

特別研究報告書

選好獲得費用が異なる場合の
2サイドマッチングプロトコルの提案

指導教員 松原 繁夫 准教授

京都大学工学部情報学科

小谷 敏也

平成 21 年 2 月 2 日

選好獲得費用が異なる場合の 2サイドマッチングプロトコルの提案

小谷 敏也

内容梗概

2サイドマッチングとは、研修医志望者と病院、入学希望者と学校など異なる二つの集合があって、一方のメンバともう一方のメンバの組み合わせを決める問題であり、それを実現する方式をマッチングプロトコルと呼ぶ。2サイドマッチングは、ゲーム理論研究の中で問題の表現や解が満たすべき性質などが議論されてきている。また、効率の求解を実現するという点で、計算機科学の分野でも議論がなされている。研修医の病院配属に関しては、米国や日本でも多くの学生・病院が参加するなど、実世界への応用という点でもマッチングプロトコルは興味深い。

マッチングに関する主要な興味のひとつはマッチングの安定性であった。マッチングが安定であるとは、互いに別の者とペアになることを望むようなペアが存在しない状態を指す。安定なマッチングが生成されない場合、その制度に対して、自発的な参加を求めることが難しくなる。

安定なマッチングを求める解法として、Gale-Shapley の遅延許諾アルゴリズムが提案されている。これは、理論的に優れた性質を持つだけでなく、実際に米国や日本の研修医配属決定にも用いられるなど、有効性が示されている。しかし、多くのゲーム理論研究と同じく、選好獲得の問題には、これまで十分注意が払われてこなかった。

従来の研究では、正確な選好リストが参加者から提示されるところから問題が議論されている。しかし、正確な選好リストを得るには、何らかの費用が生じると考えられる。情報収集にかけられる費用、例えば、見学箇所の数などに実質上の制限がある場合、病院に関する情報獲得ができず、希望候補から外れてしまうかもしれない。これは、マッチングの効率性や安定性の点で問題を引き起こす。

本研究での課題をまとめると以下となる。

選好獲得費用が存在する場合のマッチングモデルの欠如 遅延許諾アルゴリズムを用いた2サイドマッチングでは、正確な選好リストが事前に得られていることを前提としている。しかし、選好獲得費用が存在する場合には、こ

の仮定は成立しない．このような状況を扱うことが可能なマッチングの理論的モデルが必要である．

効率的なマッチングの失敗 マッチングの運営団体は，マッチングの失敗が生じるのを避けるため，できるだけ長い選好リストを提出するように勧めている．しかし，選好獲得費用が存在する場合，選好リストを長くすると，結果として不要な費用を費やすことになる．一方，選好獲得費用を節約しようとして，短い選好リストしか提出しなければ，マッチングの失敗が生じやすくなる．

これらの課題を解決するため，まず，選好獲得費用を明示的にモデルに取り込むことにする．ここで，選好獲得費用を削減するには，候補の中から，行けそうにない選択肢を早期に削除すれば良い．例えば，学生 A と病院 B が相思相愛であるという情報があれば，別の学生 C は病院 B を調査対象から外すことができる．つまり，マッチングを段階的に行い，先行する者のマッチング情報を後続の者が利用することを考える．この状況を扱うために，学生に関しては，選好順位がはじめから決まっている学生と，選好順位を正確に決めていない学生の集団に分割する．また，候補リストから選択肢の削除を可能とするため，希望を提出される側，つまり，病院に関しては，学生個々を面接するのではなく，学生が持つスキルで選好が決定されるというモデルを考える．

本研究の貢献を以下にまとめる．

選好獲得費用が存在する場合のマッチングモデルの提案 選好獲得費用が存在する場合に，学生側では，選好順位が決まっている者と決まっていない者に分割し，病院側では，学生が持つスキルによって選好順位が決まるというモデルを提案した．

効率的なマッチングプロトコルの提案 選好順位が決まっている者を先にマッチングし，そこでのマッチング情報を用いて選好順位を正確に決めていない学生の希望候補を削減する方法を提案した．提案方法は，元のプロトコルが持つ，学生側にとっては，正直に選好リストを提出するという戦略が最善という性質を引き継ぐことを示した．

Suggestion of two-sided matching protocol when being different in a preference acquisition cost

Toshiya KOTANI

Abstract

Two-sided matching problem is to decide pairs of two distinct agents, for example, candidates and schools, students and hospitals. The system which achieves that is called a protocol. In game theory, it is argued that character of two-sided matching problems. And also in computer science, it is argued to solve two-sided matching problem efficiently. In the scene of the residency matching in USA and Japan, two-sided matching protocols are used. So more and more people get interested in matching protocols.

One of the main interests about matching was stability of matching. If matching is stable, the state pair which wishes for a different person doesn't exist. If matching is not stable, participants will not participate in matching.

To get stable matching, Gale and Shapley proposed deferred-acceptance algorithm. This algorithm is theoretically excellent, and is actually used in residency matching in USA and in Japan. However, study about this problem is not argued about preference acquisition.

From the past, matching problem is argued, premised that participants show correct preference. A cost may occur to get a correct preference. When there is a constraint in a cost necessary to information collection, it's lacking in information about a hospital, so a hospital of a request sometimes comes off from preference. This causes a problem by the efficiency and stability of matching.

I will research following two problems.

The lack of matching models when preference acquisition cost exists

It is premised that candidates can acquire correct preference beforehand. But when preference acquisition cost exists, this assumption isn't formed. So we need a theoretical model of matching to handle such situation.

Failure of efficient matching A governing body of matching recommends to submit a preference list as long as possible for failure of matching to avoid forming. But when a preference acquisition cost and candidates make their preference list long, an unnecessary cost will be spent as a result. On

the other hand, trying to save a preference acquisition cost, and when submitting only a short preference, failure of matching becomes easy to form.

To solve these problems, first, I put preference acquisition cost into matching models specifically. Here, to reduce preference acquisition cost, we should eliminate the choices participants don't seem to match early. For example, if student A and hospital B want to match each other, student C need not to research hospital B. In other words, we match participants step by step, and we can use the matching result at each time. To handle these situation, I classify students into two types; students who decide preference correctly from the beginning, and students don't decide preference correctly at beginning. According to hospitals, I suppose that hospitals can decide preference by using skill of students, not interviewing with them.

The main contributions of this paper are the following.

Suggestion of the matching models when a preference acquisition cost exist

When a preference acquisition cost exist, I divide students into two types; students who have clear preference and students who does not it. And I propose model hospitals acquire preference by the skill of students.

Suggestion of efficient matching protocols Preference, a decided person, first, matching, the way to reduce the student's request candidate who has not decided preference correctly using matching information there was proposed. Suggestion method showed that the strategy to which a preference list is submitted honestly for the student side an original protocol has takes the nature as the best over.

