TRANSACTIONS ON EMERGING TELECOMMUNICATIONS TECHNOLOGIES
Trans. Emerging Tel. Tech. 2012; **00**:1–**??**
DOI: 10.1002/ett

Research Article

A Coral-Reef Optimization Algorithm for the Optimal Service Distribution Problem in Mobile Radio Access Networks

S. Salcedo-Sanz[1*], J. E. Sanchez-Garcia[1,2], J. A. Portilla-Figueras[1], S. Jimenez-Fernandez[1] and A. M. Ahmadzadeh[1]

[1]Department of Signal Theory and Communications, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain.
[2]Department of Cost Modeling and Internet Economics, WIK Consult GmbH, 53604 Bad Honnef, Germany.

ABSTRACT

Mobile technology is currently one of the main pillars of worldwide economy. The constant evolution that mobile communications have undergone in the last decades, due to the appearance of new services and new technologies such as UMTS/HSPA and LTE, has contributed to achieve this position in global economy. However, due to the crisis of the sector in the last five years, mobile operator's revenues and investments have been reduced. Thus, mobile network operators tend to exploit the existing infrastructure at maximum possible, trying to use the existing network in the most efficient way. In this paper, a novel bio-inspired algorithm, the coral reef optimization algorithm (CRO) is introduced to minimize a network deployment investment cost problem. This is carried out by means of optimizing the user demand of different services offered by mobile operators over the available technologies in the market, namely the Optimal Service Distribution Problem (OSDP). The CRO is a recently proposed meta-heuristic based on the computer simulation of corals reproduction and reefs' formation. In this paper, this algorithm has been tested on several OSDP scenarios in Spain, observing a significant reduction (up to 400 M€) on the total investment costs associated to the Radio Access Network deployment. We compare the performance of the CRO approach with that of a classical (experience-based) services distribution, and with alternative meta-heuristics techniques, obtaining good results in all cases. Copyright © 2012 John Wiley & Sons, Ltd.

***Correspondence**
Department of Signal Theory and Communications, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain. E-mail: sancho.salcedo@uah.es

1. INTRODUCTION

Mobile telecommunication market is currently one of the most relevant players in global economy. Penetration of the different mobile technologies and services has continuously grown since they were massively introduced at the beginning of the 1990's, producing a corresponding important growth of Mobile Network Operators' (MNO) revenues. Figure 1 shows the evolution of the MNO revenue's absolute values, together with the year-to-year revenues growth percentage.

One of the key points that explains the current importance of mobile telecommunications in economy is the fast development and introduction of new technologies in the market. Namely, in Europe: 1) 2^{nd} Generation, Global System for Mobile Telecommunications (GSM), 2) 3^{rd} Generation, Universal Mobile Telecommunication Systems (UMTS), and 3) 4^{th} Generation, Long Term Evolution (LTE). Furthermore, some of them have enhancements within its own generation, i.e., High Speed Data Access (HSPA) which is an evolution of 3G-UMTS also know as 3.5G. Note that the higher the technology generation, the higher the binary data rate offered to the customer, and the more efficient, in terms of spectrum use and system performance, it is. 2G-GSM was the first digital technology and it is mainly oriented to the provision of voice service. 3G-UMTS has better voice performance than 2G and, in addition, it provides data services that achieve up to 384 Kbps with global mobility cases. Moreover, HSPA provides up to 14.4 Mbps to the user with local mobility. On the other hand, 4G-LTE is designed to be the competitor to the broadband fixed access with data rates that can reach up to 100 Mbps.

At the end of the last century, MNOs deployed their 2G-GSM networks, and currently they have countrywide coverage. The deployment of 3G-UMTS started on the first years of last decade, and, in most countries, it currently reaches more than 90% of the population. The coverage status of HSPA depends on the country, but it is already deployed in most metropolitan areas and large cities. Finally, LTE is currently being deployed, and its final implementation will depend on the strategic technological view of each MNO. Therefore, a large MNO may have up to three overlaying networks, GSM, UMTS/HSPA and an incipient LTE network. Moreover, the MNO may implement a large set of services such as voice, best effort data, streaming, and even the use of advanced multicast techniques recently developed [1, 2] on their networks. With the current global crisis, the investment on new network infrastructures has been dramatically reduced [3]. In Spain (where we focus the experimental part of this work) investment has dropped from 2,439M€ in 2007 to 1,575M€ in 2011 (35% reduction) [4]. Therefore, any MNO must exploit the existing infrastructure at maximum possible, trying to use the existing network in the most efficient way. This implies to optimize the use of the radio access network resources by the set of services considered. Optimal distribution of services over the radio access network in current mobile broadband networks has been extensively studied in the literature. However previous works refer to different scheduling methods which are mainly applied in the operation of the network [5, 6] or consider a single technology [7].

In this paper, we consider a previous stage of the problem, the dimensioning of the network, and we focus on the distribution of the services over all existing technologies, from 2G up to 4G, instead of optimizing the scheduling for a single technology. Note that given a specific service that can be carried by two (maybe three) technologies, an efficient distribution of this service among the deployed technologies is critical to provide a high quality of service with an optimized cost network infrastructure. Note also that the optimal distribution of the services, which minimizes the investment costs, depends on a large set of parameters such as frequency bands, base stations' transmission power, individual traffic per service, market share, market penetration, etc. Therefore, in this paper we state and solve the Optimal Service Distribution Problem (OSDP), that is a complex problem that cannot be tackled with traditional approaches. In this case we study the possibility of using modern optimization techniques such as bio-inspired meta-heuristics. These approaches have been previously applied to solve different complex optimization problems in the area of mobile communications, such as optimal location of network facilities or base stations (BS) [8, 9, 10, 11], different versions of channel allocation problems [12, 13, 14, 15, 16, 17], load balancing problems [18, 19], etc. The majority of bio-inspired algorithms used in these works are genetic and evolutionary algorithms, particle swarm approaches or ant colony optimization algorithms.

The aim of this paper is two-fold. First, tackle the problem of optimal distribution of different services over available technologies in a mobile communications system (the OSDP), which leads to an optimized-cost network deployment. Second, we introduce a new bio-inspired algorithm to do this task, the Coral Reef Optimization (CRO) algorithm. The CRO is a bio-inspired meta-heuristic recently introduced in [20], and based on the simulation of coral reproduction and reef formation. We will show that the proposed CRO approach obtains excellent results in the OSDP, better than traditional solutions to the problem (based on the use of specific distribution of services provided by experienced users), and also better that alternative meta-heuristics such as an evolutionary algorithm and a Teaching-Based-Learning approach.

The rest of the paper is structured as follows: Section 2 provides the mathematical definition of the OSDP. Next, the proposed solution using the CRO algorithm, together with the definition of the objective function used, is described in detail in Section 3. A general description of the Strategic Mobile Network Planning Tool (SMNPT) employed to obtain the objective function in the CRO, is also introduced in this section. The experimental section (Section 4) shows the performance of the algorithm in a real scenario composed by the 5354 largest Spanish cities. Finally, conclusions and future work lines are described in Section 5.

