Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp139-144)

Design of a Universal High-Speed Telecommunication Network Using Genetic Algorithms. The Case of the Digital Divide in Andalusia

PABLO CORTÉS[1], JESÚS MUÑUZURI[2], JOSÉ GUADIX [3], NICOLÁS IBÁÑEZ[4]
Department Ingeniería de Organización
University of Seville
C/ Camino de los Descubrimientos s/n 41092 Sevilla
SPAIN
1

Abstract: - The digital revolution holds great promise for prosperity and progress. However, the Information Society evolution is supported through the telecommunication networks deployment. The high cost of such infrastructures forces the companies to focus their efforts in the profitable society sectors. For that, the promotion of performances to bridge the digital divide between rural and urban areas appears as a necessary action line. We present a case study for a high-speed telecommunication network design in which a decision support system based on a genetic algorithm is implemented providing cost effective solutions. We make use of real life data from the telecommunication industry and present different solutions separated by coverage.

Key-Words: - digital divide; genetic algorithm; case study; telecommunication network; ADSL

1 Introduction

The digital revolution holds great promise for prosperity and progress. Access to information and communication technologies (ICT) can launch small and medium-sized firms directly into the heart of regional, national and global markets. Telemedicine can provide access to up-to-date health and medical information to even the most remote communities. ICTs can facilitate low-cost distance learning. And they can empower civil society, strengthen democratic institutions and make governments more transparent and accountable.

The Scientific Community rarely deals with these aspects. An exception to the rule is [1], where the use of new information technologies was proposed to gain competitiveness in rural small businesses. However the proposals were done in terms of business solutions but not in terms of infrastructure necessities.

Within the Spaniard framework, the fact must be highlighted that private companies have focused their network extension and modernization strategic plans on the more profitable areas. With this background, the Andalusian Regional Government has been leading a series of initiatives aimed at promoting the development of fiber optic networks through the region. Especially in the less favored areas.

2 The digital divide in Andalusia

In Europe the digital divide is mainly found between the rural and urban zones. As a matter of fact the rural areas and its municipalities tend to be non profitable for the private companies. So the telecommunication networks do not reach those zones.

Following the EU Commission Staff working paper [2], the Public authorities, and particularly regional (it is the case of the Junta de Andalucía, the regional government of Andalusia) and local authorities, have a key role to play in the development of the information society by (1) using information society applications and services in the process of modernization of services provided to citizens and companies, (2) promoting the information society in the region and (3) monitoring the evolution of the communications networks and services provision on the region in order to avoid exclusion and contribute to the balanced development of regional activities.

Following these guidelines, we have divided the Andalusian region into three sets. The first, municipalities of type A, are those where exist alternative infrastructure. So there is commercial competition, basically between ADSL and coaxial cable networks. Secondly, we state municipalities of type B as those where only ADSL exists but they could have enough services demand to support another telecommunications operator. Finally, the municipalities of type C are those which has not ADSL infrastructure due to a strong lack of demand.

Furthermore we divide the municipalities of type B between B1 and B2. Municipalities of type B1 are those where the consideration of another telecommunication operator would be profitable and B2 those where probably would not be profitable. The separation between B1 and B2 was done

Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp139-144)

considering the expected Payback period, the Net Present Value (NPV) and the Internal Rate of Return (IRR) of each municipality type B. See [3] for a detailed discussion on it.

Next figure 1 represents the main characteristics of the municipalities (the demand is expressed as the expected number of equivalent ADSL lines for the final period of the horizon). Each equivalent ADSL line will be assumed of 256 Kb/s. So, final users with capacities upper than 256 Kb/s are converted in the correspondent number of equivalent ADSL lines.

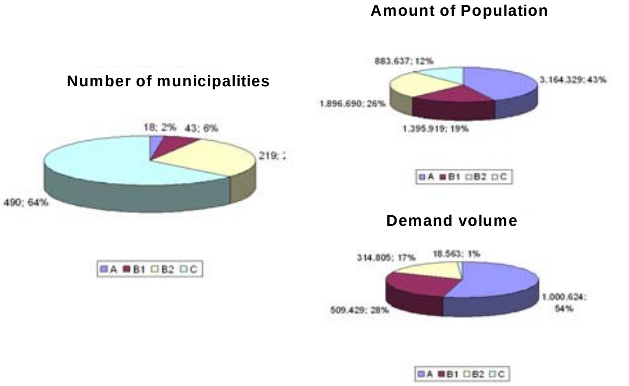

Figure 1. Municipality classification

3 The Model

To deal with the problem, we defined a graph G=(N,A). The nodes can represent municipalities (M) or simple intersections (S) of the underlying transport network where the telecommunications network will be placed. So they will be active (municipalities) or passive nodes (typically intersections of the graph not supplying or receiving information). For our case we used the road network of Andalusia. The arcs represented the roads linking two different nodes. The initial graph was supplied by the Instituto de Cartografia de Andalucia (ICA) in Arc/Info format.

