

博士論文

進化計算を用いた「合コン」割当問題の解法

黒岩 将

2013年2月23日

奈良先端科学技術大学院大学
情報科学研究科 情報処理学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
博士(工学) 授与の要件として提出した博士論文である。

黒岩 将

審査委員：

伊藤 実 教授	(主指導教員)
杉本 謙二 教授	(副指導教員)
安本 慶一 教授	(副指導教員)

進化計算を用いた「合コン」割当問題の解法*

黒岩 将

内容梗概

本論文は「合コン」における問題の科学的な解法に関する研究をまとめたものである。「合コン」は、結婚相手を見つけるため男女が交流を深めることを目的に集う会合である。近年、我が国では、少子化の進行に伴い、独身の数を減らそうという動きが活発化し、「合コン」をセッティングするサービスが出現し、趨勢となってきた。こうしたサービスにおける主要な課題は、カップル成立数を高めることである。本研究では、相性の良い（以下、好相性と呼ぶ）男女を同じ合コンに参加させるために、以下の2つの問題を同時に解くシステムの設計・開発に取り組んだ。どのような男女が好相性なのかという問題（予測問題）、また、既知となった好相性をどのように用いて、どのように合コンの参加者を構成すれば良いのか（組合せ最適化問題）、という問題である。前者の予測問題の解決には、進化計算（EC, Evolutionary Computation）を用い、後者の組合せ最適化問題には、好相性を表す解ごとに、それに似た属性情報を持つ男女を合コンの定員数ずつ集めるという単純な方法を採用し、合コン割当問題を段階的に解く構成とした。

提案システムでは、入力を、システム利用者の年齢、血液型、学歴、趣味、年収などの属性情報とし、出力を、カップル成立の可能性が高い合コン参加者名簿とする。ECの解候補は、好相性と思われる属性の組合せを表現しており、それぞれ異なった合コンに対応させる。解候補に似た属性情報を持つ男女をその解候補に対応する合コンに割り当てる。その後、合コンを行い、得られた結果（カップル成立数）を、その解候補に評価値（適応度）として与える。その後、評価値

* 奈良先端科学技術大学院大学 情報科学研究科 情報処理学専攻 博士論文, NAIST-IS-DD0861004, 2013年2月23日.

付きの解候補群に対し EC 演算を加える．以上の流れを繰り返すことで，高い評価値を持つ解が解集団内に残っていく形で好相性の準最適解集団を求める．

本論文では，一つの最適解を求める手法を提案し評価を行った後，相性空間のような多峰性関数の複数極大値算出手法として，類似の解が増えないように工夫した EC の選択法を導入し評価を行った．相性空間（男女の属性の組が取り得る値の範囲）は多峰性であり，多様なカップルを成立させるためには，EC によって，この多くの峰を求める必要があるからである．合わせて，少ない合コン回数でより良い解を得ることができるよう，実際の合コンを行わずに，過去の合コン履歴を活用して任意の解に評価値を与える方法を導入した．

提案システムを評価するため， N 個のピーク（好相性の男女の属性情報の組合せに相当）を持つ多峰性関数である N Max 問題を定義し， $N = 5$ の場合について計算機実験を行った．比較のため，カップル成立した男女の属性を好相性としてそのまま使用する手法を用意した．その結果，提案システムが，比較手法に比べて，半分の合コン実施回数で，最適解に約 90% 以上一致する解集団を求めることができ，得られた解を基にした合コン割当てにより約 2 倍のカップル成立数を達成できることを確認した．

キーワード

男女の好相性の探索，合コン，進化計算，多峰性関数，マルチ・ニッチイズ・クラウディング

Solving Matchmaking Parties Assignment Problem through Evolutionary Computation *

Sho KUROIWA

Abstract

This thesis is about the study on the scientific and mathematical method to solve the match-making parties (called Gokon, hereafter) assignment problem. Gokon is usually held by unmarried male participants and the same number of unmarried female ones, and they have a dinner party for a couple of hours together in hope of finding a marriage partner. Recently, the low birthrate has become a serious problem in Japan. So, movement for reducing number of singles is now attracting considerable attention. For example, several local governments and enterprises have provided opportunities for unmarried people to meet potential marriage partners in Japan. However, in order to assign participants so as to maximize the number of good matches of man and woman pairs, there are two problems. First, good affinity between men and women is not yet well-understood (prediction problem). Second, determining the best Gokon assignments is also difficult (combinatorial optimization problem). In this thesis, we designed and developed a system to solve these two problems. To resolve the prediction problem, Evolutionary Computation (EC) is used. To find the optimal solution of the target problem, an incredibly large number of Gokons are needed. Thus, the method using archival records to evaluate the solution instead of holding new Gokons is adopted. To resolve the combinatorial optimization problem, we set Gokons as many as possible and assign, to each Gokon, a specified number of men and women who have the attributes similar to good affinity obtained as a solution.

* Doctoral Dissertation, Department of Information Processing, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD0861004, February 23, 2013.

Inputs of the proposed system are attribute information of participants such as age, blood type, academic record, hobby and annual income. Outputs are Gokon participant lists which maximize the number of good matches. The system has concatenations of man and woman attribute information as candidate solutions of EC representing good affinity patterns. The system assigns members into the Gokon according to each candidate solution (generate participant list of Gokon) and gives results of the Gokon as a fitness value of the corresponding candidate solution. By applying EC Operation to the candidate solutions, they will improve gradually toward the optimal solutions.

Since our target prediction problem is a multimodal problem, EC has to find many peaks (good affinity patterns) in the domain of affinity patterns. First, we address how to find the peak of a unimodal problem, and propose a method for solving the problem and evaluate it. Second, we devise a method for solving multimodal problem and evaluate it by using computer simulation. Third, the method using archival records to evaluate the solution instead of holding new Gokons is introduced.

To evaluate the system, we defined the N Max problem assuming that there would be N good affinity patterns between men and women as a benchmark test. Through computer simulations for $N = 5$, we confirmed that the proposed system obtained the semi-optimal solutions with about 90% agreement rate to the optimal solutions and achieved almost twice as many good matches as a conventional method with about half the evaluation times.

Keywords:

Evolutionary Computation, Matchmaking Party, Multi-modal, Multi-niches Crowding

目 次

1. 序論	1
2. 関連研究	3
3. 合コン割当問題の定式化と進化計算によるアプローチ	6
3.1 概要	6
3.2 合コン割当問題	6
3.3 提案システム	6
3.3.1 概要	6
3.3.2 解集団と初期化	8
3.3.3 進化計算オペレータ	9
3.3.4 合コン割当オペレータ	9
3.3.5 合コン評価オペレータ	10
3.4 実験	11
3.4.1 概要	11
3.4.2 ベンチマークテスト (NMax 問題)	11
3.4.3 評価基準: 合コン割当評価値	12
3.4.4 比較手法	13
3.4.5 実験設定	14
3.4.6 結果	15
3.4.7 考察	15
3.5 まとめ	16
4. 複数の相性ピークの検出と参加者利用履歴の活用	18
4.1 概要	18
4.2 提案システム	19
4.2.1 概要	19
4.2.2 提案選択法	19
4.2.3 割当一致率を考慮した解候補評価法	21

4.2.4	ストック利用解候補評価法	22
4.3	実験	23
4.3.1	概要	23
4.3.2	ベンチマークテスト ($N\text{Max}$ 問題) の改良	24
4.3.3	評価基準: 最適解到達率	24
4.3.4	比較手法	24
4.3.5	実験設定	25
4.3.6	結果	26
4.3.7	考察	29
4.4	まとめ	30
5.	結論	31
	謝辞	33
	参考文献	34
	付録	41
A.	既発表文献との整合性について	41
B.	乱数生成された属性値とカップル成立数の関係	41

図目次

1	提案システムの全体像	7
2	解の例	8
3	合コン割当オペレータの動作	10
4	合コン評価オペレータの動作	11
5	N Max 問題 ($N = 2$)	12
6	貪欲手法	13
7	入力データの簡単化	14
8	合コン割当評価値の推移 (2Max 問題)	16
9	提案選択法の全体像	20
10	割当一致率	21
11	ストックの導入	22
12	ストックを利用した解候補への評価値の与え方	23
13	N Max 問題 ($N = 5$)	27
14	提案選択法を導入時 ($\bar{R} = 91\%$)	27
15	提案選択法を非導入時 ($\bar{R} = 64\%$)	27
16	利用者属性長に対するスケーラビリティ ($N = 5$ の時)	29
17	問題サイズ N に対するスケーラビリティ (利用者属性長 = 20 の時)	30

表目次

1	解集団に対する処理の流れ	8
2	設定値 (3 章の実験)	15
3	設定値 (4 章の実験)	25
4	提案手法と貪欲手法の比較	28

1. 序論

2005 年以降，我が国の人口増加率は 0% 以下で推移 [1] してきており，高齢化の進行と同時に，大きな問題として議論されている．この傾向は，1992 年以降，一貫して続いており，政府も，2003 年より少子化社会対策基本法を整備，内閣府に少子化対策室を設置するなど，その対策に乗り出した [2]．出生率低下の要因の一つとして，当然ではあるが，結婚しない男女の増加にあることが指摘されており，未婚，非婚の主要な理由は，国立社会保障・人口問題研究所の調査の結果，男女共に「適切な相手に巡り合わない」であることが分かっている [3]．こうした流れの中で，「合コン」をセッティングするサービスが出現し，趨勢となっている [4, 5, 6, 7]．「合コン」は，主として男女が交流を深めることを目的に集う会合であり，基本的には男女同数で行われるのが一般的である．こうしたセッティングサービスでは，カップル成立数を高めることが課題となっている．

しかし，相性の良い男女を同じ合コンに参加させるためには，どのような男女が好相性なのか（予測問題），また，既知となった好相性をどのように用いて，どのように合コンの参加者を構成すれば良いのか（組合せ最適化問題），という 2 つの問題が存在する．組合せ最適化問題とは，目的関数を満たす解が複数の離散値で与えられている問題であり，一般的にすべての組合せを調べその中で最適解を求めるためには，現在の計算機を用いても数十年かかるような難しい問題であることが多い．本論文では，上記の予測問題と組合せ最適化問題を合わせて解くシステムを考案，実現する．

前者の予測問題の解決には，進化計算（EC, Evolutionary Computation）[8] を用いる．男女の属性情報の組合せを EC の解（染色体あるいは個体）としてシステムに持たせ，カップル成否を解の評価値（適応度）としてフィードバックしながら，好相性の準最適解集団を求める．後者の組合せ最適化問題には，好相性を表す解毎に，それに似ている属性値を持つ男女を合コンの定員数ずつ集めるという単純な方法を採用し，段階的に合コン割当問題が解けるような構成にする．