2. PROBLEM DEFINITION

Let us consider a scenario E that is composed by a large set of parameters, e.g. the set of cities and villages where the network deployment has to be calculated, the frequency constraints in terms of available frequency bands (800 MHz, 900 MHz, 1800 MHz, 2100 MHz and 2600 MHz) and the total bandwidth available in each frequency band, the market situation of the operator under study, the specific hardware that the operator can deploy, and the set of individual costs related to each network element. In addition to these parameters, a set of services $\mathcal{S} = \{1 \leq i \leq |\mathcal{S}|\}$, such as voice, best effort data, guaranteed data, etc., each one characterized by a set of parameters $\mathcal{P}_i^S$, $\mathcal{P}_i^S = \{1 \leq p \leq |\mathcal{P}|\}$, is also defined. In this work, parameters defining the services, among others, are the binary rate Rb_i^S, the individual traffic per user, a_i^s and the required quality of service q_i^s.

Let us define also the set of available technologies $\mathcal{T} = \{1 \leq j \leq |\mathcal{T}|\}$, that in this work we consider to be: 2G-GSM, 3G-UMTS, 3G-HSPA and 4G-LTE. A deployment function, ϕ, which depends on the general features of the scenario E, the set of services defined by their characteristics parameters $\mathcal{S}$, together with the distribution of these services, $\mathcal{D}$, over the considered technologies. $\mathcal{D}$ stands for the specific service's traffic percentage, i, allocated to the specific technology, j, $(\mathcal{D}_{i,j})$.

The deployment function ϕ provides the number, location and type of the network resources, which are the base stations (BS) of the different technologies, required by a specific scenario E.

Thus, the problem's objective is to minimize the cost of the network deployment, $C(\phi)$, fulfilling the quality constraints of the set of services $\mathcal{S}$. The result of the deployment depends on the set of services $\mathcal{S}$, the distribution of these services over the technologies $\mathcal{D}$ and the corresponding scenario E, $\phi(E, \mathcal{S}, \mathcal{D})$. Given E and $\mathcal{S}$, the cost of the deployment exclusively depends on an optimal services' distribution over the technologies, $\mathcal{D}$.

Thus, the optimal network deployment, in terms of investment costs, is based on the estimation of the optimal services' distribution over the technologies.

$$\text{Find } \mathcal{D} \quad / \quad min(C(\phi(E, \mathcal{S}, \mathcal{D})))$$

3. ALGORITHM FOR SOLVING THE OSDP

Optimization problems arise in many fields of Science and Engineering. One possible approach to solve these problems is to apply classical optimization algorithms, however due to multiple factors such as high dimension search spaces, non-linear objective functions and constraints, these approaches do not offer a good solution. Efforts on solving non-linear optimization problems, in a more efficient way, have been made with the use of modern optimization heuristics and meta-heuristics algorithms, such as bio-inspired algorithms. In this section a novel bio-inspired meta-heuristic algorithm, based on the coral behavior [20, 21], to solve the OSDP is introduced.

3.1. The Coral Reef Optimization Algorithm

The CRO is a novel meta-heuristic search approach based on corals' reproduction and coral reefs' formation, proposed in [20] and [21]. The CRO borrows concepts from Evolutionary Computation and Simulated Annealing algorithms, but introducing new variants and concepts. The exploration phase of the algorithm is carried out by operators that simulate the sexual and asexual reproductive processes of corals, and there is a process of fight for the space in the reef, that allows the best corals (best solutions to a given optimization problem) to survive.

Basically, the CRO is based on the artificial modeling of a coral reef, Λ, consisting of a $N \times M$ square grid. We assume that each square (i, j) of Λ is able to allocate a coral (or colony of corals) $\Xi_{i,j}$, representing a solution to a given optimization problem, which is encoded as a string of numbers in a given alphabet $\mathcal{I}$. The CRO algorithm is first initialized at random by assigning some squares in Λ to be occupied by corals (i.e. solutions to the problem) and some other squares in the grid to be empty, i.e. holes in the reef where new corals can freely settle and grow in the future. The rate between free/occupied squares in Λ at the beginning of the algorithm is an important parameter of the CRO algorithm, which is denoted as ρ, and note that $0 < \rho_0 < 1$. Each coral is labeled with an associated *health* function $f(\Xi_{ij}) : \mathcal{I} \rightarrow \mathbb{R}$, that represents the problem's objective function. The CRO is based on the fact that the reef will progress, as long as healthier (stronger) corals (which represent better solutions to the problem at hand) survive, while less healthy corals perish.

After the reef initialization process described above, a second phase of reef formation is artificially simulated in the CRO algorithm: a simulation of the corals' reproduction in the reef is done by sequentially applying different operators. This sequential set of operators is then applied until a given stop criteria is met. Several operators to imitate corals' reproduction are defined, among them: a modeling of corals' sexual reproduction (broadcast spawning and brooding), a model of asexual reproduction (budding), and also some catastrophic events in the reef, i.e. polyps depredation. After the sexual and asexual reproduction, the set of larvae formed (new solutions to the problem), try to locate a place to grow in the reef. It could be in a free space, or in an occupied one, where they have to fight against the existing coral and the best survives. If larvae are not successful to locate a place to grow in a given number of attempts, they are depredated (depredation phase). Figure 2 illustrates the flow diagram of the CRO algorithm referencing the two CRO phases (reef initialization and reef formation), along with all the operators described above.

The objective of the OSDP tackled in this paper is to estimate an optimal distribution of the users' demand for every service defined in the services' profile, $\mathcal{S}$. Therefore, the first step is to determine how to encode the distribution of the user's demanded traffic in the CRO algorithm. This distribution indicates the percentage of the user's demand of every service i that is carried by one specific technology j. Thus a real encoding for this distribution is used: $D_{i,j} \in [0, 1]$. Figure 3 shows the definition of a generic coral encoded in the CRO. Using this encoding, the CRO algorithm is based on four different processes: an initialization phase, a reproduction phase, a coral larvae allocation phase and a depredation phase.

First, in the initialization phase, a fraction of the coral reef is initialized from randomly generated corals or polyps. The creation of the initial coral larvae is made by means of a pseudo-random process (note that several technology constraints related to the capabilities of each technology will be defined later – Section 4– which might affect the feasibility of individuals). Later on, these larvae are settled into the reef at randomly chosen positions. Note that after the initialization process not all the positions are occupied, this is necessary to guarantee the correct simulation of a coral reef formation in real world.

Second, the reproduction phase consists of three different types of reproduction: external sexual reproduction, implemented by means of crossover operators, internal sexual reproduction, that uses a random mutation, and asexual reproduction carried out by randomly choosing one among the best existing corals and making a copy of it.

- The external sexual reproduction or broadcast spawning, is simulated using a crossover operator

applied to a fraction of the existing corals, F_b. Our model implements the following crossover operators: one point crossover, two points crossover and N-swap crossover. The type of crossover operator used in the algorithm is set at the beginning of the process, and maintained during all the generations of it.
The one point crossover operator randomly selects two parents among the existing corals in the reef, and a random point where the crossover is performed to generate a new coral larva (Figure 4(a)). Once the parents are chosen they are no longer used for reproduction purposes at that given iteration. This new coral larva will become a coral at the beginning of the next iteration, after the larvae setting process that will be explained later.

The two points crossover operator is based on the same process as one point crossover. In this case, a random selection of two different points, where the crossover is performed, is done (Figure 4(b)).

Finally, the N-swap crossover operator follows the same parents selection procedure, and randomly chooses N blocks of services, in the encoded inhabitant, to be swapped (Figure 4(c)).

As has been previously mentioned, the distribution of the traffic demand over the different technologies has to meet specific constraints, which depends on each specific scenario, see Section 4. Thus, after the crossover takes place, a reparation phase starts. During this reparation phase, the algorithm checks whether all the constraints are fulfilled, and if not, the larva is repaired. This operation may affect one or several services within the larvae's encoding.