選好獲得費用が異なる場合の 2 サイドマッチングプロトコルの提案

目次

第 1 章	はじめに	1
第 2 章	遅延許諾アルゴリズム	3
2.1	安定結婚問題	3
2.2	遅延許諾アルゴリズム	4
2.3	安定配属問題	6
2.4	遅延許諾アルゴリズムの問題点	6
第 3 章	研修医配属問題	7
3.1	配属プロセス	7
3.2	病院の選好獲得	8
3.3	学生の選好獲得	8
第 4 章	プロトコルの提案	8
4.1	モデル	9
4.2	プロセス	11
4.3	プレマッチングの役割	14
第 5 章	分析	15
5.1	Gale-Shapley アルゴリズムの性質の保存	15
5.2	マッチングアルゴリズムの精度	15
第 6 章	評価	17
6.1	シミュレーションの設定	17
6.2	マッチング規模と候補数の関係	18
6.3	プレマッチングの効果	18
6.4	既存プロトコルと提案プロトコルの比較	18
6.5	実世界への適用	21
第 7 章	おわりに	21
	謝辞	23
	参考文献	23

第1章 はじめに

2サイドマッチングとは、研修医志望者と病院、入学希望者と学校など異なる二つの集合があって、一方のメンバーともう一方のメンバーの組み合わせを決める問題であり、それを実現する方式をマッチングプロトコルと呼ぶ。2サイドマッチングは、ゲーム理論研究の中で問題の表現や解が満たすべき性質などが議論されてきている。また、効率の求解を実現するという点で、計算機科学の分野でも議論がなされている。以下の図1は学生と病院の2つの集合のマッチング問題を表したものである。

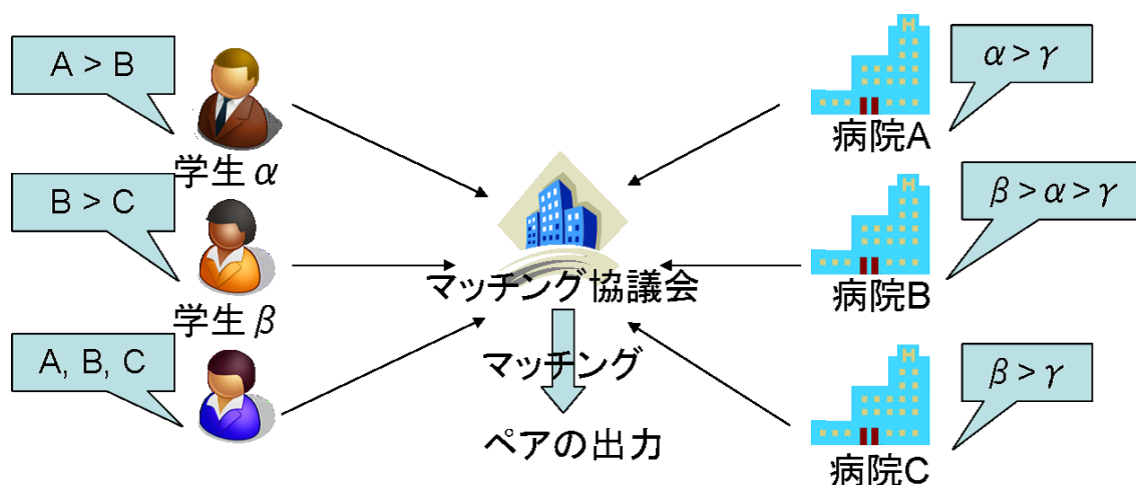


図1: マッチング問題

研修医の病院配属に関しては、米国では1952年から National Resident Matching Program という制度が運用され、近年では、およそ16,000名の米国医学校卒業生と15,000名の米国外医学校卒業生が参加し、24,000の職を競う状況となっている。日本でも2003年から研修医マッチングが始まっており、およそ8,500名の学生が参加する規模となっている。米国や日本での研修医配属決定には、ゲーム理論家が開発したプロトコルが適用されるなど、実世界への応用という点でもマッチングプロトコルは興味深い。

マッチングに関する主要な興味のひとつはマッチングの安定性である。マッチングが安定であるとは、互いに別の者とペアになることを望むようなペアが存在しない状態を指す。適用されるプロトコルが安定なマッチングを生成するものでなければ、制度に対する自発的な参加を促すことが難しくなる。実際、過

去の事例などを調べた結果，長く継続して利用される制度で使用されるプロトコルは，安定なマッチングを生成するものであることがわかっている．

安定なマッチングを求める解法として，Gale-Shapley の遅延許諾アルゴリズムが提案されている．これは，理論的に優れた性質を持つだけでなく，実際に米国や日本の研修医配属決定にも用いられるなど，有効性が示されている．しかし，多くのゲーム理論研究と同じく，選好獲得の問題には，これまで十分注意が払われてこなかった．

従来の研究では，正確な選好リストが参加者から提示されることから問題が議論されている．しかし，正確な選好リストを得るには，何らかの費用が生じると考えられる．研修医配属問題で言えば，学生は，最初，関心がある／ない，あるいは，非常に関心がある／少し関心がある／関心がないといった，2段階あるいは3段階くらいの緩い選好があって，それが先輩の声を聞いたり，現地を見学するなどして，正確な順位がつけられると考えられる．例えば，見学一つをとっても，時間のかかることであり，学生／病院双方にとって費用が生じていると言える．

情報収集にかけられる費用，例えば，見学箇所の数などに実質上の制限がある場合，病院に関する情報獲得ができず，希望候補から外れてしまうかもしれない．これは，マッチングの効率性の点で問題を引き起こす．また，マッチングの事後にその病院の情報がたまたま得られたりすると，研修先の変更を希望するなど，マッチングを不安定化させる要因となり得る．

本研究での課題をまとめると以下となる．

選好獲得費用が存在する場合のマッチングモデルの欠如 遅延許諾アルゴリズムを用いた2サイドマッチングでは，正確な選好リストが事前に得られていることを前提としている．しかし，選好獲得費用が存在する場合には，この仮定は成立しない．このような状況を扱うことが可能なマッチングの理論的モデルが必要である．

効率的なマッチングの失敗 マッチングの運営団体は，マッチングの失敗が生じるのを避けるため，できるだけ長い選好リストを提出するように勧めている．しかし，選好獲得費用が存在する場合，選好リストを長くすると，結果として不要な費用を費やすことになる．一方，選好獲得費用を節約しようとして，短い選好リストしか提出しなければ，マッチングの失敗が生じやすくなる．

この問題を解決するために，本研究では，選好順位が決まっている学生と決まっていない学生の存在に着目し，前者の情報を適切に利用することで，後者の学生が選好獲得に費やす費用を低減させるマッチングプロトコルを提案する．

以降，本稿の構成は次の通りである．まず，第2章で Gale-Shapley の遅延許諾アルゴリズムの概説を安定結婚問題を例に概説し，関連研究を紹介する．次に第3章では，本研究で扱う研修医配属問題の概説を行い，既存のプロトコルと志願者である学生と病院の選好を分析する．第4章では，志願者の選好を共有することで選好獲得にかかる費用を低減できるようなマッチングプロトコルの提案をする．第5章では提案したプロトコルを分析し，満たすべき性質について考察を行う．第6章では，提案したプロセスの性能を評価するためにシミュレーションを行い，その結果について考察する．第7章でまとめを行い，本稿を締めくくる．

第2章 遅延許諾アルゴリズム

2.1 安定結婚問題

マッチング問題は1対1のマッチングと多対1の問題に分類でき，前者は一人の男性と一人の女性を組み合わせる場面を想定して結婚問題と呼ばれ，後者は複数の入学希望者と一つの学校を組み合わせる場面を想定して配属問題と呼ばれる．遅延許諾アルゴリズムは，どちらの問題に対しても適用できる．まず，安定結婚問題に関して説明を与える．

n 人の男性の集合 M と n 人の女性の集合 W が存在するとする．男性 / 女性は女性 / 男性に対して選好を持っている．男性 m は，集合 $W \cup m$ 上で選好リスト $P(m)$ を持っているとして表現される． $P(m) = \{w_1 \succ_m w_2 \succ_m f \succ_m \dots\}$ といった形で表現される．これは，男性 m は女性 w_1 とマッチすることを最も選好し，つぎに，女性 w_2 とマッチすることを選好し，3番目の選択は，誰ともマッチしない，つまり，女性 w_1 が w_2 とマッチするのでなければ，独身に留まることを選好する，という状況を表現している．特に混乱が生じない場合には，選好関係 \succ_m を単に $>$ と表現する．女性 f についても，同様に，集合 $M \cup f$ 上で選好が定義される．

男性 m にとって，女性 w_1 と w_2 が無差別であるという状況も存在するかもしれない．後で述べる遅延許諾アルゴリズムでは，マッチメイカーが任意に，

表 1: マッチング問題の表現

$$\begin{aligned}
 m_1: w_2 > w_1 > w_3 & \quad w_1: m_1 > m_3 > m_2 \\
 m_2: w_1 > w_3 > w_2 & \quad w_2: m_3 > m_1 > m_2 \\
 m_3: w_1 > w_2 > w_3 & \quad w_3: m_1 > m_3 > m_2
 \end{aligned}$$