本論文では，まず合コン割当問題を定式化し，それに進化計算を適用して，1 つの最適解を求める手法を提案する（3 章）．しかし，相性空間（男女の属性の組が取り得る値の範囲）は多峰性であり，多様なカップルを成立させるためには，

ECによって、この多くの峰を求める必要があるという問題が存在する．この問題を解決するため、相性空間のような多峰性関数の複数極大値算出手法として、類似の解が増えないように工夫した EC の選択法を導する（4.2.2 項）．

また、対象問題の最適解を求めるためには、様々な参加者による多数の合コンを実施する必要があることが分かった．この回数をなるべく減らしたい．なぜなら、理想的な男女の属性値の組（属性ベクトルと呼ぶ）が求まるまでは、最適な合コンへの割り当てが行われないため、参加者はカップル成立数の面で不利になるからである（すなわち、参加者が「人柱」的になる．人柱とは、古来、橋や堤防などの大型建造物を造る際、安全祈願のための生贄として、生き埋めにされた人のことを指す．このように他人のために犠牲になることがしばしば人柱と表現される）．合コンを行わずに、解候補の評価を行うことができれば、利用者の時間的、経済的な負担、不利益の享受を減らすことができる．そこで、解候補を評価する際、合コンを行わずに、過去の合コン履歴を活用して行う方法も導入した（4.2.4 項）．

提案システムを評価するため、 N 個のピーク（好相性の男女の属性情報の組合せに相当）を持つ多峰性関数である $N\text{Max}$ 問題を定義し、 $N = 5$ の場合について計算機実験を行った．比較のため、カップル成立した男女の属性を好相性としてそのまま使用する手法を用意した．その結果、提案システムが、比較手法に比べて、半分の合コン実施回数で、最適解に約 90%一致する解を得ることができ、約 2 倍のカップル成立数を達成できることを確認した．

2. 関連研究

本章では、本論文の関連研究について述べる。

Evolutionary Computation (以下, EC) は, 生物の進化を模した, 対象問題の解候補群による多点探索アルゴリズムあり, 評価関数をブラックボックスのまま扱うことができるため, 幅広い分野に応用されている [8, 9]. 身近な例としては, 新幹線のぞみ 700 系統の車両先頭部の形状設計への応用がある [10]. この研究では, 形状を EC の解 (EC 用語で「染色体」あるいは「個体」) にエンコードし, 高速走行時に発生する騒音と空気抵抗を評価値 (EC 用語で「適応度」) にして最適化を行っている. また, 人間の好みや感性などのモデル化が困難なものをそのまま EC の評価系として組み込んだ対話型進化計算 (IEC: Interactive Evolutionary Computation) と呼ばれる分野の研究も広く行われている [11]. 代表的なものでは, 補聴器のフィッティング [12], 作曲支援システム [13, 14], 似顔絵生成システム [15] がある. 補聴器のフィッティングに関する研究では, 補聴器のパラメータを IEC の解として, その補聴器の使用者の聞こえを評価値として, 最適化が行われている. その結果, 従来手法に比べて, 使用者がより高い満足度を示すことが確認されている. 作曲支援システムに関する研究は, 楽譜情報を解にエンコードし, 進化計算により, 適切な解, すなわち曲を生成する研究である. 似顔絵生成システムに関する研究は, 顔の輪郭, 眉, 目といったアニメの顔のパーツを, EC の解にエンコードして, 一つの解候補で一つの顔を表現する. 使用者は, 表示された顔のリストを, 目的とする顔 (似顔絵) に近いかどうか判断しながら選択を行い, 選択された個体が高い評価値を持つように, 次世代の顔リストが生成される. この過程が繰り返されることで, システムは使用者の似顔絵を獲得することができる. しかし, これらの応用例は, いずれも 1 つの大局的最適解やパレート最適解の探索問題を扱っており, 本研究で扱う局所最適解も含めた複数解の探索問題にそのまま適用することはできない. また, EC を多峰性関数に適用する研究もいくつか存在するが, NSGA-II[16] を筆頭にほとんどがパレート最適解を求める多目的最適化を対象としており, 本論文で扱う問題 (以下, 合コン割当問題と呼ぶ) とは種類が異なる.

男女のマッチングに関する研究として, 古くから結婚安定問題が知られている

[17]. これは、複数の男女それぞれが、異性に対し、自分の好みで順位付けを行い、その順位リストの「安定性」を満たすマッチング（結婚）を求めるものであり、EC の適用例も報告されている [18]. 同問題を解くアルゴリズムを考案した Shapley[17] と、これを実世界の問題に応用（社会制度として設計）した Roth[19] は、同業績により、2012 年度のノーベル経済学賞の受賞に輝いた [20]. 我が国では、2004 年度から施行された医師臨床研修制度下の医師の研修先病院を決定するシステムに、同問題を解く手法が実装されている [21].

相性を考慮した男女のマッチングに関しては、生物学、生態学、心理学の領域で膨大な研究の蓄積があり、中でも有名なものが “Assortative mating” である. これは、男女のさまざまな属性、例えば、身長や性格、IQ 等々、男女間でその属性が似ているほど相性が良いという仮説であり、数々の実験によって、その証明が試みられてきた [22, 23]. この前提として、身長や指、腰のサイズなど体のパーツの大きさ（長さ）と異性に対する好意の相関も調べられている [24]. “Assortative mating” から着想を得た進化計算の交叉方法の改良に関する研究 [25] や、生物の個体群をモデル化し、その挙動を調べるという理論的な研究もある [26].

また、近年、ヒト白血球抗原（HLA : Human Leukocyte Antigen）という生体内のタンパク質の個人間のわずかな構造の違いが、男女のマッチングに影響を及ぼすことが分かってきた [27, 28]. HLA 型が似ていないほど相性が良いとされており、特に女性が好む男性の HLA は、その女性の父親由来の HLA 型と似ているということが分かっている [29]. 類似度が低いほど良いというのは、これは既存の仮説 [30] には反する結果であった. こうした知見から、我が国でも、それを利用したビジネスモデル特許が公開され [31], 米国では、ビジネスへの応用が始まっている [32].

合コン割当問題は、複数の男女をいくつかの合コン会場に割り当てる問題であり、これは一種のクラスタリングとみなすことができる. クラスタリングの代表的なアルゴリズムとしては、EM アルゴリズム [33] や K -means 法 [34] が挙げられる. しかし、これらは全てのデータ点間の距離が与えられた場合に適用可能な手法であり、本研究で対象とする合コン割当問題では、合コンという小さなグループ内のデータ点間の距離しか与えられないため、直接適用することはできない.

また，EC を応用したアルゴリズムとしては Fitness Based Partitioning[35, 36] をはじめ幾つか提案されており，中でも MOCK[37] は，未知のグループ数を自動的に決定する能力を有し，様々なデータセットにおいて高い分割性能を示すことで知られている．しかし，いずれも評価関数が既知であるものとしてマッチングや分割を行う研究であったり，1つの大域的最適解の探索問題であったりする等，そのまま合コン割当問題に適用することはできない．

3. 合コン割当問題の定式化と進化計算によるアプローチ

3.1 概要

本章では、この進化計算を計算基盤として採用し、大人数の利用者集合の中から、カップルの成立数の最大化を目的として、相性の良い男女を同じ合コンに割当てる問題（合コン割当問題）を定式化し、それを解くシステムを提案する。また、計算機実験によって、その有効性を検証する。

3.2 合コン割当問題

合コン割当問題とは、大規模な利用者集団から、好相性の男女混合グループ（合コン）を複数作る問題である。

入力: 問題の入力は男性利用者集団 $B = \{b_1, b_2, \dots\}$ と女性利用者集団 $G = \{g_1, g_2, \dots\}$ と定義する。男性 $b_i \in B$ と女性 $g_j \in G$ は、それぞれ k 個と l 個の属性情報 $(b_{i1}, b_{i2}, \dots, b_{ik})$, $(g_{j1}, g_{j2}, \dots, g_{jl})$ を持つ。属性情報は、例えば、身長や体重、経歴や性格に関する情報である。

出力: B と G をそれぞれ L 人ずつの部分集団に重複を許して分け、男女同数の部分集団の組（合コン）に割り当てる。得られる合コンの数を h とし、 i 組目の合コンを $[B_i, G_i]$ と表す。問題の出力は、以下の目的関数を最適化する合コン集団 $[B_1, G_1], \dots, [B_h, G_h]$ と未知の評価関数 F である。

目的関数: $[B_i, G_i]$ のカップル成立数を返す評価関数を F とし、 h 回の合コンにおけるカップル成立数の総和を最大化する目的関数を以下のように与える。

$$\text{maximize} \quad \sum_{i \in \{1, \dots, h\}} F(B_i, G_i) \quad (1)$$

3.3 提案システム

3.3.1 概要

図1に示すように、ECを用いて、合コン割当問題の評価関数 F を推定し、合コン割当名簿を出力（＝参加者を決定）するシステムを提案する。提案システム

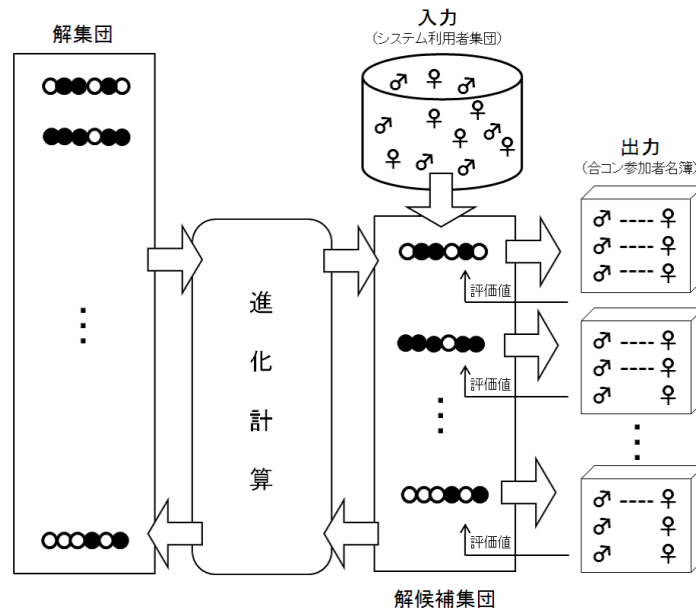


図 1 提案システムの全体像

は，男女の相性の良さを表現する**解集団**と，その解集団に対し EC を実行する**進化計算オペレータ**群（選択，交叉，突然変異），生成された**解候補集団**に対し合コン割当名簿を生成する**合コン割当オペレータ**，合コン結果を評価値として解候補に与える**合コン評価オペレータ**から構成される．解集団に対して，進化計算，合コン割当，合コン評価の各オペレータ（それぞれ 3.3.3 項，3.3.4 項，3.3.5 項で後述）を適用し，評価値付きの解候補集団を生成する．この解候補集団を用いて解集団の更新を行う．ここで，初期解集団を 1 世代， $t - 1$ 回目の更新後の世代を t 世代と呼ぶ．世代を経ながら，**最適解集団**，すなわち， F の推定を目指す．解集団に対する操作を表 1 に示す．

