

リンケージ同定を導入した遺伝的アルゴリズムによる 都市圏ネットワークの設計

辻 美和子* 棟朝 雅晴** 赤 間 清**

*北海道大学工学研究科システム情報工学専攻, **北海道大学情報メディア教育総合センター

遺伝的アルゴリズムにおいてはビルディングブロックとなる遺伝子をストリング上で密に符号化することが重要である。しかし、ネットワーク設計問題では地形、通信要求、経路などさまざまな要素が互いに複雑に影響するため、適切な符号化を行なうことは難しい。多くの既存研究はビルディングブロックの密な符号化について考慮しておらず、これを考慮していたとしても地理的な要素のみである。本論文では、遺伝子の値の振動による適応度の変化を用いて問題に関する前知識なしにビルディングブロックの位置であるリンケージを同定する手法である LIEM(Linkage Identification with Epistasis Measure) を導入し、ビルディングブロックを効率的に組み合わせ、遺伝的アルゴリズムによる効果的な解の探索を実行する。実験を行ない、本論文による手法で設計されたネットワークと幾つかの交叉手法、符号化手法による単純遺伝的アルゴリズムによって設計されたネットワークの敷設コストを比較し LIEM によるネットワーク設計の有用性を証明する。

A Genetic Algorithm using Linkage Identification for Metropolitan Area Network Design

MIWAKO TSUJI*, MASAHARU MUNETOMO** and KIYOSHI AKAMA**

*Division of System and Information Engineering, Graduate School of Engineering, Hokkaido University,

**Center for Information and Multimedia Studies, Hokkaido University

In genetic algorithms, it is important to encode strings ensuring tight linkage for their building blocks. In network design problems, however, it is difficult to encode strings appropriately because network design is dependent not only on geographical constraints but also on other complex factors such as bias on traffic demands, routing policy, and so on.

Although there's many applications of genetic algorithms to network topology design, most of them haven't paid attention to tight encoding of building blocks, or considered only geographical characteristics. In order to realize tight linkage among loci and realize effective genetic search, this paper introduces LIEM (Linkage Identification with Epistasis Measure) — a technique for identifying *linkage sets*, sets of loci tightly liked to form building blocks — to realize effective network design. Through empirical studies, we show the effectiveness of the network design with the LIEM compared to that with conventional genetic algorithms.

1. はじめに

情報通信ネットワークの爆発的な普及から、既存のサービスの多くが通信ネットワーク上でも実現されつつあり、広帯域なネットワークへの要求が高まっている。しかし、情報通信ネットワークの設計は、多くの組合せの中からコストやサービス品質に関する複数の矛盾した要求を同時に最適化する解を発見する必要があり、多項式時間で解くアルゴリズムは存在しない。

近年、良く用いられる発見的手法のひとつとして、遺伝的アルゴリズム [1] がある。遺伝的アルゴリズムでは、ビルディングブロックを構成する遺伝子がストリング上に密に配置されていない場合、交叉によってビルディングブロック破壊が起こる可能性がある。しかしながら、ネットワーク設計問題に遺伝的アルゴリズムを適用した研究例で、これらのうちでビルディングブロックの適切な符号化について考慮したものは少ない。加えて、ネットワーク設計問題の解は、地形、通信要求、それに通信経路などが複雑に関連するため、問題に関する前知識のみを用いて符号

化を行なうのでは不十分であると考えられる。よって、前知識に依存せず、適切なビルディングブロック交換を実行する手法を用いる必要がある。

本論文では、遺伝的アルゴリズムを用いたネットワーク設計に非線形尺度によって前知識によらずに同一のビルディングブロックを構成する遺伝子座の位置を同定する LIEM(Linkage Identification with Epistasis Measure) を導入することによって、適切な符号化が困難なネットワーク設計問題を遺伝的アルゴリズムで効率的に解くことを試みる。

2. ネットワーク設計問題に対する遺伝的アルゴリズムの適用

本論文で取扱うネットワーク設計問題とは、リンクのトポロジー、ノードの地理的な配置、リンクの容量といったネットワークの形状を決定する問題である。問題の目的は質の高い安価なネットワークを設計することである。しかし、コストとサービス品質の間にはトレードオフが存在する。ネットワーク設計は、規模によって指数関数的に増加

する組合せのなかから互いに矛盾する要求を同時に満足するような解を発見しなければならない。

このような問題を遺伝的アルゴリズムで解くにあたって、ビルディングブロック破壊を防ぎ、効率的に探索を進めるためには、関連する遺伝子同士を密に符号化する必要がある。しかし、ネットワーク設計問題において、複雑なネットワーク構造をストリングに符号化する際、密な符号化を行なうことは難しい。

