

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^{††}

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、遺伝子の値の摂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。実験を行い、本論文による手法で設計されたネットワークといくつかの交叉手法、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,[†] MASA HARU MUNETOMO^{††} and KIYOSHI AKAMA^{††}

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、映像や音楽配信など広帯域を必要とするサービスの多くが通信ネットワーク上で実現されつつある。このような用途に応えるためには、広帯域で安定したネットワークを都市圏に敷設することが必要である。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、よく用いられる発見的手法の1つとして、遺伝的アルゴリズム²⁾がある。遺伝的アルゴリズムの探索は、交叉を用いてビルディングブロックと呼ばれる優れた部分解を組み合わせることで実行される。しかし、符号化が適切でなく、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。このような場合、単純遺伝的アルゴリズムによる探索は必ずしも効率的に実行されない。

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は存在するが、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、

[†] 北海道大学工学研究科システム情報工学専攻

Division of Systems and Information Engineering,
Graduate School of Engineering, Hokkaido University

^{††} 北海道大学情報メディア教育総合センター

Center for Information and Multimedia Studies,
Hokkaido University

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うのでは不十分であると考えられる. よって, 前知識のみに依存せず, 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.

ここで, 同一のビルディングブロックを構成し相互依存関係のある遺伝子座の集合をリンケージと呼ぶ. 適切なビルディングブロック交換のために, リンケージをあらかじめ同定する試みがなされている. このようなリンケージ同定手法の 1 つとして, 非線形尺度に基づいてリンケージ同定を行う LIEM (Linkage Identification with Epistasis Measure)⁷⁾ が存在する. この手法では, リンケージは遺伝子の値の摂動による変化量から同定される. このため問題に関する特別な知識を用いることなく, リンケージを同定することが可能である.

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 LIEM を導入し, 前知識のみでは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

2.1 ネットワーク設計問題

本論文で取り扱うネットワーク設計問題とは, リンクのトポロジ, 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.

設計されたネットワークは

- 通信サービスの品質
 - 伝送遅延 (信号が伝搬して相手に届くまでに生じる時間の遅延)
 - スループット (単位時間あたりの処理能力)
 - 信頼性 (代替経路数などの耐故障性能)
 - パケット損失率
- 敷設コスト
- 拡張性
- 管理の容易さ

などによって評価される⁸⁾.

ネットワーク設計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. たとえば, 高速で遅延の少ないネットワークを提供するためには容量の大きな回線を使用すればよいが, その分コストは増大する. 逆に, サービスの品質をおさえることでコストをおさえることができる. また, ネットワークの敷設は地形によっても制約される.

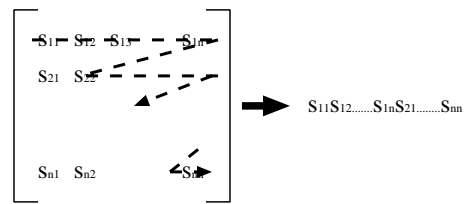


図 1 符号化例

Fig. 1 Example of encoding.

ネットワーク設計は, 規模により指数関数的に増加する組合せの中から上述のような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.

2.2 既存手法の問題点

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, 関連する遺伝子どうし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, 密な符号化を行うことは難しい.

また, 対象となるネットワークは通常 2 次元であり, これを 1 次元のストリングに割り当てるとき, その地形をつねに適切に反映することは困難である. たとえば, リンクトポロジを遺伝的アルゴリズムを用いて決定するような場合でよく使用される符号化手法として以下のようなものがある. まずノードに 1 から n (n はノード数) までの識別番号をつける. 符号化はノード (i, j) 間のリンクの敷設候補地を遺伝子座 k に割り当てることにより実行される. $n \times n$ の行列 (無向グラフなら上三角部) を右から左, 上から下といったルールに従って走査したものがストリングとされる (図 1). この手法の場合, ノード (i, j) 間の敷設候補地とノード (i', j) 間の敷設候補地は, 同じノード j に連結されているにもかかわらず, i, i' のノード識別番号に大きく差がある場合, ストリング上では離れた位置で表現されることになる. 加えてノード識別番号に関して, すべての地理的に近隣なノードどうしに近隣な遺伝子座を割り当てることは難しい.

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どうし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^{3), 6)} 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される必要があると考えられる.

