

The Positioning of Base Station in Wireless Communication with Genetic Approach

Yong Seouk Choi¹, Kyung Soo Kim¹, Nam Kim²

¹ETRI, 161 Gajeong-Dong, Yuseong-Gu, Daejeon, Korea
choiys@etri.re.kr

²Chungbuk National University, 12 Gaeshin-Dong, Heungduk-Gu, ChungJu, Korea

Abstract. This paper addresses the displacement of a base station with optimization approach. A genetic algorithm is used as optimization approach. A new representation that describes base station placement, transmitted power with real numbers and new genetic operators is proposed and introduced. In addition, this new representation can describe the number of base stations. For the positioning of the base station, both coverage and economy efficiency factors were considered. Using the weighted objective function, it is possible to specify the location of the base station, the cell coverage and its economy efficiency. The economy efficiency indicates a reduction if the number of base stations for cost effectiveness. To test the proposed algorithm, the proposed algorithm was applied to homogeneous traffic environment. Following this, the proposed algorithm was applied to an inhomogeneous traffic density environment in order to test it in actual conditions. The simulation results show that the algorithm enables the finding of a near optimal solution of base station placement and it determines the efficient number of base stations. Moreover, it can offer a proper solution by adjusting the weighted objective function.

1. Introduction

Base station placement is a highly important issue in achieving high cell planning efficiency. It is expected that third generation wireless systems will provide a great variety of services. Thus, cell planning should be carried out considering inhomogeneous traffic. The placement of base stations depends on the traffic density, channel conditions, interference scenario, the number of base stations, and the other network planning parameters; as a result, it is a very complex issue. A genetic algorithm is useful for solving this type of complex problem. This method represents feasible solutions in terms of individuals with genomes, and determines which individuals could survive in a certain criterion formulated to maximize (or minimize) a given objective function. In several studies, a genetic approach has been used to find the best possible base station placement [1, 2]. Binary string representation is applied in [1], and a hierarchical approach is considered in [2]. However, those approaches have a representation limit, and a lot of trials can not guarantee an optimum result, as the possible base station positions are discrete.

In this paper, a new representation describing base station placement is suggested, and is one which uses a real number and introduces new genetic operators. The proposed representation can determine not only the locations of the base stations but also the number of base stations.

In addition, the transmitted power of base station is considered as a factor of the proposed algorithm. To consider both coverage and the economy efficiency, an objective function with a weighted factor is established. The proposed algorithm is verified by applying it to homogeneous traffic density case as an obvious optimization problem. In addition, the approach is tested in an inhomogeneous traffic density environment.

2. Overview of Genetic Algorithm

Like other computational systems inspired by natural systems, genetic algorithms have been used in two ways: as techniques for solving technology problems, and as simplified scientific models that can answer questions about nature [3]. Genetic algorithms (GA) are evolutionary optimization approaches which are an alternative to traditional optimization methods. GA approaches are most appropriate for complex non-linear models where location of the global optimum is a difficult task. It may be possible to use GA techniques to consider problems which may not be modeled as accurately using other approaches. Therefore, GA appears to be a potentially useful approach. GA performance will depend very much on details such as the method for encoding candidate solutions, the operator, the parameter setting, and the particular criterion for success. As for any search, the way in which candidate solutions are encoded is very important. Many genetic algorithm applications use fixed-length, fixed-order bit strings to encode candidate solution. However the algorithm proposed in this paper uses real-valued encoding schema to represent solutions. In GA, feasible solutions are modeled as individuals described by genomes. A genome is an arrangement of several chromosomes, which symbolize characteristics of the individual. Population is the total amount of individuals. Some of them can survive and others will die in the next generation by their own fitness and a given selection rule. Fitness is evaluated by a given objective function. Genetic operations such as crossover and mutation are performed to produce new individuals in subsequent generations. The crossover operator defines the procedure of generating a child from its parent's genomes. The mutation is carried out chromosome by chromosome, and its exploration and exploitation helps the algorithm to avoid local optimum. If the current population accepts the given termination condition, new generation is no longer produced. Otherwise, dominant individuals are selected and genetic operators reproduce new individuals from them. The best individual of each generation is transferred over to the next generation if elitism is adopted.

