

遺伝的アルゴリズムによる重み付きグラフの多分割について

野口博範[†], 大森健児^{††}

本論文では、グラフ分割問題を一般化した重み付きグラフ多分割問題を遺伝的アルゴリズムを用いて解く方法を提案する。また、具体的な例として、小選挙区での分割問題を取り上げる。本論文の方法を用いることで現在の小選挙区の区割りより、人口格差の点で良質な結果を実現できる。

Multiple-partitioning of a Weighted-graph Using GA

HIRONORI NOGUCHI[†] and KENJI OHMORI^{††}

This paper describes a genetic algorithm for multiple partitioning of a graph with weighted-vertexes and weighted-edges so that the conventional graph partitioning problem is expanded to more general application areas. This paper also presents applicable examples of this algorithm by proposing the dividing method in order to allocate constituencies automatically. This method has successfully given us better results than the conventional method.

1. はじめに

グラフ分割問題は、解を得る方法を示すことはできるがいつ解が得られるかは分からない NP 困難問題である。代表的な近似解法には、Kernighan と Lin の KL 法¹⁾がある。この解法は比較的良好な解を与えるが、初期分割に対する依存性が強く局所解に陥ることがある。KL 法の計算時間は、グラフの頂点数を N とした場合 $O(N^2 \log N)$ である。また、別の代表的な近似解法には、Fiduccia と Mattheyses の辺数に対する線形解法²⁾ (以下、FM 法)がある。この解法も初期分割に対する依存性が高い。

従来の解法の多くは、頂点および辺の重みをすべて一定にしたグラフの 2 分割問題に対してである。しかし、多方面への応用を考えたとき、頂点および辺の重みは異なる方が柔軟であり、また 2 分割ではなく多分割が望ましい。ここでは、組合せ最適化問題の 1 つであるグラフ分割問題を、重み付きグラフ多分割問題へ一般化する。そして、これを解くために、いろいろな組合せ最適化問題で良好な解が得られている遺伝的アルゴリズムで、重み付きグラフ多分割問題を解くこと

を考え、その手法を提案する。

また、応用として衆議院小選挙区の区割りを扱う。選挙区の作成は、様々な条件を考慮して行われるため、きわめて複雑な組合せ問題であり、膨大な時間と労力をかけて現行の選挙区は定められた。そこで本論文では、各地域の人口と各地域の隣接関係を重み付き無向グラフで表し、小選挙区の区割りを行う。ここでこの手法を用いることで、現行の区割りより少ない人口格差の小選挙区案が得られた。

2. グラフ分割問題

グラフ分割問題とは、グラフ $G(V, E)$ の頂点と辺の集合が与えられたとき、各グループに属する頂点の重みの総和を許容範囲 (制約条件) に保ちつつ、グループ間の辺の重みの総和が最小 (目的) となるようにグラフを分割するものである。ここで V は頂点の集合、 E は辺の集合である。

グラフ分割問題の制約条件と目的は、次のように与えられる。

「制約条件」同一のグループ G_K に属する頂点の重みの総和 W_K は許容範囲の中にある。ただし、 W はグラフ G での全頂点の重みの総和、 D は分割個数、 α は許容範囲を制御するパラメータである。

$$(1 - \alpha)W/D \leq W_K \leq (1 + \alpha)W/D \quad (K = 1 \sim D) \quad (1)$$

「目的」グループ間にまたがる辺の重みの総和 C を最小にする。

[†] 法政大学大学院工学研究科

Graduate School of Engineering, Hosei University

^{††} 法政大学情報科学部コンピュータ科学科

Computer and Information Sciences, Hosei University

現在、中部日本電気ソフトウェア株式会社

Presently with Chubu NEC Software

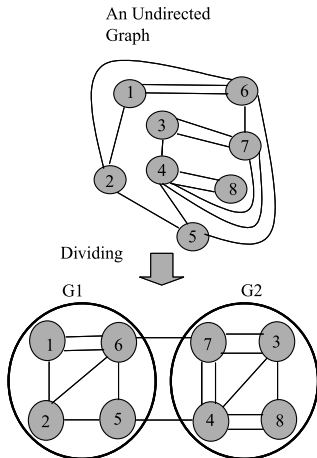


図1 重みなしグラフ2分割問題
Fig.1 Dividing a non-weighted graph.