本論文では, LIEM を用いて遺伝子座間の相互依存関係であるリンケージを同定し, 交叉の際には相互依存関係の強い遺伝子座をまとめて扱うことで効率的な解の探索を試みる. この手法は遺伝子の値の摂動によ

る適応度の変化量のみを情報として必要とする。従来の手法では、遺伝的アルゴリズムのユーザは個体がどのように適応度を得るかについて考慮して、符号化やオペレータの設計を行わなければならなかったが、本論文の手法では適応度関数というブラックボックスの中身について考慮する必要はない。

3. 都市圏ネットワーク設計問題

本論文で取り扱う問題のモデルについて述べる。本論文ではいくつかの現実的な条件を含めた単純なモデルを用いる。

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持つ。
 - 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。

目的は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。ノード位置はあらかじめ与えられ、これらを連結するリンクのトポロジと容量を遺伝的アルゴリズムにより決定する。

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、 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。

単位コスト c_i がリンク容量に比例して連続的に増加するような関数である場合、問題は比較的簡単である。しかし、実際のリンク容量は離散的に変化し、そのコストも容量に比例しない。本論文においても離散的に変化するリンク容量および単位コストを想定する。

ネットワークのトポロジは地理的に制約される。現実のネットワークにおいては、すべてのノード間で自由にリンクが敷設できることは少ない。現実の地形には、人工の、自然の、障害物が存在する。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。

また、ネットワーク中のすべてのノードは通信可能でなければならない。どのような孤立ノードも存在してはならない。

さらに、ネットワークは通信要求を満足するものでなければならない。ネットワークはすべての通信要求が存在するノードの組を連結するものでなければならない。加えて、リンクの容量は通信要求をすべて不足なく伝達できるものでなければならない。

ネットワークの通信要求はあらかじめ与えられる。ネットワーク中の任意のノードは別のノードに対して通信要求を持つ。ネットワークのノードが n であるとき、これは $n \times n$ 行列の形で表現される。

符号化

リンクが敷設可能な候補地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。リンク敷設候補地が l カ所であるとき、ストリング長は l となる。 i 番目の遺伝子座のとり値 s_i は、番号 i が割り当てられたリンク敷設候補地にリンクが敷設されるとき 1、敷設されないときに 0 である。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に制約されるため、リンク敷設候補地はこの幹線道路のみである。

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。 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 の最小化であるので、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。

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

リンク容量の決定と適応度の算出は以下のように実行される。

- (1) ネットワーク上に対応する遺伝子 i の値 s_i が $s_i = 1$ であるリンクを敷設する。
- (2) 通信が存在するすべての 2 ノード間の経路を最短経路法で決定する。経路を持たない通信要求にはペナルティを課す。

- (3) (2)でもとめた経路上の各リンクを流れるすべての通信量を合計し, 可能なリンク容量からこの合計容量を下回らない最も容量の小さなリンクを選択する.
- (4) すべての $s_i = 1$ であるリンク候補地に関して容量が決定されれば, この容量とリンクの距離から式 (1) を用いてコストが決定できる. このコストにペナルティを加え, 最小化問題の適応度 f とする.

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持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り, 次式から得られる.

$$\Delta f_i(s) = f(\dots \bar{s}_i \dots) - f(\dots s_i \dots) \quad (5)$$

$$\Delta f_j(s) = f(\dots \bar{s}_j \dots) - f(\dots s_j \dots) \quad (6)$$

$$\Delta f_{ij}(s) = f(\dots \bar{s}_i \dots \bar{s}_j \dots) - f(\dots s_i \dots s_j \dots) \quad (7)$$

また, POP は個体集団を表す. ビルディングブロック内部で線形な振舞いを見せる騙し関数のような問題も存在するため, 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うえで決定される.

ここで, e_{ij} の値がゼロのとき, 摂動による適応度の変化量は

$$\Delta f_{ij}(s) = \Delta f_i(s) + \Delta f_j(s) \quad (8)$$

という関係にあり, 遺伝子座の組 (i, j) は別個に最適化して後に組み合わせることが可能であると考えられる.

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

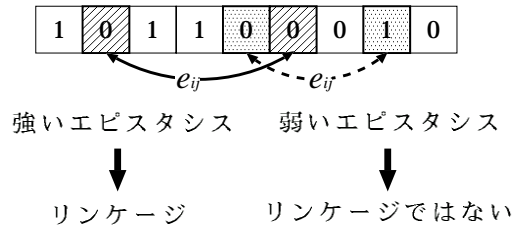


図2 エピスタシス尺度によるリンケージ同定
Fig. 2 Linkage identification with epistasis measure.

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2).

遺伝子座 i に関するリンケージ集合の要素は, e_{ij} の値が大きい順に上位 k 個の j とする. k は問題の次数であり, 遺伝的アルゴリズムのユーザにより定義される.

このような各遺伝子座 i に関して e_{ij} の値が大きい遺伝子座 j を k 個含むような l 個 (l は遺伝子長) の集合をもとに, 以下の手法で互いに素な集合を構成する.

得られた l 個の集合を $l \times l$ 行列 $A = (a_{ij})$ で示す. 行列の成分 a_{ij} は遺伝子座 i に関する集合に遺伝子座 j が属しているか否かを表す. この値が1のとき, 要素 j は集合に属し, 0のときは属さない. ただし, 対角成分はすべて1であるとする.

