A Study of the effects of clustering and local search on Radio Network Design evolutionary computation approaches

Y. Saez, F. Zazo, P. Isasi,

Departamento de Informática, Inteligencia Artificial, Universidad CARLOS III de Madrid
Ave. de la Universidad 30 Leganés 28911, SPAIN
yago.saez@uc3m.es, fernando.zazo@alumnos.uc3m.es, isasi@ia.uc3m.es

Abstract

The goal of this paper is twofold. First, we want to make a study about how evolutionary computation techniques can efficiently solve the radio network design problem. For this goal we test several evolutionary computation techniques within the OPLINK experimental framework and compare them. Second, we propose a clustering approach and a 2-OPT in order to improve the results obtained by the evolutionary algorithms. Experiments carried out provide empirical evidence of how clustering-based techniques help in improving all algorithms tested. Extensive computational tests, including ones without clustering and 2-OPT, are performed with three evolutionary algorithms: genetic algorithms, memetic algorithms and chromosome appearance probability matrix algorithms.

1 Introduction

The extraordinary success of radio mobile telecommunication systems has led service providers to dynamically adapt network infrastructures to continuously increasing demand. As the demand of the Personal Communication Services (PCS) is increasing, capacity has now become a critical issue for all of these services. In response, all of the mobile service operators are investigating different techniques such as cell splitting, allocation of new spectrum (digital switchover), alternative access architectures or dynamic channel allocation. The design of large cellular networks is a complex problem with a large impact on the deployment costs, the availability and coverage, and the quality of service. In this scenario, the optimal selection of base stations (BSs) sites and the use of optimal frequency planning methods becomes imperative. As we know, the engineering of radio mobile telecommunication networks presents basically two problems: the radio network design problem

(RND) and the frequency assignment problem (FAP) [11]. In this paper, we address the first, the RND problem.

2 The radio network problem

The RND problem, also known as the antenna or BS placement problem (APP) can be informally described as follows: given a set of candidate places for antennas, the problem consists of deciding on the best positioning of BSs in order to accomplish some objectives and constraints. Although different versions can be defined for the RND, the most usual objectives are maximizing the coverage while minimizing the number of antennas used. The constraints concern a large number of predefined candidate positions, the number of BSs needed, the population density, peculiarities of the terrain, bandwidth requirements, handover areas, etc. This problem has been tackled in literature by different approaches, [5], [13], [14], [6], [17], [9]. Only a few attempts are known whose approach is an analytic formulation of the complete RND problem, [8]. In general, the analytical models show great improvement over the classical geometry-oriented designs, however, as the size of the search space for the complete problem was enormous (exponential), it was impossible to carry out a optimization. Therefore, computational numerical methods were required, [18], i.e. software tool GRAND, [7].

3 The OPLINK domain

The instance of the RND problem used in this paper can be found in [1],[12]. It belongs to a Technology and Science Council project called OPLINK and it makes the following assumptions:

- Geographic information about the terrain: in this domain a rectangular area modelled by a grid of 4.25 x 6.4 kilometers is used. Each cell represents approximately 15 x 15 meters. The grid is extracted from real data about

IEEE
computer society

the city of Malaga, Spain. It can be downloaded from http://oplink.unex.es.

- Information about the BSs: in this domain the antenna type is omnidirectional (circular) with radio 450 meters. For the coverage calculation parameters such as transmitted power, height, tilt, azimuth, or signal propagation (strength, noise, obstacles, etc.) are not taken into account.

- Information about aims: the objective function is calculated through equation 1 and is imposed by the project in order to normalize the results obtained by different institutions.

$$Fitness = \frac{(coverage)^2}{number\_of\_base\_stations} \qquad (1)$$

As can be seen, even when relaxing the problem, there is still a complex task to solve since there are one thousand candidate positions for the BSs, which means a large search space of $2^{1000} \simeq 1.07 * 10^{301}$.

4 Evolutionary computation for radio network design problems

As mentioned before, the OPLINK environment is a complex NP-hard optimization problem with a large search space. Evolutionary Computation (EC) techniques are well-known for their performance when solving this kind of complex problems, and it is not the first time evolutionary algorithms have been applied to solve this problem, [2]. This paper concentrates on three evolutionary computation techniques: genetic algorithms (GAs), memetic algorithms (MAs) and chromosome appearance probability matrix algorithms (CAPMs). Starting from these techniques we propose two ideas which can improve the results of the techniques tested: clustering the BSs and performing a 2-OPT search.