以下，解集団の初期化及び各オペレータについて詳述する．

表 1 解集団に対する処理の流れ

-
- 1: 解集団を初期化
 - 2: 指定した世代数，以下を適用
 - 3: 進化計算オペレータ
 - 4: 各解候補に以下を適用
 - 5: 合コン割当オペレータ
 - 6: 合コン評価オペレータ
 - 7: 解候補集団を用いて解集団を更新
-

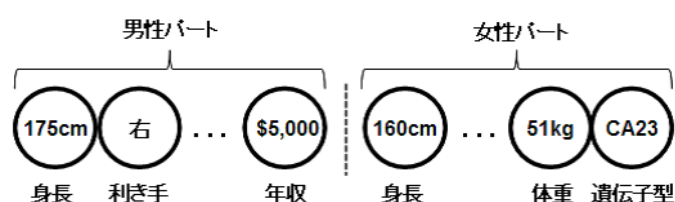


図 2 解の例

3.3.2 解集団と初期化

図 2 に示すように，解は，男性と女性の，身長や収入，学歴といった属性情報の組合せにより構成する．前半分を男性の属性，後半分を女性の属性に対応させる．これらの情報はアンケートによりシステム利用者から取得する．属性には HLA（2 章）などの先天的属性（しかし，HLA は評価関数が既知のため推定の必要がない．したがって今回は取り扱わない）と，TCI（Temperament and Character Inventory）[38] などの性格気質（後天的）テスト [39] の結果を加える．

こうした属性は，実際に，いくつかの結婚情報サービスで使用されているものである．しかし，これらは業者毎に異なり，また，あまり学術的研究も進んでいないため，どの情報がどの程度，相性に影響するかは未知である．身長や収入，

学歴なども、よく調べれば、平均より高い/低い、多い/少ない、高い/低いという分類で十分の可能性もあり、この場合は各 1 ビットで表現できる。また、TCI などは、アンケートの項目数は非常に多いが、結果は 12 パターン (4 ビット) に分類される。性格気質を 5 パターン (3 ビット) に分類できるとする研究も存在する [39]。

本章では、合コン割当問題への進化計算適用の可能性を探るため、まずは男女の各属性情報を 5 ビット (各解候補は 10 ビット) として実験を行うものとする。

本研究では、男女の属性情報によって、相性の良さが決まるという仮定を置いていている。各初期解の各要素は乱数を用いて生成する。

3.3.3 進化計算オペレータ

本オペレータは、図 1 に示すように、解集団に対して選択、突然変異という 3 つの EC 操作を加える。選択はルーレット戦略、スケーリングを $F' = 2F + 1(F_{ave} = F_{ave})$ とする線形スケーリングとした (F と F' をそれぞれスケーリング前後の評価値とする)。EC の定石である交叉は行わない。なぜなら、合コン割当問題が多峰性の問題の場合、解構造が交叉によって容易に崩壊するからである (この問題の対処法は、4 章で述べる)。

3.3.4 合コン割当オペレータ

以下に割当オペレータのアルゴリズムを示す。

Step 1. ある合コンに対して、割り当てられていない利用者の中で、その属性とその合コンの解候補の距離、例えばマンハッタン距離やハミング距離が最小であるような利用者を見つける。

Step 2. 利用者をその合コンに割り当てる。もしその合コンの最大収容人数に到達したら、その合コンを割り当て完了とする。そうでなければ、**Step 1** に戻る。(図 3)。

Step 3. **Step 1-Step 2** の過程をすべての合コンに対して繰り返す。

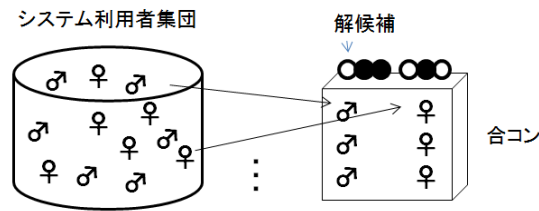


図3 合コン割当オペレータの動作

同オペレータを動作させた結果，現存の解候補集団（現時点で相性が良いと思われる男女の属性の組）に基づいた合コン割当てが実現される．この時，解候補の方も利用者の属性をある程度反映している（なぜなら，解候補に似た属性を持つ利用者をその合コンに集めたから）点が，3.3.5 項でポイントとなる．

3.3.5 合コン評価オペレータ

本オペレータは，合コンの結果を，その合コンに対応する解候補の評価値として与えるオペレータである．評価には，相手に対する満足度のような定量化が難しい主観的なものではなく，図4に示すように，その合コンでのカップル成立数という客観的な値を用いる．これは，合コン後にアンケートを取るなどして得る．例えば，男女間でカップル（両思いの場合をカップルとする）が1組成立した場合，その評価値を“1”とし，まったく成立しない場合は“0”とする．5組のカップルが成立したのであれば，その合コンに対応する解候補の評価値は“5”となる（図4）．解候補によく似た利用者を集めて合コンを作った結果，その合コンでの両思い成立数が多いなら，その解候補は，相性の良い男女の属性の組合せを表現している可能性が高い．また，この命題の裏も真である．

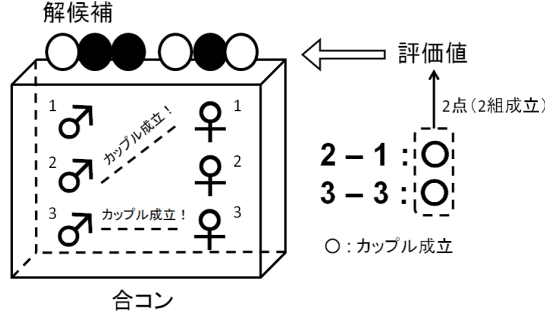


図 4 合コン評価オペレータの動作

3.4 実験

3.4.1 概要

提案システムが合コン割当問題を解くのに有効であることを計算機実験によって確認する．ここでは，解集団にとって最初未知である男女の相性の評価関数に， $N\text{Max}$ 問題をベンチマークテストとして導入し（3.4.2 項），実験を行った． $N\text{Max}$ 問題は N 個の相性の良さ（ピーク）があると仮定した多峰性関数である．実験では，相性の良い男女を同じ合コンに配置できたかどうか，について評価を行った．比較のため，単純な方式で良い合コン割当てを探す貪欲手法（3.4.4 項）と，解集団に最適解を与えて合コン割当てを求める神託手法（3.4.4 項）を用意し，実験を行った．

3.4.2 ベンチマークテスト ($N\text{Max}$ 問題)

3.3.2 項で定義した任意の男女の属性を表わすベクトル $\mathbf{x}$ と好相性の一つを表わす最適解 $\mathbf{p}$ との一致率を返す関数 $\text{lovematch}(\mathbf{x}, \mathbf{p})$ を以下のように与える．

$$\text{lovematch}(\mathbf{x}, \mathbf{p}) = \frac{\sum_{i=1}^{|\mathbf{x}|} g(i)}{|\mathbf{x}|} \quad (2)$$

$$g(i) = \begin{cases} 1 & (x_i = p_i) \\ 0 & (\text{otherwise}) \end{cases} \quad (3)$$

Peak.1 00000000000000000000
Peak.2 11111111111111111111

図 5 NMax 問題 ($N = 2$)

ここで i は、ベクトルの i 番目の要素を表す. NMax 問題では N 個の最適解 $\mathbf{p}^1, \dots, \mathbf{p}^N$ が存在すると定義する. NMax 問題を, 任意の男女の相性の良さを返す関数として, 以下のように与える.

$$f_{NMax}(\mathbf{x}) = \max_{1 \leq j \leq N} (\text{lovematch}(\mathbf{x}, \mathbf{p}^j)) \quad (4)$$

j は NMax 問題の N 個のピークのインデックスである. ビット列 $\mathbf{x}$ と, 各ピークのビット列の間で一致率が計算され, その最大値が f_{NMax} の出力値となる. NMax 問題は, 進化計算のベンチマークとしてよく知られる OneMax 問題 [42] を一般化したもので, $N = 1$, $\mathbf{p}^1 = (1, 1, \dots, 1)$ の時, OneMax 問題と同値となる. $N = 2$ の場合 (2Max 問題) の最適解の例を図 5 に示す.

3.4.3 評価基準: 合コン割当評価値

実験では, 好相性の男女を同じ合コンに割当てることができたかどうかを示すための指標として合コン割当評価値を用意する.

提案システムへの入力として男女の利用者 (属性) 集合を与えた時, 任意の合コンの評価値を返す関数 $F(B_i, G_i, \mathbf{P})$ を, 以下のように与える

$$F(B_i, G_i, \mathbf{P}) = \sum_{\mathbf{b} \in B_i} \sum_{\mathbf{g} \in G_i} f_{NMax}(\mathbf{x}, \mathbf{P}) \quad (5)$$

これを合コン割当問題の目的関数 (式 (1)) に代入した時の出力値を合コン割当評価値とする. この評価値は, カップルに成り易い男女をうまく同じ合コンに割当てることが出来たかどうかを表す.