ここで行列 A は必ずしも対称行列ではない. 準備として次の計算を行い, A を対称行列にする. ただし, 演算はすべてブール代数 B_2 の演算子で実行される.

$$a_{ij} = a_{ij} + a_{ji} \quad \forall i, j \quad (9)$$

また, 遺伝子座 i に関する集合に遺伝子座 j が含まれ, 遺伝子座 j に関する集合には遺伝子座 k が含まれるが, 遺伝子座 i に関する集合には遺伝子座 k が含まれない, というような場合を考える. このとき, $i \rightarrow j \rightarrow k$ という間接的関係が存在する. つまり, 遺伝子座 i に関する集合に含まれない要素 k でも他の集合を介して $i \rightarrow \dots \rightarrow k$ という間接的な関わりを持つことが考えられる.

$$A = A \times A \quad (10)$$

という更新を最大 $l - 1$ 回繰り返すことにより間接的な関係をすべて求めることができる.

式 (10) の計算の後は, 遺伝子座 i に関する集合に遺伝子座 j が含まれていた場合, 遺伝子座 i に関する集合と遺伝子座 j に関する集合は等しい. よって, 遺伝子座 j に関する集合は, すでに i に関する集合で表されているので, 空にすることができる.

以上の操作で, エピスタシス値から互いに素ないくつかのリンケージ集合を構成できた. これらの集合を

S_1, S_2, \dots, S_m (ただし, $m \leq n$), すべての遺伝子座の集合を S とすれば, S_1, S_2, \dots, S_m は S を直和分割している.

しかし, このようにして構成した集合は, $i \rightarrow \dots \rightarrow k$ のような間接的な関わりが多い場合, 要素数の多い大きな集合になってしまう. これを防ぐため, 式 (10) を計算する前に, tightness という尺度を導入し相互依存関係が弱い遺伝子座は集合から取り除く⁵⁾.

遺伝子座の組 (i, j) 間の tightness の値 $\text{tightness}(i, j)$ は各遺伝子座 k に関するリンケージ集合 k での i, j の同時発生割合として次式で定義される.

$$\text{tightness}(i, j) = \frac{n1(i, j)}{n1(i, j) + n2(i, j)} \quad (11)$$

ここで, $n1(i, j)$ はすべてのリンケージ集合のうちで i, j の両方が存在する集合の総数, $n2(i, j)$ はすべてのリンケージ集合のうちで i, j の一方のみが存在する集合の総数である.

LIEM により形成されたリンケージ集合に対して, $\text{tightness}(i, j)$ の値がすべての i, j ペアに関して計算され, この値が閾値 δ ($0 \leq \delta \leq 1$) より小さいものはリンケージ集合から取り除かれる.

3.3 LIEM の導入

本論文ではネットワーク設計問題を LIEM を導入した遺伝的アルゴリズムを用いて解くことを考える. 3.2 節で述べたように, エピスタシス値は個体群中のすべての個体を調べたうえで決定されるべきである. しかし, エピスタシスの計算はストリング長 l に対して $O(l^2)$ の負荷がかかることから, 計算にかかる時間が大きくなりすぎる. また, ランダムに生成したネットワークは, 孤立ノードが多く存在することから, 適応度に占めるペナルティの割合が大きく, 正確なリンケージ同定が難しい.

そこで, 本論文ではリンケージ同定にかかる負荷を軽減するために, エピスタシスの計算には初期個体群をそのまま用いずに, 単純な遺伝的アルゴリズムを実行しある程度収束した個体群から上位個体のみを用いた. 上述のようにリンケージ同定においてはノイズの影響を考慮し問題の次数から適切に見積もった個体数の個体群を用いることが重要であるが⁵⁾, 現実の問題ではすべての遺伝子座が弱い相互依存関係を持つために, 個体数は現実的でない値となる可能性がある. 逆に, 本手法のように単純遺伝的アルゴリズムにより進化させた個体群を用いると, トラップ関数のような線形性と非線形性が混在する関数の場合に正しくリンケージを同定できない恐れがある. しかし, 現実の問題においては, 単純遺伝的アルゴリズムにより収束し

た局所点は大域的な最適解でなくても, 摂動による非線形性の振舞いを見せることが多いだろう, と考えられる.

単純遺伝的アルゴリズムを実行するのはエピスタシス値を計算し, リンケージ集合を得るためであり, 得られたリンケージ集合を用いて初期個体群に対して遺伝的オペレータを適用する.