- Internal sexual reproduction or brooding, as it is less likely to occur, has been implemented by using the mutation operator over the complementary fraction, $1 - F_b$, of existing corals at iteration k. The mutation is done over a randomly selected characteristic of the coral, that is a specific percentage of a service over one technology, and it is performed by the aggregation or subtraction of a extremely low randomly generated number to the value of a specific characteristic. As it occurred in broadcast spawning reproduction, a reparation phase follows the brooding method to ensure viable corals.

- The asexual reproduction or budding, occurs with a very low probability, P_a, at each iteration k. To be performed, the existing corals are sorted by their health function in ascending order, choosing a fraction, F_a, of the corals and making a copy of one, randomly chosen, among the selection.

Using the operators above, the new coral larvae are generated and the setting phase starts. First, the CRO algorithm calculates the health function of every larva. Second, a random position, (i, j), of the reef is chosen to host the larva. On the one hand, if the chosen position is empty, the new larva settles in the reef. On the other hand, if the position is occupied by a previous larva, a comparison of the health function's value is carried out. If the health function value of the new larva is better than the existing coral, then the new larva replaces the former coral and is settled in the reef, otherwise, a new position in the reef, $(i, j)'$ is randomly chosen and the allocation phase for this larva starts from the beginning. The CRO algorithm defines a maximum number of attempts, κ, to settle each larva.

Finally, at the end of each iteration k, the CRO algorithm checks the status of the reef. Should the reef be full, a depredation phase is performed. A fraction of corals, F_d, having the worst health function is selected as candidates to be discarded. The depredation of the candidates occurs with a very low probability, P_d.

The process explained above is performed iteratively until a specific stopping condition is met. In our problem, the stop condition is defined by the number of total evaluations of the health function. Once the number of evaluations reaches the defined threshold, the algorithm stops and selects the best coral, in terms of investment cost, as the final solution of the optimization phase.

3.2. CRO health function definition

Each coral (problem solution) in the reef has its own associated health function. In our problem the health function is $C(\phi)$, as was stated in Section 2. Due to the many input parameters involved in the analytic expression of $C(\phi)$ and for the sake of simplicity, it can be summarized by dividing it up in three different parts:

$$C(\phi) = C_s + C_{BS} + C_f \tag{1}$$

where:

- C_s: Defined as the investment cost of the site, including terrain and civil working investment. This factor depends on the type of technology or technologies' combination installed in each site.

- C_{BS}: Cost of the base station. It depends on the type of technology, the type of base station considered, macro-cell, micro-cell or pico-cell, and number of sectors in each base station.

- C_f: Investment cost of the required frequency-related hardware resources. In case of 2G-GSM technology the total number of TRX installed in each BS is considered. In 3G-UMTS and 3G-HSPA, the cost is defined by the total number of carriers, radio-frequency (RF) modules, required in the dimensioning process, here it is supposed each

RF module is capable of managing only one 5 MHz frequency block. Finally, in LTE the eNodeBs considered are able to manage the whole set of LTE bandwidth possibilities, from 1.4 MHz to 20 MHz. This makes the price of this equipment to be larger than classic 3G system's NodeBs. This feature makes unnecessary the definition of a new hardware frequency-related cost parameter for LTE.

Note that although the three parameters (C_s, C_{BS} and C_f) are relevant from the point of view of network dimensioning and investment, the parameter C_{BS} has the largest influence in the health function of the CRO, since it is the cost driver for traffic-limited Node B's. These are usually located in urban areas with high traffic load. However, note also that it cannot be dissociated from the other two, C_s and C_f, because the higher the traffic the more frequency resources are required (measured by C_f) and, on the other hand, if the capacity of the NodeB is overpassed, a new site has to be deployed, and therefore the cost term C_s increases. Moreover, since terms (C_s, C_{BS} and C_f) cannot be expressed analytically, a simulation tool must be used in order to obtain each one of them, and thus calculate $C(\phi)$.

3.3. Health function evaluation

As stated before, the coral encoding determines the services' distribution over the different mobile technologies available. To obtain the health function associated to each coral, the software tool Strategic Mobile Network Planning Tool (SMNPT) [22]-[23] has been used in this work. The SMNPT, developed by the University of Alcalá, the University of Cantabria and the German company WIK-Consult GmbH, has been successfully applied before to several regulatory projects [24]-[25]. Figure 5 shows how a given coral in the CRO is passed to the SMNPT to obtain the coral's health function.

The SMNPT obtains the resources needed in a network deployment to fulfill certain services' load over certain technologies. Network infrastructure depends on the number and type of the required network elements. To minimize the investment, each base stations cell range has to be maximized; hence the number of required BS is minimized. The cell range is defined as the maximum range of a single cell able to fulfill the user demand and propagation restrictions. Thus, for each technology considered, a capacity analysis and a propagation analysis (subsections 3.3.1 to 3.3.3) have to be performed, and the most restrictive is the one that determines the cell range. Finally, the number of sites required is calculated as explained in the following subsections.

3.3.1. 2G network deployment model

2G-GSM technology is a time division multiple access (TDMA) and a hard blocking system, which means the capacity is directly related to the amount of hardware (sectors, transceivers per sector (TRX) and traffic channels per TRX) installed at the base stations. The implemented model estimates the minimum cell range, and therefore the number of second generation base stations, BTS, to be deployed in each area of a specific city. The capacity analysis applied is based on the Erlang-B formulation, the grade of service (GoS) and the user demand. The propagation model applied to perform the coverage study is based on a one-slope empirical model, the Cost231-Hata model which is commonly used in mobile network applications [26].

3.3.2. 3G network deployment model

3G-UMTS Technology 3G-UMTS technology is based on wide-band code division multiple access (W-CDMA) and therefore it is a soft blocking system, which means the capacity of the network does not depend on the hardware installed in the base station but on the amount of interference present in the system. This interference is caused by the active users both in one's cell and in neighboring cells. The maximum interference level allowed is defined by the Interference Margin, IM, and is applied to the capacity and propagation analysis. Therefore, in 3G-UMTS there is a relation between the capacity and the propagation dimensioning processes.

The model implements a multi service optimization algorithm which estimates the number of base stations (Nodes B) needed, and maximizes the Node B range according to a set of propagation and capacity constraints [27]. The propagation model applied to 3G-UMTS dimensioning process is, as in 2G, based on the Cost231-Hata model.

3G-HSPA Technology 3G-HSPA technology uses W-CDMA, and therefore there is a close relation between propagation and capacity too. The dimensioning method followed by the HSPA algorithm is based on: 1) The guaranteed user throughput, defined as the guaranteed data transfer rate offered to a subscriber at the cell edge, and 2) the terminal category. Each terminal category presents its own features in terms of modulation and coding schemes (MCS), signal to interference noise ratio, $SINR$, used in the Cost231-Hata model propagation analysis, and maximum offered throughput applied to the capacity study. The model estimates the cell range and the number of HSPA base stations to be deployed based on [22], [28].

3.3.3. 4G network deployment model

To determine the number of 4G-LTE evolved Nodes B, a propagation and capacity analysis is performed. In this technology, data from an external link level simulator, providing necessary input parameters such as the SINR and the resources required per user, is needed. Capacity study is based on the same concept as in HSPA, that is, the subscriber demanded throughput. Applying an external link level simulator based on the value of the subscriber throughput the SMNPT fixes a starting modulation and coding scheme, which in turn, sets a value for SINR and number of frequency resources required per subscriber to fulfill the demanded throughput. Propagation analysis is based on Cost231-Hata model, as in previous technologies. The model carries out both analysis, propagation and capacity, for all possible MCS according to the demanded throughput, selecting the one that minimizes the investment costs as final solution (we consider that all mobile terminals implement all possible MCS schemes).

