

TEIDE: Técnicas de Edición e Imputación de Datos Estadísticos

TIC2002-00895

Juan José Salazar González*
University of La Laguna, Tenerife

Abstract

Many agencies (e.g., statistical offices) collect and process huge amount of data from respondents and enterprises. Each provider answers a set of questions producing a subset of data named record. Each value of a record is called field and it can be a number or a categorical data. The set of all records is called microdata and it is the raw material for generating the output of the agency. Nevertheless, originally all microdata use to contain missing or wrong values because the respondent did not answer a question, or he/she did lie, or the collector made a mistake, etc. Typically a wrong record is detected by using a set of rules (called "edits") that classify a record as invalid whenever one is not violated. In order to use the bigger number of records from a microdata, the agency must try to correct each record by detecting the wrong fields and by guessing the right value, while minimizing the number of modifications. This is not a trivial task, and it use to require large amount of resource from statistical offices since there are no good automatic tools helping on it. The objective of this project is to research the problem from a Mathematical Programming point of view, and to produce new algorithm helping decision makers on this task. The state of the art is that there are some simple approaches based on statistical considerations, but nothing using Operational Research, even if it concerns with solving some problems in Combinatorial Optimization. Some Spanish statistical offices will support our research by checking that our work follows their real-world requirements.

Keywords: Editing and Imputation, Statistical Agencies, Mathematical Programming.

1 Objectives of the project

The main aim of this research project is to develop a new automatic procedure to help practitioners working in statistical agencies to clean collected data. In general collected data (named *microdata*) contains errors and/or missing values, and before processing the collected data to produce conclusions, it is very important to guarantee a minimum of quality. Otherwise, the decision taken from this data (generally by politicians and regarding the Society) would be based on wrong information. In other to guarantee this minimum of quality, the statistical office checks and modifies the collected data. For checking the data, the statistical office has designed (before collecting the data) a set of coherence rules, named *edits*. For example, an edit can be

“if the age is small that 5 then the civil status cannot be divorce”

* Email: jjsalaza@ull.es

Another example can be

“an employee cannot spend more than 50% of his/her salary in the mortgage of his house”

When the answer from a respondent does not satisfy an edit then it must be corrected. The way to correct a data is to change the minor number of individual values in the wrong record to produce another record satisfying all the edits.

The main aim of this project is to produce an automatic tool running on a computer to help practitioners in Statistical Agencies to clean their data. This software is named TEIDE, which means “Techniques for Editing and Imputation of Statistical Data” (in Spanish). The software must have good quality, not only in the internal algorithm doing the editing and imputation, but also in the external appearance since it will be used by non-expert users in statistical offices. In addition, the software must solve the internal optimization algorithm with exact or near-optimal approaches running in short time. This is mainly due to the fact that a microdata contains typically thousands of records, and the combinatorial problems must be solved for all the wrong records in a short overall computational time. To this end, this project proposed developing new algorithms in Optimization, all involved in proper friendly software. This was exactly the first objective mentioned in the proposal approved for this project.

A second aim was to joint our research group at University of La Laguna with other teams working on the same topic. Unfortunately there are not other Spanish teams conducting research on it. Therefore, this project was intended to contact other researchers in universities and Research centres outside Spain and give us the opportunity to show our contributions.

A third aim was to reduce the distance between the Spanish statistical offices and the Spanish universities. Indeed, the Society is always asking the university to close the gap between enterprises and universities, and our students require the analysis of more case studies. An objective in this project was to establish a better relation between the Statistical Department of University of La Laguna (ULL) and the Regional Statistical Office in Canary Islands (ISTAC). Indeed, the aim was to use microdata from ISTAC and produce a tool on it for ISTAC. This experience was also aimed for the Spanish Statistical Office (INE). We wish to start collaborating and understanding the real problems in our statistical offices, thus opening new research topics (in particular we have pointed out other problems in Statistical Disclosure Control, similar to Data Editing and Imputation).