リンケージに関する情報の得られた個体群に対する遺伝的アルゴリズムとして, intraGA と interGA という 2 種類の実行が提案されている⁴⁾.

intraGA は, リンケージ集合の内部でさまざまな組合せを探索し, それによって GA-easy な部分問題を解き, より適応度の高いスキーマを得るための遺伝的アルゴリズムである. 本論文でスキーマとは, ある 1 つのリンケージ集合に属する遺伝子をいうものとする. たとえば, あるリンケージ集合に属する遺伝子座を f , そうでない遺伝子座を $\#$ とした場合, リンケージ集合 $ff\#\#\#$ に属する可能性のあるスキーマは $\{11****, 10****, 01****, 00****\}$ である (* は don't care). intraGA はリンケージ集合の内部で, 単純な一様交叉と突然変異を実行する. 各リンケージ集合に対するスキーマはリンケージ集合の要素を初期個体群中の個体から切りぬくことで生成される. これらのスキーマの適応度は, 対象となるリンケージ集合に属さない遺伝子座 (そのスキーマの適応度に関して * である遺伝子座) の値にテンプレートを用いることで計算される. テンプレートとしては初期個体群中の最良個体を用いた. たとえば, はじめの 2 つの遺伝子座を要素として含む集合 $ff\#\#\#$ に関しては, 各スキーマの適応度は, ff の部分は初期個体群からそれぞれの個体の該当する遺伝子の値を切り取ってきて用い, それ以外の $\#$ の部分の遺伝子はすべて同一の値 (最良個体の遺伝子がとっている値) で埋める.

intraGA で得られたスキーマから, 後に組み合わせるビルディングブロックの候補として, 適応度の高いスキーマを選択し個体群に戻す. この際, 最良スキーマとの適応度の差があらかじめ定めた値より大きいものは, 取り除かれる.

以上の操作で得られたビルディングブロック候補を組み合わせ, 最適解を探索する遺伝的アルゴリズムが interGA である.

interGA ではリンケージ集合に基づく交叉と通常のビット反転の突然変異が実行される. 親個体の組の各リンケージ集合に関して, 交叉を行うか否かが検討される. 交叉が行われるとき, リンケージ集合に属する遺伝子座の値はまとめて交換される. interGA での交叉

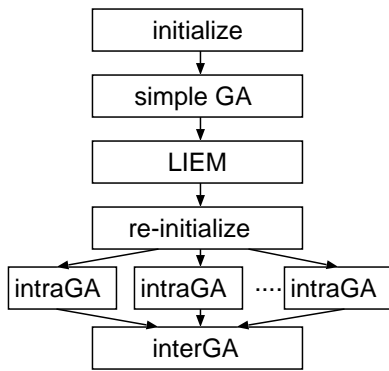


図 3 リンケージ同定遺伝的アルゴリズムによる処理

Fig. 3 A genetic algorithm using linkage identification.

オペレータによって親個体 $(p1, p2)$ から子個体 $(c1, c2)$ をつくる操作は、以下ようになる。 $p1[j]$ は個体 $p1$ の j 番目の遺伝子の値を表す。また、 $\text{random}[0, 1)$ は $0 \leq \text{random}[0, 1) < 1$ の一様乱数である。

```

for(∀ 親個体のペア (p1, p2)){
  for(∀ リンケージ集合 i){
    if(random[0,1) < 0.5){
      for(∀ j ∈ リンケージ集合 i){
        c1[j] = p1[j];
        c2[j] = p2[j];
      }
    }else{
      for(∀ j ∈ リンケージ集合 i){
        c1[j] = p2[j];
        c2[j] = p1[j];
      }
    }
  }
}
  
```

以上の処理をまとめると図 3 のようになる。

4. 実験

4.1 実験条件

4.1.1 コスト

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 1 km あたりのコストは表 1 を仮定する。表の階層は SONET (Synchronous Optical Network) で用いられたものを参考に決定した。実際のコストは不明だったため、コストは容量にあわせて $\log(T_i)$ に比例するものとした。このモデルは、現在 MAN に多く用いられるギガビットイーサなどは異なるが、離散的に変化する複数のリンク容量の選択肢を問題に与えるための仮想のモデルを与える。

4.1.2 地理的な制約

現実のネットワーク敷設に対する地理的な制約を以

表 1 回線容量の階層

Table 1 Class of link capacity.

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

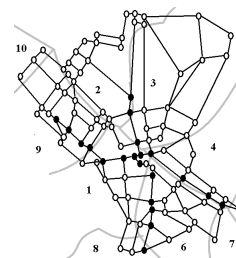


図 4 実験対象 1

Fig. 4 Test network 1.