The theoretical basis of GA relies on the concept of schema. A schema is defined as the similarity of templates describing a subset of genomes with similarities in certain chromosomes. Schemata are available to measure the similarity of individuals. John Holland's schema theorem and building-block hypothesis [4] have often been used to explain how the GA works. According to the schema theorem, short, low-order, and

above-average schemata receive exponentially increasing trials in subsequent generations. This proves that the individuals with high fitness will have a high survival probability when a suitable representation is applied. The building-block hypothesis suggests that the GA will perform well when it is able to identify above-average-fitness and low-order schemata and recombine them to produce higher-order schemata of higher fitness. In sum, individuals with similar characteristics must be represented by a similar genotype.

3. Proposed Algorithm for Base Station Placement

The processing of the proposed algorithm is implemented in a two-dimensional map; therefore representation in binary form is difficult to present for the genome which describes the number of base stations and the location of the base stations. For a good approximation, it is necessary to have a longer genotype. A real value representation is more efficient than the representation of a binary genome in this case. Consequently, in this paper the genotypes that have real value representations for the optimization algorithm were chosen. Given the allowable transmitted power of a cell site in a traffic map, this chapter introduces GA that optimizes the cell site location, the number of cell sites and the transmitted power. A GA that works well in terms of the base station placement problem is proposed. The main characteristics considered for the development of the proposed algorithm are:

- <1> The genome must represent all of the base station locations, and the genotype can describe the number of base stations as well as the position of the base station.
- <2> A chromosome expresses one base station position.
- <3> The number of possible base station locations must be unlimited; therefore, there are infinite candidates of base station locations.
- <4> Similar genotypes represent the genomes of the closely located base stations.

An algorithm satisfying the above factors is consistent with the building-block hypothesis and schema theorem.

The three things that must be defined in order to solve a problem through genetic algorithms are as follows:

- Define a representation
- Define the genetic operators
- Define the objective function

How one defines a representation, genetic operators, and objective function determines the algorithm. It is essential to design the genetic algorithm by considering <1>~<4>. The following sections explain the proposed algorithm in detail.

3.1 Representation

Fig. 1 illustrates the representation of the genomes. A genome is denoted as a vector $g = (c_1, \dots, c_K)$ where $c_k = (x_k, y_k, pwr_k)$ is the chromosome for the k -th base station position. This method fulfills <1> and <2>. K is the maximum number of base stations, and all of these can be located in the x-range $[-X_{\max}, X_{\max}]$ and y-range $[-Y_{\max}, Y_{\max}]$ with origin $(0,0)$.

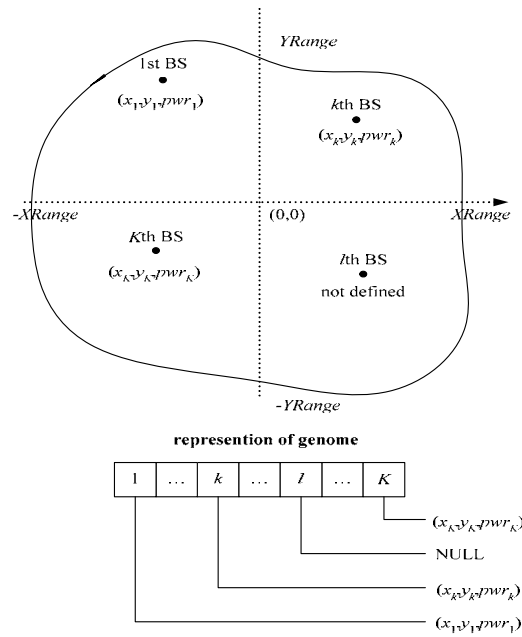


Figure 1. Representation of the genome for the placement of the base station

If the position of a base station is not defined, it is expressed as NULL. This method applies for a case in which there are fewer base stations than in K , so that it fulfills <1>. $n(g)$ is defined as the number of EXISTENCE in g . In order to satisfy <3> and <4>, x_k , y_k and pwr_k must be real numbers. M is assumed as population size.

3.2 Genetic Operators (Crossover and Mutation)

It is necessary to design an initialization and a termination method, a crossover and mutation operator, and a selection strategy in order to define the reproduction procedure.