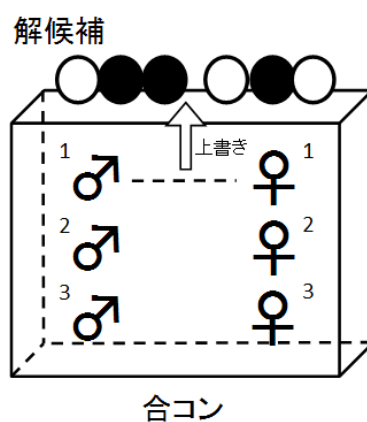


図 6 貪欲手法

3.4.4 比較手法

貪欲手法:

提案手法（以後，提案システム内の進化計算オペレータ部位を指す）と比較するため，次のような貪欲な手法を考えた．これは，図 6 に示すように，合コンの度に，その合コンでカップル成立した男女の属性ベクトルの中で，その解候補に最も類似する属性ベクトルを，解候補に上書きする方法である．良い組合せが見つかった場合，それがそのまま解として次世代に受け継がれる単純な探索手法である．

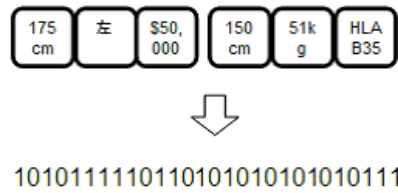


図 7 入力データの簡単化

神託手法:

神託手法は、 N Max 問題の解集団を、0 世代に一回限り解候補集団に与える手法である。この場合、各解候補は、 N Max 問題の各解にそれぞれ対応する。神託手法は、貪欲手法に比べて、高速かつ簡単に良い合コン割当てを実現する。しかし、この割当てが最適であることを保証するためには、合コン割当てオペレータが、最適な合コン割当てを算出できる必要があるが、本論文では、実現の容易性を優先し、単純な割当て方 (3.3.4 項) を採用している。最適な合コン割当ての実現は今後の課題とする。

3.4.5 実験設定

本実験では、まず、入力データの簡単化およびビット列への符号化を行った。概要を図 7 に示す。入力であるシステム利用者の属性情報をビット列で与える。既存の結婚情報サービスの利用者情報は全てビット列に符号化可能である。本実験では、相性の判定に必要な情報がビット列化できることを前提に、アルゴリズムがうまく動作するかを調べる。ただし、情報の種類によっては符号化によってビット数が増大するものもあり、これらの短縮には、相性に影響する要因の特定と 3.3.2 項で述べたような工夫が必要である。実験パラメータの設定値を表 3 に示す。カップル成立閾値は、解候補長に対してのハミング距離の割合で計算するものとする。カップル成立有無は、この値を下回る場合に成立、上回る場合に不成立というステップ関数で決まる設定とした。

表 2 設定値 (3 章の実験)

合コン	
利用者集団サイズ	120 名 ($ B = 60, G = 60$)
利用者属性長 ($ b_{ik} , g_{jl} $)	5 ビット
合コンサイズ	20 名 ($ B_i = 10, G_i = 10$)
カップル成立閾値	0.2
進化計算	
解集団サイズ	20 個
解長	10 ビット
選択	ルーレット選択
交叉率	0
突然変異率	0.2
スケーリング	2 : 1 線形

3.4.6 結果

合コン割当評価値を用い、提案手法（図内実線、**Proposed**）と貪欲手法（図内ダッシュ線、**Greedy**）の比較を、 $N = 2$ の場合について行った。横軸は、進化計算の世代（図内、**Generation**）であり、縦軸は、合コン割当評価値（全合コンのカップル成立総数）である。グラフは、10 試行の平均と標準偏差（エラーバー）で表示した。2Max 問題において、貪欲手法に対して提案手法が有意に高い合コン割当評価値を記録することを確認した。しかし、 $N \geq 3$ の場合、提案手法の有効性を確認することが出来なかった（付録 A 参照）。

3.4.7 考察

$N \geq 3$ の時、良好な結果が得られなかった理由について、幾つか理由が考えられる。1 つ目は、ある世代で良い解が見つかったも、次の世代でそれが上書きされて消えてしまい、探索が滞っている可能性である。この現象は実際に確認でき

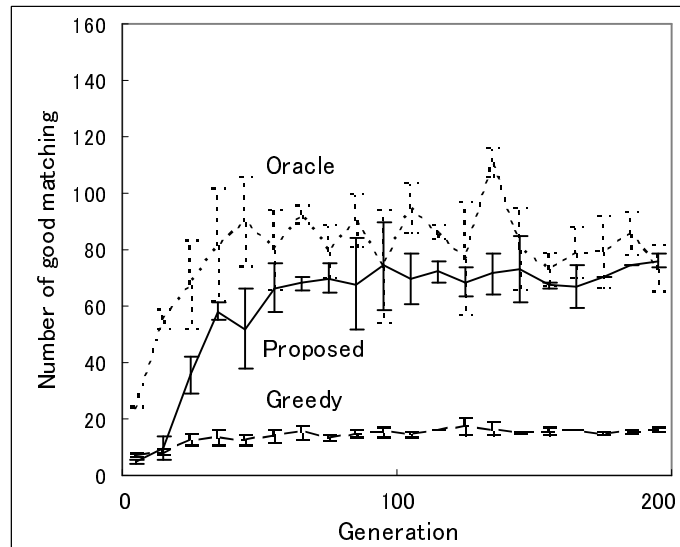


図 8 合コン割当評価値の推移 (2Max 問題)

た．2つ目は，解集団の多様性が保たれておらず，特定のピークしか検出されていない可能性である．3つ目は，割当オペレータにおける，解候補と割り当てられた利用者の属性値間の距離の存在である．なぜなら，本研究では利用者の属性値で相性が決まるという仮定を置いていおり，解候補が，カップル成立した男女の属性値と距離が離れていれば，その結果（カップル成立）が解候補の良さとは異なる（離れる）と考えられるからである．これらの3つの問題の解決が必要である．

3.5 まとめ

本章では，大人数の男女集団の中から好相性の男女グループ（合コン）を多数抽出するための合コン割当問題を定義した．またこの問題を進化計算を用いて解くシステムを提案し，計算機実験による評価を行った．その結果，ベンチマークテスト 2Max 問題において，提案手法は，比較的手法に比べて，有意に相性の良い男女を同じグループに配置できることが分かった．

本章における提案手法の課題は、解候補を評価する際、実際の合コン参加者による評価をなるべく減らすことである。なぜなら、理想的な男女の属性値の組（属性ベクトルと呼ぶ）が求まるまでは最適な合コンへの割り当てが行われないため、参加者はカップル成立数の面で不利になるからである。利用者の時間的、経済的負担、不利益享受を減らすため、できるだけ利用者を介さずに、解候補の評価を行う必要がある。

4. 複数の相性ピークの検出と参加者利用履歴の活用

4.1 概要

3章では、大人数の男女集団の中から相性の高い男女をグループ（合コン）を多数抽出するための合コン割当問題を定義し、その問題を進化計算を用いて解くシステムを提案し、計算機実験による評価を行った。その結果、ベンチマークテスト 2Max 問題において、提案手法は、比較手法に比べて、有意に相性の良い男女を同じグループに配置できることが分かった。しかし、 $N \geq 3$ の場合、良好な結果が得られなかった。

その原因は、解集団の多様性が保たれていないことである可能性が高い。相性空間（男女の属性の組が取り得る値の範囲）は多峰性であり、多様なカップルを成立させるためには、EC によって、この多くの峰（多様な解）を求める必要がある。本章では、相性空間のような多峰性関数の複数極大値算出手法として、類似の解が増えないように工夫された EC の選択法を導入する（4.2.2 項）。

本研究では利用者の属性値で相性が決まるという仮定を置いており、解候補が、カップル成立した男女の属性値と距離が離れていれば、その結果（カップル成立）が解候補の良さとは異なる（離れる）と考えられる。したがって、割当オペレータにおける、解候補と割り当てられた利用者の属性値間の距離の存在が、もう一つの原因である可能性がある。そこで、解候補と利用者属性値間の距離の指標として、**割当一致率**（4.2.3 項）を導入する。

本章では、割当一致率を活用し、合コンを行わずに、履歴データを使って解候補に評価値を与える**ストック利用評価法**（4.2.4 項）を導入する。3章においては、解候補を評価する際、実際の合コン参加者による評価をなるべく減らす必要があることが分かったが、理想的な男女の属性値の組（属性ベクトルと呼ぶ）が求まるまでは最適な合コンへの割り当てが行われないため、参加者はカップル成立数の面で不利になるからである（すなわち、参加者が「人柱」的になる）。

本章で対象とする合コン割当問題（3.2 節）は、前章と同じであるため説明を割愛する。ベンチマークテスト（3.4.2 項）に関しても、現実世界に近づけて一部変更したが、前章との重複部分は割愛する。

4.2 提案システム

4.2.1 概要

全体構成は，前章，図 1 とほぼ同じであるが，EC 中の選択方法や合コン評価オペレータ，履歴データの利用など，変更や追加を行った部分がある．それらを以下，順に説明する．

4.2.2 提案選択法

本章では，EC の選択法として，解集団の多様性を保つ手法の一つとして知られる Multi-Niches Crowding (MNC) 法 [40] を参考にした選択法（以下，**提案選択法**と呼ぶ）を提案する．MNC 法は， t 世代解集団 S^t から $t+1$ 世代解集団 S^{t+1} を生成する際，類似の解を増やさないようにする選択法である．同選択法は，従来，1 つの大域的最適解の探索問題に応用されてきたが，提案選択法では，局所最適解も含めた複数解の探索に活用する点で新規性がある．また，それぞれの解に対応する峰で，山登りのように解を改善していく必要があるため，提案選択法では，以下の Step 4.b においてエリート保存戦略の要素を取り入れ，高評価値の方のみを残すものとした．この部分は，差分進化 [41] の目標ベクトルと比較参照ベクトルを比べて良い方を残すという戦略にも通じる．

図 9 に示すように，まず， S^t から m 個 ($m \leq |S^t|$) 重複なしでランダムに選び出した集団 S^{*t} を生成する． t 世代解集団 S^t からルーレット戦略で n 個 ($n \leq |S^t|$) 選び出し，進化計算（一点交叉，突然変異）を行い，合コン参加者を割当て，評価を行う．これにより，評価値付きの解候補集団 D が生成される ($|D| = n$)．次に，全解候補 $d \in D$ と距離（両者の要素の一致率）が最短の $s^* \in S^{*t}$ を探し， d と s^* の評価値を比較する．評価値が大きい方を $t+1$ 世代解集団 S^{t+1} に代入する．以下に具体的なステップを示す．このステップは図 9 に対応している．