3.1 Previous conditions and hypothesis

Previously to the design of the network some strategic decisions were taken:

(1) The transmission media would be fiber optic. That allows to ensure a transmission with quality and capacity (speed in the order of hundred of Gbit/s).

(2) The topology should permit the rerouting of the flows among the main municipalities in case of fails. Designing a ring-tree topology can get this. The ring would include the municipalities of type A, those with a greater demand. Bi-directional self-healing rings have been the most used topology in the case of traditional public operators due to its resilient wide coverage.

(3) The ring should follow alternative fiber optic rings being property of operators different from the leader operator. This is the case of the available fiber optic in the railway and electrical networks. This decision will make possible a faster deployment of the network in the less favored areas carrying out the ring deployment in the last stage. This decision fixes the ring topology to the observed one in figure 2.

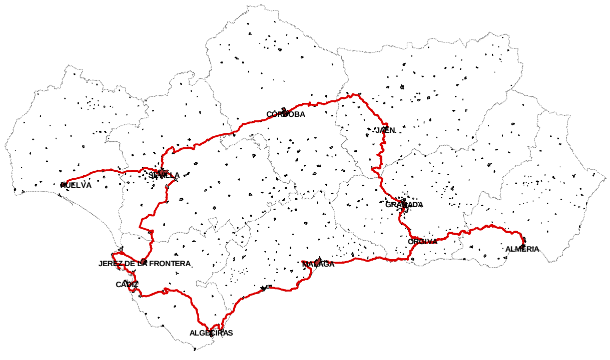

Figure 2. Resilient self-healing ring in the backbone

This ring has two characteristics. One of them is the direct connection between Seville and Huelva. This is to avoid the Doñana Natural Park, a protected biosphere reserve. The other one is the direct connection Orgiva-Almeria. This is due to the savings that produces to avoid a cyclic connection as far as Almeria. Both problems are solved by deploying a parallel double fiber optic cable.

(4) The ring would be dimensioned taking into account that the Internet gates were in Seville and Malaga.

3.2 Cost model

The relevant costs in the infrastructure deployment are classified into four categories:

1. Node costs:
   a. Fixed cost associated to the existence of the node, F_i, $\forall i \in M$. It corresponds to the terminal cable room adaptation to connect the fibers into the DSLAM.
   b. Cost depending on the capacity managed by the node. $G_i(C_i)$, $\forall i \in M$. It corresponds to the modular equipment, where the transmitters and receivers are placed.
2. Arc costs:
   a. Cost depending on the arc length, $L_{ij}(D_{ij})$, $\forall (i,j) \in A$. It corresponds to the ditches and conduits.

b. Cost depending on the arc length and its supported capacity, $C_{ij}(D_{ij},C_{max})$, $\forall(i,j)\in A$. It corresponds to the fiber optic links including transmitters, receivers and regenerators or optical amplifiers when needed.

Two different technological alternatives are considered to calculate the 2.b costs.

The first option cost expression corresponds to a high quality fiber optic link (without regenerators). It can be estimated as follows:

$$C_{ij}\left(D_{ij}^{sG},C_{\max}\right)=\frac{P_{SDHo}}{n_o}+\frac{P_{SDHd}}{n_d}+P_{fo}^{sG}\cdot D_{ij}^{sG}+C_{oam}\cdot\left\lceil\frac{D_{ij}^{sG}}{D_{\max}}\right\rceil+2C_{tr}\cdot\left\lceil\frac{C_{\max}}{C_{10Gb}}\right\rceil \quad (1)$$

The second option cost expression is for a lower quality fiber optic link (with regenerators) that can be estimated as follows:

$$C_{ij}\left(D_{ij},C_{\max}\right)=\frac{P_{SDHo}}{n_o}+\frac{P_{SDHd}}{n_d}+P_{fo}\cdot D_{ij}+C_{reg}\cdot\left\lceil\frac{D_{ij}}{D_{\max}}\right\rceil+2C_{DWDM}\cdot\left\lceil\frac{C_{\max}}{C_{10Gb}}\right\rceil \quad (2)$$

