Multiobjective Hierarchical 2G/3G Mobility Management Optimization: Niched Pareto Genetic Algorithm[1]

Timucin Ozugur [a], Anand Bellary and Falguni Sarkar

[a] ALCATEL, Research and Innovation Center, Richardson, TX, 75081 USA
Corresponding author: Tim.Ozugur@usa.alcatel.com

Abstract

In this paper, we first propose four-layer optimization for UMTS coverage area: (i) cell-oriented intra-SGSN layer, which is optimized RA areas covering the intra-SGSN signaling cost, paging cost and RA load balancing, (ii) RA-oriented intra-MSC layer, which is optimized location areas covering the intra-MSC signaling cost and LA load balancing,(iii) RA-oriented inter-SGSN layer, which is optimized SGSN coverage areas covering the inter-SGSN signaling cost, RNC and SGSN load balancing, (iv) LA-oriented inter-MSC layer, which is optimized MSC coverage areas covering the inter-MSC signaling cost and MSC load balancing. In this paper, we focus on the RA optimization, namely layers (i) and (iii). The optimization of MSC coverage areas and LAs is performed in a similar manner. We propose a schema-based niched Pareto genetic algorithm, which deals with multiple objectives by incorporating the concept of Pareto domination in its selection operator, and applying a niching pressure to spread its population out along the Pareto optimal tradeoff surface. The proposed genetic algorithm uses a schema-based partially matching crossover using tournaments of *n* size, where the crossover pairs are chosen in two steps, first based on the class ranking and then schema ranking. New offsprings are modified using the geographical footprints to converge to the optimal solution faster.

1. Introduction

Universal Mobile Telecommunication System (UMTS) is the next-generation mobile communications system, which based on the evolving GSM core network, [1]. The UMTS evolves from General Packet Radio Service (GPRS). The UMTS core network supports both circuit-switched and packet-switched domains. Circuit-switched domain connects the PSTN/ISDN. Packet-switched domain connects to IP/X.25 networks through Gateway GPRS Service node (GGSN). UMTS SGSN (Serving GPRS node) performs mobility management, session management, and implements authentication procedures, and routes data for the UMTS Packet Domain. Signaling Gateway (SG) processes the protocol layers that are involved in the transport of circuit and packet signaling between the UMTS Access Network and either the Call Server (MSC in 2G systems) or the 3G-SGSN. The GGSN provides the point of interconnection with external packet data networks (PDN) for Public Land Mobile Networks (PLMN). The GGSN stores routing information for attached UMTS users. The routing information is used to tunnel Protocol Data Units (PDU) to the current 3G-SGSN serving the MS. Radio Network Controller (RNC) provides control of the multiple radio resources. The HLR is a network database used for permanent management of mobile subscribers within a PLMN. The HLR has been enhanced to include UMTS subscription data and routing information. Call Server (MSC) is responsible for Circuit Domain call processing and circuit-switched data. The Radio Access Network consists of the Radio Network Subsystems (RNS). The RNS is a collection of one RNC and multiple Node Bs. A Node B is an abstract node similar to a Base Transceiver Station (BTS). Each RNC may control multiple Node Bs. The goal of this paper is to optimize the service coverage areas in 2G/3G networks based on user mobility and throughput. We introduce multiple optimization layers to balance the throughput among the coverage areas and reduce the signaling cost, which is due to mobility. The proposed optimization layers use a new schema-based niched Pareto genetic algorithm. The introduction of reduced signaling results in capacity gains for the wireless components due to decreasing the resource waste. Balanced throughput prevents the congestion in the coverage areas, which increases the percentage of service rejection for the mobile users. The outline of this paper is as follows: Section 2 details the routing and location area updates that mobile performs when it moves. In Section 4, we propose hierarchical optimization planes by layering RA and SGSN coverage areas, or LA and MSC/VLR coverage areas. In Section 5, we propose a new niched Pareto genetic algorithm using a new schema-based crossover procedure, which uses two ranking steps, namely class and schema ranking, to select the individuals. Section 6 concludes the paper.

2. Service Coverage Area Definitions

The UMTS network may consist of six different areas: (1) Base Station Controller (BSC) area, (2) Radio Access Controller (RNC) area, (3) Location area (LA), (4) Routing area (RA), (5) MSC/VLR service area VLR, (6) SGSN service area.