Step 1. S^{t+1} を空の集団を代入

Step 2. S^t からランダムかつ重複なしに m 個の要素を選んで S^{*t} を生成

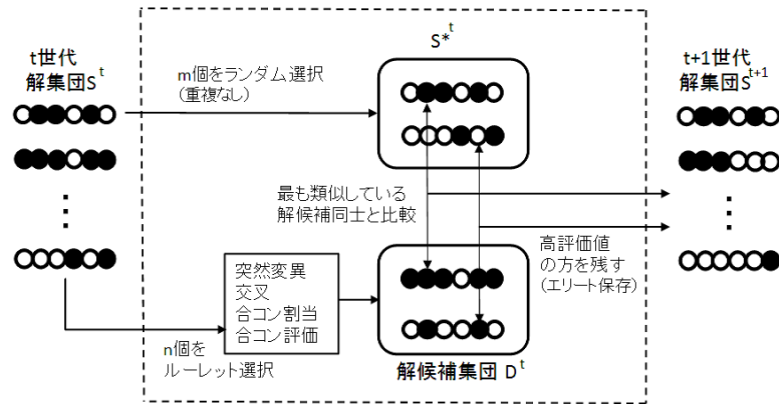


図 9 提案選択法の全体像

Step 3. S^t に対しルーレット選択，交叉，突然変異を実行し， n 個の要素から成る D を生成

Step 4. 全ての $d \in D$ に対して以下を反復

Step 4.a. d と最も距離が近い $s^* \in S^{*t}$ を探索

Step 4.b. d と s^* の評価値を比較し， d の評価値の方が大きければ S^{t+1} に d を追加．そうでなければ， S^{t+1} に s^* を追加

Step 5. t が規定回数に到達していれば終了．そうでなければ Step 1 へ

m と n は， $|S^t|$ の 8~10 割程度の大きさとする．

通常の MNC 法では，上記ステップ 2 を複数回繰り返し，その回数分のグループを生成する．またステップ 3 においては $n = 1$ とし，交叉のみしか行わない．ステップ 4.b では，ステップ 2 で生成したグループ毎に，その中で最も d と距離が近い要素を 1 個ずつ選び出し，その選び出した要素集団中，評価値が最低の要素を d で上書きする．

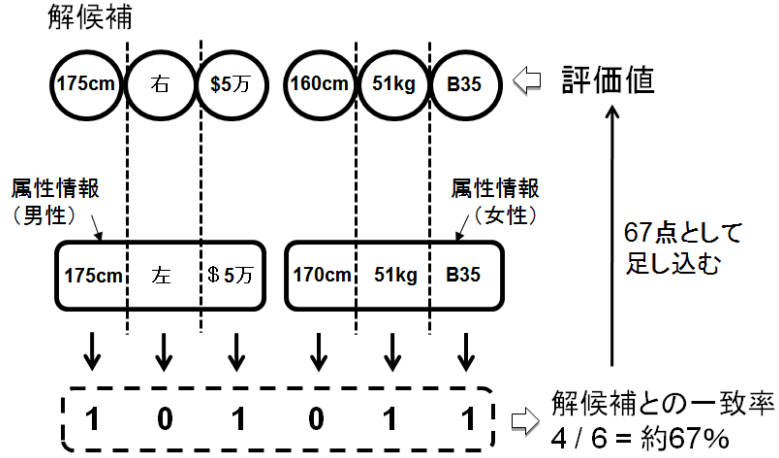


図 10 割当一致率

4.2.3 割当一致率を考慮した解候補評価法

解候補をベクトル $\mathbf{s} = (s_1, \dots, s_{k+l})$ で表わす. 男女の組 (b, g) を, b の属性情報 $(b_1, b_2, \dots, b_k)$ と g の属性情報 $(g_1, g_2, \dots, g_l)$ を組み合わせたベクトル $\mathbf{x} = (b_1, b_2, \dots, b_k, g_1, g_2, \dots, g_l) = (x_1, x_2, \dots, x_{k+l})$ で表わす. このとき, $\mathbf{s}$ と $\mathbf{x}$ の割当一致率 $C(\mathbf{s}, \mathbf{x})$ を次のように与える.

$$C(\mathbf{s}, \mathbf{x}) = \frac{\sum_{i=1}^{k+l} \text{match}(s_i, x_i)}{k+l} \quad (6)$$

ここで,

$$\text{match}(s_i, x_i) = \begin{cases} 1 & (s_i = x_i) \\ 0 & (\text{otherwise}) \end{cases} \quad (7)$$

図 10 に示すように, その合コンで成立した全カップルの属性情報ベクトルとテスト解候補との割当一致率の総和を, 解候補の評価値として与える. カップル成立数が 0 の場合は評価値を 0 とする.

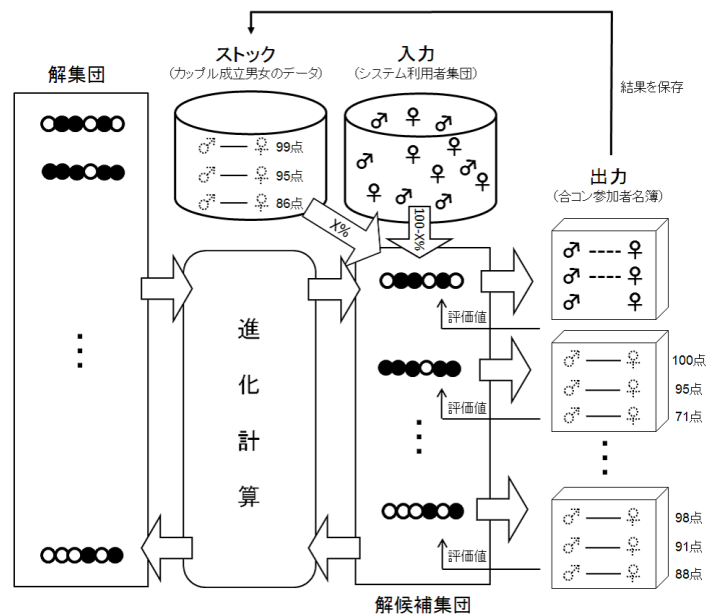


図 11 ストックの導入

4.2.4 ストック利用解候補評価法

解候補を評価する際、実際の合コン参加者による評価をなるべく減らしたい。なぜなら、理想的な男女の属性値の組（属性ベクトルと呼ぶ）が求まるまでは、参加者が人柱（1章参照）的になるため、利用者に時間的、経済的負担を与えるからである。参加者を用いずに、評価履歴データを利用して、解候補の評価を行う方法を導入する。

図 11 に示すように、この方法では、まず、カップルが成立した男女の属性ベクトルを保存する。この集団を**ストック**と呼ぶ。図 11 にストックを付け足したシステム全体像を示す。これ以降、提案システムとは、図 1 ではなく、図 11 を指すものとする。

解候補を評価する際、まず、合コン割当オペレータの全適用対象の $X\%$ をストックから取り出して割当てる（システム利用者集団からの割当ては $100\% - X\%$ ）。 X は、ストックが十分蓄積された段階で $50\sim 100\%$ の間に設定する。

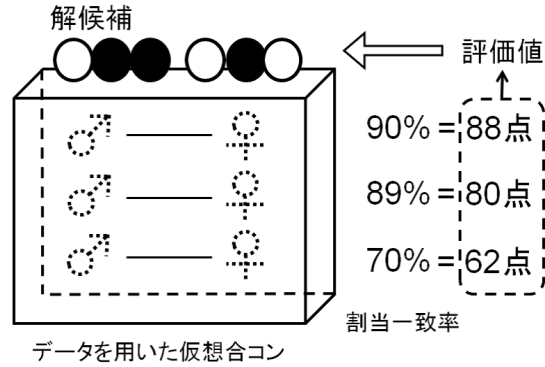


図 12 ストックを利用した解候補への評価値の与え方

解候補に距離が近い属性ベクトルを，距離が近い順に規定人数割当てする．これより，図 12 に示すように，解候補と属性ベクトルの割当一致率を計算し，その総和を解候補の評価値として与える．ストックから解候補 s に距離が近い順に L 個取り出したストック男女の集団を $\mathbf{O}$ ，ストック男女を $o \in \mathbf{O}$ ， o の s に対する割当一致率 $C(s, o)$ として， s の評価値 E_s の計算式を式 (8) に示す．

$$E_s = \sum_{o \in \mathbf{O}} C(s, o) \quad (8)$$

4.3 実験

4.3.1 概要

提案システムの有効性を確かめるため 3 つの計算機実験を行った．各実験においては，合コン割当て評価値 (3.4.3 項) に加え，最適解集団にどれだけ近い解集団を求めることができるか (最適解到達率，4.3.3 項に後述) を調べた．実験 2 では，貪欲手法 (3.4.4 項) との比較を行い，実験 3 では，利用者情報量と問題サイズを前章より増やし，提案システムの問題に対するスケーラビリティを調べた．

実験 1： 解集団の多様性を保つ提案選択法 (4.2.2 項) の有無

実験 2： スtock利用法 (4.2.4 項) の有無

実験 3： 入力情報量 (3.3.2 項) と問題 (3.4.2 項) サイズの変更

4.3.2 ベンチマークテスト (NMax 問題) の改良

本実験では、カップル成否を返す評価関数として NMax 問題 (3.4.2 項) を与えたが、より現実世界に近いものにするため、この評価関数に、確率的な要素を取り込む。なぜなら、実際の合コンでは、相性が良くてもカップルに成らない場合や、多少悪くても成る場合があるからである。そこで、カップル成立有無の評価関数で使用するカップル成立確率 P を、 f_{NMax} を用い、式 (9) として与える。

$$P = \begin{cases} 1 - \frac{1-f_{NMax}}{1-\beta} & (\beta \leq f_{NMax}) \\ 0 & (otherwise) \end{cases} \quad (9)$$

β はカップル成立の閾値であり、 f_{NMax} の出力が、 β 未満の時はカップ成立確率 P は 0 になり、 β 以上の時は P が線形に増加するように設定した。

4.3.3 評価基準: 最適解到達率

得られた解集団を評価するため、解が、最適解にどれくらい到達できているかを示す指標として最適解到達率を定義する。解 $\mathbf{s}$ の最適解到達率 R は、最適解集団 $\mathbf{p}$ (図 13 に示す) と、式 (2) の *lovematch* 関数を用いて、以下の式で与える。

$$R = \max_{1 \leq i \leq n} (\text{lovematch}(\mathbf{s}, \mathbf{p}^i)) \quad (10)$$

獲得した解集団がどれくらい相性空間を推定できているかは、この解集団の解の最適解到達率の平均値によって確認することができる。

4.3.4 比較手法

貪欲手法は前 3 章に示した手法と同じものを用いる。神託手法に関しては、前 3 章では、最適解集団を 0 世代に一回限り解候補集団に与えるものであったが、本章では、それを毎世代与え続ける方法を採用する (付録 A 参照)。その方が、より最適に近い合コン割当を出力するため比較対象として理想的だからである。

表 3 設定値 (4 章の実験)