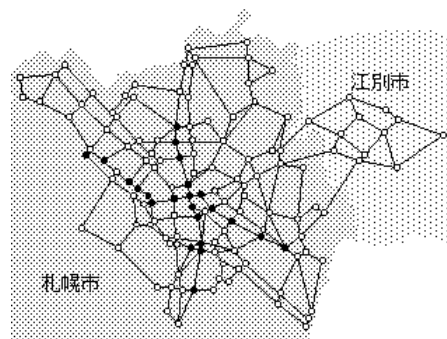


図 5 実験対象 2

Fig. 5 Test network 2.

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 4 は札幌市の中心部の幹線道路を、図 5 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は幹線道路を表す。本論文では、リンクはこれらの実線上にのみ敷設可能であるものとする。

また、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。

これらのノード位置、およびリンクの敷設可能位置は既知であるとする。

4.2 サービス品質に関する制約

ネットワーク中のノードはすべて他のノードと通信可能でなければならない、孤立ノードなどは存在してはならない。

また、ネットワーク中の任意の異なる 2 ノード間に

は通信が存在する．ネットワークはこれらすべての通信を伝達しなければならない．

4.3 通信要求の性質

通信要求は以下で述べるような仮想的なデータを用いた．

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．上述のように通信要求はあらかじめ問題に関して与えられる．これは、ネットワーク中の通信を代表するものとする．

ただし、通信がすべてのノード間で均一に行われることは考えにくい．現実のネットワークでは、個人ユーザが多く通信を他に配信することは少ない．逆に、コンテンツ配信を行う企業などがユーザに向ける通信の量は大きい．この傾向はブロードバンド化が進む将来、より顕著なものになると考えられる．これをモデル化するために以下を仮定した．

まず、ノードには個人ユーザが多いノードと法人ユーザが多いノードが存在するものとする．この場合の法人ユーザとは特にインターネットサービスプロバイタやコンテンツ配信などのネットワークサービスを行うユーザとする．法人ユーザから個人ユーザへは多くのデータが送信されるが逆は少なく、また、法人ユーザ同士は比較的多くのデータをやりとりしていると考えられる．図 4 では、個人ノード () および、法人ノード () で表現されている．法人ノードの所在に関しては便宜的に地下鉄沿線に近いノードを仮定した．

各ノードの通信量はそのノードの属する区の人口 (表 2) をもとに決定する．ある区に存在するノードは、その区の人口に比例する通信量を平均とした正規分布をとると仮定して、通信要求を生成した．ただし、個人ノードから法人ノードへの通信、および、個人ノードどうしの通信は、法人ノードどうしの通信や法人ノードから個人ノードへの通信の 1~5% 程度とした．同じ区に属するノードの通信量の比率は表 3 である．始点ノードと終点ノードが異なる区に属する場合は、2 つの区の人口の平均を用いて通信量を決定した．通信量は 50k から 2Mbps 程度の値をとる．

4.4 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った場合の一点交叉、二点交叉、一様交叉の 3 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、実験を行った．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．

ここで地形的な要素を反映した符号化とは以下のよ

表 2 各地域の人口

Table 2 Area populations.

No.	区	人口	No.	区	人口
1	中央区	183,763	6	豊平区	204,727
2	北区	261,328	7	清田区	110,736
3	東区	249,308	8	南区	156,707
4	白石区	198,517	9	西区	200,947
5	厚別区	128,044	10	手稲区	136,429
-	江別市	120,455			

表 3 通信要求量の比率

Table 3 Ratio of traffic demands.

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

うなものである．まず、ノードに $1 \sim n$ のノード識別番号をつける．このノード識別番号は、北から南、東から西、といったルールに従ってつけられる．次に、リンクを敷設できる候補地の両端のノード番号 (i, j) (ただし $i < j$ とする) に関して、 i の数値が小さい順にリンク識別番号 k をつける．ただし、 i が同じ値をとる複数のリンク敷設候補地に関しては、 j の値が小さい敷設候補地により小さいリンク識別番号を割り当てる．このリンク識別番号を用いて 3.1 節で述べた遺伝子座とリンク敷設候補地との対応づけを行う．また、ランダムな符号化はこのリンク識別番号と遺伝子座の対応づけをランダムなものに変えた符号化である．

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、優れた個体から順に次世代に残す $\mu + \lambda$ 戦略を用いた．

LIEM の問題の次数は $k = 6$ を想定した．要素間の間接的な関係からリンケージ集合が大きくなりすぎること防ぐため、閾値 0.5 で tightness detection を行った．interGA での交叉割合は $p_c = 0.9$ 、ビットごとの突然変異確率は $p_m = 0.01$ とした．interGA の前に遺伝的アルゴリズムにとって簡単な部分解を探索するために intraGA を実行した．intraGA の世代数は 2、交叉割合は $p_c = 0.9$ 、突然変異確率は $p_m = 0.05 / (\text{個体長})$ とした．また intraGA の実行後、テンプレートをを用いた適応度が各リンケージ集合の最も優れたスキーマと比べて 0.5×10^{10} 以内であるもののみをビルディングブロック候補として次の interGA へ戻した．各遺伝的アルゴリズムでの選択は他の実験と同様に $\mu + \lambda$ 戦略¹⁾ で行った．

ただし、実験 2 は実験 1 と比較して変数が多く、個

表 4 結果(実験 1)単位 10,000 円

Table 4 Result (Test network 1) unit 10,000 yen.