A proper initial population can provide a fast convergence to the optimum point. It is desirable for a user to define initial positions of base stations intuitively. The first

individual, $c_{1k} = (x_{1k}, y_{1k})$ for $k = 1, \dots, K$, is determined by a user and other individuals (for $m = 2, \dots, M$) are determined by the following rule: If $c_{1k} = \text{NULL}$, then $c_{mk} = \text{NULL}$ with probability P_n^I or $c_{mk} = (v_1, v_2)$ with probability $1 - P_n^I$, where $v_1 = U(-X_{\max}, X_{\max})$ and $v_2 = U(-Y_{\max}, Y_{\max})$. If c_{1k} is defined ($c_{1k} \neq \text{NULL}$), then $c_{mk} = \text{NULL}$ with probability $1 - P_v^I$ or $c_{mk} = (x_{1k} + \xi_1, y_{1k} + \xi_2)$ with probability P_v^I , where $\xi_1, \xi_2 = N(0, \sigma_S^2)$. $U(a, b)$ is a uniformly distributed random variable between a and b . $N(\bar{x}, \sigma^2)$ denotes a Gaussian distributed random variable with mean \bar{x} and variance σ^2 . P_n^I and P_v^I indicate the probability of producing NULL from NULL and that of producing EXISTENCE from EXISTENCE, respectively. However, it may require further trials in order to determine the global optimum if the initial value, as user defined, is close to the local optimum. When the user does not define any initial positions, it is decided that $c_{mk} = \text{NULL}$ with \tilde{P}_n^I or $c_{mk} = (v_1, v_2)$ with probability $1 - \tilde{P}_n^I$ for $m = 1, \dots, M$, where \tilde{P}_n^I denotes the probability of producing NULL.

A termination criterion is used to determine whether or not a GA is finished. Generation, convergence, or population convergence can terminate the procedure of genetic algorithm. The easiest scheme is termination upon generation. When the number of current generations is larger than the specified number of generations, the algorithm is finished. Termination upon convergence compares the previous best-of-generation to the current best-of-generation. If the current convergence is less than the requested convergence, the reproduction procedure is ceased. Termination upon population convergence compares the population average to the score of the best individual in the population.

In the proposed application, one child crossover operator is used. A single child c_k^{child} is born from its father and mother, c_k^{dad} and c_k^{mom} . Fig. 2 shows the procedure of one child crossover operation in the proposed algorithm. If one of the parents is NULL, the child receives the other parent's attributes. Otherwise, the child is generated by (1), where σ_C is the parameter of the crossover operation.

$|x_k^{\text{dad}} - x_k^{\text{mom}}|$ and $|y_k^{\text{dad}} - y_k^{\text{mom}}|$ can be used as a measure of closeness. This method is based on the fact that if the attributions of both parents are similar, the child's attributions are also similar to its parents.

Mutation is performed chromosome by chromosome with probability P_{mut} . Fig. 3 shows the procedure of the mutation operation in the proposed algorithm. The mutation is very close to the initialization scheme with the user-defined base station position. If $c_{mk} = \text{NULL}$, redefine $c_{mk} = \text{NULL}$ with probability P_n or $c_{mk} = (v_1, v_2)$ with probability $1 - P_n$. If $c_{mk} \neq \text{NULL}$, redefine $c_{mk} = (x_{mk} + \chi_1, y_{mk} + \chi_2)$ with probability P_v or $c_{mk} = \text{NULL}$ with probability $1 - P_v$, where χ_1 and χ_2 are Gaussian distributed random variables with zero mean and variance σ_m^2 . P_{mut} and σ_m^2 are the parameters of the mutation operation.

A roulette wheel method is applied for the selection scheme. This selection method chooses an individual based on the magnitude of the fitness score relative to the rest of the population. The higher the score, the more selective an individual will be. Any individual has a probability p of the choice where p is equal to the fitness of the individual divided by the sum of the fitness of each individual in the population. Therefore, the individual with a high fitness level can survive with high probability.

$$x_k^{\text{child}} = \frac{x_k^{\text{dad}} + x_k^{\text{mom}}}{2} + \zeta_1, \quad \zeta_1 = N\left(0, \left(\frac{(x_k^{\text{dad}} - x_k^{\text{mom}})\sigma_C}{2}\right)^2\right)$$

$$y_k^{\text{child}} = \frac{y_k^{\text{dad}} + y_k^{\text{mom}}}{2} + \zeta_2, \quad \zeta_2 = N\left(0, \left(\frac{(y_k^{\text{dad}} - y_k^{\text{mom}})\sigma_C}{2}\right)^2\right). \quad (1)$$