Where P_{SDHo} and P_{SDHd} are the origin and destination SDH multiplexer costs; n_o and n_d are the number of links in the SDH equipment placed in node. As one SDH multiplexer can serve diverse digital links only should be attributed the correspondent part to it; D_{ij}^{sG} is the arc distance. It is covered without regenerators; D_{ij} is the arc distance. It can be covered with regenerators; P_{fo}^{sG} is the linear cost of a high quality cable of 24 monomode fibers plus the cost of connecting it; P_{fo} is the linear cost of a standard cable of 24 monomode fibers plus the cost of connecting it; C_{oam} is the cost of an optical amplification system; C_{reg} is the cost of a regenerator; D_{max} is the amplification maximum distance. Typically 140 kilometres for option 1 and 40 kilometres for option 2; C_{tr} is the cost of DWDM interface card (transmitter/receiver) with a capacity of 10 Gb/s; C_{max} is the capacity that must be equipped.

3.3 Topological model

As stated previously, a ring-tree topology was selected. This topology consists of a hierarchical structure given by nodes of level 1 and nodes of level 2 trying to simplify the subsequent network management. See next figure 3.

The nodes of level 1 are the ten larger cities of Andalusia: Almeria, Algeciras, Cadiz, Jerez de la Frontera, Cordoba, Granada, Huelva, Jaen, Malaga and Seville. All of them are municipalities of type A. These nodes depict the main ring of the backbone as figure 3 reveals. The rest of nodes in the network must be connected at one of these level 1 nodes through a tree structure.

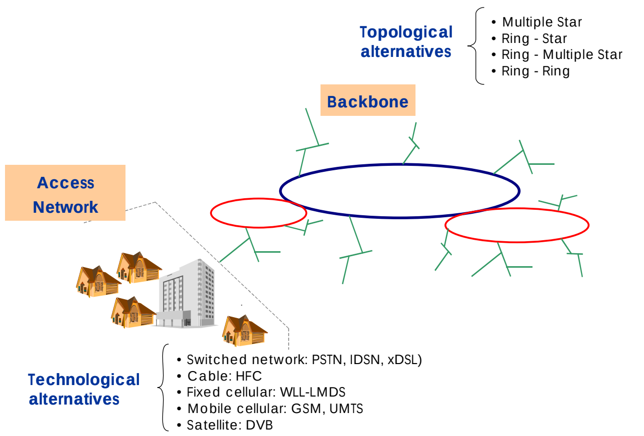

Figure 3. Access network and backbone

The rest of nodes (not class A) can be located into the ring structure or out of the ring. In the first case, the basic infrastructure has not to be constructed. There already exist the ditch and conduit, so we only have to count the cost 2.b. In the other case (out of the ring) we would have to count the costs of civil infrastructure (2.a) and transmission link (2.b). Figure 4 depicts the description.

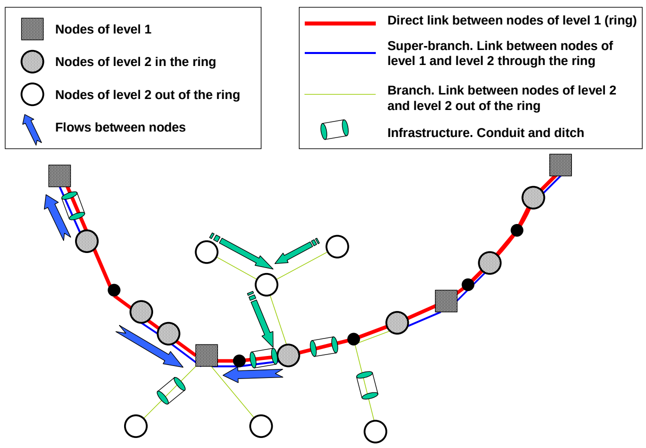

Figure 4. Topological structure

4 Decision Support System based on a genetic algorithm

We implemented a decision support system based on a genetic algorithm engine. Genetic algorithms are recently been considered to deal with different kinds of telecommunication problems. Most of them are in the scope of theoretical applications. A wide-scope genetic algorithm was presented by [4]. Works in the same way are [5] and [6]. [7] presents a topology planning tool, GenOSys based on a genetic algorithm to design network infrastructure.

Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp139-144)

4.1 Decision support system structure

The source data are the cartography supplied by the ICA and the socio economic database provided by the Instituto Estadistico de Andalucia (IEA).