$w_1 \succ_m w_2$, あるいは, $w_2 \succ_m w_1$ という関係を導入することで処理が可能である．よって, 以下本稿では無差別の関係については言及しない．

マッチング μ は, $\mu(f) \neq f$ であれば, $\mu(f) \in W$ となり, $\mu(w) \neq w$ であれば, $\mu(w) \in F$ となる, 集合 $F \cup W$ からその上への 1 対 1 写像である．本稿では, $\mu(x)$ を x のパートナーと呼ぶ．

男性 m と女性 w がマッチング μ では互いにマッチしていないが, μ におけるマッチングと比較して, 互いに選好している状況を考える．すなわち, $w \succ_m \mu(m)$, かつ, $m \succ_w \mu(w)$ という状況である．この男性と女性のペア (m, w) を, ブロッキングペアと呼ぶ．マッチング μ は, 誰にも, どの男性と女性のペアにもブロックされなければ, 安定である．マッチングが安定でなければ, 男女の組が決まっても, そこから逸脱する誘因が存在する．自発的な参加を促し, 制度として成功を収めるには, マッチングが安定である必要がある．

マッチング問題の例を表 1 に示す．ここでは, 3 名の男性 m_1, m_2, m_3 と 3 名の女性 w_1, w_2, w_3 が存在し, 例えば, 男性 m_1 は, w_2, w_1, w_3 という順に女性を選好することを表している．マッチング μ は, 男性と女性の組を決定するものであり, 例えば, $\mu = \{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$ がその一例である．

マッチング $\mu = \{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$ が与えられたとき, 男性 m_1 と女性 w_2 は, 現在の男女のペアよりも, (m_1, w_2) というペアを互いに選好する．これがブロッキングペアである．別のマッチング $\mu' = \{(m_1, w_1), (m_3, w_2), (m_2, w_3)\}$ を考えよう．この場合, 各ペアに関して, 別のパートナーを互いに選好するという状況は生じない．よって, マッチング μ' は安定である．

2.2 遅延許諾アルゴリズム

安定マッチング問題の解法として Gale-Shapley の遅延許諾アルゴリズムが知られている．その手順を以下に示す．

1. まず, 男性は最も選好する女性にプロポーズする．

2. 女性は複数のプロポーズを受けた場合，最も選好する男性からのプロポーズを仮に受け入れることとし，それ以外のプロポーズを断る．
3. プロポーズを断られた男性は，これまでにプロポーズを断られていない女性の中で，最も選好する女性にプロポーズする．ここでは，その女性に既に婚約者がいるかどうかは考慮しない．
4. 女性は最も選好する男性からのプロポーズを一つ残して，他のプロポーズを断る．
5. ステップ3, 4を繰り返し，どの男性もプロポーズを断られなくなった段階でアルゴリズムを停止する．
6. 男女は現在婚約している相手と結婚する．

上述のように，このアルゴリズムでは，男性女性が厳密な選好を持つ，つまり，パートナーの選好が完全順序として定義されることを仮定している．複数のパートナーについて無差別な関係が存在する場合でも，なんらかのタイブレーク規則，例えば，名前の50音順など，を導入することで上記アルゴリズムが適用可能となる．

この安定結婚問題に関しては，つねに安定なマッチングが少なくとも1つは存在することが証明されている．証明は，背理法を用いて，上記アルゴリズムから得られるマッチングにブロッキングペアが存在すると仮定して，矛盾を導くことでなされる．一度プロポーズを断られた女性に，再度プロポーズすることは認められていないため，このアルゴリズムは有限時間で停止する．

このアルゴリズムでは男性が女性にプロポーズをする方法をとっているが逆に女性が男性にプロポーズをするという方法を行っても安定マッチングが得られる．遅延許諾アルゴリズムにおいては，男性側がどのような順序でプロポーズを行ったとしても，同一のマッチングが導かれる．

(定理1)[Gale and Shapley, 1962] Gale-Shapley アルゴリズムにおいて男性側がどのような順序でプロポーズを行ったとしても，Gale-Shapley アルゴリズムは同一のマッチングを導く．

男性からプロポーズをする方法をとれば男性にもっとも好ましい(男性最良)，女性からプロポーズをする方法をとれば女性にもっとも好ましい(女性最良)のマッチングが得られる．

2.3 安定配属問題

安定結婚問題は1対1のマッチングを考えた。研修医配属の問題を考えれば、病院側は、複数の採用枠を持つかもしれない。つまり、1対多のマッチングへの拡張が考えられる。これは、安定配属問題と呼ばれる。

安定配属問題に関して、Gale, Shapley, Rothらによって有用な定理が証明されている。

(定理2)[Gale and Shapley, 1962] 任意の選好プロファイルについて、安定なマッチングの集合は非空である。[1]

(定理3)[Roth, 1985] 任意の安定マッチング集合について、すべての学生が任意の安定マッチングより少なくとも同順位でマッチする「学生最適安定マッチング」が唯一つ存在する。同様に「病院最適マッチング」も唯一つ存在する。[3]

遅延許諾アルゴリズムにおいて、学生側から配属を順番に申し込む場合、得られるマッチングは学生最適安定マッチングとなる。逆に、病院側から申し込む場合は病院最適安定マッチングとなる。

(定理4)[Roth, 1985] 任意の安定マッチング集合について、学生最適な安定マッチングは、病院最悪な安定マッチングである。同様に、病院最適な安定マッチングは学生最悪な安定マッチングである。[4]

遅延許諾アルゴリズムで得られるマッチングは必ず安定であるが、学生側と病院側で公平ではないことを示す。

(定理5)[Roth, 1985] 任意の選好プロファイルについて、学生本位の遅延許諾アルゴリズムによってマッチングを行うとき、学生は真の選好を表明することが支配戦略となる。[5]

つまり、どんな偽の選好を表明しても付加的な効用を得ることはできない。

2.4 遅延許諾アルゴリズムの問題点

遅延許諾アルゴリズムでは、厳密な選好リストが得られていることを仮定している。不完全なリストを持つ、つまり、すべてのパートナー候補に対して選好を申告するのではなく、一部についてのみ選好を申告するという場合への拡張も図られている。この場合でも、選好リストの存在が前提とされている。

しかし、各参加者にとって、選好リストは先験的に与えられるものではなく、獲得されるものであると考えられる。学生にとっては、正確な選好リストを作

成しようとするれば、各病院が提供する研修プログラムの内容、例えば、設備や勤務時間などを調べる必要がある。つまり、学生にとって効用に関しては、配属先の病院に対する満足度だけでなく、配属決定にかかる不効用も合わせて考えるのが妥当である。

もちろん、学生は、正確な選好リストを得るために、すべての病院を見学に行くといった行動はとっておらず、見学先を選んでいる。その場合でも、どの病院に見学に行くかの意思決定が必要である。学生にとっては、見学に割ける時間や費用も有限である。見学先に、他の学生や病院の選好によって、配属の見込みが少ない病院が多く含まれる場合、実質的な選好リストが短くなって、マッチングの失敗という事態が生じる恐れもある。一方、マッチング失敗を恐れて、見学先を増やせば、結果として、見学に無駄な費用を投資することになる。よって、見学先を決める何らかの方法が必要である。

第3章 研修医配属問題

研修医の病院配属に関しては、米国では1952年から National Resident Matching Program (NRMP)²⁾ という制度が運用され、近年では、およそ16,000名の米国医学校卒業生と15,000名の米国外医学校卒業生が参加し、24,000の研修医の地位を競う状況となっている。

日本でも病院と研修志願者の組み合わせを決定する研修医マッチング2004年より始まっており、およそ8,500名の学生とおよそ1,100の病院が参加する規模となっている。これは医師臨床研修マッチング協議会³⁾によって運営され、中央集中型のマッチングが行われている。