2 Achieved success of the project

Since the beginning of this research project (December 1, 2002) we initiated the following work:

- Develop models and algorithms for solving the editing-and-imputation problem for numerical data with linear constraints. Mainly this was a working task conducted with the member of the team, Dr. Jorge Riera Ledesma. He finished his PhD dissertation with this work and he got the “EURO Doctoral Dissertation Award”, which is a prize by the European Association of Operational Research societies to the best Ph Dissertation presented in 2003. See <http://www.euro-online.org/display.php?page=edda4&>
We published, for example, the first results in the article
J. Riera-Ledesma, J.J. Salazar-González, “Algorithms for automatic data editing”, in Statistical Journal of the United Nations Economic Commission for Europe 20 (2003) 255-265

We also continued working and we submitted in 2004 another manuscript entitled “A branch-and-cut algorithm for the Error Location Problem in Data Cleaning”, to *Management Science*. We are waiting for the answer by the editor. In this paper we propose new algorithms using *Mathematical Programming* for the editing-and-imputation problem with quantitative variables and linear edits, and we feel very optimistic about the publication of this paper because it outperforms previous results from other authors.

- We initiated collaborating with Andrea Lodi and Michela Milano, both researchers at the Engineering School, University of Bologna, Italy. They are experts of using *Constraint Programming*, which is a promising methodology for dealing with qualitative variables and logical edits. We have implemented an algorithm and are now in the process of writing the first draft with our findings.
- We also initiated the implementation of the graphical interface named TEIDE. To this end, a programmer is being included in the team to perform this implementation task. His name is Sergio Delgado Quintero, from Computer Science. Basically, it is a friendly framework for reading databases in Microsoft Access, which is the format used by ISTAC, the statistical office providing data to our experiments. The researcher started working on TEIDE since August 1, 2003, and during the last 14 months he have conducted an excellent implementation. Part of his work is available at <http://webpages.ull.es/users/istac>. We have now the first prototype that was tested on a microdata provided by ISTAC. This microdata contains information about health of some population in Canary Islands, and is about 5000 records containing about 200 fields each record. There are about 100 edits. Our experiments were very satisfactory to ISTAC, and we are now in the process of trying and improving the prototype with a larger microdata also provided by ISTAC.
- Regarding other problems from Statistical Offices, we initiated collaborating with the English statistical office to work on the Controlled Rounding Problem. This is the problem of perturbing a dataset such that the new dataset satisfy all the edits. In this cases the edits are linear equations, so the dataset is named *tabular data* or simply *table*. We have proposed new models and algorithms to solve this combinatorial problem. A new programmer was considered in the team to implement the graphical interface. This new interface makes use of Microsoft Excel, and is implemented as an “add-in” tool.
- We increased our contact with other experts, including people interested in national statistical offices. Hopefully, in the future, our team will be part of a European Project devoted to the editing-and-imputation problem.
- Finally, we have been also showing our results in different conferences and journals. Some of the results concern methodology developed for the editing-and-imputation problem, but applicable also to other optimization problems. In particular, the basic approach used in *Mathematical Programming* is a branch-and-cut algorithm which has been used for other routing and telecommunication problems.

3 Guaranties of the obtained results

The first output of our work of our work in this Research Project was the invitation to J.J. Salazar for being in the commission that evaluated a Ph. Dissertation with the title *Processing of Erroneous and Unsafe Data* (“Verwerking van foutieve en onveilige data”) defended by Ton de Waal at University of Rotterdam, The Netherlands, in June 19, 2003. This is a proof that the first discussions of the preliminary ideas in different statistical offices (mainly in Spain, Netherlands and England) were successful. In the jury of this Ph. Dissertation there were the most relevant expert in the field of Operational Research applied to the Statistical Offices. More information in http://eps.eur.nl/dissertaties/erim_phd_series/624874117/

The second important output was to receive an invitation to present our preliminary results in the “Work Session on Statistical Data Editing”, Madrid, 20-22 octubre 2003. This workshop was in Athens, November 6-9, 1995, in Voorburg, November 4-7, 1996, in Rome, June 2-4, 1999, in Cardiff (U.K.) October 18-20, 2000, and in Helsinki, May 27-29 2002. It is the most important meeting on the editing-and-imputation problem organized by the Statistical Division of the United Nations’ Economic Commission for Europe. In the 2003 meeting, J.J. Salazar presented an invited paper, which was published in J. Riera-Ledesma, J.J. Salazar-González, *Algorithms for automatic data editing*, Statistical Journal of the United Nations Economic Commission for Europe 20 (2003) 255-264. The talk at the conference is available from the webpage <http://www.unece.org/stats/documents/2003/10/sde/wp.5.e.pdf>