$$C \rightarrow \text{minimum} \quad (2)$$

図1は、頂点の数を8、頂点および辺の重みをすべて1に固定した重みのある無向グラフの2分割例である。

3. 遺伝的アルゴリズム

遺伝的アルゴリズムは、ダーウィンの進化論に基礎をおいた確率的探索・学習・最適化手法である。選択・交叉・突然変異といった3つの遺伝的操作を順に実行し、さらにこれを繰り返すことで、環境に最も適した個体群を見出すことができる。また、自然淘汰に基づく確率的な探索手法として頑健な枠組みを持ち、組合せ最適化問題のようなNP困難問題を解くための有力な方法である。

3.1 遺伝子の表現

遺伝的アルゴリズムでは、解を染色体の形で表し、染色体を遺伝子で構成する。ここでのグラフ分割問題に対しては、分割されたグラフ、すなわち、解としての案を染色体で表し、遺伝子座が頂点を、遺伝子の値がグループに含まれていく順序(優先度)を表すようにした。したがって、遺伝子は、頂点がどのグループに属するかを間接的に表す。各遺伝子のとりうる値は総頂点数までとした。各頂点がどのグループに属するかは次の操作により定まる。1) 遺伝子の値が昇順になるように頂点を並べる。もし、遺伝子の値が同じである場合、頂点の番号が昇順になるようにする(図2参照)。2) 頭から順番に頂点を取り出し、これまでに取り出した頂点の重みの総和 W_1 が式(1)を満たしたときはこれらを1つのグループ G_1 とする。3) 引き続き残っている頂点に対しこの操作を行い、次のグループ

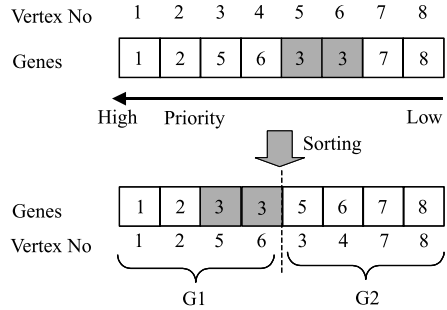


図2 8頂点グラフの2分割問題での並べ替え
Fig.2 Sorting 8 vertexes.

G_2 を作る。4) これを最後まで繰り返し G_1, \dots, G_D を作る。

3.2 アルゴリズムの詳細

アルゴリズムはステップ1からステップ10で構成される。ただし、ステップ9は、収束速度を補助的に速めるためのもので、各頂点の重みが比較的均衡しているときのみ用いる。

ステップ1「初期設定」: α の値をあらかじめ定めた値に設定する(式(1)参照)。

ステップ2「初期集団の生成」: 初期集団を作る。初期集団は、多くの個体で構成され、各個体は分割の案を示す。初期集団にランダムな分割案を持たせるために、染色体を構成する遺伝子にはランダムな値を設定する。

ステップ3「並べ替え(昇順)」(図2参照):

