithm that stopped without satisfying all the termination criteria was PCC. The BHKM was the best algorithm for *hinf* and *qap* classes while PCC was the worst. As *infd* and *infp* are infeasible problems and problems with an infeasible dual, respectively, the values in RDgap do not mean, just by themselves, that the results were not satisfactory. All the algorithms solved the *infd* problems but only BNT solved the *infp* ones. Finally, for *thetaG51* both PCC and BHKM presented similar results while, as already mentioned, due to memory problems it was not possible to obtain any results with BNT.

7 Final conclusions

From the obtained results we can conclude that, in general, the algorithm that uses the NT direction was the slowest and, due to memory problems, did not allow to obtain results for the problems with bigger dimensions. It was only better than the remaining algorithms for the *infp* problems. Although, for most of the problems, the BHKM algorithm needed more iterations than the others it was, in general, the fastest. Relatively to the classes of problems for which at least one of the algorithms ends without satisfying all the termination criteria, although PCC was the best for *control* class the BHKM was superior for *gpp*, *hinf* and *qap* classes, having both presented similar behaviour for the *thetaG51* problem. Finally, we can conclude that the version of our predictor-corrector variant with HKM direction is better than the classical predictor-corrector for the generality of the problems.

References

[1] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM Journal Optimization, 5(1) (1995), 13-51.

[2] F. J. Bastos, *Problemas de transporte e métodos de ponto interior*, Dissertação de Doutoramento, Universidade Nova de Lisboa, Lisboa, 1994.

[3] R. Bellman, K. Fan, *On systems of linear inequalities in Hermitian matrix variables*, V. L. Klee (editor), Convexity, Proceedings of Symposia in Pure Mathematics, American Mathematical Society, 7 (1963), 1-11.

[4] B. Borchers, *CSDP, a C library for semidefinite programming*, Optimization Methods and Software, 11(1) (1999), 613-623.

[5] B. Borchers, *SDPLIB 1.2, a library for semidefinite programming test problems*, Optimization Methods and Software, 11(1) (1999), 683-690.

[6] B. Borchers, *CSDP 3.2, User's Guide*, Technical Report, Department of Mathematics, New Mexico Tech, USA, 2000.

[7] B. Borchers, *Implementation issues in CSDP*, Technical Report, Department of Mathematics, New Mexico Tech, USA, 2001.

[8] C. Helmberg, F. Rendl, R. J. Vanderbei, H. Wolkowicz, *An interior-point method for semidefinite programming*, SIAM Journal Optimization, 6(2) (1996), 342-361.

[9] R. Horn, C. Johnson, *Matrix Analysis* , Cambridge University Press, Cambridge, 1991.

[10] R. Horn, C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[11] M. Kojima, S. Shindoh, S. Hara, *Interior-point methods for the monotone semidefinite linear complementary problem in symmetric matrices*, SIAM Journal Optimization, 7(1) (1997), 86-125.

[12] L. Lovász, *On the Shannon Capacity of a Graph*, IEEE Transactions of Information Theory, IT-25(1) (1979), 1-7.

[13] R. D. C. Monteiro, *Primal-dual path-following algorithms for semidefinite programming*, SIAM Journal Optimization, 7(3)(1997), 663-678.

[14] A. Nemirovskii, Y. Nesterov, *Interior-Point Polynomial Algorithms in Convex Programming*, Society for Industrial and Applied Mathematics, Philadelphia, 1994.

[15] Y. E. Nesterov, M. J. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Mathematics of Operations Research, 22(1) (1997), 1-42.

[16] M. J. Todd, *A study of search directions in primal-dual interior-point methods for semidefinite programming*, Optimization Methods and Software, 11 (1999), 1-46.

[17] M. J. Todd, K. C. Toh, R. H. Tutuncu, *On the Nesterov-Todd direction in semidefinite programming*, SIAM Journal Optimization, 8(3) (1998), 769-796.

Aplicação do algoritmo volumétrico à resolução aproximada e exacta do problema do caixeiro viajante assimétrico

Ana Maria Rocha † Edite M.G.P. Fernandes † João Soares *‡

† Departamento de Produção e Sistemas
Universidade do Minho
{arocha;emgpf}@dps.uminho.pt

‡ Departamento de Matemática
Universidade de Coimbra
jsoares@mat.uc.pt

Abstract

In this paper we present computational results with the volume algorithm, a variant of the subgradient method, when solving the linear relaxation that stems from the extended disaggregated flow formulation of the Asymmetric Travelling Salesman Problems. Computational experiments were performed on a selection of instances from the TSPLIB and some randomly generated instances according to the *Dimacs Implementation Challenge*. We have also tried ATSP heuristics within the volume algorithm. Computational experiments show moderated success on medium-scale instances.

Resumo

Neste artigo apresentamos resultados computacionais obtidos com o algoritmo volumétrico, uma variante do método do subgradiente, na resolução da relaxação linear que decorre da formulação estendida de fluxo desagregado para o problema do Caixeiro Viajante Assimétrico. As experiências computacionais foram realizadas numa selecção de instâncias da TSPLIB e num conjunto de instâncias geradas aleatoriamente de acordo com o *Dimacs Implementation Challenge*. Também experimentámos a aplicação de heurísticas durante a execução do algoritmo volumétrico. As experiências computacionais mostram sucesso moderado com instâncias de média dimensão.

Keywords: Asymmetric travelling salesman problem, Disaggregated flow formulation, Lagrangian relaxation.

*João Soares acknowledges partial financial support from Fundação para a Ciência e Tecnologia (Projecto POCTI/MAT/14243/1998).

Apdio © 2005 *Associação Portuguesa de Investigação Operacional*

Title: Application of the volume algorithm to the approximate and exact solving of the asymmetric travelling salesman problem.

---

1 Introdução

Seja $G = (V, A)$ um grafo orientado simples (*i.e.*, sem laços ou arcos múltiplos), com n vértices e m arcos, e um custo c_{ij} associado a cada arco (i, j). O problema do Caixeiro Viajante Assimétrico, mais adiante referido como PCVA, consiste em determinar em G um circuito Hamiltoniano (*i.e.*, um circuito que passa por todos os vértices uma única vez) cuja soma dos custos associados aos seus arcos seja a menor possível. Este problema pertence à classe dos problemas NP-completos. A referência [27] sumaria a investigação no PCVA em todos os seus aspectos até 1985, enquanto que [17] é uma referência do mesmo género mas mais actual.

Diversas formulações têm sido propostas para este problema. Em geral, mas nem sempre, elas incorporam as restrições de um problema de afectação (em número igual a $2n$ se excluirmos as de não negatividade) e as restrições que eliminam as soluções inteiras que não definem circuitos Hamiltonianos (que podem ser em número polinomial ou exponencial em n e m). Uma das primeiras formulações foi sugerida em [10], a formulação DFJ. Apesar de possuir um número exponencial de restrições, esta formulação permanece como sendo a formulação de base para os códigos *state-of-the-art*, como é o caso de FT-B&C [11]. Outros investigadores propuseram formulações que envolvem um número polinomial de variáveis e restrições (as formulações compactas), à custa da introdução de variáveis auxiliares, *e.g.*, [29, 12, 33, 8, 15, 16]. As diferentes formulações são comparadas em, por exemplo, [30, 14, 26]. Na língua portuguesa, recomendamos a dissertação de José Pires [31] onde todas estas formulações são analisadas.

Ainda que as formulações compactas possuam um número polinomial de variáveis e restrições, esse número pode ser o suficiente para dificultar a mera resolução da relaxação linear. Por exemplo, para um método simplex, a base pode ter dimensão de tal modo elevada que o método tem grandes dificuldades em aproximar-se da base óptima; para um método de pontos interiores, a dificuldade poderá ser o espaço de armazenamento em memória. Esta parece ser uma das razões pela menor utilização deste tipo de formulações em códigos *branch-and-cut*. Neste contexto, surge a questão que motivou o nosso trabalho, a técnica de relaxação Lagrangeana poderá definir uma base inicial mais próxima da base óptima e, deste modo, acelerar a execução do método simplex.

A aplicação de relaxação Lagrangeana ao problema do Caixeiro Viajante foi iniciada por Held e Karp [18, 19] para a versão simétrica. Em geral, aplica-se relaxação Lagrangeana quando um grupo de restrições *destrói* o facto de que um problema combinatório é estruturado e pode ser resolvido por métodos específicos muito eficientes (*e.g.*, caminho mais curto com restrições adicionais, árvore de suporte com restrições de grau, etc). Aplicar esta técnica significa minimizar uma função convexa não diferenciável num conjunto poliedral. Um dos métodos que mais tem sido usado neste tipo de problemas é o método do subgradiente. Cada iteração deste método consiste na resolução de um subproblema estruturado (árvore de peso mínimo, no caso da abordagem de Held e Karp) e na redefinição da respectiva função objectivo - essencialmente caracterizada pelos valores de variáveis duais relativamente às tais restrições

problemáticas. Uma variante do método de subgradiente que tem recentemente obtido razoável sucesso na resolução da relaxação linear de problemas de optimização combinatória é o Algoritmo Volumétrico (AV) [4].

Neste trabalho abordamos, essencialmente, a resolução da relaxação linear da formulação de fluxo desagregado para o PCVA, proposta em [8], através do algoritmo volumétrico. O subproblema estruturado a resolver em cada iteração é um problema de afectação. As soluções primais e duais assim obtidas são, então, usadas para definir uma base inicial para o método dual-simplex. Os custos reduzidos gerados no decurso do AV são também aproveitados para a execução de métodos heurísticos. Testámos a nossa abordagem em instâncias da TSPLIB e instâncias geradas aleatoriamente.