合コン	
利用者集団サイズ	6000 名 ($ B = 300, G = 300$)
利用者属性長 ($ b_{ik} , g_{jl} $)	10 ビット
利用者調整理想度*	0~100% (一様乱数)
合コンサイズ	30 名 ($ B_i = 15, G_i = 15$)
ストックサイズ	500 レコード
カップル成立閾値	0.5
最低参加回数*	5 回
進化計算	
解集団サイズ	30 個
解長	20 ビット
選択個数 m	10 個
交叉率	0.95
突然変異率	0.2
スケーリング	1:1 線形

4.3.5 実験設定

本実験パラメータの設定値を表 3 に示す。以下に、各設定値の詳細について述べる。相性を決める情報量（本論文では属性長）がどれくらいかという点については、3.3.2 項で挙げた情報に加え、既存の結婚紹介サービスが使っている血液型 (2 ビット)、長男長女かどうか (1 ビット)、家が近いか (1 ビット) などの情報も必要かもしれない。3.3.2 項で挙げた情報とこれらを合わせると 11 ビットである。相性に影響する要因を正確に突き止めれば、必要な情報量は男女合わせて 20 ビット程度に収めることができると予想している。したがって、実験 1 と 2 では、男女の属性長それぞれ 10 ビット（解長は 20 ビット）とした。

本実験では、図 13 の 5 つの好相性に近い男女が、式 (9) に従いカップル成立す

るという設定で実験を行った。そのため、利用者の属性値を一様乱数で生成すると、カップル成立数が極端に減り（付録 B 参照）、探索が進まない。したがって、本実験では、図 13 の N 個の各ピークとの類似度（利用者調整理想度）がある程度高くなるよう利用者の属性情報を生成することにした。NMax 問題のピークを利用者の属性値にコピーし、ランダムに 1~100 の間から数字 θ を選び、コピーした利用者の属性値 $\theta\%$ をビット反転させることで生成した。これを各ピークに対して順番に、利用者集団が埋まるまで繰り返した。

また、カップル成立した利用者を逐次、新しい利用者と入れ替える設定とした。ただし、ただし、入れ替えはその利用者が保持する参加回数が**最低参加回数**を越えた場合のみとした。新しい利用者も利用者調整理想度のルールに従って生成する。参加回数は、全利用者がそれぞれ保持し、初期値は 0、合コン参加によりカップルが成立した場合に 1 ずつ加算される。参加回数が最低参加回数を越えると、利用者は利用者集団から排除される。実際のお見合いパーティでも、カップリングが成功（カップリングゲームで両思いが成立）しても、もっと良い相手を探して、何回もパーティに参加し続ける人が多い（最初から何回か参加することを前提としている人も多い）。

合コン割当問題を、EC 世代毎に繰り返し解くことを、**繰り返し合コン割当問題**と定義する。繰り返し合コン割当問題は、世代毎のカップル成立の累積値を最大化することが目的関数となる。 F の正確な推定（高い最適解達成率に到達）は、合コン割当問題はもちろん繰り返し合コン割当問題の解決に貢献する。実験 3 では、利用者属性長は 3.3.2 項での考察に基づき、3 章の 2 倍の長さ（解長も倍となる）までを調べた。同じく実験 3 の N に関しては 12 個までとした。 N の値は、3.3.2 項で述べた性格パターン数の 2 乗（男女の組合せのため掛け算）で 25 程度を考えても良いが、組合せの中には、好相性ばかりではなく、相性が良くない組合せも存在すると考えられる。したがって 25 の約半分の値とした。

4.3.6 結果

図 14 と図 15 に、実験 1 の結果として、提案選択法を用いた場合と用いなかった場合の、1,000EC 世代後に得られた解集団の平均最適解到達率 $\bar{R}$ を示す。初期

```

Peak1 00000000000000000000
Peak2 00000000000111111111
Peak3 111110000001111100000
Peak4 111111111110000000000
Peak5 11111111111111111111

```

図 13 $NMax$ 問題 ($N = 5$)

集団をランダムに変更し 30 回実験を行ったところ、 $\bar{R}$ の値はそれぞれ変わらず、標準偏差値は 1% であった。

11101111111011111111 peak.5 90%	10011011110011101010 peak.4 60%
11111111110000001010 peak.4 90%	11011110010000111000 peak.4 70%
11111111110000000000 peak.4 100%	11101010010001100110 peak.3 60%
11111111111111011111 peak.5 95%	11011100110001111110 peak.5 65%
00000000000001000010 peak.1 90%	11011010110011110010 peak.3 60%
00000000001000000001 peak.1 90%	11111001111001111011 peak.5 75%
10111111111101011111 peak.5 85%	11011110111011101010 peak.5 70%
00000000000010000000 peak.1 95%	00011010010010111010 peak.1 55%
00010000000000000000 peak.1 95%	11011010111000101010 peak.4 65%
11111101110000010000 peak.4 90%	11001010110101010111 peak.5 60%
11111101100100101010 peak.4 80%	11001010110101001011 peak.4 55%
00000000001111111111 peak.2 100%	11111001011001100001 peak.3 75%
10000100000000000000 peak.1 90%	10101000111001010110 peak.1 50%
11111000001111101000 peak.3 95%	11101010011011111010 peak.3 65%
1111111111111110101 peak.5 90%	11100000010001110010 peak.1 60%
00000000000100000000 peak.1 95%	11010010111011100110 peak.3 60%
11010111111000010000 peak.4 85%	11001010011011100011 peak.3 65%
11011111010000001000 peak.4 85%	11111010110001111110 peak.5 70%
00000101000000000000 peak.1 90%	11110001100001110010 peak.3 60%
00100010000100000000 peak.1 85%	11101001011011111011 peak.5 70%

図 14 提案選択法を導入時 ($\bar{R} = 91\%$) 図 15 提案選択法を非導入時 ($\bar{R} = 64\%$)

これにより提案選択法が $\bar{R}$ を約 27% 高めることに貢献していることが分かる。提案選択法が導入されていない図 15 の獲得解集団の例では、5 つのピークのうち 1 つが発見できておらず、また $\bar{R}$ も低い。

実験 2 の結果を表 4 に示す。 $\bar{R}$ 及びカップル成立数は、合コン回数 10,000 回

表 4 提案手法と貪欲手法の比較

	最適解到達率 $\bar{R}$	カップル成立数	実合コン回数
神託手法	100 $\pm$ 0%	4494 $\pm$ 6 組	0 回
提案手法 (ス※ 0%)	91 $\pm$ 1 %	3440 $\pm$ 112 組	10,000 回
提案手法 (ス※約 50%)	90 $\pm$ 1 %	3358 $\pm$ 132 組	5,250 回
提案手法 (ス※約 50%)	91 $\pm$ 1 %	3457 $\pm$ 125 組	10,000 回
貪欲手法	64 $\pm$ 2 %	1456 $\pm$ 106 組	10,000 回

※ス: スtock利用率

(1,000 世代) 後に得られた解集団における値と、その解集団を用いて合コンを行った場合で得られた値であり、初期集団をランダムに変更し行った 30 回の試行の平均値と標準偏差値で示している。しかし、表内 3 行目はストックを 50% 利用した場合で、実際の合コン実施回数は約半分の 5,250 回後の値となっている。表内 4 行目も 50% スtockを利用した場合であるが、実際の合コン回数が 1,000 回に達するまで (1,950 世代) 実験を行った結果である。カップル成立数は、それぞれ得られた解集団に対し合コン割当オペレータを適用して得られた値である。これらの結果より、提案システムが、比較手法に比べて、約 22% 以上高い最適解到達率を求めることができ、約 2 倍のカップル成立数が多い合コンの実施に、約半分の合コン回数で至ることが確認できた。

実験 3 の結果を図 16, 図 17 に示す。いずれも縦軸に $\bar{R}$ の平均、横軸はそれぞれ利用者属性長と $NMax$ 問題の N である。実験 2 と同じく初期集団をランダムに変更した 30 回の試行の平均値をプロットした。標準偏差値は実験 3 の結果の全ての値において 1 であった。これは図中には記載していない。

これより、提案手法が、現在の倍のスケールの入力情報量に対し、 $\bar{R}$ を約 10% しか劣化させないことが分かった。また、 N の値が 2 倍以上になっても、ほぼ $\bar{R}$ が変わらないことも分かった。ただし、図 17 においては $N = 8$ の場合、2 回 (30 回中) の試行で 1 個 (8 個中) のピークが未発見、 $N = 12$ の場合、14 回 (30 回中) の試行で 1 個 (12 個中) のピークが未発見であった。

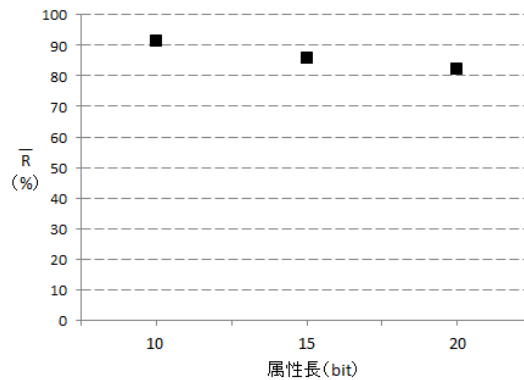


図 16 利用者属性長に対するスケーラビリティ($N = 5$ の時)

4.3.7 考察

提案手法が、最適解集団を与えた場合のカップル成立に及ばなかった理由としては、カップル成否の評価関数を確率で与えているため、解候補の評価値が変動し、たまたま良い値をつけた解候補が解集団に残ることが考えられる。とくに、カップル成立閾値が小さい場合、この問題が顕著になるはずである（偶然、カップルになる数が増加するから）。

また、利用者調整理想度が低く、最適解から遠い属性を持つ利用者ばかりの場合も、カップルが成立し難いことに加え、成立しても偶然の確率が高いため、最適解到達率は下がる。カップル成立確率をどのように与えるかによっても変る。

ストックの利用方法の改善について、今後、検討が必要である。今回の実験では、ストック利用の有無で結果に差は出ていないが、実世界の問題において、ストックに保存されていないような新しい好相性の解候補が出てきた場合、正しい評価値を与えることができない（類似の属性ベクトルを割当てることができないため）可能性がある。したがって、ストックに保存する属性ベクトル群の多様性を保つ工夫が求められる。また、このストックに、過去蓄積されてきたデータを初期データとして入れることができれば、収束までの時間も短縮できる。