(i) 各個体で遺伝子の値が昇順になるように並べ替える。

(ii) もし遺伝子の値(グループを決めるための値)が同じである場合、頂点の番号が昇順になるようにする。

ステップ4「グループ分割する際の位置決め」: 並べ替えた後での r 番目の頂点を S_r とし、ここまでの頂点の重みの総和は

$$PW_r = \sum_{i=S_1}^{S_r} w_i$$

とする。頂点 S_r がグループ NB に属し、その次の頂点 S_{r+1} がグループ $NB+1$ に属する境界を示す r は次の式により求めたものとする。

$$|(W/D) \times NB - PW_r| \rightarrow \text{minimum} \quad (3)$$

図3は頂点の数9、各頂点に重みのある無向グラフをグループ数3に分割した例である。

ステップ5「各個体の適応度の計算」: グループ間にまたがる辺の重みの総和(カット重み)を計算する。これが小さいほど良い分割を与える。

ステップ6「選択」: 次世代に残れる個体は、一部

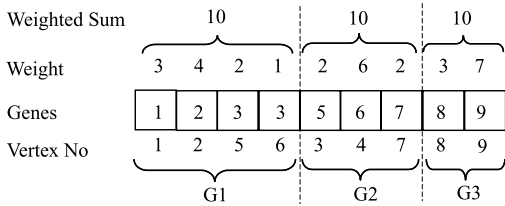


図3 重み付きグラフ3分割問題の染色体の分割位置決め
Fig. 3 Allocation of genes.

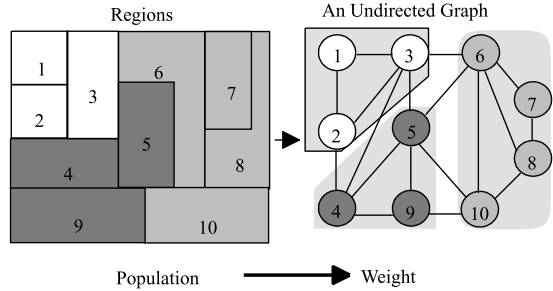


図4 無向グラフに置き換え
Fig. 4 Transformation to an undirected graph.

をエリート戦略で残りをルーレット選択で定める。エリート戦略では、適応度の高いものを無条件に残す。ルーレット選択では、ルーレットの板に占める各個体の割合を適応度に基づいて定め、ルーレットを回して当たった個体を残す。

ステップ7「交叉」: 2つの個体で染色体を交叉させ、新しい個体を作る。親となる2つの個体はルーレット選択で選ぶ。親の染色体を一樣交叉(各遺伝子ごとに乱数をとりどちらの親から継承するかを定める)させ、子としての染色体(個体)を作る。

ステップ8「突然変異」: 個体に対する遺伝子操作の1つで、遺伝子情報の一部をあらかじめ決めておいた確率で変化させる。これにより局所解から抜け出す機会ができる。

ステップ9「最大カット数頂点抽出」: グループごとに他のグループにまたがる辺の重みの和が最大である頂点を検出し、これを確率的に他のグループに移す。

ステップ10「集団の評価」: 一定世代変化がない場合、もしくは決められた世代数に達した場合終了する。この条件を満たさないときステップ3へ戻る。

4. 衆議院小選挙区区割り

本研究のアルゴリズムを小選挙区の区割りに応用し、遺伝的アルゴリズムを用いた手法の有用性を検証する。

実験では、「人口の格差を2倍以内におさえる」「飛び地はなくす」を小選挙区作成の制約条件とする。現行選挙区作成では「以前の中選挙区の区域を尊重する」「地域事情を考慮する」などの制約条件が加わるが、ここでは考慮しない。遺伝的アルゴリズムでは、染色体とその適応度を定める関数を決めるだけでよく、遺伝的アルゴリズムの部分は、どの問題に対しても共通であるので新たな制約条件を簡単に加えることが可能である。このため、区割りの制約が変わっても容易に対応できる。

4.1 衆議院小選挙区区割りのグラフ化と定式化

各地域(区市郡町村)の人口と各地域の隣接関係を重み付き無向グラフに変換する。このとき、各地域の

Population → Weight

図5 飛び地有グラフ3分割

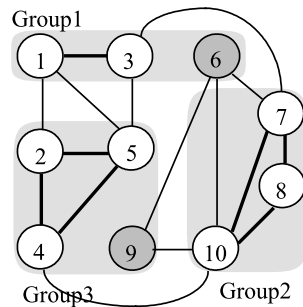


図5 飛び地有グラフ3分割
Fig. 5 Dividing with enclaves.