The BSC area is an area of radio coverage consisting of one or more cells controlled by one BSC. The boundaries of a BSC area and a location area are independent, and any of them may span the boundary between the other. The RNC area is an area of radio coverage consisting of one or more cells controlled by one RNC. The boundaries of a RNC area and a location area are independent, and any of them may span the boundary between the other. The LA is defined as an area in which a mobile station may move freely without updating the VLR. A location area may include one or several cells. The Routing Area (RA) is defined as an area in which a mobile station, in certain operation modes, may move freely without updating the SGSN. A routing area may include one or several cells. A RA is always contained within a location area. The MSC area is the part of the network covered by an MSC. An MSC area may consist of one or several location areas. An MSC area may also consist of one or several BSC areas. The VLR area is the part of the network controlled by a VLR. A VLR area may consist of one or several MSC areas. In general implementation, each MSC is employed with a VLR. In this case, the area is referred to as MSC/VLR service area. The SGSN area is the part of the network served by an SGSN. An SGSN area may consist of one or several routing areas. An SGSN area may also consist of one or several BSC areas. There need not be a one to one relationship between SGSN area and MSC/VLR area.

3. LA/RA Update

Mobile user needs to update its LA whenever it moves to a different LA. When the mobile moves to a different LA within the same

[1] This study has been performed when the authors were working at Nortel Networks.

MSC/VLR area, it updates the LA using the *intra MSC/VLR update* procedure. If new LA is located in a different MSC/VLR area, the mobile performs the *inter MSC/VLR update* procedure. The intra MSC/VLR update procedure is more frequent and less costly than the inter MSC/VLR update. When mobile moves from one RA to another, it performs intra SGSN or inter SGSN update procedure if it is still within the same SGSN coverage area or not, respectively. Assume that the mobile moves from RA to another, then it needs to perform one of the six procedure types:

(1) Intra-SGSN update (no LA update)
(2) Intra-SGSN and Intra-VLR updates
(3) Inter-SGSN update (no La update)
(4) Inter-SGSN and Intra-VLR updates
(5) Intra-SGSN and Inter-VLR updates
(6) Inter SGSN and Inter-VLR updates

Note that the procedures are itemized from more-frequently and less-costly one to less-frequently and more costly one. For example, item (6) is more costly than item (5), but it is also less frequent one. As an example, RA update types (1), (4) and (5) are depicted in Figure 1.

4. Hierarchical LA/RA Optimization

In this section, we propose a hierarchical network optimization to balance the load among the RAs, LAs, RNCs, MSCs and SGSNs, and to minimize the location management signaling cost in 2G/3G networks. The proposed LA and RA hierarchical optimization methods consist of two optimization planes, namely intra- and inter-level optimization planes. The objectives of each plane are given as follows:

- **<u>RA Optimization</u>**
  - Cell-based Intra-SGSN Level: Optimization Plane-I
    <u>Objectives</u>: Minimizing two cost functions: (*i*) Intra SGSN layer signaling cost (including paging cost), (*ii*) average RA load balancing.
  - RA-based Inter-SGSN Level: Optimization Plane-II
    <u>Objectives</u>: Minimizing three cost functions: (*i*) Inter-SGSN layer signaling cost, (*ii*) average RNC load balancing, (*iii*) average SGSN load balancing.
- **<u>LA Optimization</u>**
  - RA-based Intra MSC/VLR Level: Optimization Plane-I
    <u>Objectives</u>: Minimizing two cost functions: (*i*) Intra-MSC/VLR layer signaling cost (including paging cost), (*ii*) average LA load balancing.
  - LA-based Inter MSC/VLR Level: Optimization Plane-II
    <u>Objectives</u>: Minimizing three cost functions: (*i*) Inter-MSC/VLR layer signaling cost, (*ii*) average BSC load balancing, (*iii*) average MSC/VLR load balancing.

Note that Intra MSC/VLR Level optimization is RA-based since a RA is always contained within a location area. If only 2G system is considered, there will be no RA optimization, and Intra MSC/VLR Level optimization will be cell-based. Since LA optimization is similar to RA optimization, so that we only consider the RA optimization throughout the paper. Figure 2 illustrates the two-level hierarchical RA optimization layers, namely intra-SGSN and inter-SGSN layers. Intra-SGSN Level-Optimization Plane-I optimizes load balance among the RAs, and minimizes the intra-SGSN signaling and paging cost based on cells. Inter-SGSN Level-Optimization Plane-II optimizes the inter-SGSN signaling cost, and balances RNC and SGSN loads based on RAs. Note that inter-SGSN layer is independent of paging cost, since the RAs remains unchanged in this layer. There can be two different network-planning goals; (*i*) planning and optimizing a new network, (*ii*) optimizing an existing network.