Este artigo está organizado da seguinte forma. A secção 2 introduz a formulação de fluxo desagregado do PCVA. A secção 3 apresenta o problema dual Lagrangeano, incluindo as simplificações que ocorrem na resolução do subproblema. A secção 4 sumaria o algoritmo volumétrico, que é usado na resolução do problema dual. A secção 5 apresenta os resultados computacionais obtidos e a secção 6 sumaria as conclusões.

2 Formulação

O conjunto dos vectores característicos de circuitos Hamiltonianos de G é um subconjunto das soluções inteiras que pertencem ao poliedro P^{A} descrito pelo seguinte sistema de igualdades e desigualdades

$$\left.\begin{array}{rll} \sum\limits_{(i,j)\in\delta^+(i)} x_{ij} & = 1 & (i \in V) \\ \sum\limits_{(i,j)\in\delta^-(j)} x_{ij} & = 1 & (j \in V) \\ x_{ij} & \geq 0 & ((i,j) \in A) \end{array}\right\} \tag{1}$$

onde $\delta^-(j)$ denota o conjunto dos arcos que convergem para o vértice j e $\delta^+(i)$ denota o conjunto dos arcos que divergem do vértice i. O conjunto dos vectores característicos de circuitos Hamiltonianos de G é o subconjunto daquelas soluções inteiras x que satisfazem, para alguns vectores $y^2, y^3, \ldots, y^n \in \mathbb{R}^m$, o seguinte sistema de igualdades e desigualdades

$$\left.\begin{array}{ll} \sum\limits_{(1,j)\in\delta^+(1)} y^k_{1j} - \sum\limits_{(j,1)\in\delta^-(1)} y^k_{j1} = -1 & (k \in V \setminus \{1\}) \\ \sum\limits_{(k,j)\in\delta^+(k)} y^k_{kj} - \sum\limits_{(j,k)\in\delta^-(k)} y^k_{jk} = 1 & (k \in V \setminus \{1\}) \\ \sum\limits_{(i,j)\in\delta^+(i)} y^k_{ij} - \sum\limits_{(j,i)\in\delta^-(i)} y^k_{ji} = 0 & (k \in V \setminus \{1\}, i \in V \setminus \{1,k\}) \\ 0 \leq y^k_{ij} \leq x_{ij} & ((i,j) \in A, k \in V \setminus \{1\}). \end{array}\right\} \tag{2}$$

Esta formulação proposta por Claus [8], e inspirada em Wong [33], é conhecida por formulação estendida de fluxo desagregado. A designação "fluxo" deve-se ao facto de se usar variáveis de fluxo y^k_{ij} $((i,j) \in A, k \in V \setminus \{1\})$ que se podem interpretar como indicando a quantidade de fluxo percorrendo o arco (i,j), enviada pelo vértice 1 e com destino ao vértice k. A designação "desagregado" deve-se ao facto de haver distinção por vértice.

A projecção do poliedro definido pelos sistemas de igualdades e desigualdades (1) e (2) no espaço das variáveis $x_{ij}, (i,j) \in A$, define um poliedro denotado P_L^{PCVA} que é definido por (1) e

$$\sum_{(i,j)\in A(S)} x_{ij} \leq |S| - 1 \quad (S \subseteq V \setminus \{1\}) \tag{3}$$

onde $A(S)$ denota o conjunto dos arcos do grafo induzido $G[S]$. O interesse em usar (1) e (2) em vez de (1) e (3) tem naturalmente a ver com o número de restrições que é, respectivamente, polinomial num caso e exponencial no outro.

3 Problema Lagrangeano

O problema de optimização linear $z_L^* \equiv \min\{cx \colon x \in P_L^{\text{PCVA}}\}$ pode descrever-se do seguinte modo usando o formato matricial, que aproveitamos para exibir o dual,

$$z_L^* = \left\{ \begin{array}{lllll} \min & cx & & & \\ \text{s.a} & Dx & = & \mathbb{1} & \\ & By^k & = & b^k & (k \in V_1) \\ & x - y^k & \geq & 0 & (k \in V_1) \\ & y^k & \geq & 0 & (k \in V_1) \\ & x \geq 0 & & & \end{array} \right\} = \left\{ \begin{array}{lllll} \max & w\mathbb{1} + \sum_{k\in V_1} \pi^k b^k & & & \\ \text{s.a} & wD + \sum_{k\in V_1} \rho^k & \leq & c & \\ & \pi^k B - \rho^k & \leq & 0 & (k \in V_1) \\ & \rho^k & \geq & 0 & (k \in V_1) \end{array} \right\} \tag{4}$$

onde $V_1 \equiv V \setminus \{1\}$, D denota a matriz de incidência vértice-aresta do grafo bipartido não orientado $G' = (V \times V, A)$, $\mathbb{1}$ é um vector de uns, B denota a matriz de incidência vértice-arco do grafo G e, para cada $k \in V_1$, b^k é um vector cujas componentes 1 e k são iguais, respectivamente, a -1 e 1 e as restantes são iguais a 0. Procurando explorar a estrutura do problema de minimização em (4), consideremos o problema Lagrangeano

$$z_L^* = \max\left\{ z(\pi) : \pi = \left[\pi^k\right] \in \mathbb{R}^{(|V|-1)\times|V|} \right\} \tag{5}$$

cuja função objectivo $z \colon \mathbb{R}^{(|V|-1)\times|V|} \to \mathbb{R}$ é definida em cada π por

$$z(\pi) \equiv \left\{ \begin{array}{ll} \min & cx + \sum_{k\in V_1} \pi^k \left(b^k - By^k\right) \\ \text{s.a} & Dx = \mathbb{1} \\ & x - y^k \geq 0 \quad (k \in V_1) \\ & y^k \geq 0 \quad (k \in V_1) \\ & x \geq 0 \end{array} \right\} = \left\{ \begin{array}{ll} \max & w\mathbb{1} + \sum_{k\in V_1} \pi^k b^k \\ \text{s.a} & wD + \sum_{k\in V_1} \rho^k \leq c \\ & -\rho^k \leq -\pi^k B \quad (k \in V_1) \\ & \rho^k \geq 0 \quad (k \in V_1) \end{array} \right\}. \tag{6}$$

Uma solução óptima do problema de minimização em (6) pode ser obtida através da resolução de um adequado problema de afectação. Primeiro, observe-se que para todo x fixo, uma solução óptima $\bar{y}(x) = [\bar{y}^k(x)]$ no espaço das variáveis y é solução óptima de

$$\begin{array}{ll} \max & \displaystyle\sum_{k\in V_1} \sum_{(i,j)\in A} \left(\pi_i^k - \pi_j^k\right) y_{ij}^k \\ \text{s.a} & 0 \leq y_{ij}^k \leq x_{ij} \ ((i,j) \in A, k \in V_1) \end{array} \tag{7}$$

que admite o valor óptimo $\sum_{k\in V_1} \sum_{(i,j)\in A} \max\left(\pi_i^k - \pi_j^k, 0\right) x_{ij}$ para uma solução óptima explícita

$$\bar{y}_{ij}^k(x) = \left\{ \begin{array}{ll} x_{ij} & \text{se } \pi_i^k - \pi_j^k \geq 0, \\ 0 & \text{se } \pi_i^k - \pi_j^k < 0. \end{array} \right\} \quad ((i,j) \in A, k \in V_1)\,. \tag{8}$$

Por isso, o problema primal em (6) é equivalente ao problema de afectação:

$$\begin{array}{ll} \min & \bar{c}x \\ \text{s.a} & Dx = \mathbb{1}, x \geq 0 \end{array} \tag{9}$$

para um vector $\bar{c}$ definido componente a componente por $\bar{c}_{ij} = c_{ij} - \sum_{k \in V_1} \max\left((\pi_i^k - \pi_j^k), 0\right)$, na medida em que, se $\bar{x}$ é uma solução óptima de (9) então $(\bar{x}, \bar{y})$, com $\bar{y} = \bar{y}(\bar{x})$ definido em (8), é solução óptima para (6). A resolução de (9) pode efectuar-se através de, por exemplo, o método Húngaro em $\mathcal{O}(n^3)$ operações. Mais, se $\bar{x}$ é uma solução óptima de (9) então

$$z(\pi) = \bar{c}\bar{x} + \sum_{k \in V_1} \left(\pi_1^k - \pi_k^k\right),$$

e uma solução óptima $(\bar{w}, \bar{\rho})$ do dual em (6) é caracterizada pela solução dual $\bar{w}$ óptima para o problema de afectação (9) e $\bar{\rho}_{ij}^k = \max(\pi_i^k - \pi_j^k, 0)$, para todo $k \in V_1$ e $(i,j) \in A$. Note-se que o vector $(\pi, \bar{w}, \bar{\rho})$ é dual admissível em (4) pois, para cada $(i,j) \in A$,

$$\left[c - \left(\bar{w}D + \sum_{k \in V_1} \bar{\rho}^k\right)\right]_{ij} = c_{ij} - \left(\bar{u}_i + \bar{v}_j + \sum_{k \in V_1} \max(\pi_i^k - \pi_j^k, 0)\right) = \bar{c}_{ij} - (\bar{u}_i + \bar{v}_j) \geq 0,$$

e ainda porque, para cada $(i,j) \in A$ e $k \in V_1$,

$$\left[0 - \left(\pi^k B - \bar{\rho}^k\right)\right]_{ij} = -\left(\pi_i^k - \pi_j^k - \max(\pi_i^k - \pi_j^k, 0)\right) = \max(\pi_j^k - \pi_i^k, 0) \geq 0.$$

A próxima secção mostra como o algoritmo volumétrico é usado para resolver (5). Tal como foi mencionado anteriormente, foi dada especial atenção à forma como foi usada a relaxação Lagrangeana e às técnicas do subgradiente de tal forma que os subproblemas Lagrangeanos a resolver são reoptimizados de forma muito rápida.