人口数を各頂点の重みとし、各地域の隣接関係を各頂点間にまたがる辺とする(図4参照)。各都道府県ごとに制定された議員数分割を行う際のグループ数となる。

ここで、小選挙区区割り問題の制約条件と目的は次のようになる。

「制約条件」

(i) 人口の格差を2倍以内におさえる

$$(1 - \alpha)W/D \leq W_K \leq (1 + \alpha)W/D \quad (4)$$

$(\alpha = 1/3, K = 1 \sim D)$

(ii) 飛び地をなくす

各グループですべての頂点が結合するようにする。グループ G_K 内で分離しているサブグループの数を GN_K 、飛び地の個数を FG とすると次の式が成り立つ(図5で6と9は飛び地、グループ1と2のサブグループは2)

$$FG = \sum_{K=1}^D GN_K - D \quad (5)$$

「目的」

(i) 人口の格差をできるだけ小さくする

グループごとに頂点の重みの和を計算し、その最大値を W_{MAX} 、最小値を W_{MIN} 、それらの差を ND としたとき次の式を最小にする。

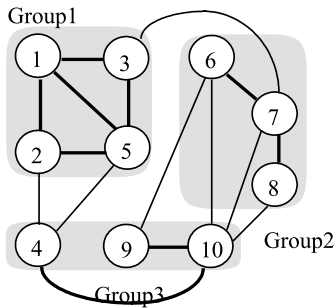


図6 飛び地なしグラフ3分割
Fig.6 Dividing without enclaves.

$$ND = W_{MAX} - W_{MIN} \rightarrow \text{minimum} \quad (6)$$

ここで、図5、図6に頂点10の無向グラフによる飛び地の有無の例を示す。

小選挙区区割りのアルゴリズムは次のようになる。
「衆議院小選挙区区割りのアルゴリズム」

ステップ5を次のように変更し、他は重み付きグラフ多分割問題と同じ。なお、ステップ9は利用しない。

ステップ5「各個体の適応度による順位」 集団の評価は、まず、式(6)を用いて、人口格差の少ないものから順に並べる。次に、式(5)を用いて、飛び地がないものを上にして並べ替える(飛び地の有無で分けられたグループ内では人口格差の順位は維持)。同様に人口格差の順位を維持しながら、式(4)を用いて格差が2倍以内のものを上にして並べ替える。

5. 実験結果

グラフ分割問題と小選挙区区割りの結果を示す。

5.1 重みなしグラフ2分割問題の実験結果

従来の近似解法FM法との比較を行うため、合計24種類のランダムグラフを作成し、評価を行った。ただし、頂点の重みは1に固定、頂点の数は100と200、グラフの分割個数は2、辺発生率(グラフでの辺の数を定める)は[0.05, 0.10, 0.15, 0.20, 0.25, 0.30]、辺の重みは1に固定と1~10のランダムな数で評価を行った。本論文の遺伝的アルゴリズム(GA法)でパラメータは次のように定め実験を行った。選択手法のエリート戦略とルーレット選択の比率は1:1、集団サイズは100、世代数は頂点の数×20、初期分割は100種類とした。

FM法、GA法での数値はそれぞれの方法で解を求めたときのグループにまたがる辺の重みの総和である。したがって、小さいほどよい。

従来の解法との比較の結果、遺伝的アルゴリズムは、多点同時探索による広域的な解の探索を行うことで局所解になるべく陥らないようにしているが、この効果

表1 本アルゴリズムの設定
Table 1 Parameter setting.

突然変異率	0.1
選択手法	エリート戦略とルーレット選択
交叉手法	一様交叉
交叉率	0.4

表2 頂点100の実験結果

Table 2 Experimental results for 100 vertexes.

頂点	辺	辺発生率	FM法	GA法	改善
1	1	0.05	26	26	0
1	1	0.10	61	60	1
1	1	0.15	106	106	0
1	1	0.20	157	155	2
1	1	0.25	200	196	4
1	1	0.30	236	236	0
1	1-10	0.05	93	84	9
1	1-10	0.10	304	293	9
1	1-10	0.15	460	445	5
1	1-10	0.20	729	715	14
1	1-10	0.25	829	829	0
1	1-10	0.30	1236	1224	12

表3 頂点200の実験結果

Table 3 Experimental results for 200 vertexes.