(i) *New* network optimization : Optimization Plane-I → Optimization Plane-II
(ii) *Existing* network optimization : Optimization Plane-II → Optimization Plane-I

In a new network, intra-SGSN level is optimized first, where the cell layouts are defined among RAs to optimize RAs. Then, the optimized RAs is given to the inter-SGSN layer, and inter-SGSN level is then optimized the layout of the RAs among SGSNs. In the existing network, since the cell distribution among RNCs has been already determined, the biggest challenge is to minimize the inter-SGSN signaling cost. In order to optimize the network, we need to define the cost functions and the constraints for each optimization plane. Then, the cost functions and constraints apply to an optimization algorithm, where genetic algorithm is very suitable for that type of optimization problems. In this paper, we also propose a schema-based niched Pareto genetic algorithm for the optimization.

5. Multi-objective Schema-based Niched Pareto Genetic Algorithm

We propose a multi-objective schema-based heuristic genetic algorithm to perform the multi-layer optimization. The proposed algorithm is a class of random heuristic search. The network optimization, which is classified as grouping problem, is an NP-complete graph problem. The exact solution procedures cannot be applied beyond a certain problem size, which is a typical of the NP-complete problem.

5.1 Foundation: Genetic Algorithm

Genetic algorithm (GA) has been extensively studied in the recent literature, [5-8]. A genetic algorithm consists of three operators, [5]: *Selection, crossover* and *mutation*. GAs are initialized with a population of guesses, which is referred to as initial population. Selection attempts to apply pressure upon the population in a manner similar to that of natural selection. Crossover allows solution to exchange information to reproduce offsprings, where 2N individuals are generated from N individuals. Mutation is used to randomly change the value of some bits within individuals to preserve the diversity.

5.2 Foundation: Multi-objective Optimization

The GA can deal with multiple objectives by incorporating the concept of *Pareto* domination in its selection operator, and applying a niching pressure to spread its population out along the Pareto optimal tradeoff surface, [8]. Non-dominated individuals, in which there are no other solutions superior in all attributes within the population constructs a surface known as *Pareto optimal front*. The goal of a Pareto GA is to find a representative sampling of solutions all along the Pareto surface. 2-D Pareto surface is given in Figure 3 for the intra-SGSN layer. Cost functions are RA load balancing and intra-SGSN signaling cost, which covers the intra-SGSN routing area update and paging cost. Figure 4 illustrates 3-D Pareto surface for the inter-SGSN layer, since there are three objective functions to minimize, namely inter-SGSN signaling cost, RNC and SGSN load balancing.

5.3 Niched Pareto GA

The use of non-dominated ranking and selection with some kind of niching is suggested to keep the GA from converting to a single point on the front. For example, fitness sharing, which is a niching mechanism, allows the GA to maintain individuals all along the non-dominated surface.