4 Algoritmo Volumétrico (AV)

O algoritmo volumétrico, proposto em [4], é uma variante do algoritmo do subgradiente tradicional que, com comparável esforço computacional, produz não só, boas aproximações para as variáveis duais como também para as variáveis primais. A convergência global de uma implementação específica desse algoritmo é analisada em [3]. Experiências computacionais com a aplicação do algoritmo volumétrico na resolução da relaxação linear de problemas de optimização combinatória incluem o problema de partição [4, 1, 5], o problema da árvore de "Steiner" [2] e o problema de localização [6]. A resolução de problemas de optimização combinatória por *branch-and-cut*, onde os modelos lineares são resolvidos pelo algoritmo volumétrico, incluem os problemas da árvore de "Steiner" e de corte máximo [7].

Uma adaptação do algoritmo volumétrico para o PCVA é descrita na Figura 1. O método possui algumas semelhanças com os métodos de feixe [3]. Em particular, usa a ideia de, em cada iteração, usar uma direcção de deslocamento que combina uma nova direcção com a anterior direcção de deslocamento. Deste modo, procura-se evitar o comportamento de ziguezague, típico do algoritmo do subgradiente, de modo a acelerar a convergência.

Entrada: $\pi^0 \in \mathbb{R}^{(|V|-1)\times|V|}$.

Inicialização: Resolver (6) com $\pi = \pi^0$ para obter a solução (x^0, y^0).
Definir $v^0 = b - By^0 \in \partial z(\pi^0)$.
Iniciar $\bar{\pi}^1 = \pi^0$, $(\bar{x}, \bar{y}) = (x^0, y^0)$, $\bar{w} = v^0$, $j = 1$ e $l = 1$.

Iteração Genérica j:

Passo 1: Para algum comprimento do passo $s_j > 0$, definir $\pi^j = \bar{\pi}^l + s_j\bar{w}$.
Passo 2: Obter (x^j, y^j) uma solução óptima de (6).
Definir $v^j = [v^{k,j}] \in \partial z(\pi^j)$ com $v^{k,j} = b^k - By^{k,j}$.
Passo 3: Para algum $\alpha_j \in [0, 1]$, definir

$$(\bar{x}, \bar{y}) = \alpha_j \left(x^j, y^j\right) + (1 - \alpha_j)\,(\bar{x}, \bar{y}) \tag{10}$$

$$\bar{w} = \alpha_j v^j + (1 - \alpha_j)\,\bar{w}. \tag{11}$$

Passo 4: Se $z(\pi^j) > z(\bar{\pi}^l)$, então definir $\bar{\pi}^{l+1} = \pi^j$ e fazer $l \leftarrow l + 1$.
Passo 5: Testar critério de paragem. Fazer $j \leftarrow j + 1$ e voltar para *Passo 1.*

Figura 1: Algoritmo Volumétrico adaptado para o PCVA.

A descrição do algoritmo volumétrico deixa em aberto, em cada iteração j, a definição do tamanho do passo s_j e o escalar α_j envolvido na redefinição da solução primal $\bar{x}$. As nossas escolhas para definição destes dois escalares são essencialmente empíricas e semelhantes às usadas em [4]. O escalar α_j é escolhido por forma a forçar a primal admissibilidade de $(\bar{x}, \bar{y})$. Conforme [4], primeiro define-se α_{opt} como

$$\alpha_{\text{opt}} \equiv \arg\min \left\|\alpha v^j + (1 - \alpha)\,\bar{w}\right\|^2 = \arg\min \sum_{k\in K} \left\|b^k - B\left(\alpha y^{k,j} + (1 - \alpha)\,\bar{y}^k\right)\right\|^2. \tag{12}$$

Depois, é atribuído um valor a $\alpha_{\max}$ (configurado pelo parâmetro `alphainit=0.01`) e define-se

$$\alpha_j = \begin{cases} \alpha_{\max}/10 & \text{se } \alpha_{\text{opt}} \leq 0, \\ \min(\alpha_{\text{opt}}, \alpha_{\max}) & \text{se } \alpha_{\text{opt}} > 0, \end{cases}$$

onde $\alpha_{\max}$ é um limite superior para o valor de α.

Após um certo número de iterações (configurado pelo parâmetro `alphaint=80`), é testado o progresso da função dual $z(\pi)$. Caso tenha aumentado menos de 1% e se $\alpha_{\max}$ for superior a um certo valor (configurado pelo parâmetro `alphamin=0.0001`) então $\alpha_{\max}$ é multiplicado por uma dada quantidade (configurado pelo parâmetro `alphafactor=0.5`).

Tal como no método do subgradiente [20], o comprimento do passo s_j é feito através da seguinte fórmula:

$$s_j = \lambda_j \frac{T - z(\bar{\pi}^l)}{\|\bar{w}\|^2} \tag{13}$$

com $\lambda_j \in (0, 2]$ e T é um valor estimado para o valor óptimo de (5). Inicialmente é usado um valor pequeno para T (por exemplo, $T = z(\pi^0)$). Cada vez que $z(\bar{\pi}^l)$ está a 5% de T, o valor de T é aumentado em 5%.

A determinação do valor de λ, em cada iteração, obedece aos seguintes passos. O valor inicial de λ é 0.1 (`lambdainit=0.1`), depois o valor de λ difere consoante a iteração é *red*, *yellow* ou *green.*

- Quando $z(\pi^j) \leq z(\bar{\pi}^l)$, não se verifica um melhoramento no valor dual e a iteração diz-se *red.* Uma sequência de **40** (`redtestinvl=40`) iterações *red* consecutivas sugere a necessidade de um comprimento do passo menor. Assim, o valor de λ é multiplicado por 0.67, se $\lambda \geq 0.0005$, e é mantido constante, caso contrário.
- Quando $z(\pi^j) > z(\bar{\pi}^l)$, obteve-se um melhoramento no valor dual e efectua-se o seguinte cálculo
$$d = \bar{w}\,(b - B\bar{y})\,.$$
  - Se $d < 0$, a iteração chama-se de *yellow*, e um passo mais longo na direcção $\bar{w}$ origina um valor menor para $z(\pi^j)$. Assim, após **2** (`yellowtestinvl=2`) iterações *yellow* consecutivas λ é multiplicado por 1.1.
  - Se $d \geq 0$, a iteração diz-se *green*, e é sugerido um passo maior. Nestas condições, após **2** (`greentestinvl=2`) iterações *green* consecutivas λ é multiplicado por 2. Se o resultado for superior a 2 passa a ter o valor de 2.

é necessário que uma das seguintes duas condições seja verificada para fazer a paragem do algoritmo:

1. Uma delas baseia-se na verificação do limite máximo de iterações
$$\text{n}^o \text{ máximo de iterações} \leq 500/1000/5000/10000 \ (\texttt{maxsgriters})$$
2. A outra é composta pela conjunção de duas condições
   (a) o valor máximo absoluto da violação das restrições
$$\|\bar{w}\| \leq 0.01 \ (\texttt{primal_abs_precision})$$
   e
   (b) a diferença relativa ou absoluta entre o limite inferior $z(\bar{\pi}^l)$ e o valor da função objectivo primal $c\bar{x}$
$$\begin{aligned}(\frac{\left|c\bar{x} - z(\bar{\pi}^l)\right|}{|z(\bar{\pi}^l)|}) &\leq 0.01 \ (\texttt{gap_rel_precision}) \quad \text{ou} \quad \left|c\bar{x} - z(\bar{\pi}^l)\right| \leq \\ &\leq 0.05 \ (\texttt{gap_abs_precision}).\end{aligned}$$

5 Resultados computacionais

Nesta secção descrevemos as experiências computacionais que efectuámos. Todos os códigos foram implementados em C++ e todas as experiências foram efectuadas num computador com processador Pentium 4 a 2.66Ghz com 512 Mb de memória RAM.

A Tabela 1 contém as principais características das instâncias seleccionadas da TSPLIB[32]. As primeiras colunas da tabela identificam o problema e as últimas colunas sumariam o comportamento do método dual-simplex (CPLEX DUALOPT), do método dos pontos interiores (CPLEX BAROPT) e do método *branch-and-bound* (CPLEX MIP) que decorre da execução do CPLEX 7.0 com base na formulação de fluxo desagregado. A execução de CPLEX MIP foi interrompida ao fim de 8 horas, sendo nesse caso apresentado '*' na respectiva entrada da Tabela 1. Nas colunas associadas à descrição do problema são listados o nome da instância, o número de vértices e o número de arcos, seguido do valor óptimo inteiro e do valor óptimo fraccionário (*i.e.*, o valor óptimo da relaxação linear). Nas restantes colunas listamos o número total de iterações do CPLEX e o tempo gasto em segundos.