To show the relevance of the topic to the Spanish National Statistical Office (INE), there was a meeting in Madrid (September 15-19, 2003) to give some basic instruction to practitioners in the Regional Statistical Offices in the different regions of Spain. A researcher of our team (Sergio Delgado Quintero) did attend the meeting, and contributed to disseminate our research in the area. The title of the course was “Depuración de datos categóricos”, and more details can be downloaded from <http://www.ine.es/ine/ceaapp/escuela.htm>

After two years working on similar problems in Statistical Offices, we have also produced different articles in the area of Statistical Disclosure Control. In particular we have been invited to show part of our work in different statistical offices in Europe, like in England, Netherlands, Norway and Spain. In particular we addressed the problem of the Cell Suppression Methodology and the problem of Controlled Rounding Methodology. For the first one we have collaborated with other researchers in the area, like Anco Hundepool (CBS, Netherlands), Sarah Guessing (Germany) and Luisa Franconi (ISTAT, Italy), and we have develop an automatic tool. A first prototype of this software can be downloaded from the webpage <http://neon.vb.cbs.nl/casc/TAU.html>

Thanks to a new programmer (Markus Schoch) we have also produced a simpler automatic tool for solving the Controlled Rounding problem. This tool was implemented in Microsoft Visual C++ and it is an “add-in” tool for the Microsoft Excel software. We choose this software because, unfortunately, many practitioners in statistical offices in Europe use this software for working with data. Therefore, they appreciated a simple tool inside this familiar software. Our tool has a branch-and-cut algorithm for solving the combinatorial problem, and it makes use of a robust linear programming solver. We produced three releases depending on the linear programming solver.

TIC2002-00895

“A new tool for applying Controlled Rounding to a Statistical Table in Microsoft Excel”, Lecture Notes in Computer Science 3050 (2004) 44-57. More details in the webpage <http://webpages.ull.es/users/istac/CRP/index.htm>

Another new programmer (Sergio Delgado Quintero) was devoted to the implementation of the software calling the editing-and-imputation algorithms. This software is named TEIDE, and it is a complex tool to read and manage large microdata. The input to this code is a Microsoft Access database, as suggested by our close collaboration with practitioner of the Regional Statistical Office in Canary Islands (ISTAC). It deals with categorical and numerical data, no matter the structure of the edits. For details, see <http://webpages.ull.es/users/istac/TEIDE/index.htm>

In year 2003, we integrated a huge European project named “Network of Excellence” (NoE) calling for a project inside the VI Framework. The name was “Network for Improving Public Statistics (NIPS)”, but very unfortunately it was rejected. We are now in the process of asking for a new project but with a small number of partners. Unfortunately, recent problems in EUROSTAT have created a lot of difficulties for succeeding in the approval of research projects in the area. These problems have appeared in newspapers, with bad articles like “Two high-ranking officials in Eurostat, the EU statistical office, have been removed from their jobs following a series of corruption charges which OLAF is investigating. Yet while the EU's anti fraud unit investigators pick through the case, no one expects them to look too deeply into whether the alleged activities were controlled or influenced by others, perhaps outside Eurostat. The Sprout examines the case and unearths a theme, leaving a dark cloud looming over other more powerful Luxembourg-based institutions.” See, for example, <http://www.thesprout.net/010/graft/graft06.htm>

The excellent work conducted in this work gave us also the opportunity of presenting our work to the most important audience in Spain: the meeting of the Regional Statistical Offices in Spain. Indeed, we were invited to present a plenary session in the “Jornadas de Estadística de las Comunidades Autónomas”, Toledo, June 2-5, 2004. This was an excellent occasion used by J.J. Salazar to present the main results obtained from this research project on editing and imputation. A summary of the presentation is available through internet, together with the main conclusions of the conference, in <http://www.ies.jccm.es/jecas/paginas/conclusiones.htm>

The team (at University of La Laguna, Tenerife) have received the visit of several experts in the topic of this research project. For example, David Brown (ONS, English statistical office), Jorge Saralegui (INE, Spanish statistical office), Fayna Alama (ISTAC, statistical office in Canary Islands), among other practitioners, have been involved in working meetings with the team in Tenerife. Also other professors from many other places have visited the team developing this project. For example, Daniele Vigo (Italy), François Louveaux (Belgium), Martine Labbe (Belgium), Frederic Semet (France), Gilbert Laporte (Canada) and Giovanni Andreatta (Italy). On the other hand, researchers of our team have been also outside. For example, Jorge Riera Ledesma was three months in Edinburgh, and Inmaculada Rodríguez was one month in Montreal.