提案手法を、より大きなスケールで有効に作用させるためには、より豊富で大きなシステム利用者集団と、より大きな解集団サイズが必要と考えられる。

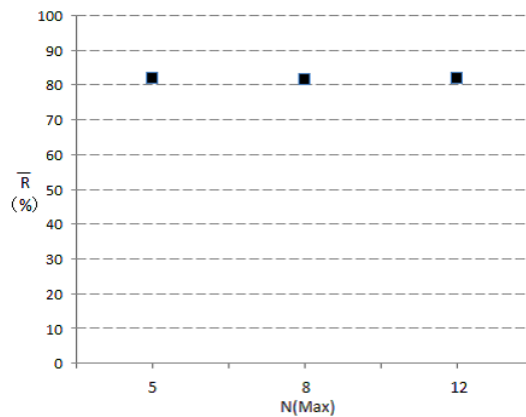


図 17 問題サイズ N に対するスケーラビリティ(利用者属性長 = 20 の時)

4.4 まとめ

本章では、実際に実施されている合コンセッティングサービスで、ECによって複数の未知の相性を予測し、カップル成立数を増やすシステムを提案した。計算機実験を用いて、最適解に平均 90%以上一致する解集団を求めることができた。また、一般的に考え得る単純な貪欲手法に比べて、半分の合コン実施数で、約 2 倍のカップルを誕生させることが可能なことを示した。

提案手法の高精度化、高速化には、解集団の多様性を保つ手法として有名なシェアリング [44] や資源共有法 [45] 等の導入が必要な可能性がある。また、属性情報に欠損がある（利用者がアンケートの一部に未回答）場合、それを如何にエンコードするのも課題である。感性検索エージェントの研究において用いられている、「嫌い」を N (Negative), 「好き」を P (Positive), 「どちらでもない」を D (Don't care) とする NPD 表現 [46] 等を参考に、「不明」という情報をそのままエンコードする手法の有効性を調べる必要がある。さらに、評価関数 F が変化した場合の適応性の強化も課題である。

5. 結論

本論文では、実際に実施されている合コンセティングサービスで、ECによって複数の未知の相性を予測し、カップル成立数を増やすシステムを提案した。計算機実験を用いて、最適解に平均 90% 一致する解集団を求めることができた。また、一般的に考え得る単純な貪欲手法に比べて、半分の合コン実施数で、約 2 倍のカップルを誕生させることが可能なことを示した。

「合コン」問題のような非線形問題には人工ニューラルネットワーク (NN, Neural Network) の適用が考えられるが、NN の素子数やネットワークトポロジの決定は難しい問題であると共に、学習後の NN から好相性を理解することが困難である場合が多い [47]。そのため今回は採用を見合わせた。他方、進化計算であれば、解集合に対してクラスタリング手法を適用したり、統計的解析をかけることで、好相性を抽出できる可能性が高い。筆者は NN のネットワークトポロジを進化計算によって最適化する NEAT [48] 分野での研究を行っており [49]、今後はこれを「合コン」問題に適用したいと考えている。

主要な今後の課題は、実用までのギャップの解消である。今回の結果を得るには 5,000 から 10,000 回の合コンの実施が必要であるが、この回数は、既存の合コンセティングサービスにおいては現実的な回数である。たとえばツヴァイ社 [5] は年間約 2,000 回、日本サプライズ社 [6] は年間約 1,200 回の合コンを実施しており、年間約 15,000 回の実施回数を報告している業者 [7] も存在する。これら数社のデータを統合することができれば、数ヶ月から数年というオーダーでの提案システムの適用が現実的となってくる。

また、相性以外のカップル成立要因にも目を向けることも重要である。欧州では、数世紀前から「お見合いおばさん」(イエンタ) と呼ばれる結婚仲介人が重要な役割を担ってきた。結婚したい男女とその親を顧客とし、入手可能な情報 (たとえば、家族関係、年齢、居住地区、収入、体重、身長、学歴、写真、目と髪の毛の色、価値観、性格、生活スタイル、恋愛観、結婚相手に求めるもの、食習慣等) を聞き出し、その情報を元につり合いそうな相手を探し出し、花嫁花婿候補とその家族を引き合わせるのが仕事である。今日、欧米社会にイエンタは見られないし、また、見合い結婚もほとんど行われていない [50]。しかし、我が国では、

お見合いおばさんが、今でも継続的に活躍している。

3.3.2 項で提起したように、相性の決定に必要なシステム利用者の属性情報は未だよく分かっていない。また、たとえ分かったとしても、たとえば、話し方、ユーモアのセンス、美醜に関わる判断基準、目の輝き、手の動かし方、笑い方等、エンコードが困難な属性も多い。お見合いおばさんは、このような属性情報を、経験的に把握して活用している可能性がある。オンライン・デーティング・システムにおいて、実際に相手とデートする前に、バーチャル環境（“Chat circles”，MIT メディアラボの Viegas と Donath が開発）[51] 上で事前にデートすることが、そのデートでの相手を気に入る確率を、それをしない場合の2倍に高めるという研究がある[52]。こうしたツールも、オンライン上で会話しながら相手の反応を観察することを通して、エンコードされなかった情報を把握することに役立つと思われる。これと本論文で提案した手法を組み合わせることで、より高いカップル成立数を実現できる可能性が高い。

本研究のようなデータと計算機を用いた相性予測に関しては、近年、米大手オンライン・デーティング・サービスである eHarmony 社が、重回帰分析等を用いた手法を考案、活用している[53]。また、上記の MIT グループの研究を引き継ぐ形で、カルフォルニア大学バークレイ校 Fiore（Facebook 社データ・サイエンティスト兼任）が、オンライン・デーティング・システムにおける利用者の行動についての研究を進めている[54, 55, 56]。我が国の研究のキャッチアップが望まれる。利用者の選好を予測し、それに基づいて利用者同士を適切に結びつけるシステムは、デーティング・サービスだけでなく、たとえば、不動産市場（買いたい人と売りたい人のマッチング）、推薦入学市場（進学希望者と学校のマッチング）にも応用可能であり、研究意義のある分野であると言える。ただし、合コンでもそうであるが、マッチングに際しては2章で紹介した安定結婚問題等が発生する。これらを同時に解決するシステムが今後は求められる。

謝辞

指導教官である伊藤実 教授には，快適な研究環境を与えて頂くとともに，会社経営と二足のわらじをはいたことによる，長期に渡った在学を許容して下さいましたことを，ここに深く感謝致します。

安本慶一 教授には，本研究の副指導教官として，本論文の作成および研究にわたり有益な情報や御助言を賜わり，その御指導は幾度となく深夜に及びました。ここに深く感謝致します。

杉本謙二 教授には，本研究の副指導教官として，中間発表時等に有益な情報や御助言を賜りました。ここに深く感謝致します。

村田佳洋 准教授（現在，広島市立大学）には，特に本論文の研究の萌芽部にご指導，ご尽力頂きました。心より感謝致します。

柴田直樹 准教授，木谷友哉助教（現在，静岡大学）には，本研究を進めるにあたり，多くの御助言を頂きました。ここに深く感謝致します。

また，研究室の生活において手助けして頂いた孫為華 助教に深く感謝します。

尾川恵理前事務補佐員，金岡恵事務補佐員には，研究費管理などの事務手続き等をお世話して頂きました。ここに深く感謝致します。

研究室の皆様（とくに武兵君）には，日頃から公私に渡りお世話になりました。ここに深く感謝致します。

最後に，研究生生活を支えて下さった家族，友人に感謝します。

参考文献

- [1] 国立社会保障・人口問題研究所：人口統計資料集 (2011),
<http://www.ipss.go.jp/syoushika/>
- [2] 内閣府少子化対策室：子ども・子育て白書 (2011).
<http://www8.cao.go.jp/shoushi/whitepaper/index-w.html>
- [3] 内閣府：少子化社会白書 (2004).
<http://www8.cao.go.jp/shoushi/whitepaper/w-2004/html-g/indexg.html>
- [4] 奈良県こども・子育て応援県民会議: なら出会いセンター
<http://www.naradeai.pref.nara.jp/>
- [5] 株式会社ツヴァイ
<http://www.zwei.com/>
- [6] 株式会社日本サプライズ社.
<http://www.surprise-japan.co.jp/>
- [7] 株式会社エクシオジャパン
<http://www.exeojapan.com/>
- [8] Foster, J.A.: “Evolutionary Computation,” *Nature Rev. Genet.*, Vol. 2, pp. 428–436, (2001).
- [9] Cagnoni, S., Poli, R., Smith, G.D., Corne, D., Oates, M., Hart, E., Lanzi, P.L., Willem, E.J., Li, Y., Paechter, B. and Fogarty, T.C.(Eds.): “Real-World Applications of Evolutionary Computing,” *Lecture Notes in Computer Science*, Vol. 1803, (2000).
- [10] 伊藤一: “N700 系新幹線車両の概要～遺伝的アルゴリズムを用いた先頭車両形状設計～,” 計測自動制御学会北陸支部講演会, (2009).

- [11] Takagi, H.: “Interactive Evolutionary Computation: Fusion of the Capacities of EC Optimization and Human Evaluation,” *Proc. of IEEE*, Vol. 89, No. 9, pp. 1275–1296, (2001).
- [12] Takagi, H. and Ohsaki, M.: “Interactive Evolutionary Computation-Based Hearing-Aid Fitting,” *IEEE Trans. on Evolutionary Computation*, Vol. 11, No. 3, pp. 414–427, (2007).
- [13] Tokui, N. and Iba, H.: “Music composition with interactive evolutionary computation,” *Proc. of the 3rd Annual Int’l Conf. on Generative Art*, pp. 215–226, (2000).
- [14] Tanji, M. and Iba, H.: “ConBreO: a music performance rendering system using hybrid approach of EC and automated evolution,” *Proc. of GECCO*, pp. 1275–1282, (2010).
- [15] 佐藤嘉洋, 栢菅彩, 有田隆也: “適応度予測に基づく対話型進化計算とその似顔絵生成への応用,” 第32回知能システムシンポジウム論文集, pp. 199–204, (2005).
- [16] Deb, K., Agrawal, S., Pratab, S. and Meyarivan, T.: “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *Trans. on IEEE Evolutionary Computation*, Vol. 6, No. 2, pp. 182–197, (2002).
- [17] Gale, D. and Shapley, L.S.: “College admissions and the stability of marriage,” *Amer. Math. Monthly*, Vol. 69, pp. 9–15, (1962).
- [18] Nakamura, M. Onaga, K. Kyan, S. and Silva, M.: “Genetic algorithm for sex-fair stable marriage problem,” *Proc. of IEEE Int’l Symp. of Circuits and Systems*, Vol. 1, pp. 509–512, (1995).
- [19] Pathak, P. A., Abdulkadiroglu, A. and Roth, A.: “The New York City High School Match,” *American Economic Review*, Vol. 95(2), pp. 364–367, (2005).