Tabela 1: Instâncias do PCVA e soluções obtidas com o CPLEX.

PCVA					**CPLEX (dualopt)**		**CPLEX (baropt)**	**CPLEX MIP**	
ID	$\|V\|$	$\|A\|$	Int. óptimo	Frac. óptimo	Simplex It.	Tempo (seg.)	Tempo (seg.)	Simplex It.	Tempo (seg.)
br17	17	272	39	39.0	3055	2.2	0.4	27349	162.2
ftv33	34	1122	1286	1286.0	24231	42.7	18.2	24226	56.8
ftv35	36	1260	1473	1457.3	31627	76.2	23.5	288329	2094.9
ftv38	39	1482	1530	1514.3	41240	121.9	32.2	2403409	25536.7
ftv44	45	1980	1613	1584.9	68004	278.9	59.1	*	*
ftv47	48	2256	1776	1748.6	100650	627.1	106.5	*	*
ftv55	56	3080	1608	1584.0	137580	1035.9	237.7	*	*
ftv64	65	4160	1839	1807.5	245467	2762.7	511.2	*	*
ftv70	71	4970	1950	1909.0	316196	3719.2	791.0	*	*
p43	43	1806	5620	5611.0	131426	2117.6	60.9	*	*
ry48p	48	2256	14422	14289.3	98816	600.2	107.0	*	*
ft53	53	2756	6905	6905.0	127702	827.1	175.2	*	*
ft70	70	4830	38673	38652.5	291086	3657.6	753.4	*	*

Nas instâncias maiores, como é o caso da instância ftv70, o método dual-simplex é muito mais lento do que o método de pontos interiores. O comportamento superior do método de pontos interiores nestas instâncias era esperado porque o número de variáveis e restrições é muito elevado. Destacamos o comportamento do método de pontos interiores com a instância p43, requerendo apenas 2.9% do tempo despendido pelo método dual-simplex.

Motivados pela dificuldade do método dual-simplex em identificar a base óptima para estas instâncias ocorreu-nos a ideia de usar a solução proveniente do algoritmo volumétrico para definir uma base inicial mais próxima da base óptima. Também, averiguaremos a adequação do AV como mecanismo de identificação de arcos que não participam na solução óptima inteira e aproveitaremos os custos reduzidos gerados no decurso do AV para obter soluções inteiras pela via heurística.

5.1 Resolução do modelo relaxado

Começamos por analisar a estratégia de usar o par de soluções primal-dual encontrada pelo algoritmo volumétrico para definir a solução inicial para o método dual-simplex. Usámos a

implementação do algoritmo volumétrico disponível em `http://www.coin-or.org`, a página-web do projecto COIN-OR. Notámos que, à semelhança do algoritmo de subgradiente clássico, também o algoritmo volumétrico tem um comportamento muito diferenciado para diferentes valores dos seus parâmetros. Os valores para os parâmetros que utilizámos foram já descritos na Secção 4. Em cada iteração do algoritmo volumétrico, o problema de afectação (9) é resolvido pelo método dual-simplex. Experiências computacionais preliminares permitiram concluir que essa estratégia é superior, em tempo de execução, à invocação do método primal-simplex para redes.

A Tabela 5.1 sumaria os nossos resultados computacionais. Cada uma das linhas dessa tabela corresponde à resolução da relaxação linear de uma instância. Cada instância é resolvida duas vezes pelo método dual-simplex, consoante o número máximo de iterações imposto ao algoritmo AV. Nas instâncias de menor dimensão optou-se por um número máximo de 500 e 1000 iterações, nas instâncias intermédias esse número foi de 1000 e 5000 e nas instâncias de maior dimensão o mesmo número foi de 5000 e 10000 iterações. A informação inicial fornecida ao CPLEX é efectuada através de uma chamada da função `CPXcopystart(env,lp,NULL,NULL,x, NULL,NULL,y)`, onde `env` e `lp` são apontadores para o ambiente CPLEX e para o problema, respectivamente, `x` designa o vector primal e `y` designa o vector dual tal como foram obtidos no final do algoritmo AV.

Na primeira coluna da Tabela 5.1 é identificada a instância do PCVA. Nas cinco colunas seguintes são apresentados os resultados obtidos com o AV: o número máximo de iterações; o limite inferior obtido, com base em informação dual; o valor na função objectivo da última solução primal encontrada; o valor da violação máxima das restrições nessa solução; e, o tempo gasto em segundos. Nas duas colunas seguintes apresentamos: o número de iterações executadas pelo método dual-simplex; e, o tempo gasto em segundos. As duas últimas colunas incluem uma coluna com a soma dos tempos de execução dos dois algoritmos e uma coluna que mostra, em percentagem, o ganho de tempo entre esta estratégia combinada (AV+CPLEX) e o tempo do CPLEX (dualopt) apresentado na Tabela 1.

Observamos que, à excepção da instância br17, a um maior número de iterações do algoritmo volumétrico corresponde sempre um menor número de iterações do método dual-simplex. A redução é significativa nalguns casos. Por exemplo, 5000 iterações do AV não têm influência na resolução da instância ftv55. No entanto, depois de duplicar o número de iterações para 10000, o CPLEX necessitou apenas de 20 iterações executadas em 16.8 segundos para chegar à solução óptima. Outro exemplo significativo é o da resolução da instância ftv70 que após 5000 iterações do AV o tempo que dominava na determinação da solução era o tempo despendido pelo CPLEX. Quando duplicamos o número de iterações do AV, passámos a obter um ganho de tempo de 87.8%.

Se compararmos os tempos de resolução das Tabelas 1 e 5.1 observamos, quase sempre, uma vantagem clara em usar o AV antes de invocar o método dual-simplex do CPLEX. Nalguns casos, como ftv70, a abordagem combinada é superior ao método de pontos interiores. No entanto, as instâncias br17 e p43 contrariam esta observação. A Figura 2 permite uma outra visualização da superioridade da estratégia combinada naquele conjunto de instâncias.

Desde o início, pensámos que, atendendo ao papel das restrições dualizadas nesta abordagem Lagrangeana, o número de subcircuitos na solução óptima de (6) diminuiria à medida que o algoritmo volumétrico decorre - tornando-as cada vez mais parecidas com circuitos. Esta hipótese não foi completamente validada embora, de facto, se observe uma tendência para a

Tabela 2: Resultados da aplicação do AV seguido do CPLEX.

PCVA ID	VOLUMÉTRICO					CPLEX		Tempo Total	Tempo Ganho
	AV It.	L (dual)	P (primal)	Max. violação	Tempo (seg.)	Simplex It.	Tempo (seg.)		
br17	500	33.5	40.4	0.00416	0.7	1208	2.3	3.0	-34.1%
br17	1000	38.3	40.0	0.00237	1.4	1208	2.4	3.8	-70.5%
ftv33	500	1246.2	1304.1	0.04462	2.8	4421	19.1	21.9	48.8%
ftv33	1000	1281.9	1311.1	0.03612	5.5	2700	13.8	19.3	54.7%
ftv35	500	1435.6	1464.2	0.13534	3.2	12449	59.4	62.6	17.8%
ftv35	1000	1454.6	1466.0	0.11960	6.4	2190	13.4	19.8	74.1%
ftv38	500	1496.1	1513.5	0.11292	3.6	8999	59.6	63.2	48.2%
ftv38	1000	1512.8	1514.6	0.08486	7.1	2926	20.8	27.9	77.1%
ftv44	1000	1577.1	1586.0	0.10168	10.8	8217	82.5	93.3	66.5%
ftv44	5000	1584.7	1585.8	0.05515	52.4	2408	27.8	80.3	71.2%
ftv47	1000	1740.3	1750.5	0.09320	14.1	27997	347.9	362.1	42.3%
ftv47	5000	1746.9	1749.8	0.05170	70.6	12957	174.6	245.2	60.9%
ftv55	5000	1583.9	1585.6	0.04191	101.7	38867	940.5	1042.1	-0.6%
ftv55	10000	1583.9	1584.7	0.01919	205.7	20	16.8	222.5	78.5%
ftv64	5000	1807.3	1813.9	0.03699	161.8	18252	584.4	746.2	73.0%
ftv64	10000	1807.3	1810.6	0.01694	309.7	9876	349.9	659.5	76.1%
ftv70	5000	1908.4	1910.3	0.02956	198.0	36051	1215.8	1413.8	62.0%
ftv70	10000	1908.4	1909.6	0.01351	409.9	35	42.2	452.2	87.8%
p43	1582	5578.3	5633.8	0.00717	28.8	175194	6424.6	6453.4	-204.7%
p43	5000	5580.1	5625.1	0.00420	76.7	86710	3041.5	3118.3	-47.3%
ry48p	3328	14280.1	14422.9	0.00455	47.0	13181	187.7	234.8	60.9%
ry48p	5000	14280.1	14393.0	0.00351	66.3	6625	87.9	154.3	74.3%
ft53	5000	6904.0	6948.9	0.01414	82.5	1572	38.9	121.4	85.3%
ft53	7243	6904.0	6945.0	0.01000	127.2	74	14.7	141.9	82.8%
ft70	5000	38640.1	8831.2	0.03572	208.9	32930	983.4	1192.3	67.4%
ft70	10000	38640.1	38735.1	0.01637	432.3	236	46.3	478.6	86.9%

redução no número de subcircuitos. Por exemplo, para o problema ftv70, exibimos o número de subcircuitos encontrados ao longo das 1000 iterações do AV na Figura 3.