3.3.4. Sites estimation process

The SMNPT carries out a propagation and a capacity analysis for all considered technologies in every area type (urban, suburban and rural) within every city, also considering all possible types of BS. Note that for each BS's technology there are several makes and models available in the market. For each type of BS considered, the SMNPT obtains a set of different cell ranges, and therefore a total number of network elements to install. The total number of BS required in order to meet the users' demand and propagation restrictions depends on A_t, the total area of each type of terrain, urban, suburban or rural, of a specific city, and A_{BS}, that is the area covered by the BS. Therefore the total number of sites required, N_{S_t}, is obtained as follows:

$$N_{S_t} = \lceil A_t / A_{BS} \rceil \tag{2}$$

where t is the type of area within a city, urban, suburban and rural.

4. EXPERIMENTS

This section is focused on the definition of the experiments performed and the results obtained to show the performance of the proposed CRO. To define the experiment, input parameters determining the scenario (parameters used to determine the deployment study), and parameters characterizing the optimization process have to be chosen.

4.1. Coral-Reef optimization parameters definition

In this subsection, values considered for the CRO algorithm are provided in Table I. The population size used for each different crossover operator (1 point, 2 points, 2-Swap and 3-Swap) in the experiments carried out in this work is set to 75 corals. A total of 100 executions were performed, each one involving 7500 evaluations (calls to the SMNPT tool).

4.2. Experimental scenarios description

The scenarios defined in this work describe the dimensioning of a Spanish nationwide network deployment as stated in the Spanish Telecommunication Market Commission 2011 annual report [29]. In this report, considered services' figures are shown in Table II. User's demand of every service, except for the Mobile Broadband Access Service (MBAS), are referred to the voice service, and thus data services demand values are expressed in voice-equivalent mE (miliErlangs). MBAS value is referred to the minimum guaranteed transfer data rate, in Kbps, at the cell edge, offered to any user in a specific cell. Note that multicast services are considered in MBAS service, that covers these and other similar services with high binary rate demand.

Several considerations have been made to deal with some computation time problems and site's mobile technology combinations. In this work, Spain was chosen as test bed, a country made up of 5354 different districts (cities and villages). Due to the high number of districts involved, the computation time of one health function evaluation (one call to the SMNPT software) for the set of services shown in Table II takes up to forty minutes. For the values considered in Section 4.1, each execution (iteration) of the coral based optimization process needs to compute 7500 evaluations, and a total of 100 iterations are considered, which is not a feasible working scenario. For this reason, instead of carrying out the optimization process for the whole country, the CRO algorithm was applied to a subset of districts. A fraction, $F_c = 10$, of those 5354 districts was chosen, which implies an average total computational time of seven hours.

The districts to be part of the coral-based optimization process are selected according to a population density criterion. First, maximum and minimum population density of the districts are calculated, Γ_M and Γ_m respectively. Second, the population density step is obtained ($(\Gamma_M - \Gamma_m)/F_c$), and different groups are formed. The process seeks one representative city that has every type of area (urban, suburban and rural), and meets the population density constraints of each of the defined groups. However, if no district fulfills the first condition, the selection process chooses the district only based on population density criterion. Following this criterion, most Spanish main cities such as Madrid, Barcelona, Vitoria, A Coruña, etc., are selected.

Once the most representative districts have been chosen, the CRO is applied to obtain the best service distribution for this set-up. Next, the cost associated to the nationwide deployment is estimated using this same service distribution. Note that in this nationwide deployment only one evaluation is needed as the service distribution is fixed. For a given site, several technology combinations are possible (varying from *pure* sites (only 2G or only 3G technology on the site) to *hybrid* sites (that can combine 2G+3G, 2G+4G, 3G+4G or even 2G+3G+4G).

In this work three different scenarios, A, B and C, are defined to test the performance of the designed algorithm. These scenarios range from an ideal setup (Scenario A), to a real setup based on the frequency constraints that apply in Spain and for a 25% operator market share (Scenario C). In all of them, the hybrid option where all available technologies can be combined on the same site is chosen, due to the fact that it leads to cost reduction.

- Scenario A estimates the optimal service distribution of the user demand considering no restrictions, such as user's demand allocation over a specific technology or real spectrum allocation.
- Scenario B estimates the optimal service distribution of the user demand considering that a minimum percentage of the user demand has to be allocated to a specific technology. Note that this scenario reduces the degree of freedom, and therefore the search space, as well as it reflects a real case scenario where a MNO needs to allocate traffic over the different technologies owned. Also, in regulatory projects, restrictions are assigned to each operator based on previous investment and depreciation. Data considered in this scenario are shown in Table III. Traditional second generation services such as voice, SMS and MMS, keep a 25% of the user demand over the GSM network, higher data rate services such as streaming, guaranteed-data and best-effort are forced to be carried by UMTS at least a 40% of the user demand, and mobile internet is carried only by HSPA and LTE technology, due to the low number of commercial LTE networks, a 75% of the user demand of MBAS is routed over HSPA technology.
- Scenario C estimates the optimal service distribution based on the restrictions considered in scenario B, together with additional frequency constraints fixed by the current spectrum allocation in Spain [30].

The frequency resources available at each scenario are shown in Table IV.

4.3. Scenarios A and B results

In this subsection, the results obtained for scenarios A and B are shown. The cost estimation for every crossover operator defined in the CRO algorithm is obtained, and finally a comparison between this cost and the cost due to a classical, non-optimized, service distribution is carried out.

Table V shows the results after the CRO execution when different crossover operators are applied. The result gives information about, on the one hand, the optimized investment costs for the subset of districts selected for the optimization process, and on the other hand, the total investment cost of a nationwide deployment on which an hypothetical new entrant will incur. Figure 6 shows the graphical representation of the total investment costs.

The total investment cost, for both the subset of districts and the nationwide deployment, depends directly on an optimal distribution of the user demand. An optimal distribution of the user demand is understood as the distribution that makes the most of the base stations installed for every technology, and it drives the demand of each service through the most efficient technology able to carry it. The obtained results show that the performance of the CRO, in terms of investment costs, is better when any of the N-Swap crossover operator is applied than when a single or double point crossover is applied. Table VI compares the distribution of the user demand when 2-points crossover and 2-Swap crossover operators are applied. It can be seen how the 2-Swap crossover operator minimizes the percentage of the data traffic load allocated to the least efficient technologies carrying data among the ones considered.

The application of the CRO algorithm explained in this paper leads to a reduction of the total investment in comparison to a classical distribution of the user demand, and therefore of the interconnection charges. Table VII shows the economic impact of the CRO algorithm in comparison toa classical, experience-based user demand distribution, see Table VI.

4.4. Scenario C results

As the performance of the N-Swap crossover operator has been found to be the best, only the results of this operator are presented for Scenario C. The result is compared to a classical, experience-based distribution of the user demand over the technologies considered. Table VIII shows the different results for both cases, optimal and experience-based distribution. Note that a reduction of more than 400 M€ in a nationwide network deployment is observed when the optimal solution from CRO is considered. Moreover, despite the fact of having worse coverage conditions than in Scenario B, LTE spectrum is allocated to the 2600 MHz frequency band instead of the 2100 MHz one. The CRO algorithm solves this disadvantage by routing more traffic through LTE technology in order to make the most of the eNodeB capacity and due to the fact that it is the most efficient technology among the ones considered to transfer data, see Table IX. The commented effect is not observed in the experienced-based user distribution case, where there is no user demand carried by LTE, because no commercial LTE network is installed in Spain yet (although there are announcements from Vodafone and Yoigo of an incipient deployment in 2014), and therefore the total investment cost is slightly lower than in Scenario B.