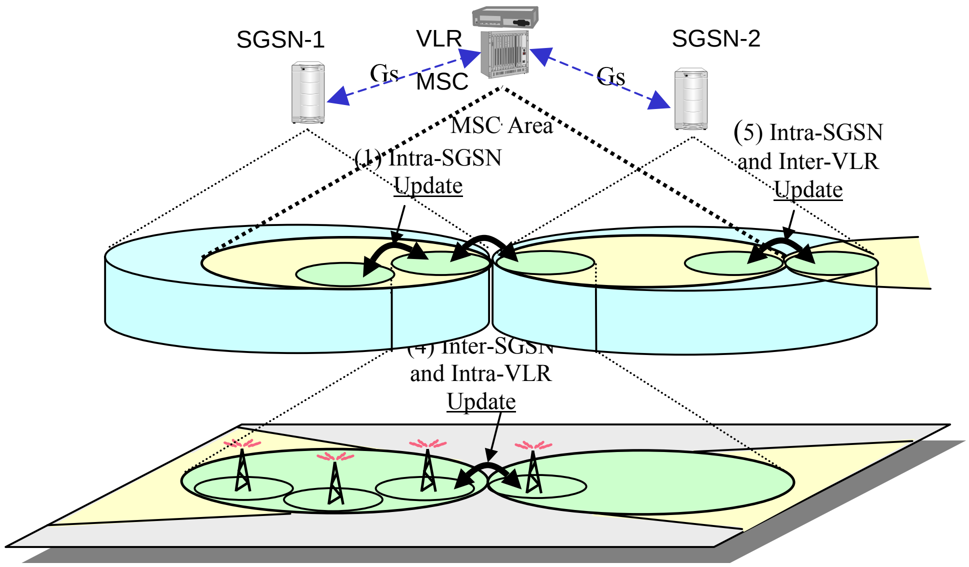

Figure 1. Illustrations of three RA updates; (1) Intra-SGSN update, (4) Inter-SGSN and Intra-VLR update, (5) Inter-SGSN and Intra-VLR update.

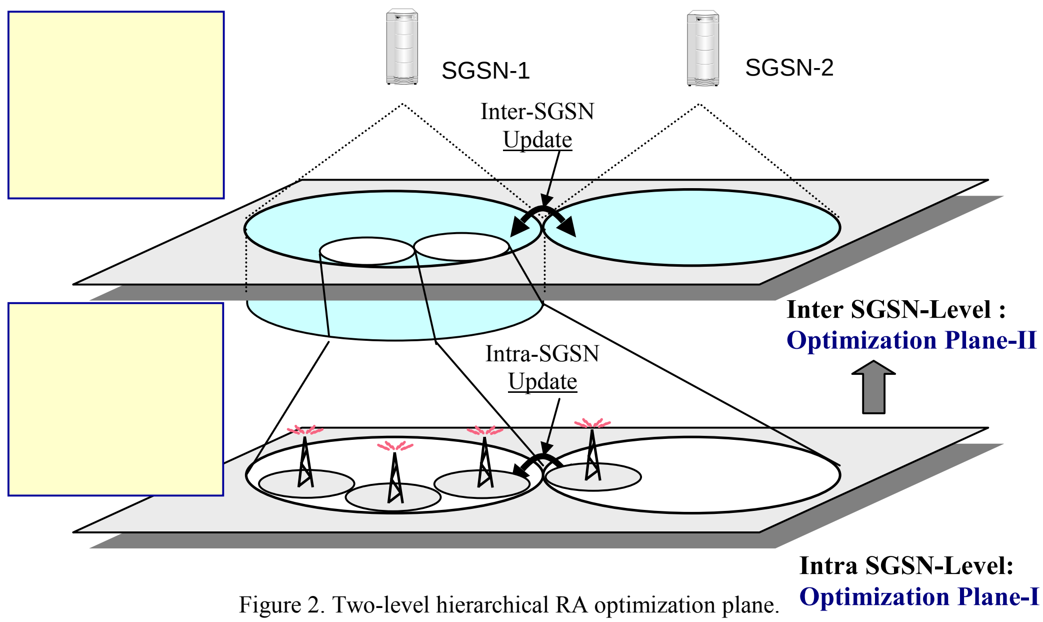

Figure 2. Two-level hierarchical RA optimization plane.

The species of the niched Pareto GA are localized to implementation of selection for the genetic algorithm. One of the widely used selection techniques for GAs is *tournament selection*, [8]. In tournament selection, a set of individuals is randomly chosen from the current population and the best of this subset is placed in the next population. The selection pressure is controlled by the sample size, $t_{dom.}$ The performance of niched Pareto GA is sensitive to amount of selection pressure. A good value of t_{dom} is around 10% to provide a tight and complete distribution, and decrease the convergence period comparing to no sampling. Higher values of t_{dom} cause the algorithm to converge to a small portion of the front.

The species of the niched Pareto GA are localized to implementation of selection for the genetic algorithm. One of the widely used selection techniques for GAs is *tournament selection*, [8]. In tournament selection, a set of individuals is randomly chosen from the current population and the best of this subset is placed in the next population. The selection pressure is controlled by the sample size, $t_{dom.}$ The performance of niched Pareto GA is sensitive to amount of selection pressure. A good value of t_{dom} is around 10% to provide a tight and complete distribution, and decrease the convergence period comparing to no sampling. Higher values of t_{dom} cause the algorithm to converge to a small portion of the front.