5.2 Resolução do modelo inteiro

Nesta subsecção, procuramos averiguar a adequação do AV como mecanismo de identificação de arcos que não participam na solução óptima inteira. Mais precisamente, com a solução resultante da aplicação do algoritmo volumétrico foram fixadas a zero todas as variáveis cujo valor da solução primal seja igual ou inferior a 0.0005 e cujo custo reduzido associado à restrição dual correspondente seja superior ou igual a 0.5. Depois, executou-se o CPLEX MIP com base na formulação de fluxo desagregado. Aplicámos esta ideia apenas a algumas instâncias estudadas anteriormente. Os resultados são apresentados na Tabela 3.

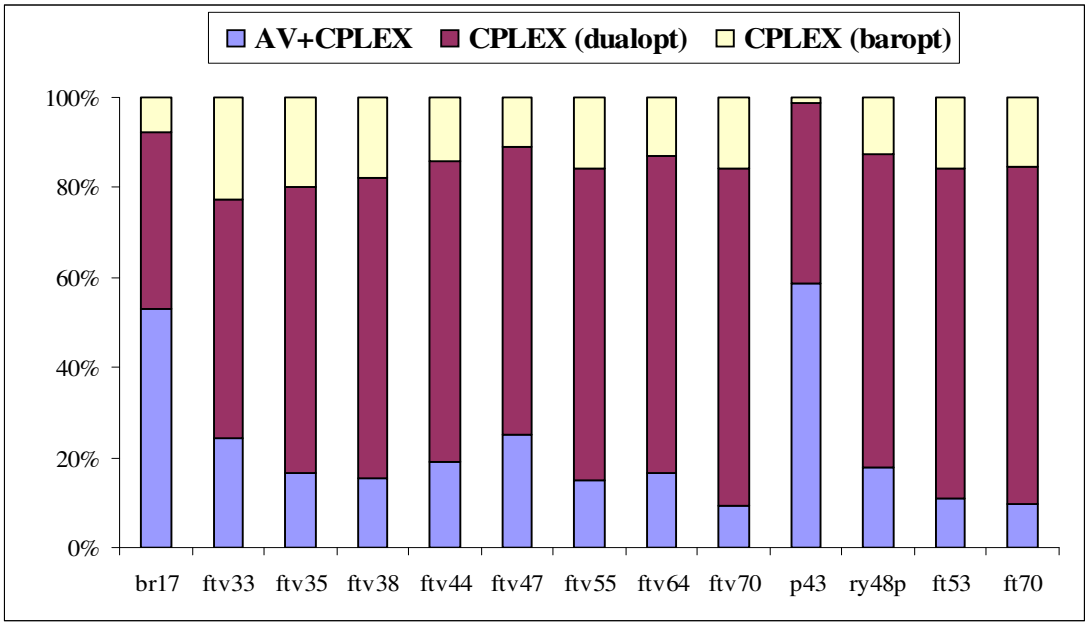


Figura 2: Comparação de tempos referidos nas Tabelas 1 e 5.1.

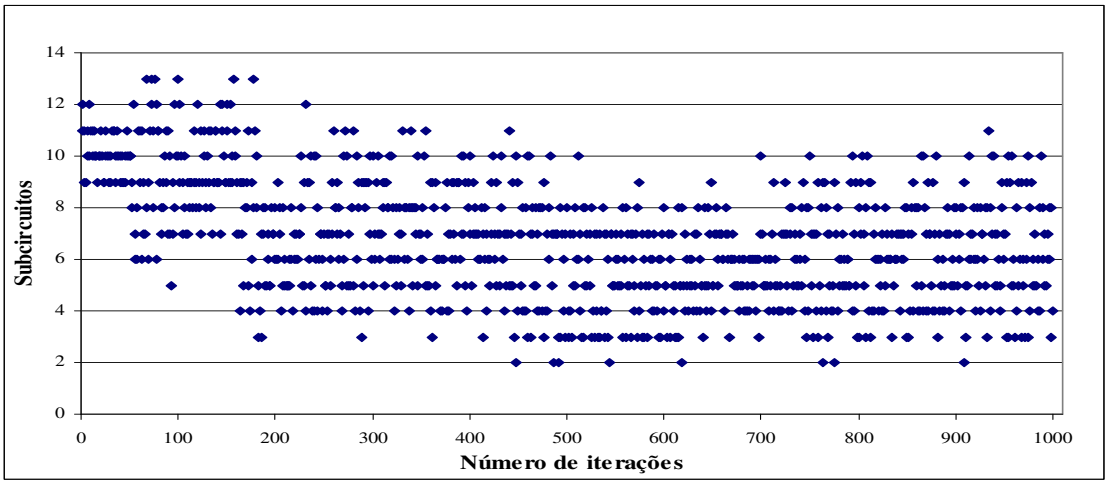


Figura 3: Número de subcircuitos para o Problema ftv70.

A primeira parte da Tabela 3 repete informação apresentada na Tabela 5.1. A segunda parte inclui os resultados da aplicação do CPLEX MIP ao problema reduzido, respectivamente, o número de iterações executadas e o tempo de resolução. A penúltima coluna diz respeito ao tempo total no cálculo da solução inteira, ou seja, a soma do tempo na aplicação do AV adicionado do tempo da aplicação do CPLEX MIP. A coluna final indica o ganho de tempo relativamente à aplicação do CPLEX MIP apenas. Em todas estas instâncias obteve-se o mesmo valor óptimo que o problema original (sem a fixação prévia de variáveis).

Comparando estes resultados com os obtidos na Tabela 1, podemos verificar que para o

Tabela 3: Resultados obtidos com o CPLEX MIP após a execução do AV.

PCVA ID	VOLUMÉT.		CPLEX MIP		Tempo Total	Tempo Ganho
	AV It.	Tempo (seg.)	Simplex It.	Tempo (seg.)		
br17	500	0.7	48837	158.4	159.1	1.9%
br17	1000	1.4	35331	121.4	122.8	24.3%
ftv33	500	2.8	2508	1.1	3.9	93.1%
ftv33	1000	5.5	2879	1.4	6.8	88.0%
ftv35	500	3.2	5028	27.9	31.1	98.5%
ftv35	1000	6.4	4106	24.4	30.8	98.5%
ftv38	500	3.6	5997	36.9	40.5	99.8%
ftv38	1000	7.1	11662	64.0	71.2	99.7%

problema br17 o tempo despendido pelo CPLEX MIP na obtenção da solução inteira foi grande relativamente ao requerido pelo AV. Mesmo assim, após 1000 iterações do AV e fixando 1071 variáveis (dum total de 4624 variáveis) conseguiu-se um ganho de tempo de 24.3%. Para os outros problemas conseguiram-se ganhos de tempo na ordem dos 90%, como podemos verificar de seguida. Para o problema ftv33, após fixar a zero 10709 dum total de 38148 variáveis, foi encontrada a solução inteira óptima num tempo total de 3.9s, que é inferior ao tempo despendido quando se utiliza apenas o CPLEX MIP que foi de 56.8s. Isto mostra, novamente, que o AV acelerou o processo de encontrar a solução. Porém, como era de esperar, na resolução da mesma instância com 1000 iterações do AV, que resultou em fixar a zero 7779 variáveis, não se verificou uma melhoria no tempo total. Para o problema ftv35, com 500 iterações do AV (fixando a zero 5990 dum total de 45360 variáveis) foi encontrada a solução óptima inteira. Para a instância ftv38, executando 500 iterações do AV (fixando a zero 6972 dum total de 57798 variáveis) foi encontrada a solução óptima inteira. De notar, que o tempo total despendido na obtenção desta solução foi muito menor que o obtido através da execução apenas do CPLEX MIP no problema completo.

5.3 Resolução por métodos heurísticos

Nesta subsecção, descrevemos experiências computacionais obtidas com a aplicação de várias heurísticas aproveitando os custos reduzidos gerados no decurso do AV. Além das instâncias apresentadas anteriormente, foram consideradas outras de maior dimensão, umas obtidas na TSPLib e outras geradas aleatoriamente através dum código disponível na página web de *DIMACS Implementation Challenge* [24]. As instâncias geradas pertencem à classe denominada *Random Asymmetric Matrices Closed under Shortest Paths* (TMAT). Nesta classe as instâncias são geradas através do gerador de matrizes de distâncias assimétricas aleatórias que escolhe cada distância $d(i,j)$ como um valor inteiro aleatório x, tal que $0 \leq x \leq 10^6$. Depois, para cada distância se $d(i,j) > d(i,k) + d(k,j)$ então $d(i,j) = d(i,k) + d(k,j)$ e repete-se até não se poderem fazer mais trocas. De notar, que estas instâncias consideram a desigualdade triangular das distâncias. Na Tabela 4, encontra-se uma descrição das instâncias retiradas da TSPLib (kro124p e ftv100 até ftv170), incluindo o valor da solução óptima inteira conhecida, e das instâncias geradas aleatoriamente (tmat100 até tmat200).

Considerámos uma redefinição dos custos $cr_{ij} \equiv c_{ij} - (u_i + v_j)$, onde u e v denotam as

variáveis duais óptimas no problema de afectação de uma dada iteração do AV. De 100 em 100 iterações, executamos uma heurística que vai permitir obter um circuito Hamiltoniano com base nessa estrutura de custos. No final do AV, observamos o melhor limite superior encontrado. Testámos heurísticas de construção e a heurística de Lin-Kernighan tal como é implementada em [21].