3.1 配属プロセス

日本における研修医配属の手順を次に示す。

1. 志願者である学生がいくつかの病院の面接を受ける。
- 2a. 志願者が面接を受けた病院について登録した順に順序をつけて選好リストをマッチング協議会に提出する。
- 2b. 病院側も同様に面接した志願者の順位をつけ定員と選好リストを協議会に提

²⁾ National Resident Matching Program(<http://www.nrmp.org/>)

³⁾ 医師臨床マッチング協議会 (<http://www.jrmp.jp/>)

出する。

3. マッチング協議会は得られた選好リストに対して遅延許諾アルゴリズムを適用し，結果を各志願者と各病院に通知する。

そのプロセスを以下の図2に示す。

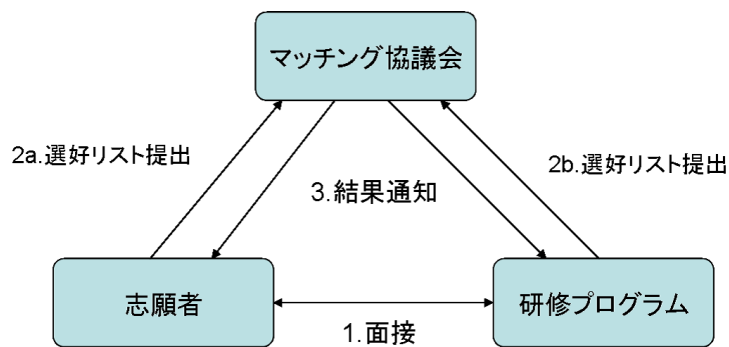


図2: 研修医配属プロセス

3.2 病院の選好獲得

参加病院は、募集定員（研修プログラム毎の募集定員の合計）の全てを、研修医マッチングで募集する。志願者である学生と面接を行い，選好リストをマッチング協議会に提出する。

3.3 学生の選好獲得

志願者である学生はいくつかの病院と実際に面接を行い，正確な選好リストを提示することが前提となっている。しかし，志願者には他者の影響を受けず，最初から明確に応募したい病院とその順位が決まっている者もいれば，漠然とした候補しか持たない者もいると考えられる。漠然とした候補しか持たない原因としては，他者の選好が不明な点と，病院自体に関する情報が不足している点が挙げられる。いずれにせよ，やみくもに見学を行って，その志願者に結果として不要なコストが生じることになる。

第4章 プロトコルの提案

本章では，問題を解決するための2サイドマッチングプロトコルに要求される条件について考察し，選好獲得費用が存在する場合のマッチングプロトコルの提案を行う。

4.1 モデル

本研究で扱うモデルについて説明する．学生の集合 S と病院の集合 H が存在する．学生数は n 人，病院数は m とする．実際の研修医配属では，一つの病院が，例えば，内科と外科など複数の研修プログラムを持つ場合がある．本稿では，説明を簡単にするため，各病院が提供する研修プログラムの数は1つとする．

学生は病院に対して選好を持つ．仮にすべての病院を見学したとすると，学生 s は，集合 $H \cup \{s\}$ 上で選好リスト $P(s)$ を定義することができる．選好リストは，安定結婚問題と同じく， $P(s) = \{h_1 >_s h_2 >_s f >_s s \dots\}$ といった形で表現される．これは，学生 s は病院 h_1 とマッチすることを最も選好し，つぎに，病院 h_2 とマッチすることを選好し，3番目の選択は，誰ともマッチしない，つまり，病院 h_1 が h_2 とマッチするのでなければ，研修プログラムに参加しないことを選好する，という状況を表現している．特に混乱が生じない場合には，選好関係 $>_s$ を単に $>$ と表現する．

ここで，学生を2つのタイプに分ける．1つは，選好順位の決まっている学生であり，1つは選好順位の決まっていない学生である．前者は，見学といった情報収集活動をしなくても，選好順位を決めている学生である．例えば，研修内容そのものでなく，著名な先生がそこに所属しているから，あるいは，知人がそこで働いているから，といった理由で選好順位を決めている学生である．これらの学生は，仮に病院を見学したとしても，その選好順位を変更することはない．

一方，選好順位の決まっていない学生は，見学に行くなど情報収集をすれば，上で述べた厳密な選好リストを獲得することができる．本稿では，選好のあいまい性の原因は，自己の情報不足のみによるもので，他者の選好が未知であることによるとは考えない．つまり，他者の選好を知って，人気がある／ない，などといった情報を得ても，それは自己の選好順位の決定に影響を与えないとする．この，学生間の情報の非対称性については，既に一部議論がなされている．

選好順位の決まっていない学生は，情報収集をしなければ，厳密な選好順位を決められないが，検討対象の候補は保持しているとする．情報収集をする前は，検討対象候補には順序関係が設定されていない．これは，無差別であることを意味するのではないことに注意されたい．情報収集活動をすることで，検

表 2: マッチング問題の表現

$s_1: h_2 > h_1 > h_3$	$h_1: s_1 > s_3 > s_2$
$s_2: h_1 > h_3 > h_2$	$h_2: s_3 > s_1 > s_2$
$s_3: h_1 > h_2 > h_3$	$h_3: s_1 > s_3 > s_2$

選好順位の決っている学生: s_1, s_2

選好順位の決まっていない学生: s_3

学生 s_3 の検討対象集合: $\{h_1, h_2\}$

討対象の候補に順序関係が与えられる。検討対象の候補であっても、情報収集を行うことで、志望リストから外れることは生じ得る。一方、検討対象以外の候補に関しては、仮に情報収集活動を行っても、志望リストに含まれることはないと仮定する。検討対象候補と非対象の2分割でなく、非常に興味がある / 興味がある / 興味がないの3分割するといった拡張も可能であるが、本稿では、2分割できるとして、議論を進める。

選好順位の決まっていない学生が、情報収集する場合、費用が生じる。旅費などといったものを考えれば、複数の見学先の地理的近接関係に影響されるが、本稿では、病院1箇所の見学当たり費用 c が生じるものとする。これは、選好順位の決まっていない学生に関しては、同一の値を取ると考える。

一方、病院についても、同様に、集合 $S \cup \{h\}$ 上で選好が定義される。ただし、通常安定結婚問題と異なり、各病院は学生のスキルを引数にする選好関数 p_h を持ち、その値によって、学生が順序付けられるとする。すなわち、個々の学生を面談しなくても、学生集合が与えられれば、例えば、履歴書と成績表を元に、学生に対する選好順位を定義できるということである。

マッチングやブロッキングペアの定義については、安定結婚問題の説明で与えた定義と同じである。

マッチング問題の例を表2に示す。ここでは、3名の学生 s_1, s_2, s_3 と3つの病院 h_1, h_2, h_3 が存在し、例えば、学生 s_1 は、 h_2, h_1, h_3 という順に病院を選好することを表している。

4.2 プロセス

本節では、選好順位が決まっていない学生が、結果として不要な情報収集活動を削減し得る、新たなマッチングプロトコルを提案する。基本となる考え方は、初期段階では、明確に順序付けされていない、あいまいな選好リストの提出を認め、選好順位が決まっている学生の選好情報を利用して、徐々に選好を明確化していくというものである。

本研究においては例示の簡単化のために各病院の定員を1名と設定する。また、マッチングアルゴリズムは実際に使われている通り、学生から病院に申し込みを行うという学生本位の Gale-Shapley アルゴリズムとする。提案プロトコルを以下に示す。