頂点	辺	辺発生率	FM法	GA法	改善
1	1	0.05	141	145	-4
1	1	0.10	287	282	5
1	1	0.15	451	447	4
1	1	0.20	709	705	4
1	1	0.25	793	791	2
1	1	0.30	1012	1011	1
1	1-10	0.05	618	627	-9
1	1-10	0.10	1355	1355	0
1	1-10	0.15	2386	2378	8
1	1-10	0.20	3110	3076	34
1	1-10	0.25	3794	3794	0
1	1-10	0.30	4413	4405	8

のため、FM法よりも良い解が得られた。

頂点の数100の実験の結果を表2に示す。

頂点の数200の実験の結果を表3に示す。

5.2 衆議院小選挙区区割りの実行結果

衆議院小選挙区の区割りでの重み付きグラフ多分割問題における実験結果を示す(表4参照)。

パラメータは表1のように定めた。集団サイズは200、世代数は1000、20種類の初期分割で実験を行った。人口データは、平成7年国勢調査のデータを用いた。ここでは10県について実験を行った。実験1は、現行小選挙区での区割りの中で、行政単位の異なる市と郡を離したりくっつけたりしてもよいこととした。さらに、同一の郡に属しているながら分離している町村についてもくっつけたり離したりしてもよいことにし

表 4 衆議院小選挙区区割りの人口格差実験結果

Table 4 A constituency proposal by experimental results.

県名	区割り数	現在	実験 1	実験 2
滋賀県	3	1.53	1.13	1.01
富山県	3	1.53	1.04	1.04
山口県	4	1.26	1.03	1.03
青森県	4	1.44	1.21	1.21
長野県	5	1.71	1.05	1.03
岐阜県	5	1.38	1.08	1.07
福島県	5	1.63	1.09	1.07
栃木県	5	1.63	1.15	1.14
茨城県	7	1.70	1.15	1.15
広島県	7	1.36	1.17	1.17

表 5 各選挙区の人口数

Table 5 Population for each allocation.

選挙区	現在選挙区	実験 (1)	実験 (2)
1 区	509,581	393,103	393,103
2 区	434,212	434,212	445,087
3 区	446,610	446,610	406,588
4 区	365,586	412,602	445,505
5 区	299,978	387,103	387,103
6 区	502,087	435,276	440,175
7 区	397,458	446,606	437,951
人口格差倍率	1.70	1.15	1.15
人口差	209,603	59,507	58,402

表 6 衆議院小選挙区区割り茨城県

Table 6 Proposed constituencies for Ibaraki Prefecture.

市区町村名	人口数	現在	実験 (1)	実験 (2)
北茨城市	52,074	5	5	5
高萩市	35,604	5	5	5
多賀郡	13,059	5	5	5
日立市	199,241	5	5	5
久慈郡	47,580	4	5	5
那珂郡	131,708	4	4	4
常陸太田市	39,545	4	5	5
ひたちなか市	146,753	4	1	1
水戸市	246,350	1	1	1
東茨城郡 A	24,111	1	4	4
西茨城郡 C	2,621	1	4	4
笠間市	30,336	1	4	4
西茨城郡 A	23,488	1	4	4
真壁郡	80,500	1	4	4
下館市	66,062	1	4	4
下妻市	36,113	1	7	4
東茨城郡 B	114,427	2	2	2
西茨城郡 B	50,566	2	2	4
鹿島郡	134,022	2	2	2
鹿嶋市	60,671	2	2	2
行方郡	74,526	2	2	2
新治郡 A	85,580	6	6	6
石岡市	52,712	6	6	2
土浦市	132,246	6	6	6
つくば市	156,009	6	6	6
筑波郡	40,493	6	7	7
稲敷郡 A	26,318	6	7	3
稲敷郡 B	126,799	3	3	3
牛久市	66,340	3	3	6
龍ヶ崎市	69,161	3	3	3
北相馬郡	99,835	3	3	3
取手市	84,475	3	3	3
水海道市	42,689	7	7	7
岩井市	44,325	7	7	7
結城郡	57,705	7	7	7
猿島郡	139,877	7	7	7
古河市	59,086	7	7	7
結城市	53,776	7	4	7
新治郡 B	8,729	6	6	2

表 7 市区町村の詳細

Table 7 Details of counties.

