Metropolitan Area Network Design Using GA Based on Hierarchical Linkage Identification

Miwako Tsuji, Masaharu Munetomo, and Kiyoshi Akama

Hokkaido University, North 11, West 5, Sapporo, 060–0811, Japan
{m_tsuji,munetomo,akama}@cims.hokudai.ac.jp

Network design is a difficult combinatorial optimization problem which searches an optimal network configuration that satisfies geographical constraints, user traffic constraints etc. from a huge number of candidates. In solving the problem with GAs, building block (BB) destructions occur frequently which misleads the algorithms to unfavorable local optima, because appropriate encoding only with prior knowledge is not usually an easy task. To overcome this difficulty, linkage identification techniques that identify linkage groups – sets of loci tightly linked to form BBs – must be employed. In addition, some real-world problems, especially large and complex problems, seem to have interactions not only among loci but among linkage groups in a hierarchical manner. Therefore we propose a hierarchical version the LIEM^2 [1] (Linkage Identification with Epistasis Measure considering Monotonicity) and apply it to the network design problem.

Previous studies [3, 2] showed that to solve hierarchical problems efficiently it is essential to encode strings appropriately and to maintain diversity over schemata. For the first condition, we decompose a problem into sub-problems hierarchically as the following manner: in the initial step, our method identifies sets of loci tightly linked to make linkage groups with the LIEM^2. An *epistasis* measure – tightness of linkage between a pair of loci – is defined by the amount of violation of the monotonicity condition along perturbations of their variables ($0 \to 1$ or $1 \to 0$). In the following steps, it checks interdependence of the merged groups and this process continues recursively. The epistasis between linkage groups is calculated by BB-wise perturbations like $\mathrm{BB}_1 \to \mathrm{BB}_2$ or $\mathrm{BB}_2 \to \mathrm{BB}_1$, where BB_1 and BB_2 are schemata that with the largest and the second largest frequencies in the population. For instance, if linkage groups $\{0, 1\}$ and $\{3, 4\}$ have BB candidates `11***,00***` and `***01,***11` respectively and fitness changes along the perturbations `00*01` $\to$ `11*01`$\to$ `11*11` or `00*01` $\to$ `00*11`$\to$ `11*11` cause larger epistasis than other pair of linkage groups, then $\{0, 1\}$ and $\{3, 4\}$ should be merged. A specified number of strings are selected from the original population and the amounts of violation are calculated, and then epistasis is obtained as the maximum value of the amounts. To obtain optimal solutions in higher levels, multiple trials of the combination of the good schemata from lower levels are necessary. For this second condition, niching is introduced to maintain diversity over schemata.

We employ this hierarchical linkage identification technique to design Metropolitan Area Networks (MANs). The purpose of the network design problem is to construct a network with the lowest cost under some constraints. In our problem, geographical constraints that limits where we can lay down links (we

E. Cantú-Paz et al. (Eds.): GECCO 2003, LNCS 2724, pp. 1616–1617, 2003.

Table 1. cost (unit : JAY 10,000)

	average	standard deviation	minimum
uniform crossover	147753	911	146137
single layer LIEM^2	147290	1304	145292
hierarchical LIEM^2	145540	682	144517

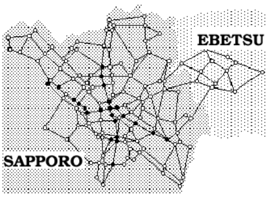

Fig. 1. test network

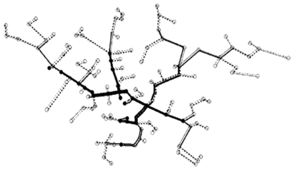

Fig. 2. best solution

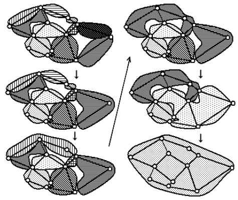

Fig. 3. linkage groups

assume that links can be laid down only along major roads in Fig.1) and user traffic constraints must be satisfied. The i-th allele s_i of string s define whether the link is constructed or not in the i-th candidate location. Table. 1 shows results of the experiment in 20 runs and the GA with hierarchical linkage identification performed the best. The solution obtained from the proposed method is shown in Fig.2. Fig.3 shows the process of merging linkage groups in a part of the test network. Variables (candidate locations for laying down a link) merged with neighborhoods at first levels gradually become large groups that cover large areas. Although GAs only with single layer linkage identification can improve some quality of solution, GAs with hierarchical linkage identification can solve MAN problems with large solution space and complex natures more effectively.

The employed test problem is simple and real problems may have many additional matters such as future expansion, server allocations, more detailed geographical constraints, routing policies, and so on. At this point if one or more of these conditions are changed, GAs tailored to a specific network problem need to change their encoding and genetic operators. On the other hand, the proposed GA should be applicable to various kinds of network design problems because it employs information only on strings and their fitness values – general information of the problems – and doesn't depend on problem domains.

References

1. Munetomo, M. : Linkage Identification with Epistasis Measure Considering Monotonicity Conditions. In Proc. of SEAL 02 (2002), 550–554
2. Pelikan, M., Goldberg, D.E. : Escaping Hierarchical Traps with Competent Genetic Algorithms. In Proc. of GECCO-2001 (2001), 511–518
3. Watson, R.A., Hornby, G.S., Pollack, J.B. : Modeling Building-Block Interdependency. In Proc. of PPSN V (1998), 97–106