1. 学生は成績や履歴書などの個人をランク付けする情報と共に、選好リストをマッチング協議会に提出する。この際にはあいまいな選好リストも認める。病院側は学生に対する順位づけを行うための選考基準をマッチング協議会に提示する。
2. マッチング協議会は各病院から提示された選考基準と学生個人の情報から、各病院を志願している学生に対して順位付けを行う。この際の順位付けには選好順位がきまっている学生と選好順位がきまっている学生を共に含める。
3. ここで、選好順位が決まっている学生と病院のマッチングを行う。このマッチングをプレマッチングと呼ぶ。
4. ステップ3で得られたマッチング結果に関して、マッチング協議会は選好順位が決まっていない学生よりステップ2における順位付けが上位でマッチしている選好順位が決まっている学生がいるか検査する。もし順位付けにおいて上位でマッチする者がいれば、その病院に他の学生はマッチできないものと判断する。
5. この情報を選好順位が決まっていない学生に提供する。そうすると選好順位が決まっていない学生はわざわざペアとなる可能性のない病院に関して情報収集する必要がなくなる。選好順位が決まっていない学生に関して、候補にペアとなる可能性のない病院が含まれているならば、それを除く。
6. 病院側から見て、プレマッチングでペアとなった学生より上位に順位付けされている選好順位がきまっていない学生がいれば、このペアは一旦解消する。なお、学生は選好順位が決まっている学生であるため、ここで選好

順位の変更は生じない。

7. 病院側の選好順位は，ステップ6においてペアが解消しない場合は確定しているものとして，ここでは考えない．またそれ以外の病院についてはマッチング1で確定した学生を除き，新たに見学にきた「選好順位が決まっていない学生」と「選好順位が決まっている学生」を合わせた志願者に対して再度評価を行い，新たに順位づけをする．
8. プレマッチングで確定した学生と病院を除いた，全志願者と病院についてマッチングを行う．この際に得られたペアとプレマッチングで確定したペアを最終的なマッチング結果として提示する．

以下の図3にマッチングの流れを示す．左側が学生のマッチングまでの流れで，右側が病院のマッチングまでの流れである．

図3．提案プロトコルのフローの学生側フローチャートにおいて，選好リストが決まっている学生を学生1，選好リストが決まっていない学生を学生2と呼ぶ．

研修プログラム側のフローチャートにおいて，選考基準の提示とは，志願者である学生をスキルなどで順位付けするための基準である．また，選好リストの更新とは，プレマッチング以前には選好リストに含まれていなかった学生2も追加した選好リストとなる事を表す．

また，現行のプロトコルと提案するプロトコルをシーケンス図で比較する．それぞれ図4，図5の通りである．

図4で学生と病院から提出される選好リストは，お互いに面接を行った相手に対し順位付けが行われているものである．

図5のシーケンス図において，学生1とは選好リストが決まっている学生を表し，学生2は選好リストが決まっていない学生を表す．また，学生2が最初に提出する選好リストは，正確に順序付けされたのものではないあいまいな選好リストである．病院側の選好は，選好リストではなく，各病院が志願者に求めるスキルなどによって志願者を順位付けするための基準であり，ここでは選好関数とよぶ．すなわち，病院側のランク付けを行っているのは実際にはマッチング協議会である．