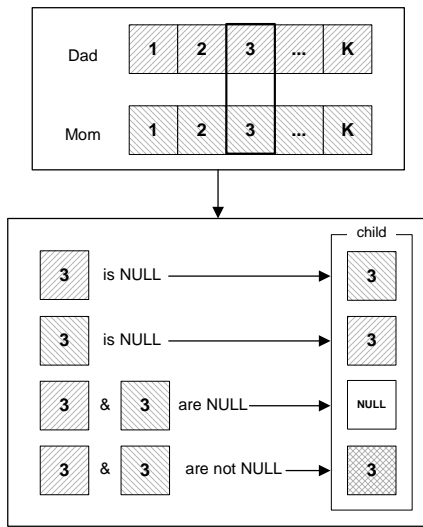


Figure 2. One child crossover operation

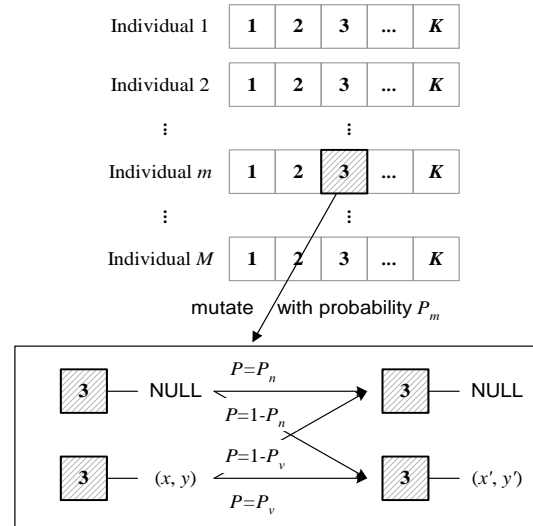


Figure 3. Mutation operation

3.3 Fitness Evaluation

Fig. 4 illustrates the fitness evaluation procedure composed of an evaluator and an objective function. The evaluator calculates the covered traffic by using a propagation model, traffic map, and map for a path loss prediction. Cell area covered by the base

stations is evaluated, and the covered traffic is then obtained. Considering coverage, power and economy efficiency, the objective function is defined as

$$f(G) = \omega_t \cdot f_t(G) + \omega_p \cdot f_p(G) + \omega_e \cdot f_e(G)$$

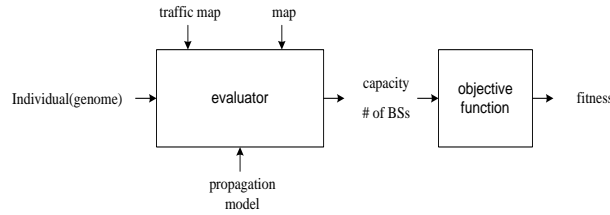


Figure 4. Fitness evaluation

where f_t , f_p and f_e are the objective functions for coverage, power and economy respectively, and these are defined as :

$$f_t(G) = \text{traffic coverage rate} = \frac{\text{covered traffic}}{\text{total traffic}}$$

$$f_p(G) = \text{BS power fitness} = \frac{\text{Available Maximum BS power} - \text{Used BS power}}{\text{Available Maximum BS power}}$$

$$f_e(G) = \text{economic fitness} = \frac{\text{Available Maximum BSs} - \text{Used BSs}}{\text{Available Maximum BSs}}$$

As the covered traffic area widens corresponding to the given propagation model, $f_t(G)$ increases. Conversely, $f_e(G)$ increases when fewer base stations are placed. Total fitness is calculated with w_t , w_p and w_e subject to $w_t + w_p + w_e = 1$. The weights are determined by the user's preference. If coverage is more important, then one may choose a large w_t . Otherwise, a large w_e may be chosen to be more desirable using fewer base stations. Therefore, the purpose of optimization in this paper is to determine the maximum traffic coverage with the minimum number of base stations and minimum amount of power.

This paper uses Hata's model to obtain the coverage of the base station.

It is possible for each individual can have K(the maximum number of base stations). To achieve the cell coverage, it is necessary to compute the path loss K times. If the population is large, the computing power required becomes very large. In this paper, to reduce processing time, Hata's model was used, which is fast for computing the path loss with height information.