We have started a new collaboration with a team at University of Bologna for using Constraint Programming in the editing-and-imputation problem. The Italian researchers are Andrea Lodi and Michela Milano, both internationally known for their contributions to Constraint Programming. The use of this technique in editing-and-imputation is new and we are planning to have a draft of a manuscript very soon.

We have participated in the following activities:

- “Joint ECE/Eurostat work session on statistical data confidentiality”, Luxembourg, 7-9 abril 2003.
- “27 Congreso Nacional de Estadística e Investigación Operativa”, Lleida, 8-11 abril 2003
- Conferencia en Oficina de Estadística de Noruega, Oslo (Noruega), 3-4 junio 2003
- “ECCO XVI” (European Chapter on Combinatorial Optimization), Molde (Noruega) 5-7 junio 2003.
- Reunión de trabajo, Oficina de Estadística de Holanda, Voorburg, 18 junio 2003
- Participación en Tribunal de Tesis Doctoral sobre “Editing and Imputation”, Rotterdam (Holanda) 19 junio 2003
- “EURO/INFORMS 2003”, Estambul (Turquía) 6-10 julio 2003
- Reunión de trabajo, Instituto Canario de Estadística, Las Palmas de Gran Canaria, 17 julio 2003
- Conferencias en la Oficina de Estadística del Reino Unido, Londres (UK) 2-11 agosto 2003
- “54th International Statistical Institute 2003”, Berlín (Alemania) 13-20 agosto 2003
- “18th International Symposium on Mathematical Programming”, Copenhagen (Dinamarca) 18-22 agosto 2003
- “6th International Conference of Operations Research”, Havana (Cuba) 15-19 septiembre 2003
- Conferencia invitada, Universidad de Sevilla, 25-28 septiembre 2003
- “Work Session on Statistical Data Editing”, Madrid, 20-22 octubre 2003
- Reunión de trabajo, Instituto de Estadística Holandés, Voorburg (Holanda), 2-6 noviembre 2003
- Reunión SADERYL, Universidad Politécnica Cataluña, Barcelona 13-15 noviembre 2003
- Reunión de trabajo, Instituto de Estadística Inglés, Londres (U.K.), 15-19 diciembre 2003
- “8rd Aussois Workshop on Combinatorial Optimization”, Aussois (Francia) 5-9 enero 2004
- Reunión de trabajo, Universidad de Bolonia (Italia), 6-21 marzo 2004
- Conferencia invitada, Universidad de Florencia (Italia), 16 marzo 2004
- “Combinatorial Optimization 2004”, Lancaster (U.K.), 29-31 marzo 2004
- “I Jornadas SEIO-RSME sobre Programación Matemática”, Elche, 6-7 mayo 2004
- “Jornadas de Estadística de las Comunidades Autónomas”, Toledo, 2-5 junio 2004
- “Privacy in Statistical Databases PSD2004”, Barcelona, 9-11 junio 2004
- “Triennial Symposium on Transportation Analysis TRISTAN V”, Guadalupe (Francia), 14-18 junio 2004
- “Applied Modelling APMOD2004”, Londres, 21-23 junio 2004
- “EURO XX conference on Operational Research”, Rodas (Grecia), 4-7 julio 2004
- “Optimization 2004”, Lisboa (Portugal) 26-28 julio 2004
- Reunión de trabajo, Instituto de Estadística Inglés (ONS), Londres 7-8 octubre 2004

We have published the following papers:

- J.J. Salazar González, *Extending Cell Supresión to Protect Tabular Data against Several Attackers*, Inference Control in Statistical Databases (edited by J. Domingo-Ferrer) **Lecture Notes in Computer Science** 2316 (2002) 34-58
- M. Fischetti, J.J. Salazar, *Partial Cell Suppression: a New Methodology for Statistical Disclosure Control*, **Statistics and Computing** 13 (2003) 13-21
- M. Fischetti, G. Romanin Jacur, J.J. Salazar, *Optimisation of the interconnection network of a UMTS radio mobile telephone system*, **European Journal of Operational Research** 144 (2003) 56-67
- H. Hernández Pérez, J.J. Salazar González, *A Branch and Cut Algorithm for the Traveling Salesman Problem with Pickups and Deliveries*, **Discrete Applied Mathematics** 144 (2004) to appear
- G. Laporte, J. Riera Ledesma, J.J. Salazar González, *A Branch-and-Cut Algorithm for the Undirected Traveling Purchaser Problem*, **Operations Research** 51/6 (2003) 940-951
- A. Corberán, E. Mota, J.J. Salazar González, *Some recent contributions to routing and location problems*, **Networks** 42 (2003) 109-113
- G. Laporte, J.J. Salazar González, F. Semet, *Exact Algorithms for the Job Sequencing and Tool Switching Problem*, **IIE Transactions** 35 (2003) 1-9
- J.J. Salazar González, *The Steiner cycle polytope*, **European Journal of Operational Research** 147(2003) 671-679
- J. Riera-Ledesma, J.J. Salazar-González, *Algorithms for automatic data editing*, **Statistical Journal of the United Nations Economic Commission for Europe** 20 (2003) 255-264
- J.J. Salazar González, *Mathematical models for applying cell suppression methodology in statistical data protection*, **European Journal of Operational Research** 154 (2004) 740-754
- M. Labbe, G. Laporte, I. Rodríguez Martín, J.J. Salazar González, *The Ring Star Problem: Polyhedral Analysis and Exact Algorithm*, **Networks** 43 (2004) 177-189
- H. Hernández Pérez, J.J. Salazar González, *Heuristics for the one-commodity Pickup-and-Delivery Traveling Salesman Problem*, **Transportation Science** 38 (2004) 245-255
- M. Labbe, I. Rodríguez Martín, J.J. Salazar González, *A Branch-and-Cut Algorithm for the Plant-Cycle Location Problem*, **Journal of the Operational Research Society** 55/5 (2004) 513-520
- J. Riera Ledesma, J.J. Salazar González, *A Heuristic Approach for the Traveling Purchaser Problem*, **European Journal of Operational Research** (2004) to appear
- J. Riera Ledesma, J.J. Salazar González, *The Biobjective Travelling Purchaser Problem*, **European Journal of Operational Research** 160/3 (2004) 599-613
- J. Riera Ledesma, J.J. Salazar González, *Solving the Asymmetric Traveling Purchaser Problem*, **Annals of Operations Research** (2004) to appear.
- J.J. Salazar, P. Lowthian, C. Young, G. Merola, S. Bond, D. Brown, *Getting the Best Results in Controlled Rounding with the Least Effort*, "Privacy in Statistical Databases" (editado por J. Domingo-Ferrer) **Lecture Notes in Computer Science** 3050 (2004) 58-72
- J.J. Salazar-González, M. Schoch, *A new tool for applying Controlled Rounding to a Statistical Table in Microsoft Excel*, "Privacy in Statistical Databases" (editado por J. Domingo-Ferrer) **Lecture Notes in Computer Science** 3050 (2004) 44-57

TIC2002-00895

- J.J. Salazar (capítulo en libro “Sociedad, Ciencia, Tecnología y Matemáticas”, editado por M.I. Marrero y R. Trujillo) *Optimización Matemática: Ejemplo y Aplicaciones* 2004 I.S.B.N. 84-7756-555-4
- I. Rodríguez Martín, J.J. Salazar González, *Decomposition Approaches for a Capacitated Hub Problem* “IBERAMIA 2004” (editado por C. Lemaître, C.A. Reyes y J.A. Gonzalez) Springer **Lecture Notes in Artificial Intelligence** 3315 (2004) 154-163
- M. Labbe, G. Laporte, I. Rodríguez Martín, J.J. Salazar González, *Locating Median Cycles in Networks*, **European Journal of Operational Research** 160/2 (2005) 457-470
- A. Caprara, J.J. Salazar González, *Laying Out Sparse Graphs with Provably Minimum Bandwidth*, **INFORMS Journal on Computing** (2004) to appear
- J. J. Salazar González, *A Unified Mathematical Programming Framework for different Statistical Disclosure Limitation Methods*, **Operations Research** (2005) to appear

We also have other manuscript under the revision process for other international journals:

- J. Riera-Ledesma, J. J. Salazar González, *A branch-and-cut algorithm for the Error Location Problem in Data Cleaning*, **Management Science**.
- H. Hernández-Pérez, J. J. Salazar González, *The one-commodity Pickup-and-Delivery Travelling Salesman Problem: inequalities and algorithms*, **INFORMS Journal on Computing**.
- J. J. Salazar González, *Controlled Rounding and Cell Perturbation: Statistical Disclosure Limitation Methods for Tabular Data*, **Mathematical Programming**
- R. Baldacci, M. Dell’Amico, J. J. Salazar González, *The Capacitated m -Ring Star Problem*, **Operations Research**
- A. Letchford, J. J. Salazar González, *Projection of Flow Variables for Vehicle Routing*, **Mathematical Programming**

We are now in the final process of sending some manuscript to international journals:

- M. Iori, J. J. Salazar González, D. Vigo, *An exact approach for capacitated vehicle routing problems with two-dimensional loading constraints*, to be submitted to **Operations Research**.
- C. Feremans, M. Labbé, A. Letchford, J. J. Salazar González, *The Generalized Subgraph Problem: Valid inequalities and separation*, to be submitted to **Mathematical Programming**
- C. Feremans, M. Labbé, A. Letchford, J. J. Salazar González, *The Generalized Subgraph Problem: Complexity and Approximability*, to be submitted to **Mathematical Programming**

A member of the research team (Hipólito Hernández Pérez) will present his Ph. Dissertation at University of La Laguna in December 2004. The title of the dissertation is “The travelling salesman problem with pickups and deliveries”, and the supervisor of this dissertation was Juan José Salazar González. We have already published some articles from this dissertation, but we still have some manuscript that we want to finish and submit for publication in international journals.

In addition, Juan José Salazar González is also supervising the Ph. Dissertation of two other students visiting University of La Laguna from University of Cumaná (Venezuela). They are Manuel Centeno and Fernando Marval, and their works concern solving hard combinatorial problems. We plan that they will finish and present their dissertation by the end of year 2005.

4 Tasks to be done in the next months

We are planning to apply for a research project in the VI framework in the next call (April 2004). The main coordinator will be Anco Hundepool (CBS, Netherlands), and the research team developing this National project will be a partner of the European project.

We have been invited to participate in two relevant conference:

- Special Topic Contributed Paper Meeting on editing at the Sydney ISI in April 2005. This is the largest conference organized by the International Statistical Institute. <http://www.tourhosts.com.au/isi2005/>
 - UNECE Work Session on Statistical Data Editing, which will be held 16-18 May 2005 in Ottawa, Canada. <http://www.unece.org/stats/documents/2005.05.sde.htm>
-

We also plan to attend other national and international conference concerning also closely related topic, like the EURO and INFORMS conference. In particular we mention two of them:

- [INFORMS Computing Society Conference 2005](#) January 5-7, 2005
 - [17th Triennial Conference of the International Federation of Operational Research Societies 2005](#), July 11-15, 2005
-

Our main task for 2005 will be to conclude the prototype of TEIDE. We have already conducted experiments on real data for ISTAC, but we still have many ideas in mind to experiment. We have incorporated in the research team some young programmers (Sergio Delgado Quintero and Markus Schoch) which are now following the doctoral program of the “Departamento de Estadística, Investigación Operativa y Computación” (University of La Laguna), and the current aim is to go on with their Ph. Dissertations. We want to emphasize here that a member of our team (Jorge Riera Ledesma) got the EURO Doctoral Dissertation Award during the meeting of the society in Rhodas, July 2004. The supervisor of this Ph. Dissertation by Jorge Riera Ledesma was the main researcher of this team, Juan José Salazar González. This European award is a clear measure that the research that we are conducting at University of La Laguna is being successful. More details of the award can be obtained from <http://www.euro-online.org/display.php?page=edda3>

We are now also working with other young researcher, Susana Pérez García. She is working on new combinatorial problems, and we hope that during the last year of the project we will start obtaining new results, and in the future she will defend a Ph. Dissertation on the topic.

During July 2004 we started collaborating with a research team from “Universidad Miguel Hernández” (Elche), and in particular with Prof. Dr. Domingo Morales. He is investigating the very interesting problem of estimations in small areas. Even if the original formulation of this problem is the statistical version, we hope to contribute to the problem from the point of view of the Operational Research and from the Computer Science. We are considering some mobility plans between our teams during year 2005.

We thank the “Plan Nacional de Ciencia y Tecnología” (“Ministerio de Educación y Ciencia”) for supporting our research at University of La Laguna (Tenerife) with this project.