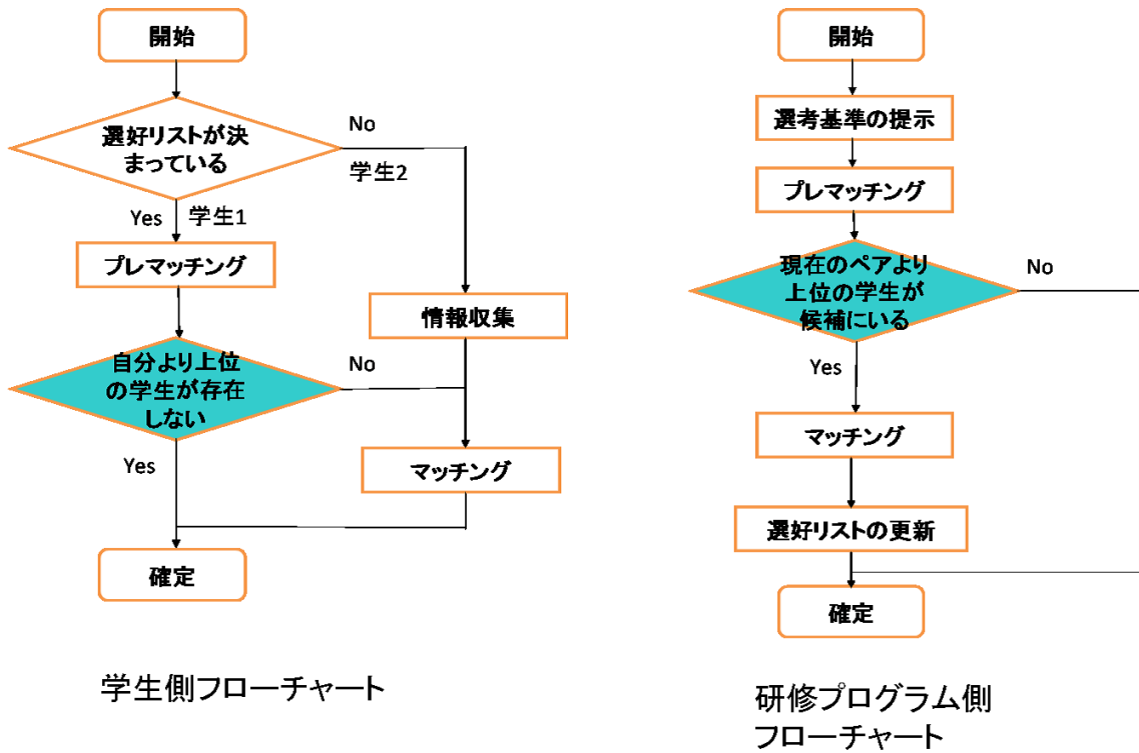


図 3: 提案プロトコルのフロー

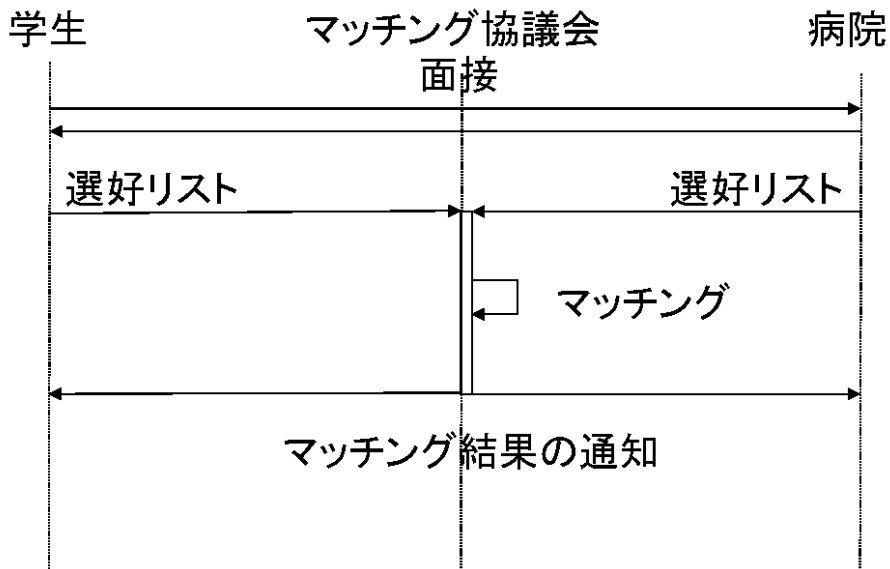


図 4: 既存プロトコルシーケンス図

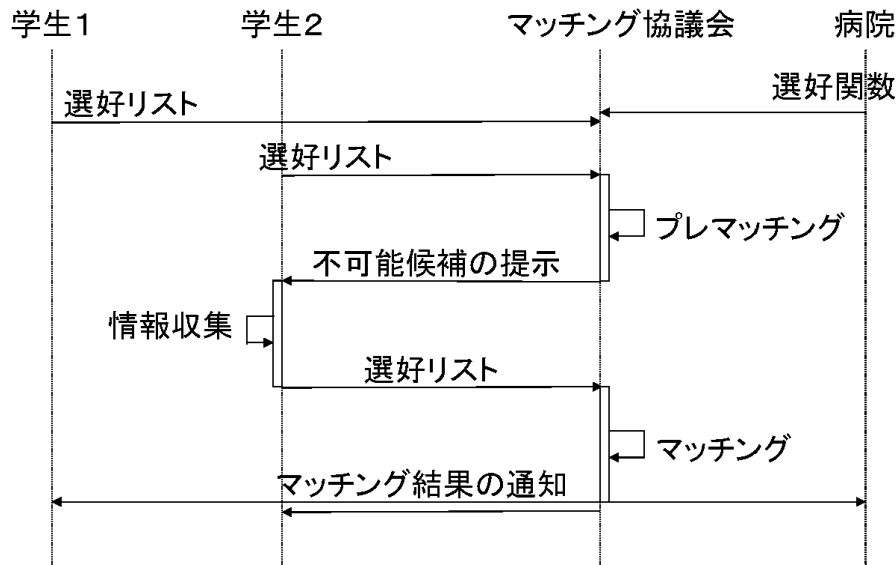


図5: 提案プロトコルシーケンス図

4.3 プレマッチングの役割

提案プロトコルの適用例を考えてみることによりプレマッチングの説明をする。学生 α は選好リストが決まっていて病院 A,B の順にいきたいと考えている。学生 γ は病院 A,B,C のどこかに行きたいと考えているが順番は決まっていない。また、病院 A は 2 人の学生 α と γ に対して順位付けを行ってある。まず最初に学生と病院はマッチング協議会に選好リスト提示する (図6の1. 提示)。この際には学生 γ のように選好リストに順位づけが行われていなくてもよい。次に選好リストがきまっている学生 α と病院 A とのマッチング (プレマッチング) を行う (図6の2. プレマッチング)。その結果の情報を学生に返す (図6の3. 結果の通知)。学生 α は病院 A とマッチすることが決まり、学生 γ は病院 A とはマッチングする可能性のないことがわかった。それによって学生 γ は病院 B と C についての情報収集を行い、その 2 つの病院に対して順位付けを行う (図6の4. 選好リストの獲得)。学生 γ は情報収集を行って正確になった選好リストを再びマッチング協議会に提出する (図6の5. 明確になった選好リストの提示)。以上がプレマッチングである。その後、マッチング協議会は学生と病院のマッチングを再度行い、結果を各学生と病院に返す。以下の図6は提案プロトコルにおいて、プレマッチングの役割を示す図である。

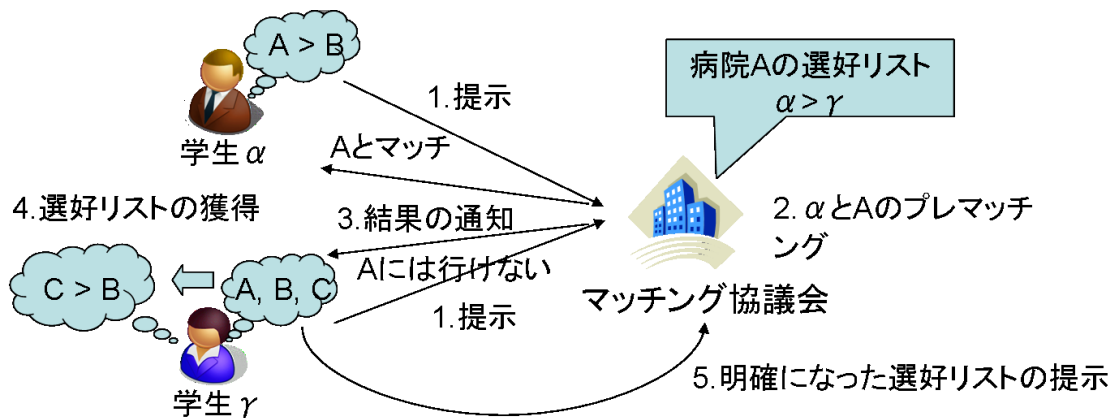


図 6: プレマッチングの役割

第5章 分析

本章では，前章で提案したプロトコルを分析し，既存プロトコルが満たしている Gale-Shapley アルゴリズムの性質が保存されているかについて考察する．

5.1 Gale-Shapley アルゴリズムの性質の保存

提案プロトコルでは，先に選好リストが決まっている学生をプレマッチングし，その結果を選好リストが決まっていない学生と共有することによりマッチングする可能性のない候補を削除するという仕組みである．よって選好リストが決まっている学生は，選好リストの持ち方によっては選好リストが決まっていない学生に不利益となるような選好リストをプレマッチングの際にマッチング協議会に提出することも考えられる．しかしながら，選好リストが決まっていない学生に不利益な選好リストを提出すると，プレマッチングにおいても Gale-Shapley アルゴリズムを用いているので，自分も本意ではない病院とマッチングしてしまう危険性がある．ゆえに学生にとって，虚偽の申告をしても得をすることは無いという性質は保存されている．

5.2 マッチングアルゴリズムの精度

提案プロトコルと既存のプロトコルのマッチング結果について考察する．提案プロトコルは学生を，選好リストが決まっている学生と選好リストが決まっていない学生の2つに分類して，プレマッチングとその後のマッチングの2段

学生の順位別マッチング人数

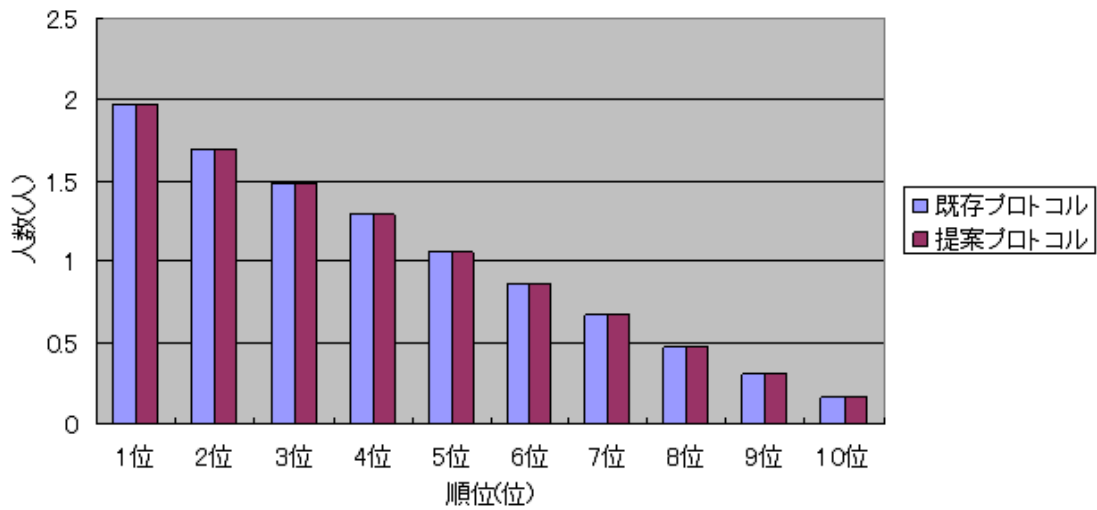


図 7: 学生の順位別平均人数

階のマッチングを行っているといえる。しかし、実際にプレマッチングで行われている操作としては、選好リストが決まっていない学生が、どのように候補に順位をつけたとしてもマッチする可能性のない病院を除いているだけなので、最終的なマッチングの結果は既存のプロトコルと提案プロトコルで変わらないはずである。

そこで、既存プロトコルと提案プロトコルに同じ選好リストを与え、10000回マッチングを行った。学生10人と病院10箇所とをマッチングした。提案プロトコルにおいては、選好が決まっている学生と選好が決まっていない学生がそれぞれ5人ずつである。次の図7、図8はそれぞれ1回あたりのマッチングの学生の順位別平均人数と病院側の順位別平均病院数を表している。これらの結果から、既存プロトコルと提案プロトコルでは同じマッチング結果が得られるといえる。