4.5. Comparison with alternative meta-heuristics

The results obtained with the proposed CRO have been compared to that of an Evolutionary Algorithm (EA) [31], and a Teaching-Based-Learning Optimization (TBLO) [32]. We have chosen these two meta-heuristics for a comparison with the CRO since they are state-of-the-art global methods for optimization, one well established and known (EA) and the other one newly proposed, with high potential (TBLO), that have shown very good results when tackling alternative optimization problems. In order to carry out this comparison, the best solution found by the CRO algorithm is considered, so the 2-swap and 3-swap crossover and the brooding operators are used. To have a fair comparison, these operators have been directly applied to EA, and the TBLO has been also adapted to solve the OSDP. The parameters of all the algorithms have been set to be comparable, basically population size and number of function evaluations. Table X shows the results obtained by the different algorithms compared, in the three scenarios considered. Note how the CRO obtains better results than the EA and TBLO algorithms. The average differences between the CRO and the algorithms for comparison are (in terms of investment reduction) 2 million Euros with respect to the EA solution and 5 million Euros with the TBLO solution. Note that in all scenarios, the results have been calculated considering a fraction $F_c = 10$, and then nationwide deployment cost has been estimated. The CRO is the best among the algorithms tested, which indicates its excellent performance in this engineering optimization problem. Moreover, an interesting point is that the three meta-heuristics are able to improve the distribution from an experienced user, so it seems that the use of this type of approaches is a very good option in this problem of network deployment.

5. CONCLUSIONS

In this paper the Optimal Service Distribution Problem (OSDP) has been tackled. In this problem two different aspects are considered: first, the optimal deployment of a radio access network is solved. The optimization of this deployment needs to fulfill a set of restrictions according to user demand and type of technology to be deployed. Second, the optimal distribution of the input users' demand, over the different technologies available, is tackled. We propose a novel bio-inspired optimization algorithm, the Coral Reef Optimization Algorithm (CRO), based on coral reproduction and reef's formation. This optimization algorithm tries to obtain the optimal distribution of the users' demand that minimizes the associated investment costs of the solution. The cost estimation process uses of the data provided by a network planning software. In the experimental section of the paper, the CRO has been tested on three different scenarios, implementing real-world restrictions. The results obtained show a significant reduction of the total investment cost of more than 400 M€ when comparing the CRO solution to that of a network deployment based on experience-based service's distribution values, in Spain. We have also compared the performance of the CRO approach to that of alternative meta-heuristics techniques, obtaining good results. Future work is focused on evaluating the performance of the CRO algorithm for different spectrum allocations, mainly for LTE technology, etc.

ACKNOWLEDGEMENT

This work has been partially supported by Spanish Ministry of Economy, under project number ECO2010-22065-C03-02.

REFERENCES

1. H. V. Zhao and W. Su, "Cooperative wireless multicast: performance analysis and power/location optimization," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 2088-2100, 2010.
2. Y. Zhou, H. Liu and Z. Pan and L. Tian, "Two-Stage Cooperative Multicast Transmission with Optimized Power Consumption and Guaranteed Coverage," *IEEE Journal on Selected Areas in Communications*, in press, 2013.
3. J. Gillet, "Mobile operators rein in costs to ride-out recession; prepare to ramp-up network investments," *Wireless Intelligence*, GSMA Intelligence, 2011.
4. Spanish National Regulatory Authority (CMT, Comisión del Mercado de las Telecomunicaciones), "Global Annual Report," http://cmtdata.cmt.es/, 2012.
5. Z. Zheng, A. Dowhuszko, and J. Hämäläinen, "Interference management for LTE-Advanced Het-Nets: stochastic scheduling approach in frequency domain," *Transactions in Emerging Telecommunications Technology*, vol. 24, pp. 4-17, 2013.
6. C. Huppert, B. Stender, "Resource allocation with minimum rates for OFDM broadcast channels," *European Transactions on Telecommunications*, vol. 18, pp. 563-572, 2007.
7. A. M. Ahmadzadeh, J. E. Sánchez-García, B. Saavedra-Moreno, A. Portilla-Figueras and S. Salcedo-Sanz, "Capacity estimation algorithm for simultaneous support of multi-class traffic services in Mobile WiMAX," *Computer Communications*, vol. 35, no. 1, pp. 109-119, 2012.
8. J. Munyaneza, A. Kurien, and B. Van Wyk, "Optimization of antenna placement in 3G networks using genetic algorithms," *Broadband Communications, Information Technology & Biomedical Applications, 2008 Third International Conference on. IEEE*, 2008.
9. R. S. Rambally, and A. Maharajh, "Cell planning using genetic algorithm and tabu search," *Applications of Digital Information and Web Technologies, 2009. ICADIWT'09. Second International Conference on the. IEEE*, 2009.
10. S.M. Gaber, M. E. El-Sharkawi, and M. Nour Eldeen. "Traditional Genetic Algorithm and Random-weighted Genetic Algorithm with GIS to Plan Radio Network," *URISA Journal*, vol. 22, no. 1, pp. 33, 2010.
11. J. Y. Wang, J. B. Wang, X. Song, M. Chen, and J. Zhang, "Network planning for distributed antenna-based high-speed railway mobile communications,"

Transactions on Emerging Telecommunications Technologies, in press, 2012.