それぞれの遺伝子座の対応する地点が地理的に近いもの同士が同一のビルディングブロックを構成するであろうと仮定して交叉を実行する遺伝的アルゴリズム [2,4] も存在するが、解は地理的な要素のみから決定されるとは限らず、適応度は通信量や経路など多くの要素から総合的に決定されるであろうと考えられる。

3. 都市圏ネットワーク設計問題

3.1 問題の定式化

本論文では、以下のような問題を扱う。

- 敷設のためのコストを最小化することを目的とする
- 制約条件として以下をもつ
 - すべてのノード間で通信可能
 - 地理的制約
 - すべての通信要求を満足する

また、以下を仮定する。

- ネットワークの通信要求は既知である
- ノード位置は固定であり、既知である

ネットワーク全体のコスト C は次式のような各リンクの距離と単位コストの積の総和である。

$$C = \sum_{vi \in L} c_i d_i x_i \quad (1)$$

ここで、 c_i はリンク i の単位あたりのコストであり容量に対して離散的に変化すると想定する。 d_i はリンク敷設候補地 i の距離、 $x_i \in \{0, 1\}$ はリンクを敷設候補地に敷設するか否かの決定変数であり、リンクを敷設するときは 1 を、そうでないときは 0 を、それぞれとる。

ネットワークのトポロジーは地理的に制約され、リンクが敷設しやすいのは、すでに電話線などの導管が存在する場所である。このような導管の多くは道路に沿って敷設されているため、本論文では道路沿い以外にはリンクは敷設できないという条件のもとで問題を解く。

符号化

ネットワーク中の候補地にリンクを敷設するか否かが 0, 1 のビット列により表現される。上述のように、本実験ではネットワークのリンクは幹線道路沿いのみを制約される。このリンクが敷設可能な幹線道路に識別番号のラベルを貼り、一列に符号化したものを個体として用いる。

適応度

適応度は、式 (1) により算出されるコスト C と制約条件を満足しない場合に課されるペナルティ P の和とする。

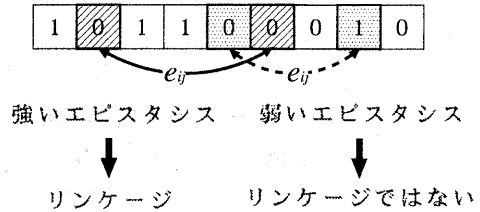


図 1 エピスタシス尺度によるリンケージ同定

P は次式により得られる。

$$P = m \times p \quad (2)$$

ここで、 m は伝達できない通信要求の総数、 p は伝達できない通信要求 1 つに対するペナルティである。

ストリングはリンクを敷設するか否かのみを決定し、適応度の算出に必要な容量に関する情報を持たない。リンク容量はこのストリングとあらかじめ与えられる通信要求から一意に決定される。適応度 f は次式のように、総敷設コストとペナルティの和である。問題の目的は f の最小化である。よって、この値がより低いものがより多くの子個体を残すことができる。

$$f = C + P \quad (3)$$

3.2 LIEM によるリンケージ同定

LIEM(Linkage Identification with Epistasis Measure) は、同一のビルディングブロックを構成し、強い相互依存関係をもつ遺伝子の組を同定するアルゴリズムである。本論文では、このような遺伝子の遺伝子座の位置をリンケージと呼び、同一のビルディングブロックを構成する遺伝子座の集合をリンケージ集合と呼ぶ。

同一のビルディングブロックを構成するような遺伝子座の間には非線形の相互依存関係があるであろう、と仮定する。この非線形性の大きさを遺伝子の値の摂動による適応度の変化量により定義する。LIEM ではこの値をエピスタシス値と呼び、遺伝子座の組 (i, j) のエピスタシス値 e_{ij} は次式で定義される。

$$e_{ij} = \max_{s \in \text{POP}} |\Delta f_{ij}(s) - (\Delta f_i(s) + \Delta f_j(s))| \quad (4)$$

ここで $\Delta f_{ij}(s), \Delta f_i(s), \Delta f_j(s)$ は個体 s の適応度と s の遺伝子の値を摂動させた個体 $\bar{s}_i = 1 - s_i$ 適応度の変化量である。また、POP は個体集団をあらわす。 e_{ij} は適切な大きさの初期個体集団中のすべての個体を調べた上で決定される。ただし、本論文では計算機負荷を軽減するために、初期個体群中の全個体ではなく、単純遺伝的アルゴリズムにより進化させた個体群から一部の適応度 f の低い個体のみを用いた。

式 (4) では、2 つの遺伝子座の値を同時に摂動させたとき適応度に与える影響が、これらの遺伝子座を別々に摂動させたとき適応度に与える影響の和とどれほど離れているかが計測される。 e_{ij} の値が大きいものは相互依存関係が強いとしてまとめて扱われ、そうでないものは別々に最適化が可能であるとされる (図 1)。