	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	96442
二点交叉	95348.8	724.4	94426
一樣交叉	95496.0	958.9	94047
リンケージ	94855.6	384.5	94045

表 5 結果(実験 2)単位 10,000 円

Table 5 Result (Test network 2) unit 10,000 yen.

	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
一樣交叉	148367	1012	146283
リンケージ	147487	1242	144644

体長が長い場合、単純遺伝的アルゴリズムの突然変異割合、および interGA のビットごとの突然変異確率を $p_m = 0.007$ とした。ネットワークの規模が大きくなっても、問題の本質的な難しさには変化がないと考えられるので、それ以外のパラメータは実験 1 と同様のものを用いた。

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 4、表 5 に示す。

実験 1 における各手法のうち最も優れた適応度を示した個体のトポロジを図 8 ~ 13 に示す。実験 2 で得られた最良のトポロジを図 15 ~ 20 に示す。

ランダムな符号化ではビルディングブロックを有効に組み合わせることができず、望ましくない局所解に収束している。このことから、符号化が単純遺伝的アルゴリズムの探索性能に大きく影響することが分かる。

ネットワークの規模が大きい場合、一樣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樣交叉に利点があるからであると考えられる。一樣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、この影響を受けにく

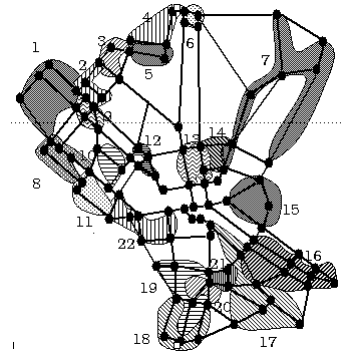


図 6 リンケージ集合(実験 1)

Fig. 6 Linkage groups (Test network 1).

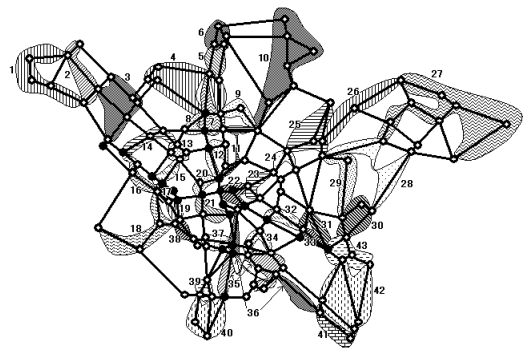


図 7 リンケージ集合(実験 2)

Fig. 7 Linkage groups (Test network 2).

く、比較的優れた解を得たと考えられる。

両方の実験でリンケージ同定を導入した遺伝的アルゴリズムは、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。本実験で導入した手法は符号化に依存しない手法であり、どのような符号化が行われた場合でもほぼ同等の性能を示すであろうことは明らかである。

得られたリンケージ集合を図 6、図 7 に示す。図から、比較的規模の小さい実験 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分かる。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敷設候補地どうしが強い相互作用を示す例も見られ、規模の大きいネットワークのコストはより複雑な変数どうし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、交叉や突然変異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

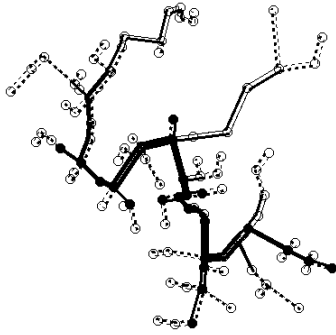


図 8 ランダムコーディング 一点交叉
Fig. 8 One-point crossover with random encoding.

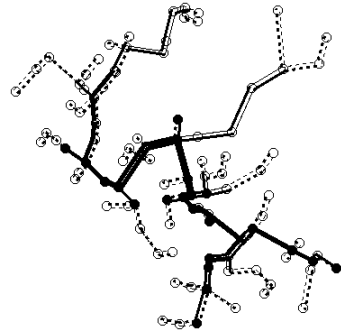


図 9 ランダムコーディング 二点交叉
Fig. 9 Two-point crossover with random encoding.

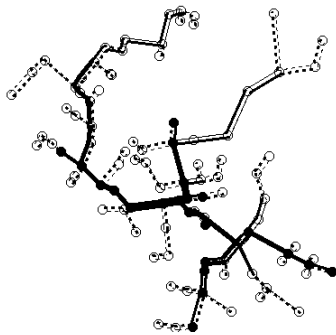


図 10 一点交叉
Fig. 10 One-point crossover.