12. C. Y. Ngo, and V. O. Li, "Fixed channel assignment in cellular radio networks using a modified genetic algorithm," *IEEE Transactions on Vehicular Technology*, vol. 47, no. 1, pp. 163-172, 1998.
13. L. Wang, S. Arunkumaar, and W. Gu, "Genetic algorithms for optimal channel assignment in mobile communications," In *Proc. of the 9th IEEE International Conference on Neural Information Processing*, vol. 3, pp. 1221-1225, 2002.
14. F. Jaimes-Romero, and D. Muñoz-Rodriguez, "Evolutionary searching in cellular radio systems planning," *European transactions on telecommunications*, vol. 10, no. 1, pp. 85-96, 1999.
15. X. Fu, A. G. Bourgeois, P. Fan, and Y. Pan, "Using a genetic algorithm approach to solve the dynamic channel-assignment problem," *International Journal of Mobile Communications*, vol. 4, no. 3, pp. 333-353, 2006.
16. K. N. Sivarajan, R. J. McEliece, and J. W. Ketchum, "Channel assignment in cellular radio," In *Proc. of the 39th IEEE Vehicular Technology Conference*, pp. 846-850, 1989.
17. H. Kamal, M. Coupechoux, and P. Godlewski, "Tabu search for dynamic spectrum allocation in cellular networks," *Transactions on Emerging Telecommunications Technologies*, vol. 23, no. 6, pp. 508-521, 2012.
18. E. Mulyana and U. Killat, "Load balancing in IP networks by optimising link weights," *European transactions on telecommunications*, vol. 16, no. 3, pp. 253-261, 2005.
19. Z. Liu, M. Kwiatkowska, and C. C. Constantinou, "A self-organised emergent routing mechanism for mobile ad hoc networks," *European transactions on telecommunications*, vol. 16, no. 5, pp. 457-470, 2005.
20. S. Salcedo-Sanz, J. Del Ser, S. Gil-López, I. Landa-Torres and J. A. Portilla-Figueras, "The Coral Reefs Optimization Algorithm: A new metaheuristic algorithm for hard optimization problems," In *Proc. of the 15th International Conference on Applied Stochastic Models and Data Analysis*, (ASMDA), Mataró, Barcelona, 2013.
21. S. Salcedo-Sanz, A. Pastor-Sánchez, D. Gallo-Marazuela and J. A. Portilla-Figueras, "A novel Coral Reefs Optimization algorithm for multi-objective problems," In *Proc. 14th International Conference on Intelligent Data Engineering and Automated Learning*, (IDEAL'2013), Heifei, China, October, 2013.
22. J. E. Sánchez-García, A. M. Ahmadzadeh, S. Jiménez-Fernández, S. Salcedo-Sanz, and A. Portilla-Figueras, "Impact of the HSPDA-Based Mobile Broadband Access on the Investment of the 3G Access Network," *Mobile Lightweight Wireless Systems*, pp. 303-311, 2012.
23. J. E. Sánchez-García, A. Portilla-Figueras, S. Salcedo-Sanz, and S. Jiménez-Fernández, "2G/3G connect: An educational software for teaching 2G/3G mobile communications to engineering students," *Computer Applications in Engineering Education*, 2012.
24. K. D. Hackbarth, D. Ilic, W. Neu and J. E. Sánchez-García, "Erstellung von Bottom-up Kostenrechnungsmodellen zur Ermittlung der Kosten der Zusammenschaltung in Festnetzen und Mobilnetzen," 2012.
25. K. D. Hackbarth, W. Neu, and D. Ilic, "Analytisches Kostenmodell fr ein Mobilfunknetz," 2012.
26. J. N. Boucher, and D. A. Conner, *The cellular radio handbook: a reference for cellular system operation*, J. Wiley & Sons., 2001.
27. A. Portilla-Figueras, S. Salcedo-Sanz, K. D. Hackbarth, F. López-Ferreras, and G. Esteve-Asensio, "Novel heuristics for cell radius determination in WCDMA systems and their application to strategic planning studies," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, 2009.
28. J. E. Sánchez-García, *Extensión del algoritmo de despliegue UMTS multibanda y algoritmo HSPA para redes de acceso móvil*, Master Thesis, Universidad de Alcalá, 2012.
29. Spanish National Regulatory Authority (CMT, Comisión del Mercado de las Telecomunicaciones), "Global Annual Report," 2011.
30. Ministerio de Industria Turismo y Comercio de España, "Consulta pública sobre actuaciones en materia de espectro radioeléctrico: Refarming en bandas 900 MHz y 1800 MHz, dividendo digital y banda 2.6 GHz," 2010.
31. A. E. Eiben and J. E. Smith, "Introduction to evolutionary computing", Springer-Verlag, 2003.
32. R. V. Rao, V. J. Savsani and D. P. Vakharia, "TeachingLearning-Based Optimization: An optimization method for continuous non-linear large scale problems," *Information Sciences*, vol. 183, pp. 1-15, 2012.

Table I. CRO optimization parameters.

	CRO
Reef size	5×15
ρ	0.8
κ	3
F_b	0.98
F_a	0.05
P_a	0.001
F_d	0.05
P_d	0.01

Table II. Services definition. mE stands for miliErlangs and VE-mE stands for Voice Equivalent miliErlangs.

Traffic A (mE)	Voice (mE)	Video Call (VE-mE)	Streaming (VE-mE)	Guaranteed Data (VE-mE)	Best Effort (VE-mE)	SMS (VE-mE)	MMS (VE-mE)	MBAS (Kbps)
90	13	2.48	2.48	1.08	1.08	2.48	2.48	428

Table III. Scenario B: minimum service distribution to be allocated over the different technologies.

Service	GSM	UMTS	HSPA	LTE
Voice	0.25	0	0	0
Video-Call	0	1	0	0
Streaming	0	0.4	0	0
Guranteed-Data	0	0.4	0	0
Best-Effort	0	0.4	0	0
SMS	0.25	0	0	0
MMS	0.25	0	0	0
MBAS	0	0	0.75	0

Table IV. Frequency resources definition.

Scenario	A / B			C		
Technology	Freq. Band (MHz)	Mode	BW (MHz)	Freq. Band (MHz)	Mode	BW (MHz)
GSM/EDGE	900/1800	Dual	8.75/18.7	900/1800	Dual	8.75/18.7
UMTS	2100	Mono	5	2100	Mono	5
HSPA	2100	Mono	5	2100	Multi	10
LTE	2100	-	5	2600	-	15

Table V. Scenarios A and B investment costs for the optimal service distribution obtained by the CRO algorithm.

Crossover Operator	Type of Deployment	Investment Cost (M€)	
		Scenario A	Scenario B
1 Point	Subset of districts	39.026	80.174
	Nationwide	1832.036	2246.009
2 Points	Subset of districts	37.983	79.887
	Nationwide	1824.681	2240.157
2-Swap	Subset of districts	27.91	79.382
	Nationwide	1470.313	2236.716
3-Swap	Subset of districts	27.46	79.586
	Nationwide	1472.093	2235.265

Table VI. Scenario B: CRO solution for the OSDP, considering different crossover operators vs Experience-based service distribution.

	2 Points Crossover operator				2-Swap Crossover operator				Experienced-based			
	GSM	UMTS	HSPA	LTE	GSM	UMTS	HSPA	LTE	GSM	UMTS	HSPA	LTE
Voice	0.25	0.75	0	0	0.25	0.75	0	0	0.765	0.235	0	0
Video-Call	0	1	0	0	0	1	0	0	0	1	0	0
Streaming	0	0.4	0.266	0.334	0	0.4	0.432	0.168	0.2	0.8	0	0
Guaranteed-Data	0.017	0.4	0.376	0.207	0	0.4	0.436	0.164	0.2	0.8	0	0
Best-Effort	0.068	0.4	0.225	0.307	0	0.4	0.432	0.168	0.04	0.06	0.9	0
SMS	0.252	0.017	0.180	0.550	0.25	0	0.595	0.155	0.775	0.225	0	0
MMS	0.250	0.009	0.261	0.480	0.25	0	0.595	0.155	0.538	0.142	0.32	0
MBAS	0	0	0.75	0.25	0	0	0.75	0.25	0	0	1	0

Table VII. Scenario B: Investment Costs (M€) for Optimal vs. Experience-based service distribution (Nationwide values).

	CRO Optimized				Experience-based deployment
	1-Point	2-Points	2-Swap	3-Swap	
Investment Cost (M€)	2246.009	2240.157	2236.716	2235.265	2779.792

Table VIII. Scenario C: Investment Costs (M€) for Optimal vs. Experience-based service distribution.

	CRO Optimized 2-Swap	Experienced-based deployment
Investment Cost (M€)	2357.994	2768.270

Table IX. Scenario C: CRO solution for the OSDP considering different frequency allocation figures.

	2600 MHz				2100 MHz			
	GSM	UMTS	HSPA	LTE	GSM	UMTS	HSPA	LTE
Voice	0.25	0.75	0	0	0.25	0.75	0	0
Video-Call	0	1	0	0	0	1	0	0
Streaming	0	0.4	0.083	0.517	0	0.4	0.432	0.168
Guranteed-Data	0	0.402	0.171	0.427	0	0.4	0.436	0.164
Best-Effort	0	0.4	0.083	0.517	0	0.4	0.432	0.168
SMS	0.25	0.014	0.237	0.499	0.25	0	0.595	0.155
MMS	0.25	0.014	0.245	0.491	0.25	0	0.595	0.155
MBAS	0	0	0.75	0.25	0	0	0.75	0.25

Table X. CRO, EA and TBLO comparison in all Scenarios considered.