After processing the cartography and considering the technological model and the strategic decisions, we get the topological model, in terms of a graph. Due to the high grade of precision in the cartography, some simplifications were required. The result was a graph with 2,796 arcs and 1,920 nodes. 770 of those nodes are municipalities, the rest are simple intersections. In the other side, the socio economic data of the municipalities plus the historical demand data are used to estimate the telecommunication demand for each of the municipalities. The demand values plus the financial analysis of each urban node allows us to classify the municipalities between profitable and non-profitable. This allows identifying the digital divide. These three components are introduced in the algorithm that provides the solution. See figure 5.

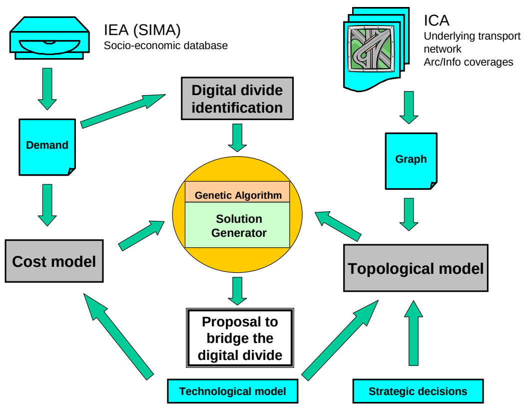

Figure 5. Decision Support System structure

4.2 The genetic algorithm

The dimensioning of the network as well as the multiple topological alternatives to deploy such network makes the problem of great computational complexity. The problem can be proved as NP-Hard by simple reduction to the Steiner problem. Furthermore, given a feasible solution is easy to compute its evaluation (it can be done in polynomial time), so the problem is NP-Complete too. The utilization of genetic algorithms has been proved as an efficient way to obtain good solutions in problems with high number of complex constraints as this one.

The algorithm strategy consists of generating diverse feasible solutions (easy of evaluating) and recombining them without losing feasibility.

A feasible solution of our problem must include the predefined ring plus the tree (super-branches and branches, see figure 4).

The complete procedure to give the final solution follows the next three steps:

(1) Constructing a supra-node that corresponds to the ring. The ring is conceptualized in one only node formed by all the nodes in the ring (including municipalities and passive nodes) and with so many incident arcs as incident arcs has each of the nodes in the ring.

(2) Constructing a tree that contains the supra-node and all the municipalities. The passive nodes (intersection of the transport network) can be adequately used if necessary. The genetic algorithm generates the network proposal.

(3) Dimensioning the capacity of the ring.

4.2.1 Characteristics of the genetic algorithm

The individuals genetic encoding consists of an array with so many registers as arcs exist in the graph. If the register is set to one then the arc is included in the solution, otherwise the arc is not in the solution (see figure 6). Additionally, the genome must ensure the feasibility of the solution. So non-feasible solutions are not accepted.

After testing with different sizes, we considered an initial population of 20 individuals. To generate them we used the Kruskal algorithm to solve the minimum spanning tree (MST) problem (note that the ring is now a supra-node). Given 20 different sets of fictitious cost values to the arcs, we solve the correspondent MSTs and afterwards the solutions are cleaned by erasing the passive nodes located in a terminal position.

The population size is a major factor in the effectiveness of genetic algorithms. Our experiments show that increasing the population size beyond 20, although increasing the computational effort, is not rewarded by a corresponding increase of performance.

We implemented two different genetic operators (see figure 6): crossover and mutation operators. The parent selection mechanism is random in order to enrich the genetic variety. It includes a mechanism of incest prevention to avoid the crossover of genetically similar parents.

Uniform crossover operator: if the genome of both parents includes the arc in their solution (i.e. it has the same register set to 1, or set to 0), then the arc is maintained for the offspring. Otherwise (only one parent has the arc) the arc is maintained with a probability equal to ½.

Mutation operator: the operator changes a register from 1 to 0, or viceverse. I.e. once a register is randomly selected if the arc belongs to the solution, then the arc is erased. Otherwise, if the arc does not belong to the solution, then the arc is included.

The application of the genetic operators can lead to unfeasible solutions. The possible infeasibilities are the following. (1) After applying a crossover operator two trees are connected. This situation generates a new graph and cycles appear. The cycles are broken by erasing a randomly selected arc of each cycle. (2) After applying a mutation operator two situations can appear: (a) if an arc is included in the solution, a cycle can appear. Then, the cycle is broken by erasing a randomly selected arc of the cycle; or (b) if an arc is erased in the solution, the tree is disconnected. Then, the tree is reconnected by means of the shortest path algorithm.