4. Testify Algorithm

To test the proposed algorithm, a one-tiered hexagonal cellular environment is considered, where traffic is distributed uniformly in each hexagonal cell whose radius is 2.5 km. In this case, the optimum position of the base station is in the center of hexagon, and the optimum number of base stations is obviously seven. A path loss prediction is carried out using the equation $L = L_0 \times (d/d_0)^4$, where $L_0 = 140$ dB and $d_0 = 2.5$ km. As the generation increases, the base stations tend to be placed where they are optimum, and the number of base stations is also converged automatically. After the 1000th generation, a base station placement that guarantees 99.78% coverage can be determined.



Figure 5. Homogenous traffic density for verification

The maximum number of base stations depends on the width of the target area. The wider the target area, the more likely a greater amount of computing time for convergence is needed. Population size is the solution set. If the population size is large, the convergence of the solution can be quickest. However, in this case the total computing time is larger, as a processing of the propagation model will be needed for each individual in the population. As the individuals with low fitness values are removed, the initial values of base station's maximum number and location are not related to the entire performance. Therefore, a null-to-null probability and Pos-to-Pos probability is loosely coupled with the fitness relationship, and the mutation probability in a real-value representation is the main factor in speeding the convergence.

Fitness with various mutation probabilities in each generation is shown in Fig. 6. The higher the mutation probability, the better the fitness. However, too high a mutation probability has a tendency to downgrade the performance, as it has a frequently changing possible solutions set. In the given homogenous traffic in Fig. 5, it is known that the best performance is shown when the mutation probability is 0.1. (Fig. 6)

Fig. 7 shows that a high deviation of mutation will be good for performance.

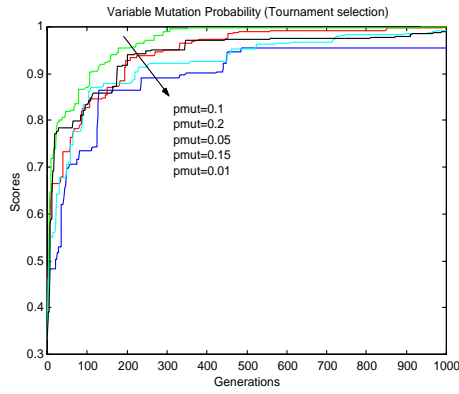


Figure 6. Fitness in various mutation probabilities

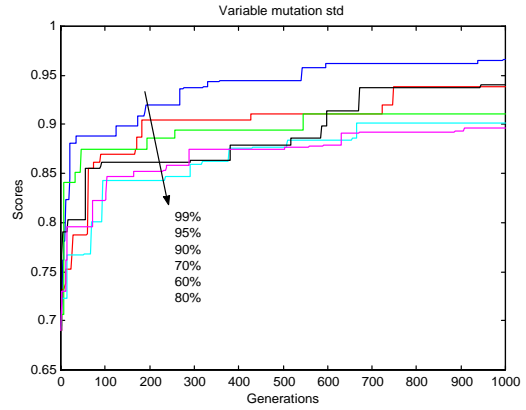


Figure 7. Fitness in various mutation deviations

Figs. 8 to 12 show the optimization processing of base station displacements. Fig. 8 shows the initial random location of the base stations, and in this case five base stations have covered 69% of the target area. In Fig. 9, seven base stations have covered 92% of target area with uniform selection, but it is still not optimized. Fig. 10 is the result of a Roulette wheel selection, and this is an improvement over the uniform selection. It covers 93.85% of the target area. The Rank selection covers 97.90%; this is a very good result. The Tournament selection offers 99.78% coverage. This is approximately at the optimization level. As fitness is sensitive in terms of selection schemes, optimization processing needs appropriate selection schemes.