Scenario A	
CRO	1472.0
EA	1474.8
TBLO	1478.1
Scenario B	
CRO	2235.2
EA	2237.0
TBLO	2241.8
Scenario C	
CRO	2357.9
EA	2359.4
TBLO	2463.5

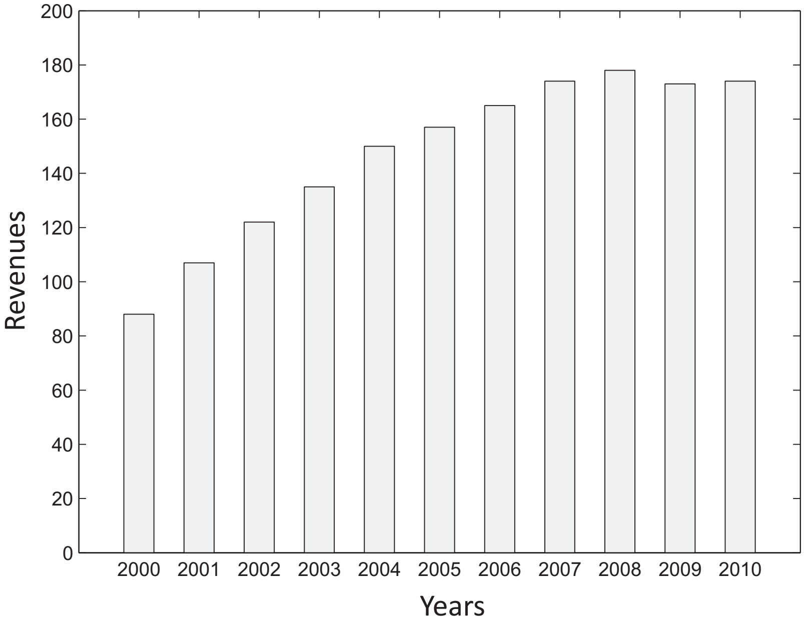


Figure 1. Growth in European Mobile Operators' Total Revenues

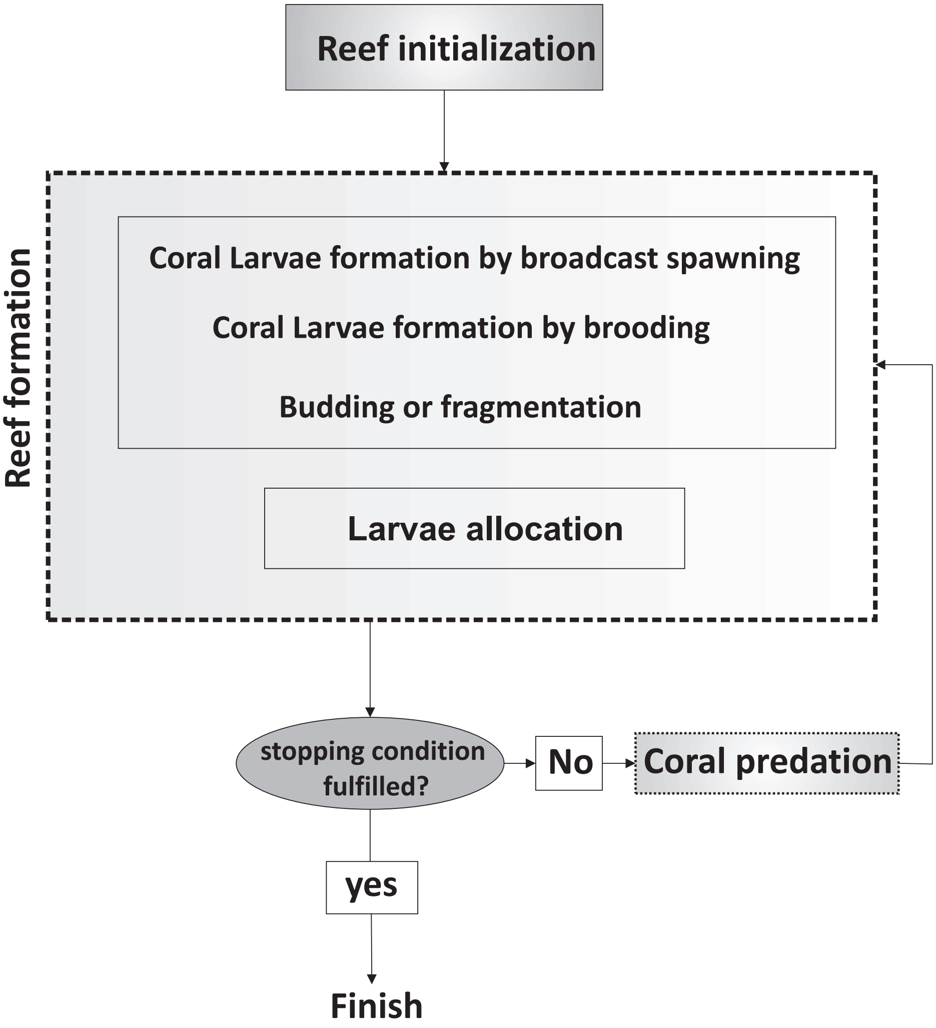


Figure 2. Flow diagram of the proposed CRO algorithm.