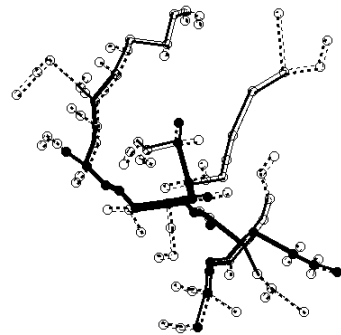


図 11 二点交叉
Fig. 11 Two-point crossover.

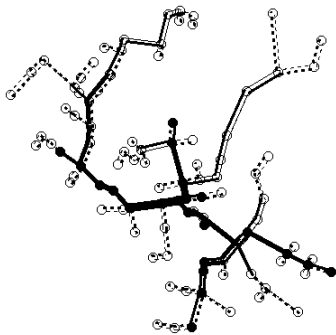


図 12 一様交叉
Fig. 12 Uniform crossover.

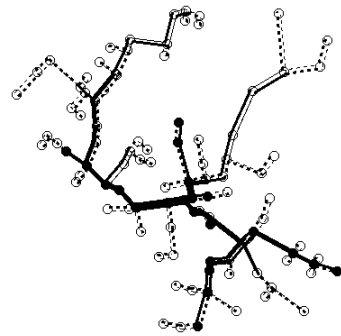


図 13 リンケージ交叉 (LIEM)
Fig. 13 Linkage crossover.

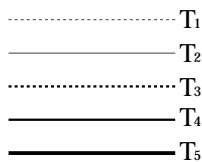


図 14 凡例
Fig. 14 Link class.

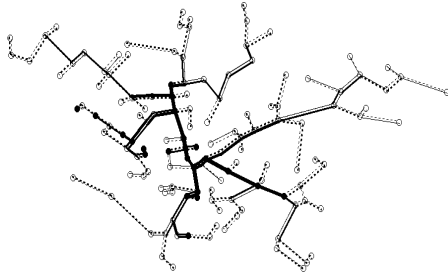


図 15 ランダムコーディング 一点交叉

Fig. 15 One-point crossover with random encoding.

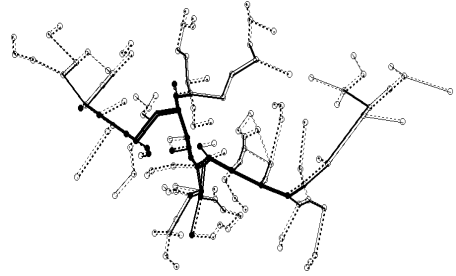


図 16 ランダムコーディング 二点交叉

Fig. 16 Two-point crossover with random encoding.

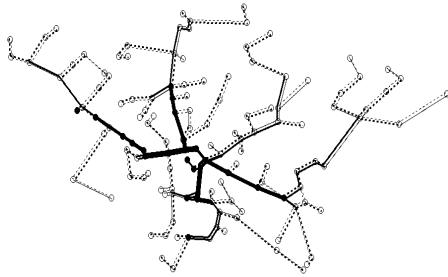


図 17 一点交叉

Fig. 17 One-point crossover.

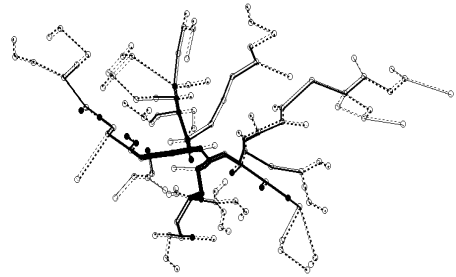


図 18 二点交叉

Fig. 18 Two-point crossover.

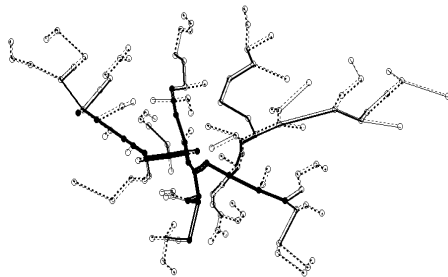


図 19 一様交叉

Fig. 19 Uniform crossover.

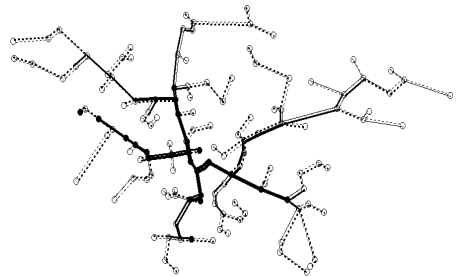


図 20 リンケージ交叉 (LIEM)

Fig. 20 Linkage crossover.