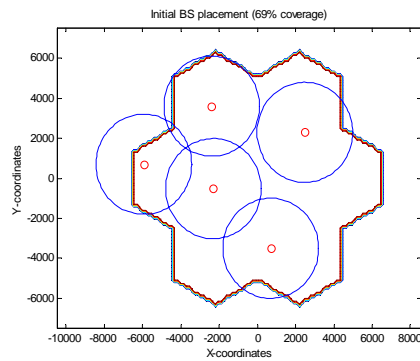


Figure 8. Initial base station location

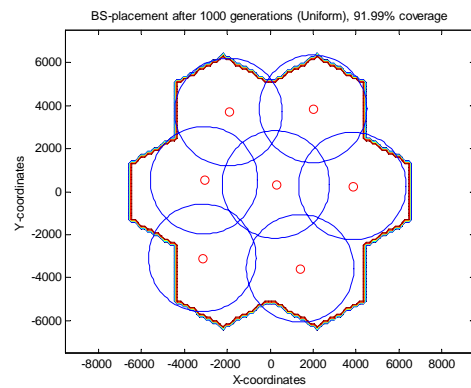


Figure 9. After the 1000th generation, base station location with uniform selection

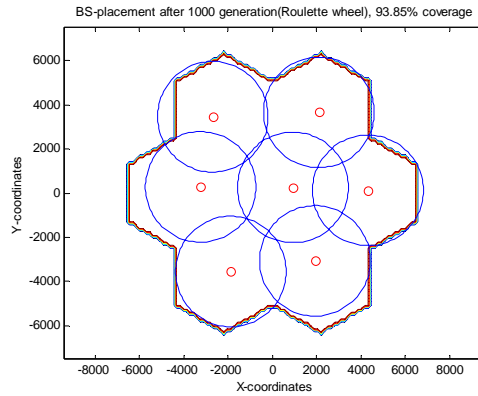


Figure 10. After the 1000th generation, base station location with Roulette wheel selection

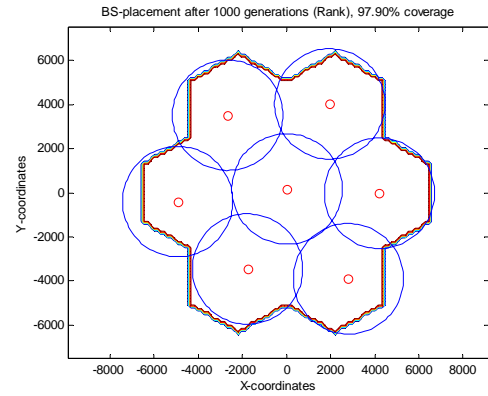


Figure 11. After the 1000th generation, base station location with Rank selection

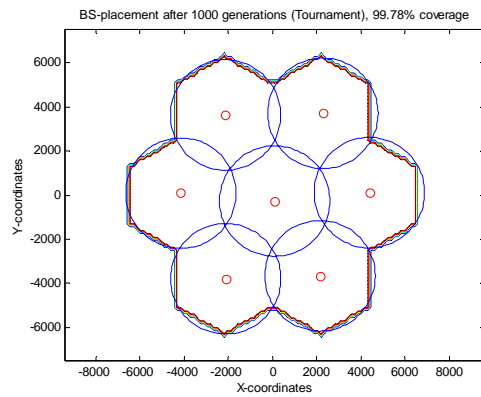


Figure 12. After the 1000th generation, base station location with Tournament selection

5. Simulation Results

To demonstrate if the proposed algorithm determines which positions match optimum location, a simulation was conducted on areas similar to that in Figs. 13 and 14 (inhomogeneous traffic). The actual valued representations in this paper, as mentioned above, consist of the candidate location of the base station's transmit power. Fig. 13 shows the altitude map of the target areas, and Fig. 14 shows the traffic density map. The traffic density is inhomogeneous and the target area for simulation is an urban pattern. The width of the area for simulation is 12Km x 12Km and the size of the bin is 120m x 120m. Therefore, the total number of bins is 10,000.