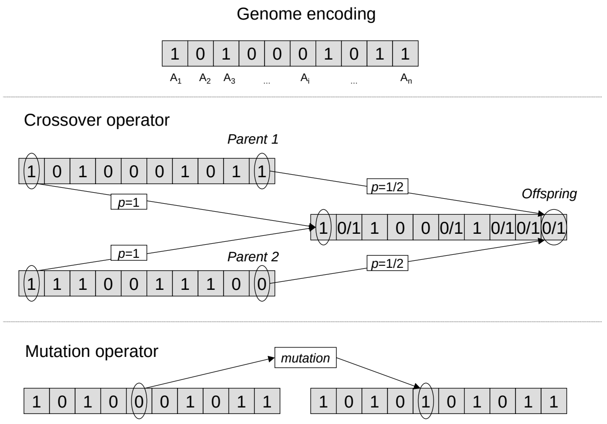

Figure 6. Genome encoding and operators

Tests were carried out with different probabilities for crossover and mutation operators. It was found that varying the crossover probability from 50% to 100% had little effect on performance, with a value of 80-90% being marginally optimal for the tests carried out. A value of 80% was used in the main runs. For mutation, values between 5% and 20% were seen to be giving better results than typically smaller values. A value of 20% is used in the main runs in order to enrich the genetic variety of the population. However, it is to be noted that the genetic algorithm is robust in the sense that the test problem solutions were achieved on the whole with a wide range of parameter values, and with no fine-tuning required to achieve efficiency.

The substitution of new individuals (offspring) by the old individuals (in previous population) is made by a ranking based replacement. We used a hypergeometric function allowing more probability of replacement for individuals with worse fitness and less probability of replacement for individuals with better fitness. So, the individual in ranking position-i, had a replacement probability equal to $q\cdot(1-q)^i$, where q is the replacement probability of the worst individual. The value of parameter q is progressive. It starts with a low value of q in order to explore the feasibility region and then the value is progressively increased to gain in exploitation of the niche quality.

Additionally to the replacement rule, we incorporated an individual duplicity control in the population generation.

The number of generations of the algorithm can be a critical parameter when we try to reach efficiency of the solution and short time execution. We selected a number of 50 iterations. The learning curve of the algorithm is depicted in figure 7.

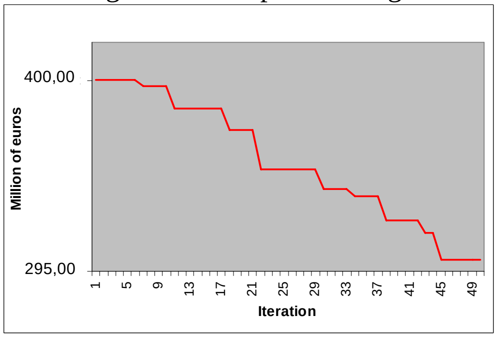

Figure 7. Learning curve

Table 1 summarizes the value of parameters selected.

Table 1. Genetic algorithm parameters

Population size, N	20
Crossover operator probability (mutation operator)	0,8 (0,2)
Ranking based replacement probability, q	0,2/0,4/0,6/0,8[a]
Maximum number of generations	50

[a]Progressive increase of q each ten generations.

4.2.2 Final dimensioning of the ring capacity

As stated, we assumed a bi-directional self-healing ring. This allows rerouting the demand in case of fail. The problem is to determine the capacity of the ring that allows rerouting all the flows in case of fail. As the ring capacity is directly related to the ring cost, we minimize the capacity that guarantees the demand flows. The generic case is depicted in figure 8.

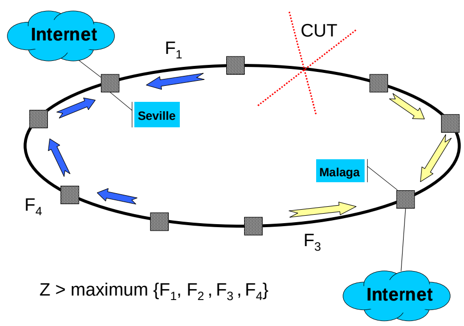

Figure 8. Generic rerouting process in the ring

We follow a similar procedure to [8]. There are four flows: F1, F2, F3 and F4. F1 and F2 are imposed, and F3 and F4 are selected in order to process the half in each of the directions. The ring capacity must be equal in all the links of the ring. So it is defined as the total demand flow in the most saturated arc. This must be calculated for all possible cuts in the network.