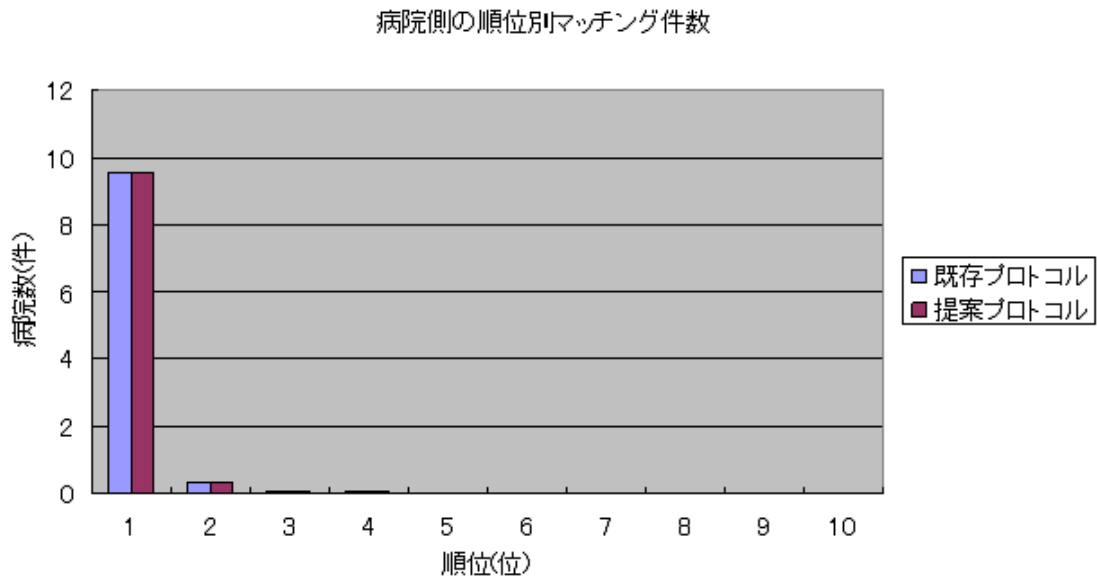


図 8: 病院側の順位別平均院数

第6章 評価

本章では、提案したプロトコルの性能を評価するためにシミュレーションを行った結果を示し、それについて考察する。

6.1 シミュレーションの設定

まずシミュレーションの設定について説明する。説明をわかりやすくするために、研修医配属の場面を次のように仮定する。(学生数):(病院数) = 1:1とする配属を考える。つまり、各病院の定員は1人である。また学生の内わけであるが(選好が決まっている学生):(選好が決まっていない学生) = 1:1とする。選好リストはランダムで与え、提案プロトコルと既存プロトコルは同じ選好リストとなっている。マッチングはそれぞれの人数において1000回行った。また、ここでは費用上の制約から10箇所病院しか見学することができない場合を想定した。

このシミュレーションの設定においては学生数と病院数が同じであるので、マッチングの規模は学生数で測るものとする。

マッチング規模と候補数の関係

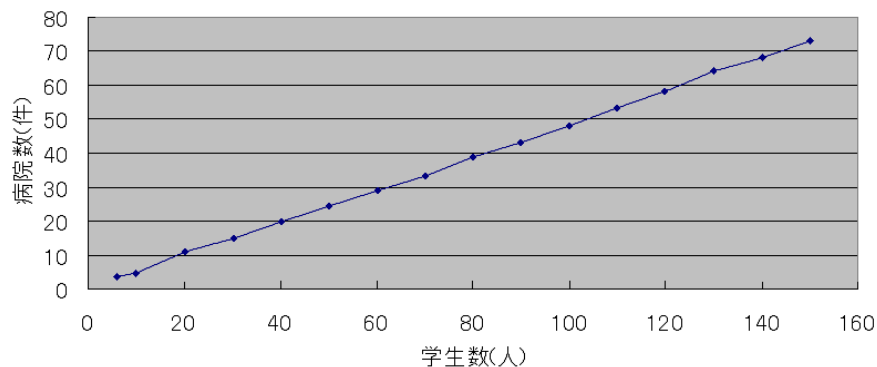


図 9: マッチング規模と候補数の関係

6.2 マッチング規模と候補数の関係

学生が選好リストを獲得する際に、マッチングの規模が大きくなる、すなわち候補となる病院数が多くなったとき、学生が本当は見学にいきたいと思う病院の数もそれに比例して増えると考えられる。そこで、シミュレーションによって学生が選好リストを得る際のマッチング規模と候補数の関係を調べた。以下の図 9 がマッチング規模と候補数の関係を表す。この図から学生は、マッチング相手である病院の数が増えれば、見学に行きたいと思う候補も比例して増えることがわかった。

6.3 プレマッチングの効果

提案プロトコルにおいて、プレマッチングを用いることにより選好リストがきまっていない学生が、自分の持っていた順序付けされていない選好リストから何箇所の候補をマッチする可能性のないものとして削除できるのかを調べた。以下の図 10 はマッチングの規模と、選好リストがきまっていない学生一人あたりががプレマッチングによって候補から削られた病院数の関係を表す。マッチングの規模が大きくなるにつれて削除できた病院数も増えることがわかった。

6.4 既存プロトコルと提案プロトコルの比較

次に費用上の制約から、見学に行くことが可能な病院数が 10 箇所と制限されている場合に、より正確な選好リストが得られる確率という点から既存プロトコルと提案プロトコルの比較を行った。ここでいう正確な選好リストというの

マッチング規模と一人あたりの病院候補削除数の関係

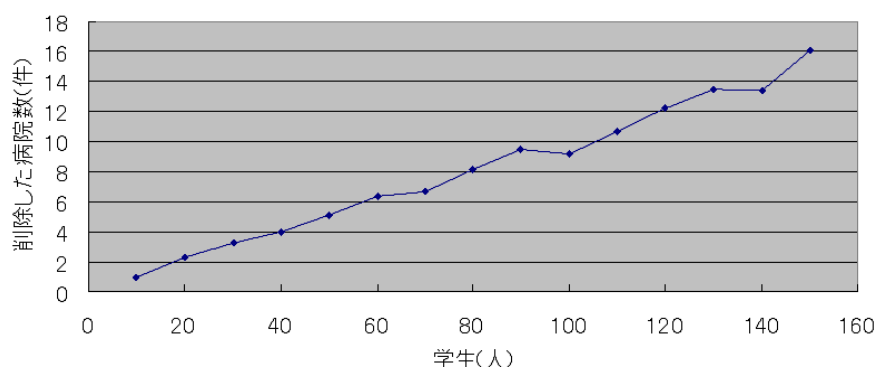


図 10: マッチング規模と学生一人あたりの候補削除数の関係

は、真に行きたい候補が費用上の制約となる 10 箇所に含まれている確率のことを言うとする。例えば 20 箇所候補があって、10 箇所しか見学できないと、本来 1 位の病院が見学先に含まれる確率は $1/2$ 、20 箇所が 15 箇所に絞られていれば、1 位の病院が見学先に含まれる確率は $2/3$ で、増加する。また、例えば見学に行きたい候補が 20 箇所あって、本来の 1,2,3 位と、14-20 位の 10 箇所を見学したとする。この場合、選好リストの正確さは 30% になるが、1,2,3 位のどこかとマッチングできれば、志望者にとって満足度は高いと考えられる。そこで正確な選好リストが得られる確率を以下のように定めた。

$$(\text{正確な選好リストが得られる確率}) = 10 \div (\text{見学に行きたい病院の候補数}) \times 100$$

この式をもとにして、既存プロトコルと提案プロトコルをシミュレーションした。その結果が以下の図 11 である。既存プロトコルと提案プロトコルにおいてマッチング規模が小さいときに 100% となっているのは、見学したいと思う候補数が 10 箇所以下であるからである。提案プロトコルではマッチングの規模に関わらず、既存プロトコルよりも正確な選好リストが得られているといえる。