市区町村名	詳細の市区町村名
東茨城郡 A	御前山, 桂, 常北
東茨城郡 B	茨城, 小川, 美野里, 内原, 大洗
西茨城郡 A	岩瀬
西茨城郡 B	岩間, 友部
西茨城郡 C	七会
新治郡 A	出島, 八郷, 千代田, 新治
新治郡 B	玉里
稲敷郡 A	笠崎
稲敷郡 B	江戸崎, 美浦, 阿見, 新利根, 河内, 桜川, 東

た。実験 2 においては、さらに、同じ郡内の町村でも地理的に隣接していないときは、分離してもよいものとした。栃木県を例にとると、現行選挙区は次に示すとおりである。

- 1 区：宇都宮市，河内郡（上三川・南河内町）
- 2 区：日光市，今市市，鹿沼市，上都賀郡，塩谷郡，河内郡（上河内・河内町）
- 3 区：大田原市，矢板市，黒磯市，那須郡
- 4 区：小山市，真岡市，芳賀郡，下都賀郡
- 5 区：足利市，栃木市，佐野市，安蘇郡

上記のようになっているので、1 区と 2 区に分離されている河内郡だけは離れてもよいものとして、これら 20 の市と郡を実験 1 では分離することにする。しかし、実験 2 では、分離している郡、すなわち、河内郡（上三川・南河内町）河内郡（上河内・河内町）那須郡（南那須町，烏山町，馬頭町，小川町，那須町，湯津上村，黒羽町）那須郡（塩原町，西那須野町）下都賀郡（壬生町，石橋町，国分寺町，都賀町）下都賀郡（大平町，岩舟町，藤岡町，野木町）は分離してもよいものとして、実験を行った。

例として茨城県の実験結果を表 5，表 6 に示す。なお、郡部の詳細は表 7 に示す。

茨城県の現行の区割りを実験 1 による区割りを図 7 に示す。

最後に、Pentium4 2.0 GHZ，512 MB メモリ，Win-

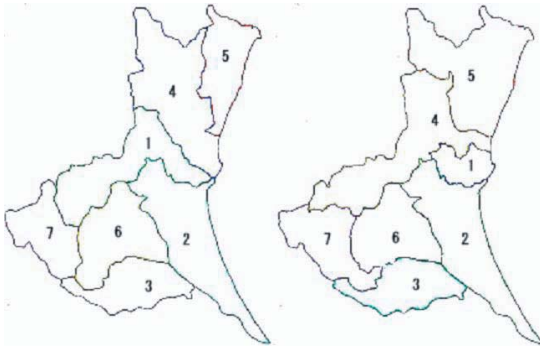


図7 茨城県の現行選挙区と実験1の結果

Fig. 7 The current and proposed constituencies of Ibaraki Prefecture.

表8 頂点200での処理時間の比較
Table 8 Comparison of processing time.

頂点重み	辺重み	辺発生率	FM法	GA法
1	1	0.05	1	33
1	1	0.10	1	34
1	1	0.15	1	33
1	1	0.20	1	42
1	1	0.25	1	43
1	1	0.30	1	42
1	1-10	0.05	1	35
1	1-10	0.10	1	39
1	1-10	0.15	1	40
1	1-10	0.20	1	42
1	1-10	0.25	2	43
1	1-10	0.30	2	45

dows2000 Professional SP3 を用いての処理時間の比較を示す(表8)。なお、頂点100の場合には、FM法が1秒に対し、本論文の遺伝的アルゴリズム(GA法)は、10から14秒であった。このことから、FM法よりは十数倍の処理時間がかかるが、多くの場合において、遺伝的アルゴリズムのほうが良い結果を与えることが判明した。また、衆議院小選挙区区割り問題では、目的にあった解を3~10秒の処理時間で得ることができた。