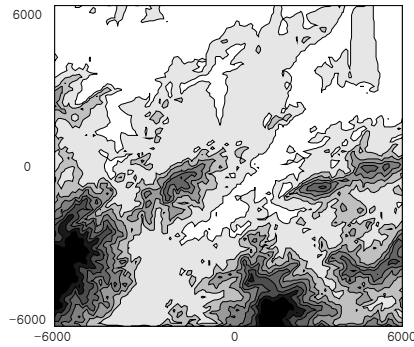


Figure 13. Altitude Map

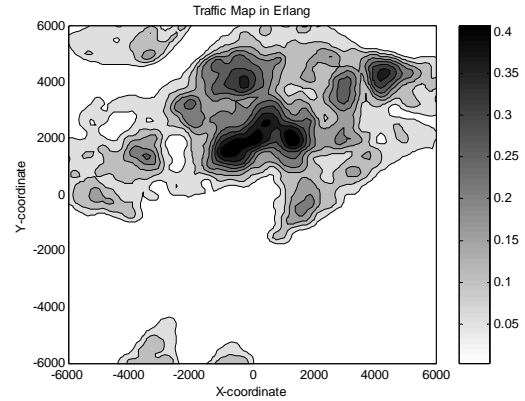


Figure 14. Traffic density map

Figs. 15 and 16 show the location of the base station from one generation to 500 generations, when the weighting condition of their object function is $(\omega_t, \omega_p, \omega_e) = (0.9, 0.0, 0.1)$. The assigned transmit power range of each base station is from 22.63 dBm to 39.36dBm, and its mean value is 33.84 dBm. In this case, the coverage rate is 82.62% and the fitness value is 0.74258.

In the case where the condition of object function is $(\omega_t, \omega_p, \omega_e) = (0.8, 0.1, 0.1)$, the results are shown in Figs. 17 and 18. The coverage rate is 77.47%, and the fitness value is 0.663181. The assigned transmit power range of each base station is from 21.1752 dBm to 38.57794dBm, and its mean value is 32.3230 dBm. As the traffic capacity is limited, the cell boundaries of the high traffic density are less than those of the low traffic density. The coverage rate is decreased according to the changing weight of the traffic factor, from 0.9 to 0.8. As the weight of the power factor increases, the actual assigned transmit power value decreases. In the results shown in Fig. 17, the overlapped base station is clearly shown. The cause of this is the decrease of the weighted economy factor. The traffic map that was used for the simulation consisted of high traffic density areas and very low traffic density areas such as mountains and rivers. Therefore, traffic is scattered in all directions on the map; consequently, the search space becomes larger. To obtain a better coverage rate, the population size can be enlarged or the mutation probability can be increased. Additionally, it is necessary to process more generations.

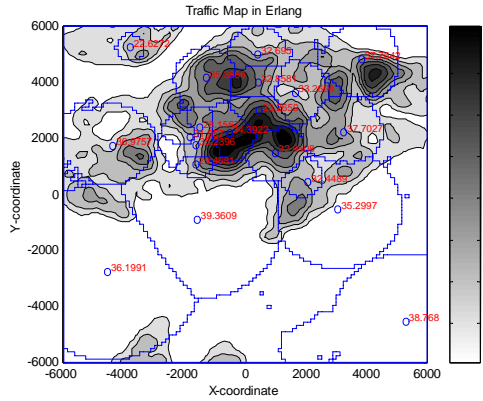


Figure 15. After 500 generations, the location of the base stations, $(\omega_t, \omega_p, \omega_e) = (0.9, 0.0, 0.1)$

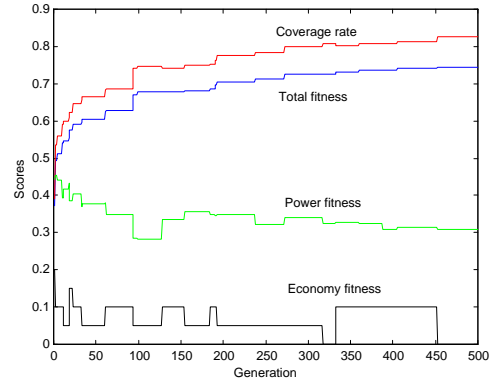


Figure 16. Fitness Value, $(\omega_t, \omega_p, \omega_e) = (0.9, 0.0, 0.1)$

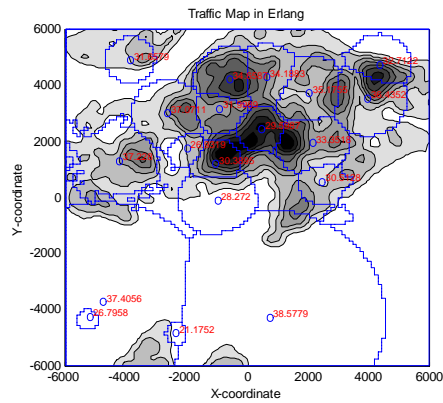


Figure 17. After 500 generations, the location of the base stations, $(\omega_t, \omega_p, \omega_e) = (0.8, 0.1, 0.1)$