5.3.1 Selection Process

The selection process is summarized as follows: Two candidates for selection are picked at random from the population. A comparison set of individuals, which is defined by the sampling rate, is also picked randomly from the population. Each of the candidates is then compared against each individual in the comparison set. If one candidate is dominated by the comparison set, and the other is not, the non-dominated candidate is selected. If neither or both are dominated by the comparison set, there are two ways to pick the winner: (*i*) A candidate with the less crowded neighborhood, where the neighborhood is defined by a radius r_{share}, is selected; (*ii*) the degradation of individual's fitness is applied according to a sharing function, which decreases as the number of neighbors increases within r_{share}. In our algorithm, we define r_{share} as

$$r_{share} = 0.1 \times \min\{c_i^j\}, \quad j = 1, \ldots, k$$

where c_i^j is the cost function of individual *i*, and *k* is the total number of cost functions. We choose 10% of the minimum cost function as the neighborhood indicator.

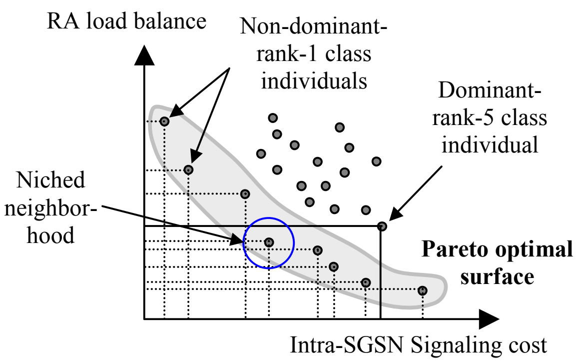

Figure 3. Pareto optimal surface and dominant-class for the intra-SGSN layer.

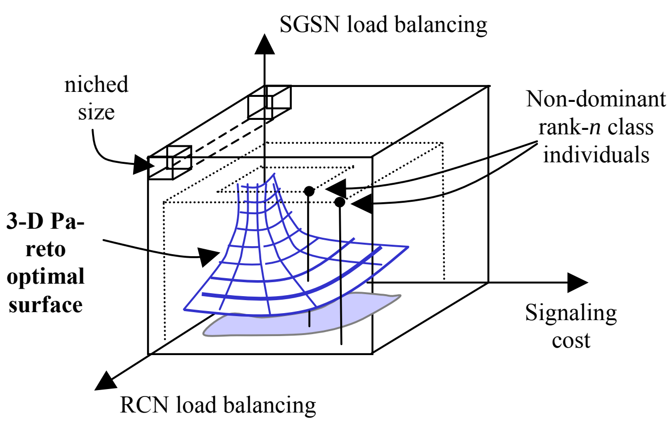

Figure 4. 3-D Pareto optimal surface for the inter-SGSN layer, where the objective functions are inter-SGSN signaling cost, RNC and SGSN load balancing.

5.3.2 A New Schema-based Crossover for Niched Pareto GA

In this section, we propose a new schema-based crossover procedure to optimize intra- and inter-SGSN layers. Note that we use the niched Pareto GA for the selection mechanism. In general, two individuals, whom basically have binary strings as the chromosome, exchange one or two bits to generate two new offsprings. Partially matching crossover with preprocessing is a well-known method for grouping, where a mapping section is selected randomly from two chromosomes, and the genes of the section is exchanged with attention to ensure that the resulting chromosomes will remain valid. Instead of selecting the two chromosomes randomly for crossover, we select one of the chromosomes randomly, and the second one is selected from a group of chromosomes based on schema.

Definition: *Schema*

A schema (plural schemata) is a fixed template describing a subset of strings with similarities at certain defined positions, [5]. Thus, strings, which contain the same schema, contain similar information to some degree.

In our algorithm, a schema degree between two individuals is based on the number of cells within the RAs of the two individuals. A network cell-layout of two individuals is given in Figure 5. Referring to the figure, RA-1 of parent A and RA-2 of parent B cover nine same cells within their routing areas. RA-2 of parent A and RA-3 of parent B, and RA-3 of parent A and RA-1 of parent B have six and thirteen same cells, respectively. The total same cell coverage is 28 out of 42 cells. Therefore, the schema degree of these two individual network layouts is 28.