- [20] The Royal Swedish Academy of Sciences: “Stable allocations? from theory to practice,” *Press Release*, (2012).
- [21] 厚生労働省: 新たな医師臨床研修制度
<http://www.mhlw.go.jp/topics/bukyoku/isei/rinsyo/index.html>
- [22] Del, T. and Barbara, G.: “Human assortative mating and genetic equilibrium: An evolutionary perspective,” *Ethology and Sociobiology*, Vol. 1, No. 2, pp. 111–140, (1980).
- [23] Aurelio, J. F., Jon, A. S. and Daniel, N. J.: “The ideal romantic partner personality,” *Personality and individual differences*, Vol. 41, No. 3, pp. 431–441, (2006).
- [24] Csatho, A., Osvath, A., Bicsak, E., Karadi, K., Manning, J. and Kallai, J.: “Sex role identity related to the ratio of second to fourth digit length in women,” *Biological Psychology*, Vol. 62, pp. 147–156, (2003).
- [25] De, S., Pal, S.K. and Ghosh, A.: “Genotypic and phenotypic assortative mating next term in genetic algorithm,” *Information Sciences*, Vol. 105, No. 1–4, pp. 209–226, (1998).
- [26] Ihara, Y. and Marcus, W. F.: “Evolution of disassortative and assortative mating preferences based on imprinting,” *Theoretical Population Biology*, Vol. 64, No. 2, pp. 193–200, (2003).
- [27] Weekind, C. and Furi, S.: “Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity,” *Proc. of Royal Society of London B*, Vol. 264, pp. 1471–1479, (1997).
- [28] Roberts, S. C., Little, AC., Gosling, LM. and Jones, BC., Perrett, DI., Carter, V. and Petrie, M.: “MHC-assortative facial preferences in humans,” , *Biol Lett.*, Vol. 1(4), pp. 400–403, (2005).
- [29] Jacob, S. et al.: “Paternally inherited HLA alleles are associated with women ’ s choice of male odor,” *Nature Genetics*, Vol. 30, pp. 175–179, (2002).

- [30] Phillippe, R. and Ian, R. N.: “Genetic similarity theory, intelligence, and human mate choice,” *Ethology and Sociobiology*, Vol. 9, No. 1, pp. 45–58, (1988).
- [31] 日立ソフトウェアエンジニアリング株式会社: HLA 情報を用いた相性判定システム.
特許公開番号 2007–34519 (2007).
- [32] Love Sciences LLC, US: ScientificMatch.com
<http://www.scientificmatch.com>
- [33] Dempster, A. P. and Laird, N. M. and Rubin, D. B.: “Maximum likelihood from incomplete data via the EM algorithm,” *J. of Royal Statistical Society Series B*, Vol. 39, pp. 1–38, (1977).
- [34] MacQueen, J. B.: “Some methods for classification and analysis of multivariate observation,” *Proc. of Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, (1967).
- [35] Rui Li, Jeroen Eggermont, Michael T. M. Emmerich, Ernst G. P. Bovenkamp, Thomas Back, Jouke Dijkstra, Johan H. C. Reiber: “Towards Dynamic Fitness Based Partitioning for IntraVascular UltraSound Image Analysis,” *EvoWorkshops*, pp. 391–398, (2007).
- [36] Jeroen Eggermont et al.: “Optimizing Computed Tomographic Angiography Image Segmentation Using Fitness Based Partitioning,” *Proc. of Applications of Evolutionary Computing, EvoWorkshops, LNSS*, 4974, pp. 26–28, (2008).
- [37] Julia, H. and Joshua K.: “Evolutionary Multiobjective Clustering,” *Trans. on IEEE Evolutionary Computation*, Vol. 11, No. 1, pp. 1081–1091, (2004).
- [38] Cloninger, C.R., Svrakic, D.M., and Przybeck, T.R.: “A Psychobiological Model of Temperament and Character,” *Archives of General Psychiatry*, Vol. 50, pp. 975–990, (1993).

- [39] 宗像恒次, 田中京子, 小林由美: “SAT 気質コーチングによる人間関係のコラボレーション, ” ヘルスカウンセリング学会年報, Vol. 13, pp. 1–11, (2007).
- [40] Cedeno, W., Vemuri, V. and Slezak, T.: “Multi-Niche Crowding in Genetic Algorithms and its Application to the Assembly of DNA Restriction-Fragments,” *Evolutionary Computation*, Vol. 2, No. 4, pp. 321–345, (1995).
- [41] Storn, R and Price, K.: “Differential Evolution - A Simple and Efficient Heuristic for global Optimization over Continuous Spaces,” *J. Global Optimization*, Vol. 11, pp. 341–359, (1997).
- [42] Schaffer, J.D. and Eshelman, L.J.: “On Crossover as an Evolutionarily Viable Strategy,” *Proc. of Int’l Conf. Genetic Algorithms*, pp. 61–68, (1991).
- [43] Sho Kuroiwa, Yoshihiro Murata, Keiichi Yasumoto, and Minoru Ito: “A Method for Assigning Men and Women with Good Affinity to Matchmaking Parties through Interactive Evolutionary Computation,” *Proc. of Simulated Evolution and Learning (SEAL 2008)*, pp. 645–655, (2008).
- [44] Horn, J., Nafpliotis, N. and Goldberg, D.E.: “A niched Pareto genetic algorithm for multiobjective optimization,” *Proc. of Int’l Conf. Evolutionary Computation*, pp. 82–87, (1994).
- [45] Juill, H. and Pollack, J. B.: “Coevolutionary Learning: a Case Study,” *Proc. of Int’l Conf. Machine Learning*, pp. 251–259, (1998).
- [46] 袴田隼毅, 徳美雄大, 徳丸正孝: “対話型進化計算を用いたレコメンドシステムフレームワーク, ” 日本感性工学会論文誌, Vol. 11, No. 2, pp. 281–288, (2012).
- [47] 石川眞澄: “ソフトコンピューティングと情報統合, ” システム制御情報学会誌, Vol. 43, No. 4, pp. 174–179, (1999).
- [48] Stanley, K. and Miikkulainen, R.: “Evolving Neural Networks Through Augmenting Topologies,” *Evolutionary Computation*, Vol. 10, pp. 99–127, (2002).

- [49] 黒岩将, 吉田久: “コンピュータシミュレーションによる NGL 獲得構造 NN の学習能力に関する比較,” 電子情報通信学会総合大会, (2006).
- [50] Ariely, D.: “The Upside of Irrationality,” p. 352, *HarperCollins*, (2010).
- [51] Viegas, F. and Donath, J. S.: “Chat Circles,” *Proc. of Int’l Conf. Human Factors in Computing Systems*, pp. 9–16, (1999).
- [52] Frost, J. H., Chance, Z., Norton, M. I. and Ariely, D.: “People are experience goods: improving online dating with virtual dates,” *J. of Interactive Marketing*, Vol. 22, No. 1, pp. 51–61, (2008).
- [53] Carter, S. and Snow, C.: “Helping Singles Enter Better Marriages Using Predictive Models of Marital Success,” *Annual Convention of the American Psychological Society*, (2004).
- [54] Fiore, A.T., and Donath, J. S.: “Homophily in Online Dating: When Do You Like Someone Like Yourself?,” *Proc. of ACM Computer-Human Interaction*, pp. 1371–1374, (2005).
- [55] Fiore, A. T., Shaw Taylor, L., Zhong, X., Mendelsohn, G. A. and Hearst, M. A.: “Assessing Attractiveness in Online Dating Profiles,” *Proc. of ACM Computer-Human Interaction*, pp. 797–806, (2008).
- [56] Fiore, A. T., Shaw Taylor, L., Zhong, X., Mendelsohn, G. A. and Cheshire, C.: “Who’s right and who writes: People, profiles, contacts, and replies in online dating,” *Proc. of Hawaii Int’l Conf. System Sciences*, pp. 1–10, (2010).

業績リスト

論文誌

1. 黒岩将, 安本慶一, 村田佳洋, 伊藤実: “進化計算を用いた「合コン」問題の解法と評価,” 情報処理学会論文誌 数理モデル化と応用, Vol. 6, No. 1, pp. 47-56, (2013).
(4 章)

国際会議 (査読有り)

1. Sho Kuroiwa, Yoshihiro Murata, Keiichi Yasumoto, and Minoru Ito: A Method for Assigning Men and Women with Good Affinity to Matchmaking Parties through Interactive Evolutionary Computation, *Proc. of Simulated Evolution and Learning (SEAL 2008)*, pp. 645–655 (2008).
(3 章)
2. Sho Kuroiwa, Keiichi Yasumoto, Yoshihiro Murata and Minoru Ito: Finding Good Affinity Patterns for Matchmaking Parties Assignment through Evolutionary Computation, *Proc. of Parallel Problem Solving from Nature (PPSN 2012)*, Vol. 7492, pp. 498–507 (2012).
(4 章)

付録

A. 既発表文献との整合性について

3.4.4 項に示した神託手法では、解集団に 0 世代に一回限り最適解集団を与える方法を採用しているが、この方法では、1 世代目以降の解集団の更新により、0 世代目に与えた最適解が劣化していく現象が観察された。そのような理由から、同 3 章に対応する内容として著者らが発表した文献 [43] では、 $N \geq 3$ の場合でも、提案手法が神託手法と同等のパフォーマンスを出しているように見える。

しかし、より最適に近い合コン割当を出力する神託手法としては、最適解集団を毎世代与え続ける方法が考えられ、これを用いた場合は、3 章の結論 3.4.6 項で述べた通り $N \geq 3$ の場合、提案手法の有効性を確認できなかった。4 章の神託手法では、この改善版の神託手法を採用した (4.3.4 項)。

B. 乱数生成された属性値とカップル成立数の関係

乱数で利用者を生成する際、解候補長が増えに従って好相性にマッチする利用者が極端に減る理由は、以下の確率によって示される。一様乱数を用いて、 n ビットの任意のベクトル $\mathbf{p}$ に $\alpha\%$ 一致する $\mathbf{x}$ を生成できる確率 $P(n, \alpha)$ は次式で与えられる。

$$P(n, \alpha) = {}_nC_{\alpha n} \left(\frac{1}{2}\right)^{\alpha n} \left(1 - \frac{1}{2}\right)^{[n(1-\alpha)]} \quad (11)$$

$P(5, 0.8) = 0.156$, $P(15, 0.8) = 0.014$, $P(25, 0.8) = 1.58 \times 10^{-3}$ と指数的に減少する。