4.1 Clustering

The imperative objective in this domain is to maximize the coverage and to minimize the number of antennas needed. Therefore, if we make an antenna clustering before the search, and we permit only one active antenna per cluster, we are minimizing the overlapping between antennas without causing negative effects in the solutions. The clustering technique used is the simple k-means [10] because it is simple to use and very well-known. In the Figure 1 a real distribution of 70 clusters is represented over the city of Malaga. To compare the results found by the simple k-means an expectation maximization(EM) technique has been also developed. The distribution of clusters with the EM is very similar and no relevant differences were found between both methods.
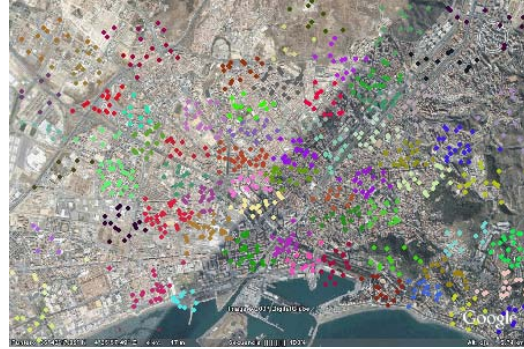

Figure 1. Distribution for 70 clusters with simple k-means (Malaga city)

4.2 2-OPT Search

All the proposed algorithms give their best solution for the problem. But we consider it useful if a lightweight algorithm can check if it is possible to improve the fitness of the final solution. The 2-OPT search is a well-known local search algorithm first proposed by Croes in 1958 [3]. Basically it consists of looking for neighbor solutions in a specific layout of transmitters. This search is done selecting two antennas under defined proximity criteria and testing all possible combinations around their neighbors for each active antenna. The neighbor vector stores a number of neighbor active antennas for each active antenna. The neighbors are calculated with a distance matrix. Once the active antennas are selected, all the possible combinations of the neighbor vector are evaluated. If the solution is improved then the search starts from the new individual found. Once the neighbor vector is completed, then the stop criteria is checked (fitness has improved or the number of iterations desired has been completed). To show the importance of how these ideas can help to all these algorithms, we present two set of experiments in Section 5, one with the raw EC techniques (bit representation) and the other with the EC techniques adapted to our proposals (clustered versions).

4.3 Genetic Algorithm

In the GA with cluster representation (GAC) an individual has the same number of genes as clusters. A gene is made of a list of transmitters in the cluster and the number of the active transmitter. If there is a transmitter which is inactive it is selected by -1. The first step in the algorithm is to determine the clusters. We have used different clustering methods but for experimental results the simple k-means method was chosen. Once the clusters are found, the population is randomly created and evaluated for the first time.

In each successive generation, part of the population is selected to breed the new offspring. The crossover does not need to be changed: it is the same as in the canonical GA: simple and uniform. The mutation operator now changes the active transmitter of a cluster randomly.

4.4 Memetic Algorithm

The memetic algorithm (MA) is a combination between local search (LS) techniques and evolutionary algorithms (EA). It is based on the concept of meme introduced by Dawkins, [4]. The key idea of a meme is that an individual can change his genetic code during his life thus improving the evolution process. To simulate this concept of meme an LS on the reproduction operators (crossover and mutation) is included. For many domains the MAs enhance the performance of the GAs. In contrast, LS methods are less generic and difficult to customize. They redefines the crossover and mutation operators proposed by the standard GA:

- The crossover operator is similar to the one used in the GA, but an LS is done in order to find the best possible crossover. The developed LS selects a predefined percentage of individuals randomly and then finds the best crossover for those individuals. At this point, when crossover is done, both parents are marked in order to avoid identical crossovers within the same iteration.

- The mutation operator is the same as the one used in the GA plus an LS. In this case, na LS is done to try to find the best possible mutation. The operator first calculates the number of mutations for each individual and then makes an LS limited to blocks of 2 genes. The size of the block can be changed and is determined depending on an admissible computational cost.

The other decisions about evaluation, selection and stop criteria are the same as the ones used for the GA.