We propose a schema-based partially matching crossover using tournaments of *n* size. The proposed crossover is as follows:

A first chromosome is first chosen randomly from the population. Then, a set of chromosomes is picked randomly, which is referred to as *n*-size tournament. The size of *n* is the user parameter, and may be picked similar to value t_{dom}. There are two steps, namely class ranking and schema ranking, to pick the second chromosomes from the selected tournament:

i. Class Ranking:
The set of chromosomes to pick the second chromosome is chosen from the selected tournament using an exponentially decreasing function versus the ranking classes in the set. The ranks are given according to the status of non-dominance and dominance class.

ii. Schema Ranking:
The second chromosome to perform the crossover is chosen from the selected set of dominance class individuals using an exponentially decreasing function versus the schema ranking (degree) of the chromosomes in the selected set.

Once the second chromosome is picked, the randomly selected RAs of the first chromosome are overlapped over the second chromosome. The number of RAs to perform the crossover is user dependent, however, it cannot be a high percentage. The core area is chosen according to the highest schema between the RA of the first parent and the overlapped RAs of the second parent. The new RA in the generated offspring is optimized by the hill climbing method using the RA-based geographical footprint boundaries in the intra-SGSN layer and SGSN-based geographical footprint boundaries in the inter-SGSN layer, which generate many smaller areas around the core RA or core SGSN.

RA-based or SGSN-based geographical footprints include boundary cells between RAs and boundary RAs between SGSNs, respectively. These boundary cells and RAs are important in terms of mobility and routing area update cost. They also play a role to reshape the cost functions for the next recursive step of genetic algorithm.

5.3.2.1 Example: Intra-SGSN Layer Crossover

Figure 5 illustrates the crossover between two chromosomes for the intra-SGSN layer optimization, where the chromosomes have three RAs. The RA-1 of Parent A is picked randomly, and overlapped over Parent B. The core area is chosen as the overlapping area between RA-1 of Parent A and RA-2 of Parent B, and this area is referred to as the core RA in the offspring.

Figure 6 shows the core area and overlapping areas of the six offspring candidates generated by overlapping RA-1 of Parent A to Parent B. The offspring candidates differ from each other using the RA-based geographical footprint boundaries, where RA-based geographical footprint consists of neighboring cells between tow RAs. The best fit of the offspring candidates is chosen as the winner. The neighboring RAs in the offspring are adjusted accordingly.

The crossover from Parent B to parent A is performed in the same manner, where an RA of Parent B is chosen randomly for crossover. In this way, two offsprings are generated from two parents in order to generate 2*N* individuals from *N* individuals.

5.3.3 Mutation

Single cell mutation is performed within the new generated offsprings. The mutation percentage is user dependent. In our algorithm, we only mutate one cell in the intra-SGSN layer optimization, where the mutated cell is located in the RA-based geographical footprint boundary. One RA, which is a member of the SGSN-based geographical footprint boundary, is mutated in the inter-SGSN layer optimization

5.3.4 Property: Dynamic Algorithm

A network candidate is selected from non-dominant ranked individuals after both layers are optimized. However, in the case of mobility pattern or throughput changes, this network layout may not provide the best solution. In order to guarantee the best solution, we propose the following method to adjust the network dynamically in the case of network condition changes:

i. The cost functions of the selected individual should be calculated periodically for certain periods.
ii. If the cost functions of the selected individual are 10% or more different than its original value for a certain number of times, the network should be optimized again. This provides a dynamic adjustment property to the proposed optimization algorithms.

6. Conclusions

In this paper, we propose a hierarchical network optimization to balance the load among the RAs, LAs, RNCs, MSCs and SGSNs, and to minimize the location and routing area update signaling cost in 2G and 3G networks. We propose a multi-objective schema-based niched Pareto Genetic Algorithm for the optimization. The proposed genetic algorithm is based on schema-based crossover procedure for the niched Pareto Genetic Algorithm using two ranking steps, namely class and schema ranking, to select the individuals for crossover. Calculation of cost functions and simulation results for the intra-SGSN and inter-SGSN layers will be given in the future studies.