表 1 回線容量の階層

種類	容量 b Mb/s	コスト /km
T_1	~51.84	1.0×10^6
T_2	~155.52	2.5×10^6
T_3	~622.08	4.4×10^6
T_4	~2488.32	6.2×10^6
T_5	~9953.28	8.0×10^6



図 2 実験対象 1

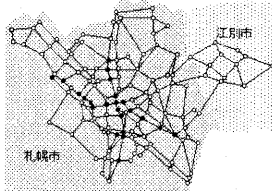


図 3 実験対象 2

表 2 通信要求量の特徴 (比率)

		終 点	
		法人	個人
始 点	法人	1	1
	個人	0.05	0.01

すべての遺伝子座の組 (i, j) に関してエピスタシス値 e_{ij} が計算され、各遺伝子座 i に関するリンケージ集合が形成された後、これらをまとめて幾つかの互いに素な集合をつくる [3]。探索はこれにより得られた相互依存関係の情報をもとに実行される。

4. 実 験

4.1 実験条件

本論文の目的は複数の制約条件を満足するコスト最小ネットワークの設計である。具体的な実験対象として札幌市を考える。

ネットワークのコストは式 (1) から算出される。この式で用いる容量 b のリンクの 1km あたりのコストは表 1 を仮定する。

現実のネットワーク敷設に対する地理的な制約を以下のように問題に組み込む。

実験は、規模の異なる 2 種類の地形により実行する。図 2 は札幌市の中心部の幹線道路を、図 3 は札幌市および隣接する江別市の幹線道路の配置を、それぞれモデル化したものである。それぞれの図における実線が幹線道路をあわらす。ネットワークのノード位置として実線の交点となる点を用いる。

4.2 通信要求

通信要求は問題にとって既知であるとする。ノード数が n のとき、ノード i からノード j への通信は $n \times n$ 行列で表現される。

現実の問題においては、個人ユーザとネットワークサービスを行なう法人ユーザとは通信量が大きくことなることが考えられるため、通信要求行列は表 2 のように偏りのあるものを仮定した。法人ノードは便宜的に地下鉄沿線

表 3 結果 (実験 1) 単位 10,000 円

実験 1	平均	標準偏差	最小
一点交叉 (ランダムコーディング)	98779.6	1151.2	97111
一点交叉	96246.6	1202.5	94650
二点交叉 (ランダムコーディング)	98268.1	1628.9	964420
二点交叉	95348.8	724.4	94426
一様交叉	95496.0	958.9	94047
リンケージ	94855.6	384.5	94045

表 4 結果 (実験 2) 単位 10,000 円

実験 2	平均	標準偏差	最小
一点交叉 (ランダムコーディング)	169647	5501	157170
一点交叉	163782	3573	158822
二点交叉 (ランダムコーディング)	162993	3681	157129
二点交叉	157689	2793	153838
一様交叉	148367	1012	146283
リンケージ	147487	1242	144644

沿いに存在するものとし、図 23 では●で表現される。それ以外の個人ノードは○で表現される。

4.3 実験結果

交叉が解に与える影響を調査するため、地理的な要素を反映した符号化を行なった場合の一点交叉、二点交叉、一様交叉の三種類の交叉と、符号化をランダムなものに変えた場合の一点交叉、二点交叉について、単純遺伝的アルゴリズムによる実験を行なった。続いて、同じ問題に対してリンケージ同定遺伝的アルゴリズムを適用し、その性能を比較した。

個体数 400 とした。単純遺伝的アルゴリズムは、交叉割合 $P_c = 0.9$ 、突然変異確率 $P_m = 0.01$ で実行された。選択には、親個体と子個体とを混合してソートし、 f の値が小さい個体から順に次世代に残す $\mu + \lambda$ 戦略を用いた。

LIEM で用いる問題の次数は $k = 6$ を想定した。リンケージ集合が過指定になるのを防ぐために、これにより得られた各集合 k での同時発生割合を計算し、値が 0.5 以下のものは集合から取り除いた。リンケージ集合ごとの交叉を行なう前に各リンケージ集合内部で GA-easy な部分解を発見するために集合の要素のみの一様交叉を行なう 2 世代の小さい単純遺伝的アルゴリズムを実行した。この後、単純遺伝的アルゴリズムと同様の交叉割合、突然変異確率、選択手法でリンケージ同定遺伝的アルゴリズムを実行した。