Tabela 4: Problemas de dimensões superiores.

PCVA			
ID	$\|V\|$	$\|A\|$	Int. óptimo
kro124p	100	9900	36230
ftv100	101	10100	1788
ftv110	111	12210	1958
ftv120	121	14520	2166
ftv130	131	17030	2307
ftv140	141	19740	2420
ftv150	151	22650	2611
ftv160	161	25760	2683
ftv170	171	29070	2755
tmat100	100	9900	
tmat110	110	11990	
tmat120	120	14280	
tmat130	130	16770	
tmat140	140	19460	
tmat150	150	22350	
tmat160	160	25440	
tmat170	170	28730	
tmat180	180	32220	
tmat190	190	35910	
tmat200	200	39800	

5.3.1 Heurísticas de construção

Testámos quatro das heurísticas que estão disponíveis no código TSP_SOLVE [22], nomeadamente, as heurísticas *addition*, *assign*, *loss* e *patching*. A heurística *addition*, descrita em [23], funciona de modo semelhante à heurística do vizinho mais próximo. As heurísticas *assign* e *patching* funcionam de modo semelhante, nomeadamente, usam uma solução do correspondente problema de afectação, que define um conjunto de subcircuitos, para definir um circuito Hamiltoniano de forma heurística. A heurística *patching*, foi inicialmente apresentada por Karp e Steele em [25], é também denominada por *Karp-Steele Patching* (KSP) em [13]. A heurística *loss* é uma implementação do método de Loss ([9]), que foi inicialmente proposta para o problema do caixeiro viajante simétrico e posteriormente adaptada para o caso assimétrico. A Tabela 5 reproduz os resultados obtidos para todos os problemas da TSPLIB e da classe TMAT, anteriormente apresentados. Nessa tabela apresentamos, para cada instância, o número total de iterações do AV, e para cada heurística o melhor limite superior, denotado LS, encontrado. Indicamos, também, a primeira iteração em que se encontrou esse LS. Os tempos de execução das heurísticas são negligenciáveis.

Tabela 5: Soluções obtidas com as heurísticas do TSP_SOLVE.

PCVA	AV	addition		assign	loss		patching	
ID	It.	LS	It.	LS	LS	It.	LS	It.
br17	1000	39	100	39	39	100	39	100
ftv33	1000	1482	200	1286	1372	600	1409	200
ftv35	1000	1491	100	1473	1508	300	1489	200
ftv38	1000	1634	200	1530	1547	700	1546	100
ftv44	5000	1733	3400	1613	1673	300	1699	100
ftv47	5000	1793	700	1776	1787	1600	1846	300
ftv55	5000	1781	3400	1608	1747	200	1657	100
ftv64	5000	2054	500	1839	1890	500	1871	900
ftv70	5000	2168	400	1950	2074	100	2004	600
p43	5000	5623	800		5638	1100	5640	200
ry48p	5000	14939	200	14422	15254	1000	14857	100
ft53	5000	8088	300	6905	7383	500	7847	100
ft70	5000	40566	500	38673	39065	1900	39197	100
kro124p	5000	40524	200	36230	41121	500	40106	100
ftv100	5000	2119	3000	1788	1969	900	1878	400
ftv110	5000	2335	800	1958	2156	4400	2036	500
ftv120	5000	2618	3000	2166	2402	100	2244	1100
ftv130	5000	2900	3300	2307	2519	100	2402	1000
ftv140	5000	2965	5000	2420	2590	300	2500	1200
ftv150	5000	3166	3400	2611	2971	200	2691	1500
ftv160	5000	3318	4200	2683	2839	100	2729	200
ftv170	5000	3509	4500	2755	2938	300	2809	600
tmat100	5000	1628486	1700	1377025	1379385	3500	1395601	500
tmat110	5000	1579919	200	1362004	1371800	700	1367459	100
tmat120	5000	1684696	2800	1390478	1396470	600	1419536	100
tmat130	5000	1716897	4500	1429864	1428864	3400	1444402	300
tmat140	5000	1719640	4800	1433486	1443556	300	1449585	100
tmat150	5000	1598899	3500	1364821	1364821	100	1389641	300
tmat160	5000	1855854	1300	1484992	1486164	2200	1487017	100
tmat170	5000	1674560	1600	1405055	1405273	4400	1405172	800
tmat180	5000	1941760	4400		1558217	500	1562008	100
tmat190	5000	1670791	500		1411658	1000	1424933	200
tmat200	5000	1968195	1400		1548893	400	1547919	100

Os LS encontrados para a heurística *assign* coincidem todos com os valores óptimos inteiros conhecidos e foram todos encontrados ao fim de 100 iterações do AV. Queremos também revelar que executando esta heurística com os custos originais do problema, os valores dos LS também coincidem com os valores óptimos inteiros. Para o problema p43, com a heurística *assign*, não se conseguiram obter quaisquer soluções após 5 horas de tempo de execução.

Para as heurísticas *addition, loss* e *patching*, só ao fim de um certo número de iterações é que se encontraram os melhores LS. Salientamos, também, que para estas heurísticas os LS obtidos com os custos originais do problema são piores que os encontrados com os custos reduzidos. Podemos observar, pela tabela, que os valores dos melhores LS, para a heurística

addition ficam, em geral, mais afastados dos valores determinados com as outras heurísticas.

Relativamente aos problemas da classe TMAT, verificamos que os melhores LS são obtidos com a heurística *assign*, apesar de não se ter conseguido obter resultados para as instâncias tmat180 até tmat200, porque, neste caso, o programa não permite resolver instâncias do PCVA com mais de 175 vértices. Nesta classe de instâncias, também se verifica que a heurística que apresenta piores resultados é a *addition*.

Na Figura 4, encontra-se um gráfico que mostra os diversos LS encontrados, com a heurística *addition*, para o problema ftv44, durante a execução de 5000 iterações do AV. Neste caso, o melhor LS foi encontrado quando se invocou a heurística na iteração 3400 e voltou a este valor na iteração 4800. De notar, que na iteração zero do AV, em que os custos reduzidos coincidiam com os custos originais do problema, o LS tinha o valor de 1829. Por isso, podemos concluir que os custos reduzidos, gerados pelo AV, ajudam a encontrar melhores LS.

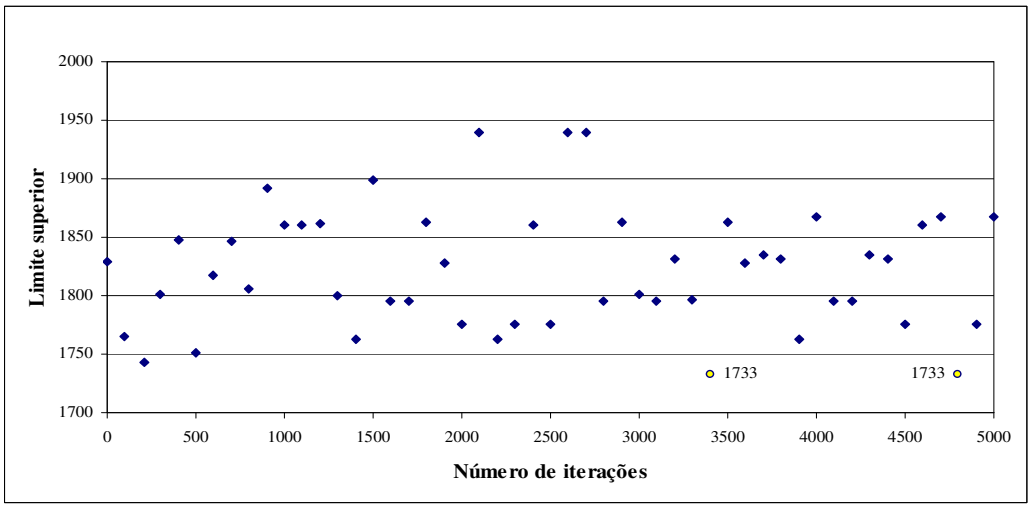


Figura 4: Limites superiores determinados pela heurística *addition* para o problema ftv44.

5.3.2 Heurística de Lin-Kernighan modificada por Helsgaun

Keld Helsgaun descreve em [21] a implementação de uma versão modificada do algoritmo de Lin-Kernighan [28]. A nova implementação difere da original em vários detalhes. A principal diferença encontra-se na estratégia de procura. O novo algoritmo usa passos de procura maiores e mais complexos que o original. Usa também a análise de sensibilidade para direccionar e restringir a procura. As experiências computacionais que efectuámos vieram confirmar o reconhecido sucesso desta heurística. Nomeadamente, os circuitos encontrados com esta heurística para os problemas da TSPLIB (usando os custos originais) são sempre óptimos, excepto para o problema p43 que encontra um circuito de valor 5621.

Tal como fizemos para as heurísticas da subsecção anterior, executámos esta heurística com os custos reduzidos, em cada 100 iterações do AV. Apresentamos na Tabela 6 os valores dos LS encontrados. Para as instâncias da TSPLIB, os circuitos encontrados pelo heurística são sempre óptimos. Para os vários problemas, os melhores LS foram encontrados na iteração 100

do AV, excepto para os problemas ftv150, ftv170 e p43, que foram encontrados, respectivamente, nas iterações 400, 500 e 400. Tal como nos casos anteriores, os tempos de execução da heurística são negligenciáveis.