4.5 Chromosome Appearance Probability Matrix Clustered

The Chromosome Appearance Probability Matrix method (CAPM) is GA modified in order to deal with micropopulations. Basically, it consists of a GA which uses a probability matrix which drives the mutation towards the solution, speeding up the convergence during the first generations. It includes a statistical-driven mutation operator and a clone remover operator. The CAPM method is less accurate than GAs, however, because on average, it converges faster. This is because the CAPM method has more selective pressure and reduces the effect of randomness in the mutation operator. All steps are explained in detail in [16] and [15]. In this approach, the probability matrix will have the same dimensions as the number of clusters determined by the simple k-means. This number

depends on the problem, but for the city of Malaga instance no more than 100 clusters were found in all executions.

5 Experimentation and Results

The first set of experiments with the raw EC techniques (bit representations) shows that the problem is not easy to solve. The Genetic Algorithm (GA) achieves better results than the Memetic Algorithm (MA) and the Chromosome Appearance Probability Matrix (CAPM) is not a feasible approach. The results achieved by both algorithms could be improved with more generations. The MA has a slower evolution than the GA because it has a high local search cost. If we observe Table 1, we can see how the average fitness is better for the GA during all iterations until the 2-OPT is run. The standard deviation (σ) is also lower for the GA; this is because in the MA, the convergence is sometimes premature. Finally, all the results given by the algorithms are refined with a 2-OPT search. The 2-OPT search improves the fitness and the standard deviation in almost all tested cases. In these experiments the 2-OPT improved the results of the MA surpassing the best results found by the GA with 2-OPT. The CAPM is impossible to develop with binary representation because the length of the chromosome is $1,000$ bits and the size of the chromosome matrix is not tractable.

Eval.	GA Average	σ	MA Average	σ
25000	**138.72**	**2.037**	101.31	3.961
50000	**140.74**	**1.280**	107.92	3.914
100000	**142.85**	**1.730**	120.29	4.257
200000	**145.57**	**1.313**	132.42	2.459
300000	**147.13**	**1.840**	137.67	1.678
400000	**148.24**	**1.735**	140.09	1.814
500000	**149.47**	**2.155**	141.54	1.885
750000	**151.33**	**1.784**	143.98	1.940
1000000	**152.65**	**1.474**	145.31	1.763
2500000	**156.27**	**1.447**	151.63	1.666
4750000	**157.45**	**1.150**	155.05	2.009
(2-OPT)	161.38	1.028	**162.54**	**0.532**

Table 1. Genetic algorithm and memetic algorithm results.

The second set of experiments with the clustered EC techniques achieves better results than the raw techniques, see Table 2. This is mainly because the search space when working with clusters is significantly lower. The Genetic Algorithm Clustered (GAC) improves all the results obtained by the GA. The Memetic Algorithm Clustered (MAC) has the same problem as the MA. The MAC does not achieve very good results because of the local search cost. It needs more evaluations to improve the fitness achieved by the GA and the GAC. The Chromosome Appearance Probability Matrix Clustered (CAPMC) is the best during the first $1,000,000$ evaluations, because the statistical oriented mutation performs surprisingly well. In fact, it

sometimes surpasses the results obtained by the GAC. Taking into account that the CAPM works only with 10 individuals its evolution and results are remarkable.

Eval.	GAC		MAC		CAPMC	
	Avg.	σ	Avg.	σ	Avg.	σ
25000	151.12	2.715	106.22	**1.579**	**155.50**	2.523
50000	153.99	2.768	110.54	**1.738**	**156.13**	2.327
100000	156.03	2.138	118.02	**1.785**	**156.92**	2.186
200000	**157.58**	2.463	129.59	**1.944**	157.29	1.945
300000	**158.13**	2.408	137.04	**1.885**	157.76	1.977
400000	**158.59**	2.178	142.05	**1.745**	158.01	1.895
500000	**158.69**	2.155	145.15	**1.629**	158.21	1.822
750000	**158.79**	2.097	149.73	**1.521**	158.41	1.866
1000000	159.00	2.172	152.12	**1.665**	**159.11**	1.813
2500000	**160.06**	2.028	156.34	**1.561**	159.79	1.774
4750000	160.68	1.994	157.68	**1.504**	**160.78**	1.543
2-OPT	**163.67**	1.086	161,75	**0,961**	162.88	1.325

Table 2. Comparison of GAC, MAC and CAPMC results (clustered versions).

The 2-OPT search improves all the results achieved by the algorithms with the exception of MAC. This happens because the local search on the MAC does a 2-OPT search on each generation. The other algorithms outperform its results with a low cost technique. We recommend the 2-OPT search as an option to refine the solutions once the optimization algorithm has finished. If we could only perform $1,000,000$ evaluations, the best algorithm would be the CAPMC. The algorithm achieves high fitness at the beginning of the evolution. The best algorithm for $5,000,000$ evaluations is GAC, as it gets the highest average fitness and in combination with 2-OPT, it achieves the best overall result in all executions (164.72).