ただし、実験 2 は実験 1 と比較して変数が多く、個体長が長い場合、突然変異確率を $p_m = 0.007$ とした。

すべての実験は乱数の種を変えて、20 回繰り返し実行した。それぞれの手法における最終的なコストの平均、標準偏差、および最小値を表 3, 4 に示す。

ネットワークの規模が大きくなる場合、一様交叉が単純遺伝的アルゴリズムのうちでは優れた性能を示している。これは、より符号化の難しくなる規模の大きいネットワークで

は、スキーマの組合せ数が制限される多点交叉では探索が難しく、多くの組合せを試みることができる一様交叉に利点があるからであると考えられる。一様交叉はビルディングブロック破壊が起こりやすい手法であるが、本実験で用いた $\mu + \lambda$ 戦略は交叉によりビルディングブロックが破壊されても親個体は生き残ることができるため、この影響をうけにくく、比較的優れた解を得たと考えられる。

各実験でリンケージ同定遺伝的アルゴリズムにより得られたトポロジーを図6および図7に示す。

リンケージ同定を導入した遺伝的アルゴリズムは、両方の実験ですべての単純遺伝的アルゴリズムと比較して優れた性能を示した。

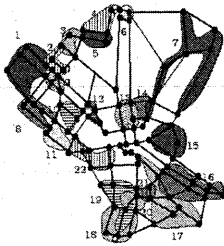


図4 リンケージ集合(実験1)

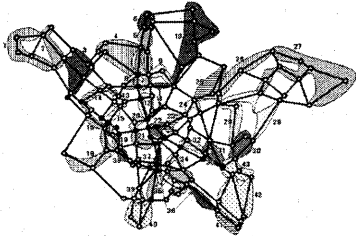


図5 リンケージ集合(実験2)

得られたリンケージ集合を図4,5に示す。図から、実験1では地理的な要素をある程度反映した集合を形成していることがわかる。より規模の大きい実験2では、地理的な要素を反映しながらも、集合35,36のように離れたリンク数候補補地同士が強い相互作用す例も見られ、規模の大きいネットワークのコストはより複雑な変数同士の相互作用により決定されていると考えられる。

リンケージ同定遺伝的アルゴリズムでは、交叉や突然変移による探索に加えて、前処理としてストリング長 l に対して $O(l^2)$ の計算機負荷がかかるエピスタシス値の計算を実行する。このため計算負荷は単純遺伝的アルゴリズムに比べて大きい。よって、特に対象とするネットワークの規模が小さい場合は、単純遺伝的アルゴリズムを複数解実行し同一の計算機コストで比較すると、リンケージ同定遺伝的アルゴリズムのほうが優れた性能を示している、とは言い切れない。しかし、規模の異なる2種類の実験の結果から、ネットワークの規模が大きくなるにつれて、適切に交叉を実行することが困難になると考えられる。よって、

より大きなネットワークに関してリンケージ同定遺伝的アルゴリズムは、一様交叉や多点交叉の単純遺伝的アルゴリズムの複数回の実行と比較した場合も、より優れたネットワーク設計のための手法であると言える。

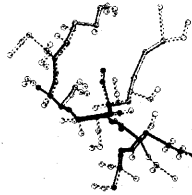


図6 結果: 実験1

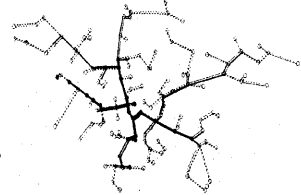


図7 結果: 実験2

5. おわりに

本論文では、交叉の際のビルディングブロック破壊によって個体群が望ましくない局所解に収束するのを防ぐために、あらかじめビルディングブロックの遺伝子座位置を同定する手法である LIEM を導入し、都市圏ネットワークの設計を行なった。

実験により、リンケージ同定を導入した遺伝的アルゴリズムが単純遺伝的アルゴリズムと比較して低コストなネットワークを発見したことを確認した。LIEM は、問題に関する前知識によらずリンケージを同定する手法であるため、事前にビルディングブロックを密に符号化することが難しいネットワーク問題においても効率的な探索を実現した。

今後の課題としては、リンケージ同定のための計算機負荷の軽減、解空間の分割可能性があいまいな現実の問題に対応可能なより柔軟なリンケージグループ交叉の設計などが挙げられる。

参考文献

- 1) D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Publishing Company, Inc., Reading, MA, 1989.
- 2) A. Kumar, R. M. Pathak, Y. P. Gupta, and H. R. Parsaei. A genetic algorithm for distributed system topology design. *Computers and Industrial Engineering*, 28(3):659-670, 1995.
- 3) M. Munetomo and D. E. Goldberg. Linkage identification by non-monotonicity detection for overlapping functions. *Evolutionary Computation*, 7(4):377-398, 1999.
- 4) M. C. Sinclair. NOMaD: Initial architecture of an optical network optimisation, modelling and design tool. In *Proceedings of the 12th UK Performance Engineering Workshop*, pp. 157-167, 9 1996.