6. おわりに

遺伝的アルゴリズムによる重み付きグラフ多分割問題の解法と応用例を紹介した。従来の解法と比較した結果、グラフ2分割問題において、FM法よりも遺伝的アルゴリズムは良い解を探索できた。また、応用例としての衆議院小選挙区区割り問題では、10県すべてについて良好な結果を得た。選挙区作成に本研究のアルゴリズムを用いることで、高速かつ正確に目的に合った処理ができ、短い時間で解を得ることができた。また、初期値や制約条件を変えることで多種多様な区

割り案を示すことができ、容易に比較検討することもできる。これによって本論文の重み付き無向グラフのグループ多分割によるアルゴリズムの有効性が示された。このアルゴリズムを用いることで衆議院選挙区区割り問題だけでなく、日常生活で起きる現実的な問題に対処できる。すなわち、与えられた問題を重み付き無向グラフに置き換え、そのグラフを定められたグループに分割するように問題を置き換え初期値やその問題に応じた制約条件を定めれば、多くの問題に本論文のアルゴリズムが適用可能である。

参考文献

- 1) Kernighan, B.W. and LIN, S.: An efficient heuristics procedure for partitioning graphs, *Bell Syst. Tech. J.*, Vol.49, No.2, pp.291-307 (1970).
- 2) Fiduccia, C.M. and Mattheyses, R.M.: A linear-time heuristics for improving network partitions, *Proc. 19th Design Automation Conf.*, pp.175-181 (1982).
- 3) Bui, T.N. and Moon, B.R.: Genetic Algorithm and Graph Partitioning, *IEEE Trans. Computers*, Vol.45, No.7, pp.841-855 (1996).
- 4) Johnson, D.S., Aragon, C., McGeoch, L. and Schevon, C.: Optimization by Simulated Annealing, *An Experimental Evaluation, Part1, Graph Partitioning*, Operations Research, Vol.37, pp.865-892 (1987).
- 5) 斎藤孝之, 武藤佳恭: ニューラルコンピューティングによる小選挙区割手法, pp.588-596, 情報処理学会 (1996).
- 6) 藤沢克樹, 久保幹雄, 森戸 晋: Tabu Searchのグラフ分割問題への適用と実験的解析, T. IEE. Japan, pp.430-443 (1994).
- 7) 丸山 勉, 小長谷明彦, 小西弘一: 遺伝的アルゴリズムを用いた並列グラフ分割アルゴリズム, pp.556-567, 情報処理学会 (1993).
- 8) 総務庁統計局: 国勢調査報告, 日本統計協会 (1995).
- 9) 総務庁統計局: 全国都道府県市区町村別人口, 日本統計協会 (1995).
- 10) 宮川隆義: 小選挙区ハンドブック, 政治広報センター (1996).

(平成12年5月23日受付)

(平成15年7月3日採録)



野口 博範(正会員)

昭和 50 年生。平成 10 年法政大学工学部経営工学科卒業。平成 12 年法政大学大学院工学研究科システム工学専攻修士課程修了。同年中部日本電気ソフトウェア(株)入社。現

在、医事、オーダ、電子カルテ等の医療支援ソフトのテクニカル業務担当。



大森 健児(正会員)

昭和 20 年生。昭和 44 年東京大学工学部計数工学科卒業。昭和 47 年カリフォルニア大学バークレイ校 EECS 修士課程修了。昭和 44 年日本電気(株)中央研究所入社。昭和

60 年法政大学工学部経営工学科教授。平成 12 年法政大学情報科学部コンピュータ科学科教授。同学部長。並列コンピュータアーキテクチャ、オブジェクト指向計算、パターン認識、遺伝的アルゴリズム等の研究に従事。情報処理学会より昭和 59 年論文賞、創立 25 周年記念論文。