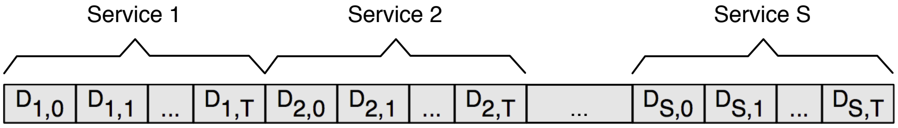
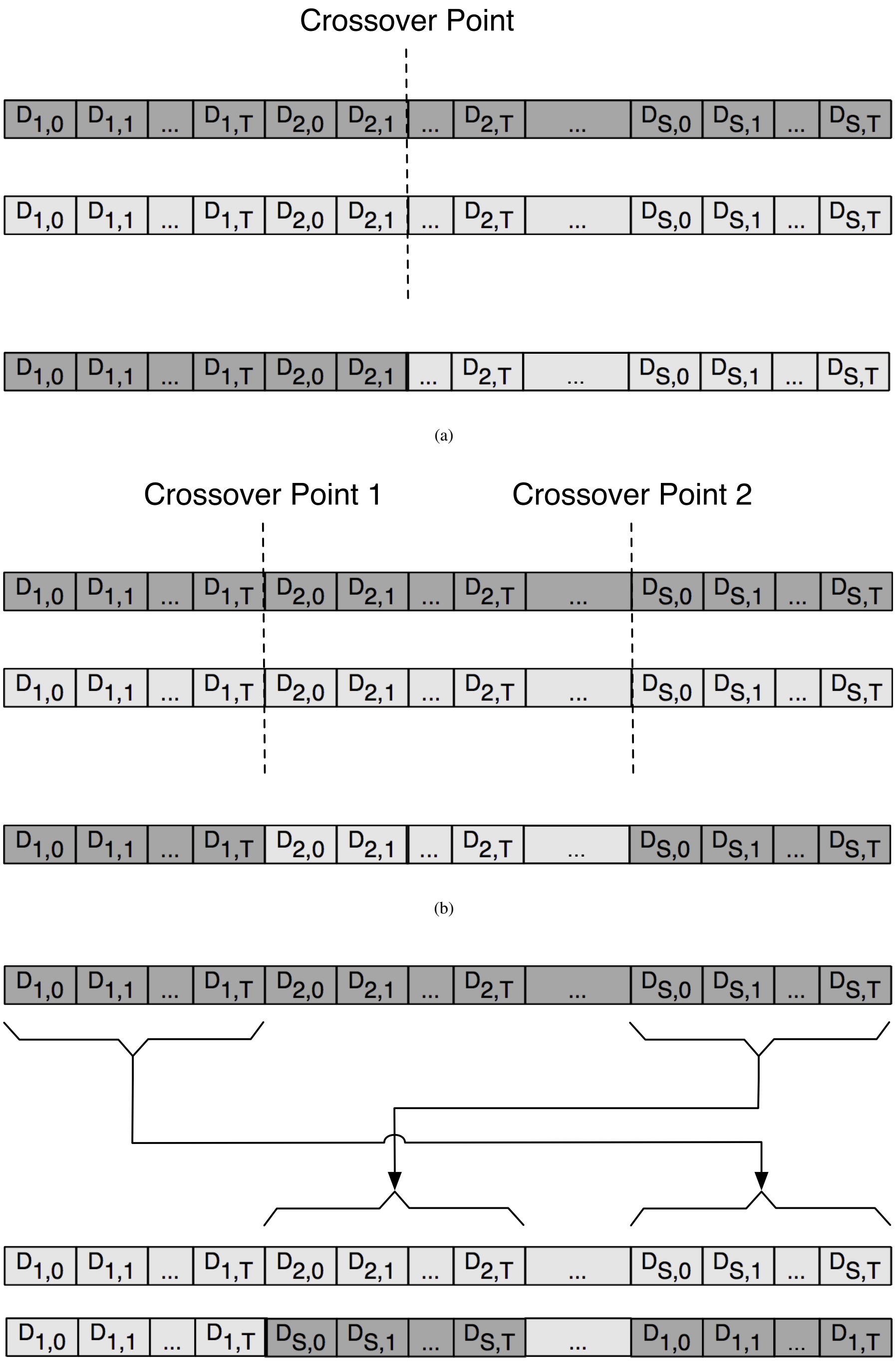


Figure 3. Services distribution on a coral larva



(a)

(b)

(c)

Figure 4. Crossover Operators: (a) 1 Point; (b) 2 Points; (c) N-Swap

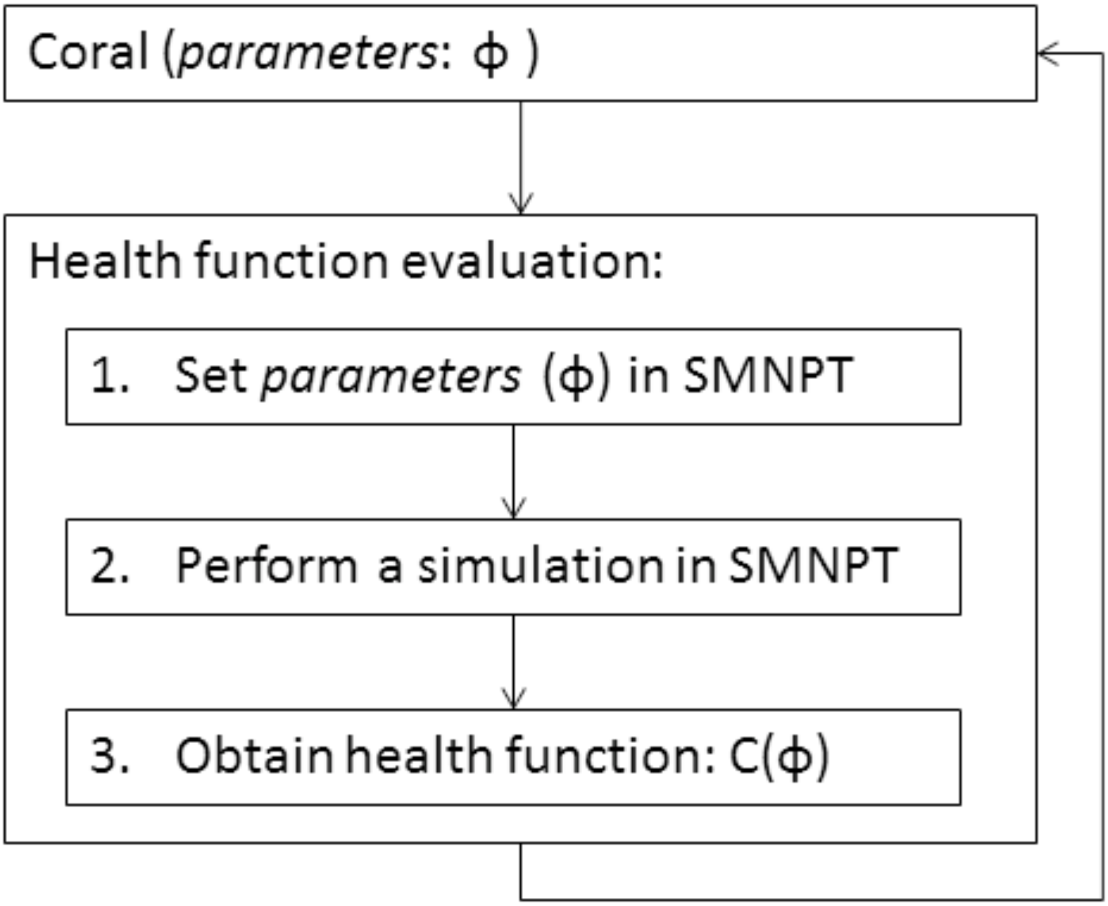
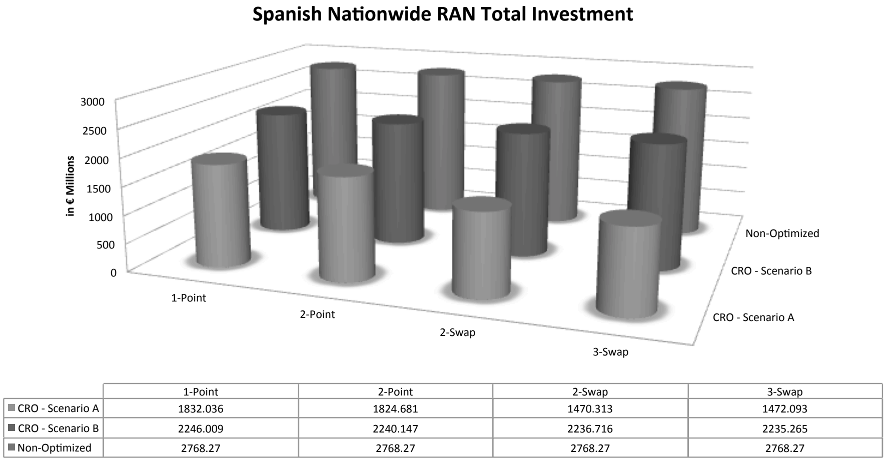


Figure 5. Outline of the corals' health function calculation using the SMNPT tool.



	1-Point	2-Point	2-Swap	3-Swap
■ CRO - Scenario A	1832.036	1824.681	1470.313	1472.093
■ CRO - Scenario B	2246.009	2240.147	2236.716	2235.265
■ Non-Optimized	2768.27	2768.27	2768.27	2768.27

Figure 6. Optimized vs. Non-Optimized Nationwide RAN Investment Costs