References

[1] ITU - UMTS Standards.
[2] I. A. Cimet, ``How to Assign Service Areas in a Cellular Mobile Telephone System,'' *Proc. of ICC'94*, p.197-200, 1994.
[3] P. R. Gondim, ``Genetic Algorithms and the Location Area Partitioning Problem in Cellular Networks,'' *Proc. of VTC'96*, p.1835-1838, 1996.
[4] P. Demestichas et al. ``Control of the Location Update and Paging Signalling Load in Cellular Systems by means of Planning Tools,'' *Proc. of VTC'99-Fall*, p.2119-2123, 1999.
[5] D. A. Coley, *An Introduction to Genetic Algorithms for Scientists and Engineers*, World Scientific, 1999.
[6] G. Magyar, M. Johnsson, O. Nevalainen, ``An Adaptive Hybrid Genetic Algorithm for the Three-Matching Problem,'' *IEEE Trans. on Evo. Comp.*, Vol. 4, No2, p.135-146, July 2000.
[7] C. M. Fonseca, P. J. Fleming, ``Multiobjective optimization and multiple constraint handling with evolutionary algorithms. I. A unified formulation,'' *IEEE Transactions on Systems, Man and Cybernetics*, Part A, , Vol. 28, No. 1, p.26-37, Jan 1998.
[8] J. Horn, N. Nafpliotis, D. E. Goldberg, ``A Niched Pareto Algorithm for Multiobjective Optimization,'' Proc. of the First IEEE Conference on Evolutionary Computation, p.82-87, 1994.

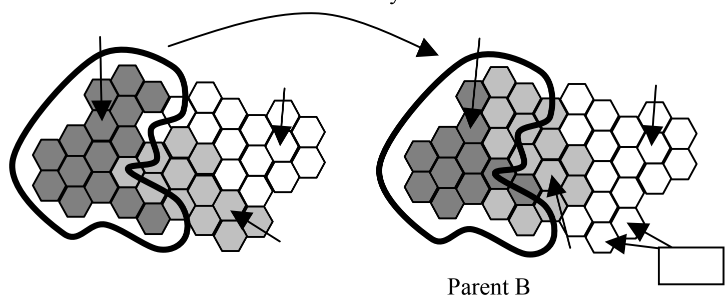

Figure 5. Two parents performing a crossover; RA-1 of primary Parent A is carried over RA-2 of secondary Parent B. Routing area to carry over between two parents are chosen randomly. The core area is chosen according to the highest schema between the exchanged RA of the Parent A and the overlapped RAs of Parent B.

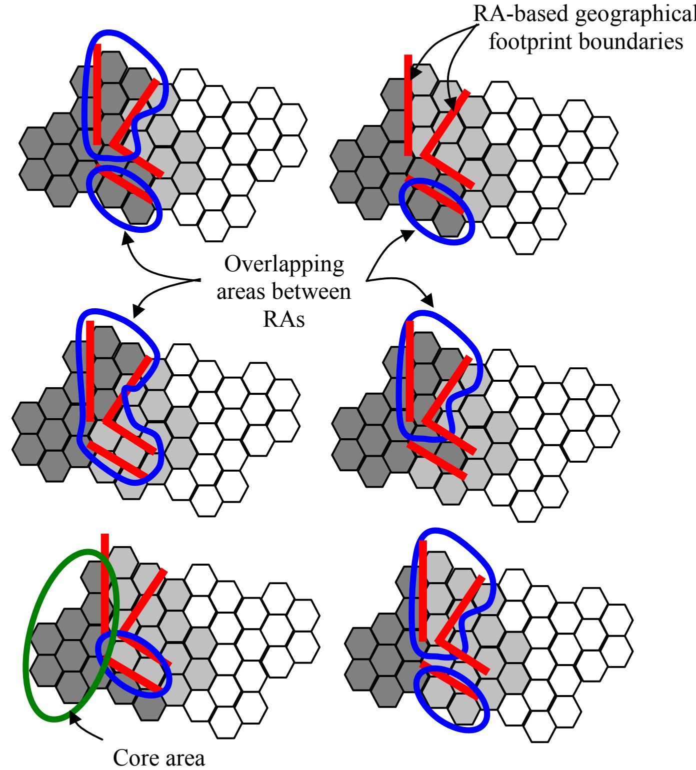

Figure 6. Six offspring candidates generated according to the crossover from primary Parent A to secondary Parent B: Lines shows the RA-based geographical footprints, circles shows the difference from the secondary Parent B. One of six candidates is the winner based on the fitness. Six candidates differ from each other by RA-based geographical footprints.