6 Conclusions and Future Work

Analyzing the RND problem in the selected instance and the behavior of the proposed algorithms, we conclude that the clustering technique allows us to reduce the effective search space without causing negative effects on solutions. In addition, the 2-OPT technique performs a final search which may redefine the solution with low cost. The results obtained confirm that the algorithms tested perform better when combined with clustering techniques, and that the 2-OPT is recommendable in all cases. However, for further developments, other important issues must be considered, such as, different types of antenna, real propagation models, the demand of the spot, the quality of service, or the hangover problem.

References

[1] E. Alba. Evolutionary algorithms for optimal placement of antennae in radio network design. *ipdps*, 07:168, 2004.

[2] E. Alba, M. A. Vega-Rodríguez, J. A. Gómez-Pulido, D. Vega-Pérez, S.Priem-Mendes, and G. Molina. Different evolutionary approaches for selecting the optimal number and locations of omnidirectional bts in a radio network. In *Alexis Quesada-Arencibia, et al. (Eds.). Eleventh International Conference on Computer aided Systems Theory. EUROCAST*, pages 292–294, Feb. 2007.

[3] G. Croes. A method for solving traveling salesman problems. *Operations Research 6 (6)*, pages 792–812, 1958.

[4] Dawkins, R. The selfish gene. *Oxford University Press (New York)*, 1976.

[5] J. Engel. The effects of cochannel interference on the parameters of a small-cell mobile telephone system. *IEEE Transactions on vehicular technology, 1968*, 18(3):110–116, 1969.

[6] R. Frenkiel. A high-capacity mobile radiotelephone system model using a coordinated small-zone approach. *Vehicular Technology, IEEE Transactions on*, 19(2):173–177, May 1970.

[7] A. Gamst, R. Beck, R. Simon, and E.-G. Zinn. An integrated approach to cellular radio network planning. *Vehicular Technology Conference, 1985. 35th IEEE*, 35:21–25, 21-23 May 1985.

[8] Gamst, A. Remarks on radio network planning. *Vehicular Technology Conference, 1987. 37th IEEE*, pages 160–165, February 1987.

[9] Gamst, A., Zinn, E.-G. , Beck, R., and Simon, R. Cellular radio network planning. *IEEE Aerospace and Electronic Systems Magazine (1)*, February 1986.

[10] MacQueen. Some methods for classification and analysis of multivariate observations. *Proc. 5th Berkeley Symp. Math. Statisi. Prob., 1:281-297*, 1967.

[11] H. Meunier, E.-G. Talbi, and P. Reininger. A multiobjective genetic algorithm for radio network optimization. In *Proceedings of the 2000 Congress on Evolutionary Computation CEC00*, pages 317–324, La Jolla Marriott Hotel La Jolla, California, USA, 6-9 2000. IEEE Press.

[12] A. Nebro, E. Alba, G. Molina, F. Chicano, F. Luna, and J. Durillo. Optimal antenna placement using a new multi-objective chc algorithm. In ACM, editor, *In Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, pages 876–883, 2007.

[13] N. Nomura, T.; Yoshikawa. Multiple radio zone plans in mobile radio systems. *Vehicular Technology, IEEE Transactions on*, 25(3):68–74, Aug 1976.

[14] C. Rypinski. Economic design of interference limited radiotelephone systems. *Vehicular Technology Conference, 1983. 33rd IEEE*, 33:332–340, 25-27 May 1983.

[15] Y. Saez, P. Isasi, and J. Segovia. Interactive evolutionary computation algorithms applied to solve rastrigin test functions. In Springer-Verlag, editor, *In Proc. of Fourth IEEE Int. Workshop on Soft Computing as Transdisciplinary Science and Technology (WSTST 05)*, pages 682–691, 2005.

[16] Y. Saez, P. Isasi, J. Segovia, and J. Hernandez. Reference chromosome to overcome user fatigue in iec. *New Generation Computing*, 23(2):129–142, 2005.

[17] J. Wells. Cellular system design using the expansion cell layout method. *Vehicular Technology, IEEE Transactions on*, 33(2):58–66, May 1984.

[18] J. Whitehead. Cellular system design: An emerging engineering discipline. *Communications Magazine, IEEE*, 24(2):8–15, Feb 1986.