Tabela 6: Soluções obtidas com a heurística de Helsgaun.

PCVA ID	LS	PCVA ID	LS	PCVA ID	LS
br17	39	kro124p	36230	tmat100	1377025
ftv33	1286	ftv100	1788	tmat110	1362004
ftv35	1473	ftv110	1958	tmat120	1390478
ftv38	1530	ftv120	2166	tmat130	1429864
ftv44	1613	ftv130	2307	tmat140	1433486
ftv47	1776	ftv140	2420	tmat150	1364821
ftv55	1608	ftv150	2611	tmat160	1484992
ftv64	1839	ftv160	2683	tmat170	1405055
ftv70	1950	ftv170	2755	tmat180	1554371
p43	5620			tmat190	1411658
ry48p	14422			tmat200	1545211
ft53	6905				
ft70	38673				

6 Conclusões

Este estudo incidiu, essencialmente, na resolução da relaxação linear subjacente à formulação de fluxo desagregado para o problema do caixeiro viajante assimétrico através do algoritmo volumétrico.

O nosso estudo mostrou que vale a pena utilizar o algoritmo volumétrico como mecanismo de identificação de uma base inicial para posterior aplicação de um método dual-simplex. Em geral, um número maior de iterações do algoritmo volumétrico corresponde a um menor número de iterações do método dual-simplex (CPLEX). Quase sempre o tempo total de execução desta estratégia combinada é bastante inferior ao tempo despendido na execução do método dual-simplex. Os casos em que este comportamento não foi observado devem-se a uma interrupção demasiado prematura do algoritmo volumétrico.

O algoritmo volumétrico foi bastante eficiente na identificação dos arcos que não participam na solução óptima inteira, para as instâncias pequenas que se testaram. Com a consequente redução no tamanho do problema o método *branch-and-bound* (CPLEX MIP) requereu muito menos tempo de execução. Para o futuro, pretendemos testar esta estratégia com instâncias de dimensão maior da TSPLib.

Os custos reduzidos gerados no decurso do algoritmo volumétrico ajudaram a melhorar o comportamento de certas heurísticas. Verificámos, ainda, que para a instância p43, a abordagem combinada permitiu encontrar uma solução inteira de melhor qualidade do que aquela que é obtida com a implementação de Helsgaun da heurística de Lin-Kernighan (usando apenas os custos originais), que é considerada como um heurística muito eficiente.

Com base nas experiências efectuadas, pensamos poder afirmar que o algoritmo volumétrico permite acelerar consideravelmente a resolução do problema do caixeiro viajante assimétrico.

Agradecimentos

Os autores agradecem os comentários e sugestões do revisor que contribuíram para melhorar e clarificar alguns aspectos do artigo.

7 Referências

[1] R. Anbil, J. J. Forrest, and W. R. Pulleyblank. Column generation and the airline crew pairing problem. *Documenta Mathematica*, Extra Volume ICM III:677–686, 1998.

[2] L. Bahiense, F. Barahona, and O. Porto. Solving steiner tree problems in graphs with lagrangian relaxation. Journal of Combinatorial Optimization, 3(7):259–282, 2003.

[3] L. Bahiense, N. Maculan, and C. Sagastizábal. The volume algorithm revisited: Relation with bundle methods. Mathematical Programming, 94:41–70, 2002.

[4] F. Barahona and R. Anbil. The volume algorithm: Producing primal solutions with a subgradient method. Mathematical Programming, 87:385–399, 2000.

[5] F. Barahona and R. Anbil. On some difficult linear programs coming from set partitioning. Discrete Applied Mathematics, 118(1-2):3–11, 2002.

[6] F. Barahona and F. A. Chudak. Near-optimal solutions to large scale facility location problems. Technical report, IBM Watson Research Center, 1999.

[7] F. Barahona and L. Ladanyi. Branch and cut based on the volume algorithm: Steiner trees in graphs and max-cut. Technical report, IBM Watson Research Center, 2001.

[8] A. Claus. A new formulation for the travelling salesman problem. SIAM Journal of Algebraic and Discrete Methods, 5:21–25, 1984.

[9] P. V. D. Cruyssen and M. Rijckaert. Heuristic for the asymmetric travelling salesman problem. Journal of the Operational Research Society, 29(7):697–701, 1978. 18A. M. Rocha, E. M. G. P. Fernandes, J. Soares / Investigação Operacional, 25 (2005) 1-19

[10] G. Dantzig, D. Fulkerson, and S. Johnson. Solution of a large-scale traveling salesman problem. Operations Research, 2:393–410, 1954.

[11] M. Fischetti and P. Toth. A polyhedral approach to the asymmetric traveling salesman problem. Management Science, 43(11):1520–1536, 1997.

[12] B. Gavish and S. Graves. The traveling salesman problem and related problems. Technical report, Operations Research Center, MIT, 1978. Working Paper OR-078-78.

[13] F. Glover, G. Gutin, A. Yeo, and A. Zverovich. Construction heuristics for the asymmetric TSP. European Journal of Operational Research, 129:555–568, 2001.

[14] L. Gouveia and J. M. Pires. Uma análise comparativa de formulações para o problema do caixeiro viajante assimétrico. Investigação Operacional, 16:89–114, 1996.

[15] L. Gouveia and J. M. Pires. The asymmetric travelling salesman problem and a reformulation of the miller-tucker-zemlin constraints. European Journal of Operational Research, 112(1):134–146, 1999.

[16] L. Gouveia and J. M. Pires. The asymmetric travelling salesman problem: on generalizations of disaggregated Miller-Tucker-Zemlin constraints. Discrete Applied Mathematics, 112(1-3):12–145, 2001.

[17] G. Gutin and A. P. Punnen, editors. The traveling salesman problem and its variations, volume 12 of Combinatorial Optimization. Kluwer Academic Publishers, Dordrecht, 2002.

[18] M. Held and R. M. Karp. The traveling-salesman problem and minimum spanning trees. Operations Research, 18:1138–1162, 1970.

[19] M. Held and R. M. Karp. The traveling-salesman problem and minimum spanning trees: Part II. Mathematical Programming, 1:6–25, 1971.

[20] M. Held, P. Wolfe, and H. P. Crowder. Validation of subgradient optimization. Mathematical Programming, 6:62–68, 1974.

[21] K. Helsgaun. An effective implementation of the Lin-Kernighan traveling salesman problem. European Journal of Operations Research, 126:106–130, 2000. Code available at http://www.akira.ruc.dk/ keld/research/LKH/.

[22] C. Hurwitz and R. Craig. GNU Tsp_solve. Version 1.3.8, 1994. Available at http://www.cs.sunysb.edu/ algorith/implement/tsp/distrib/tsp_solve/.

[23] D. Johnson and C. Papadimitriou. Performance guarantees for heuristics. In E. Lawler, J. Lenstra, A. R. Kan, and D. Shmoys, editors, The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, pages 145–180. John Wiley & Sons, Chichester, 1985. A. M. Rocha, E. M. G. P. Fernandes, J. Soares / Investigação Operacional, 25 (2005) 1-1919

[24] D. Jonhson, L. McGeoch, F. Glover, and C. Rego. Website for the DIMACS implementation challenge on the traveling salesman problem. http://www.research.att.com/ dsj/chtsp.

[25] R. Karp and J. Steele. Probabilistic analysis of heuristics. In E. Lawler, J. Lenstra, A. R. Kan, and D. B. Shmoys, editors, The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, pages 181–205. John Wiley & Sons, Chichester, 1985.

[26] A. Langevin, F. Soumis, and J. Desrosiers. Classification of travelling salesman problem formulations. Operations Research Letters, 9:127–132, 1990.

[27] E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, and D. B. Shmoys, editors. The traveling salesman problem. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Ltd., Chichester, 1990. A guided tour of combinatorial optimization, Reprint of the 1985 original, A Wiley-Interscience Publication.

[28] S. Lin and B. W. Kernighan. An effective heuristic algorithm for the traveling salesman problem. Operations Research, 21:498–516, 1973.

[29] C. Miller, A. Tucker, and R. Zemlin. Integer programming formulation of traveling salesman problems. Journal of ACM, 7:326–329, 1960.

[30] M. Padberg and T. Sung. An analytical comparison of different formulations of the travelling salesman problem. Mathematical Programming, 52:315–357, 1991.

[31] J. M. O. Pires. Formulações para o problema do caixeiro viajante assimétrico e sua aplica ção a um problema de desenho de redes com topologia em forma de anel. PhD thesis, Universidade de Lisboa, Setembro 2001.

[32] G. Reinelt. TSPLIB - a traveling salesman problem library. ORSA Journal on Computing, 3:376–384, 1991. Available at http://www.iwr.uniheidelberg. de/groups/comopt/software/TSPL IB95.

[33] R. Wong. Integer programming formulations of the travelling salesman problem. pages 149–152. Proceedings of the IEEE International Conference of Circuits and Computers, 1980.

REVISTA INVESTIGAÇÃO OPERACIONAL

Pol tica Editorial

Investigação Operacional (IO) é a revista científica da *APDIO - Associação Portuguesa de Investigação Operacional.* A política editorial da IO é publicar artigos originais e de elevada qualidade que contribuam para a teoria, metodologia, técnicas e software de Investigação Operacional e a sua aplicação a diferentes campos. A Revista também publica artigos com revisões relevantes de temas de IO. Casos de sucesso na aplicação a problemas práticos são especialmente bem vindos.