5 Network solution

The decision support system supplied us a cost effective solution whose topological shape can be observed in figure 9. We show three stages for the deployment named as S1, S2 and S3. They will correspond to the 80%, 90% and 100% of the population respectively.

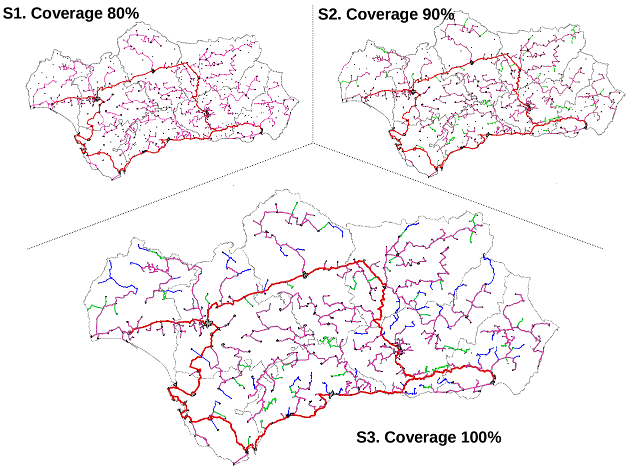

Figure 9. Graphical solution: Deployment by stages

Solution S1 includes 616 municipalities. The stage S2 deploys the network over 700 municipalities and finally S3 over 770 municipalities (the overall of Andalusia). The data corresponding to S3 implies the following values: 7,101 kilometres of ditches and conduits (civil engineering). The ring has 1,158 kilometres from the total. And 14,698 kilometres of fiber optic (digital link). It corresponds at 3,953 kilometres to the ring.

Table 2 summarizes the main economical results detached by type of cost and by type of network section. A total amount upper than 295 million of euros will be needed to deploy the infrastructure of a telecommunication network capable of providing universal access of ITCs and bridging the digital divide that the rural zones suffer today.

Table 2. Main economical results corresponding to S3

Costs	Civil engineering	Digital link	Node location	Total
Ring	22,935,553 €	41,353,871 €	1,045,000 €	65,334,424 €
Rest	117,671,808 €	84,723,572 €	27,466,000 €	229,861,380 €
Total	140,607,361 €	126,077,443 €	28,511,000 €	295,195,804 €
%	47.63%	42.71%	9.66%	100.00%

6 Conclusions

Private companies are not responsible of the digital divide experimented between diverse parts of the society. They cannot provide new high-speed telecommunication network infrastructures as far as not profitable points of the rural areas. For these aspects the governments (especially the local and regional governments) have to assume their responsibilities to avoid the digital divide between the rural and urban zones, as the European Commission and the United Nations have often adverted. We have proposed a decision support system based on a genetic algorithm decision engine to construct alternative fiber optic networks that permits the progressive inclusion of the rural areas of Andalusia in the Information Society. The system was revealed as an adequate instrument to do so, and the proposals are carrying out to practice.

References:

[1] Premkumar, G., Roberts, M. Adoption of new information technologies in rural small businesses, *Omega*, 27 (4), 1999, 467-484.

[2] Commission of the European communities, 2003. Commission staff working paper: Guidelines on criteria and modalities of implementation of structural funds in support of electronic communications.

[3] Onieva, L. Vozmediano, J.M., Cortes, P., Estepa, R. 2003. Estudio sobre el despliegue universal de los servicios avanzados de telecomunicaciones en Andalucía. Technical Report, DUSAT-ESI. Seville: Dirección General de Sistemas de Información y Telecomunicaciones, Junta Andalucía.

[4] White, A.R.P., Mann, J.W., Smith, G.D. Genetic algorithms and network ring design, *Annals of Operations Research*, 86, 1999, 347-371.

[5]Armony, M., Klincewicz, J.G., Luss, H., Rosenwein, M.B. Design of stacked self-healing rings using a genetic algorithm, *Journal of Heuristics*, 6 (1), 2000, 85-105.

[6] Ghashghai, E., Rardin, R.L. Using a hybrid of exact and genetics algorithms to design survivable networks, *Computers & Operations Research*, 29 (1), 2002, 53-66.

[7] Poon, K.F., Conway, A., Wardop, G., Mellis, J. Successful application of genetic algorithms to network design and planning, *BT Technology Journal*, 18 (4), 2000, 32-41.

[8] Vachani, R., Shulman, A., Kubat, P. Multicommodity flows in ring networks, INFORMS Journal on Computing, 8 (3), 1996, 235-241.