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、とはいきれない。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て、より優れたネットワーク設計のための手法であるといえる。

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。

本論文で用いた問題のモデルは、実際のネットワーク設計を簡略化したものである。現実の問題では、長期にわたるネットワーク利用、拡張性などより多くの要件が考えられる。また、経路に関しても本論文では最短経路法を用いたが、実際には別の手法が用いられる可能性もあり、また、代替経路性を要求されることもありうる。条件が1つ変わると符号化を変える必要性が生じるため、従来問題依存の遺伝的オペレータを用いた手法では、そのオペレータが使用不能になってしまう可能性がある。本手法は適応度という一般的な情報のみを用いて交叉を設計をしているため、適応度関数というブラックボックスの中身について考慮する必要はほとんどなく、さまざまなネットワーク問題に対応できると考えられる。

今後の課題としては、リンク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クージグループ交叉の設計、拡張性や通信量の動的な変化への対応など要件の異なるネットワーク問題に対する手法の柔軟性の調査などがあげられる。また応用として、一度設計されたネットワークを利用状況の変化に合わせて改善する場合、この変化量が小さい場合には、変化に関連するであろうと考えられるリンクージ集合をピックアップし、この部分だけを再最適化することが可能であると考えられる。変化がごく小さい場合は、この変化の影響を受けない部分は固定してテンプレートとして用いる intraGA を用いるか、それが現実的な場合には数えあげなどを用いて、再最適化することが可能であると考えられる。変化が比較的大きい場合は、関連するリンクージ集合をすべてピックアップし、マージして再び変化後の利用状況に合わせた適応度関数からリンクージ同定を行って再最適化する必要がある。この際、状況の変化の影響がどの範囲まで及ぶかを完全に予測することはおそらく不可能であろうと考えられるが、前の利用状況における e_{ij} や $tightness(i, j)$ の値からある程度の推測が可能であると考えられる。

参 考 文 献

- 1) Bäck, T., Fogel, D.B. and Michalewicz, Z.: *Handbook of Evolutionary Computation* (1997).
- 2) Goldberg, D.E.: *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing Company, Inc., Reading, MA (1989).
- 3) Kumar, A., Pathak, R.M., Gupta, Y.P. and

Parsaei, H.R.: A Genetic Algorithm for Distributed System Topology Design, *Computers and Industrial Engineering*, Vol.28, No.3, pp.659-670 (1995).

- 4) Munetomo, M. and Goldberg, D.E.: Designing a Genetic Algorithm Using the Linkage Identification by Nonlinearity Check, Technical Report IlliGAL Report No.98014, University of Illinois at Urbana-Champaign (1998).
- 5) Munetomo, M. and Goldberg, D.E.: Linkage Identification by Non-monotonicity Detection for Overlapping Functions, *Evolutionary Computation*, Vol.7, No.4, pp.377-398 (1999).
- 6) Sinclair, M.C.: NOMaD: Initial Architecture of an Optical Network Optimization, Modelling and Design Tool, *Proc. 12th UK Performance Engineering Workshop*, pp.157-167 (1996).
- 7) 棟朝雅晴：エビスタシス尺度に基づくリンクージ同定手法の提案, MPS シンポジウム論文集(進化的計算シンポジウム 2001), pp.99-106 (2001).
- 8) 滝根哲也, 伊藤大雄, 西尾章治郎：ネットワーク設計理論, 岩波書店 (2001).

(平成 14 年 4 月 11 日受付)

(平成 14 年 5 月 30 日再受付)

(平成 14 年 6 月 14 日採録)



辻 美和子

昭和 53 年生。平成 14 年北海道大学工学部情報工学科卒業。同年同大学大学院工学研究科システム情報工学専攻入学。



棟朝 雅晴

昭和 43 年生。平成 8 年北海道大学大学院工学研究科情報工学専攻博士後期課程修了。同年同大学院工学研究科システム情報工学専攻助手。平成 10 年から平成 11 年イリノイ大学基礎工学科遺伝的アルゴリズム研究室客員研究員。平成 11 年北海道大学情報メディア教育研究総合センター助教授。博士(工学)。遺伝的アルゴリズム, ネットワークシステム, 分散処理システムに関する研究に従事。IEEE, ISGEC (International Society of Genetic and Evolutionary Computation) 各会員。



赤間 清 (正会員)

昭和 48 年東京工業大学工学部制御工学科卒業。昭和 50 年同大学大学院修士課程修了。昭和 54 年同大学院博士課程単位修得退学。同年同大学助手。昭和 56 年北海道大学文学部講師。平成元年同大学工学部助教授。平成 11 年より同大学情報メディア教育研究総合センター教授。人工知能，知識処理，等価変換に基づく問題解決の研究に従事。工学博士。ソフトウェア科学会会員。