Processo de Aceitação

Todos os manuscritos submetidos para publicação são revistos e aceites apenas com base na avaliação da sua qualidade, importância e adequação à política editorial. Será responsabilidade do Editor interpretar a avaliação dos revisores. A contribuição de cada artigo deve estar claramente evidenciada na Introdução. Critérios como a relação com literatura existente, comprimento e estilo do artigo são tidos em consideração. Uma indicação clara da viabilidade de aceitação do artigo é habitualmente dada na primeira fase de revisão do artigo.

Será requerido aos autores de um artigo aceite que transfiram os direitos de autoria para a APDIO, que assegurará a mais ampla disseminação possível de informação. Os volumes da Revista são publicados em papel, e distribuídos a todos os associados da APDIO, e em formato electrónico na rede SciELO - Scientific Electronic Library Online.

Resumos dos Artigos indexados em

IAOR - International Abstracts in Operations Research

Instruções aos Autores

1. *Submeter artigos* para publicação ao editor principal, de preferência por e-mail em Microsoft Word ou "Portable Document Format" (PDF) para `jfo@fe.up.pt`, ou por correio normal (quatro cópias) para o seguinte endereço: Prof. José Fernando Oliveira, Departamento de Engenharia Electrotécnica e de Computadores, Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 PORTO, Portugal.

2. *Língua.* Os artigos devem ser escritos em Português, Inglês ou Espanhol.

3. *Os Manuscritos* devem ser impressos. Numerar as páginas consecutivamente.

4. *A primeira página* do manuscrito escrito em português ou em espanhol deve ter a seguinte informação: (a) Título; (b) nome, e-mail e afiliação institucional dos autores; (c) um resumo; (d) palavras-chave; (e) título em inglês (f) um resumo em inglês; (g) palavras-chave em inglês; (h) identificação do autor correspondente. Se o manuscrito for escrito em inglês, a primeira página deve ter a seguinte informação: (a) Título em inglês; (b) nome, e-mail e afiliação institucional dos autores; (c) um resumo em inglês; (d) palavras-chave em inglês; (e) identificação do autor correspondente.

5. *Agradecimentos*, incluindo informação sobre apoios, dever ser colocados imediatamente antes da secção de referências.

6. *Notas de rodapé* devem ser evitadas.

7. *Formulas* que são referenciadas devem ser numeradas consecutivamente ao longo do manuscrito como (1), (2), etc. do lado direito.

8. *Figuras*, incluindo grafos e diagramas, devem ser numerados consecutivamente em numeração árabe.

9. *Tabelas* devem ser numeradas consecutivamente em numeração árabe.

10. *Referências.* Citar apenas as mais relevantes e listar só as que são citadas no texto. Indicar as citações no texto através de parênteses rectos, e.g., [4]. No final do artigo listar as referências alfabeticamente por apelido do primeiro autor e numerá-las consecutivamente, de acordo com o seguinte formato: Artigos: autore(s), título, nome e volume da revista (ou livro, mas neste caso incluir o nome dos editores), ano e páginas. Livros: Autor(es), título, editor, ano.

11. *Artigos aceites* devem ser enviados pelo autor ao editor, de preferência na forma de um ficheiro fonte em LaTeX com ficheiros EPS para as figuras, juntamente com um ficheiro PDF ou Postscript. Em alternativa, ficheiros fonte em Word são também aceites. Para garantir uma boa qualidade gráfica, as figuras devem ser em formato vectorial; formatos raster como JPG, BMP, GIF, etc. devem ser evitados.

12. *Provas dos artigos* serão enviadas por e-mail como ficheiros PDF para o autor correspondente. Corrigir as provas cuidadosamente, e restringir as correcções apenas aos pontos em que as provas diferem do manuscrito. Desvios à versão aceite pelo editor são apenas possíveis com a autorização prévia e explícita do editor. Trinta separatas de cada artigo são enviados gratuitamente ao autor correspondente.

Informação sobre a Publicação

Investigação Operacional (ISSN 0874-5161) está registada na Secretaria de Estado da Comunicação Social sob o número 108335. Os volumes da Revista são publicados em papel, e distribuídos a todos os associados da APDIO, e em formato electrónico na rede SciELO - Scientific Electronic Library Online. O preço da assinatura anual é de 25 euros. Os volumes são enviados por correio normal. Informação adicional sobre assinaturas pode ser solicitada ao Secretariado da APDIO- CESUR, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 LISBOA, Portugal. Tel. +351 218 407 455 - www.apdio.pt - `apdio@cesur.civil.ist.utl.pt`

JOURNAL *INVESTIGACÃO OPERACIONAL*

Editorial Policy

Investigação Operacional (IO) is the scientific journal of *APDIO - Associação Portuguesa de Investigação Operacional* (the Portuguese Operational Research Association). The editorial policy of IO is to publish high quality and original articles that contribute to theory, methodology, techniques and software of Operational Research (OR) and its application to different fields. It also publishes articles with relevant reviews of OR subjects. Cases of successful application of OR to practical problems are specially welcome.

Acceptance Process

All manuscripts submitted for publication are refereed and accepted only on the basis of its quality, importance and adequacy to the editorial policy. It will be the responsibility of the Editor to interpret the referee's assessment. The contribution of each paper should be clearly stated in the introduction. Criteria such as relationship with existing literature, length and style are taken into account. A clear indication on the suitability of a manuscript is usually provided after the first round of refereeing. The authors of an accepted paper will be asked to transfer its copyright to the publisher, which will ensure the widest possible dissemination of information. The volumes of the journal are published in hardcopies, which are distributed to all APDIO associates, and in electronic format in SciELO - Scientific Electronic Library Online.

Articles are abstracted/indexed in

IAOR - International Abstracts in Operations Research

Instructions to Authors

1. *Submit papers* for publication to the main editor, preferably by e-mail in Microsoft Word or "Portable Document Format"(PDF) to `jfo@fe.up.pt`, or by ordinary mail (four copies) to the following address: Prof. José Fernando Oliveira, Departamento de Engenharia Electrotécnica e de Computadores, Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 PORTO, Portugal.

2. *Language.* Papers must be in written in Portuguese, English or Spanish.

3. *Manuscripts* should be typewritten or typeset. Number the pages consecutively.

4. *The first page* of the manuscript written in English should contain the following information: (a) Title; (b) names, e-mails and institutional affiliations of the authors; (c) an abstract; (d) keywords (f) identification of the corresponding author.

5. *Acknowledgements*, including support information, should be placed prior to the references section.

6. *Footnotes* should be avoided.

7. *Formulas* that are referred to should be numbered consecutively throughout the manuscript as (1), (2), etc. on the right.

8. *Figures*, including graphs and diagrams, should be numbered consecutively in Arabic numbers.

9. *Tables* should be numbered consecutively in Arabic numbers.

10. *References.* Cite only the most relevant references and list only those cited in the text. Indicate citations in the text by bracketed numbers, e.g., [4]. At the end of the paper list the references alphabetically by the surname of the first author and number them consecutively, according to the following formats: Articles: author(s), title, name and number of the journal (or book, but in this case include the editors names), year, pages. Books: Author(s), title, publisher, year.

11. *Accepted papers* are to be sent by the author to the editor, preferably in the form of a source file in LaTeX and EPS files for the figures together with a PDF or postscript file. Alternatively, source files in Word are also accepted. To ensure good publishing quality the figures should be in vector formats; raster formats like JPG, BMP, GIF, etc. should be avoided.

12. *Page proofs* will be e-mailed as a PDF file to the corresponding author. Correct proofs carefully, and restrict corrections to points at which the proof is at variance with the manuscript. Deviations from the version accepted by the editor are only possible with the prior and explicit approval of the editor. Thirty offprints of each paper are supplied free of charge to the corresponding author.

Publication information

Investigação Operacional (ISSN 0874-5161) is registered in the *Secretaria de Estado da Comunicação Social* under number 108335. The volumes of the journal are published in hardcopies, which are distributed free of charge to all APDIO associates, and in electronic format in SciELO - Scientific Electronic Library Online. Subscription price is 25 euros. Issues are sent by standard mail. Additional subscription information is available upon request from APDIO Secretariat - CESUR, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 LISBOA, Portugal. Tel. +351 218 407 455 - www.apdio.pt - `apdio@cesur.civil.ist.utl.pt`

Revista Investigação Operacional
Volume 25 - Número 2 (Dezembro 2005)

ÍNDICE

R.V. Vidal
The art and science of problem solving 157

A.P. Barroso, V.H. Machado
A gestão logística dos resíduos em Portugal 179

R.A. Monteiro, A.J. Pascoal
A model and a heuristic for the primary health care coverage planning problem in Portugal - application to "Cova da Beira" case 195

F.R. Leta, J.S. Mello, E.G. Gomes, L.A. Meza
Métodos de melhora de ordenação em DEA aplicados à avaliação estática de tornos mecânicos 229

J.S. Mello, L.A. Meza, E.G. Gomes, L.B. Neto
Limites para eficiências não radiais em DEA: uma abordagem vectorial 243

F. Bastos, A. Teixeira
An extension of a variant of a predictor-corrector primal-dual method form linear programming to semidefinite programming 253

A.M. Rocha, E.M. Fernandes, J. Soares
Aplicação do algoritmo volumétrico à resolução aproximada e exacta do problema do caixeiro viajante assimétrico 277