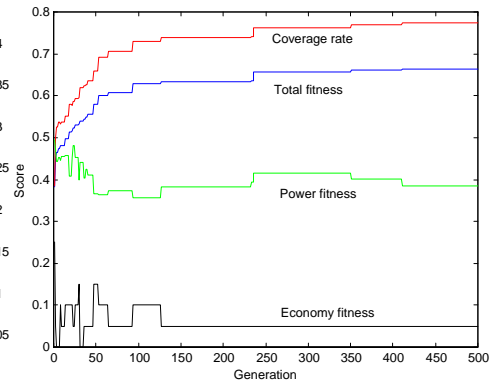


Figure 18. Fitness values, $(\omega_t, \omega_p, \omega_e) = (0.8, 0.1, 0.1)$

6. Conclusion

In this paper, given inhomogeneous traffic information and the map for the propagation model, a new algorithm was proposed that enables the optimization of the locations and transmitted power of a base station. In addition, this algorithm includes an economic factor (the number of base stations). Good use was made of the genetic algorithm and it was excellent for obtaining a solution of complex problems. Genetic operators using the real valued representation are also suggested, and the objective

function is defined in consideration of the coverage, the transmitted power of base station and the economy efficiency through an adjustment of crossover and mutation. The selection, input parameters and scaling are shown to be tightly coupled with the algorithm performance. Therefore, there is a need for these to be harmonized. From a simulation, the proposed algorithm was verified.

References

- [1] X. Huang, U. Nehr, and W. Wiesbeck, "Automatic Base Station Placement and Dimensioning for Mobile Network Planning," Proc. IEEE VTC 2000 Fall, vol. 4, pp. 1544-1549, 2000.
- [2] Melanie Mitchell, *An Introduction to Genetic Algorithms*, The MIT Press, 1996.
- [3] Holland J.H., *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, 1975
- [4] Thomas Back, Frank Hoffmeister, Hans-Paul Schwefel, "A Survey of Evolution Strategies", Proceedings of the 4th International Conference on Genetic Algorithms, pp.2-9, San Diego, CA, July, 1991
- [5] Xuemin Huang, Ulrich Behr, Werner Wiesbeck, "Automatic Base Station Placement and Dimensioning for Mobile Network Planning", IEEE Vehicular Technology Conference, October, 2000
- [6] A. H. Wright, *Genetic Algorithms for Real Parameter Optimization*, in *Foundations of Genetic Algorithms* (Ed. J. E. Rawlins), Morgan Kaufmann, 1991
- [7] Hata. M, "Empirical Formula for Propagation Loss in Land Mobile Radio Services", IEEE Transactions on Vehicular Technology, Vol.VT-29, No.3, pp.317-325, August 1980
- [8] C.M. Fonseca and P.J. Fleming, "An Overview of Evolutionary Algorithms in Multiobjective Optimization," *Evolutionary Computation* 3(1), Massachusetts, MA: MIT-Press, 1995
- [9] Aracena,J, Lamine.SB, Mernet,MA, Cohen,O and Demongeot,J "Mathematical modeling in genetic networks: relationships between the genetic expression and both chromosomal breakage and positive" *IEEE Transactions on Systems, Man and Cybernetics*, 2003
- [10]Hirasawa.K, Okubo.M, Katagiri.H, Hu.J, Murata.J, "Comparison between Genetic Network Programming (GNP) and Genetic Programming (GP)", Proceedings of the 2001 Congress on Evolutionary Computation, 2001.
- [11]Nuaymi,L., Godlewski,P., "Association of uplink power control and base station assignment in cellular CDMA systems" Proceedings ISCC 2000 Fifth IEEE Symposium on Computers and Communications, 2000.
- [12]Amaldi,E., Capone,A., Malucelli,F., "Optimizing UMTS radio coverage via base station configuration", The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2002.
- [13]Lee,C.Y., Kang,H.G., "Cell planning with capacity expansion in mobile communications: a tabu search approach", *IEEE Transactions on Vehicular Technology* 2000.
- [14]Santiago,R.C., Lyandres,V., "A sequential algorithm for optimal base stations location in a mobile radio network" 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2004.
- [15]Yufei Wu, Pierre,S., "Base station positioning in third generation mobile networks", *IEEE CCECE* 2003.
- [16]Hurley,S., "Automatic base station selection and configuration in mobile networks", *IEEE VTS-Fall VTC* 2000.