次に提案プロトコルを用いた際の既存プロトコルからの改善の割合をグラフ化した。以下の図 12 に示されるような結果となった。

最初、候補となる病院が 11 病院や 12 病院しかなければ、10 箇所見学できれば、行きたいと思う上位にランクされる病院がそこに含まれるの確率は高いので、提案プロトコルで改善できる場所は少ない。病院数が増えれば、候補を絞

10箇所しか見学できない場合の正しい選好リストが得られる確率

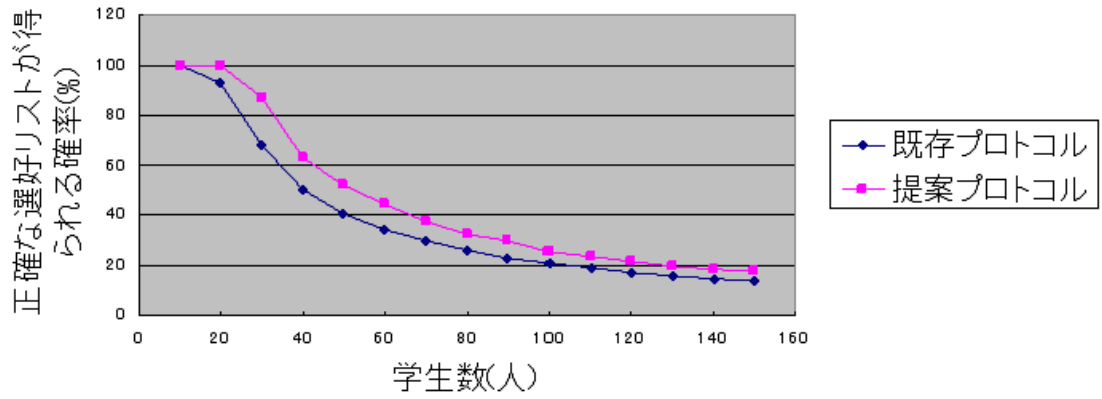


図 11: 既存プロトコルと提案プロトコルの比較

既存プロトコルからの改善の割合

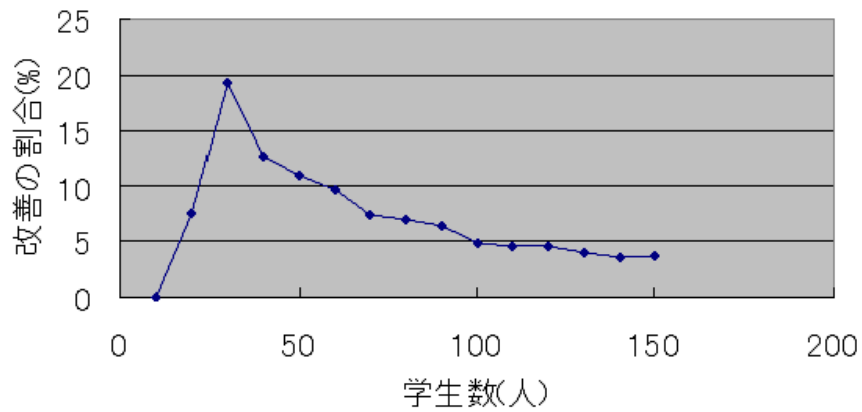


図 12: 既存プロトコルからの改善の割合

ることによる効果が大きくなる。しかし、病院数がさらに多くなると、志望者間での志望先の重なりが少なくなるため、候補の絞込みが十分でなくなる。そのため、改善の効果が小さくなるといえる。

6.5 実世界への適用

シミュレーションの結果から提案プロトコルを実世界で適用する際の有効性について考察を行う。図 12 の結果より、費用上の問題から 10 箇所しか見学にいけないことを想定した場合、マッチングの規模が学生が 30 人程度の際にもっとも有効となることがわかる。実際に日本で行われている研修医配属の規模は学生数は 8,500 人という規模である。しかし、志願者である学生の研修医プログラムにおける専門分野や、実際の地域性を考慮すると提案プロトコルが有効となる領域に近づけるのではないかと考えられる。よって実世界での適用も現実的であると考えられる。

第 7 章 おわりに

本研究では、研修医配属問題などの 2 サイドマッチングにおける以下のような問題について扱った。

選好獲得費用が存在する場合のマッチングモデルの欠如 遅延許諾アルゴリズムを用いた 2 サイドマッチングでは、正確な選好リストが事前に得られていることを前提としている。しかし、選好獲得費用が存在する場合には、この仮定は成立しない。このような状況を扱うことが可能なマッチングの理論的モデルが必要である。

効率的なマッチングの失敗 マッチングの運営団体は、マッチングの失敗が生じるのを避けるため、できるだけ長い選好リストを提出するように勧めている。しかし、選好獲得費用が存在する場合、選好リストを長くすると、結果として不要な費用を費やすことになる。一方、選好獲得費用を節約しようとして、短い選好リストしか提出しなければ、マッチングの失敗が生じやすくなる。

これらの課題を解決するため、まず、選好獲得費用を明示的にモデルに取り込むことにした。選好獲得費用を削減するには、候補の中から、行けそうにない選択肢を早期に削除すればよく、マッチングを段階的に行い、先行する者の

マッチング情報を後続の者が利用することを考えた，新たなプロトコルの提案を行った．

また，提案したプロトコルを評価するためにシミュレーションを行い，実際に研修医配属の場面で用いられているプロトコルとの比較を行った．その結果，選好獲得に必要な費用が多くかかる場合に，既存のプロトコルに比べて提案プロトコルは選好獲得の費用を低減することが確認された．

本研究の主な貢献は以下の2点である．

選好獲得費用が存在する場合のマッチングモデルの提案 選好獲得費用が存在する場合に，学生側では，選好順位が決まっている者と決まっていない者に分割し，病院側では，学生が持つスキルによって選好順位が決まるというモデルを提案した．

効率的なマッチングプロトコルの提案 選好順位が決まっている者を先にマッチングし，そこでのマッチング情報を用いて選好順位を正確に決めていない学生の希望候補を削減する方法を提案した．提案方法は，元のプロトコルが持つ，学生側にとっては，正直に選好リストを提出するという戦略が最善という性質を引き継ぐことを示した．

上記の2点の貢献により，研修医配属や就職活動などのマッチングの場面において，選好獲得費用が存在する場合に，他者の選好を利用することによって効率的なマッチングプロトコルを設計することが可能となった．

提案したプロトコルでは，プレマッチングとその後のマッチングの2ステージに分かれているといえる．情報収集費用を削減するには，これをさらに細かく多段のステージに分割することも考えられる．ただし，その場合，マッチングの安定度は低下することになる．このあたりは，適用領域ごとに調整されることになる．

また，シミュレーションにおいては，選好が決まっている学生と選好が決まっていない学生が半数ずつの場合において示したが，両者の割合が変化したときに，削除される候補数やコストにおいてどのような変化が表れるかを示す必要がある．

学生のモデル化についても本研究では選好が決まっている学生と選好が決まっていない学生の2つに分類したが，さらに詳細化を行い，分類の種類を増加させた場合も考えられる．この場合においても本稿で提案したアイデアが有効となるのか検討する必要がある．

本稿においては例示の簡単化のために、既存プロトコルにおいても提案プロトコルにおいても学生と病院の数が等しい場合で扱った。しかし、実際には病院には1人以上の定員があり、また学生の総数よりも病院の総定員数のほうが多くなっている。また、研究室配属や就職活動など研修医配属以外の配属問題においても同様に総定員数の方が多い場合が多々ある。こういった場合に提案する手法が有効であるかを検討する必要がある。

謝辞

本研究を行うにあたり、熱心な御指導と御助言を賜りました松原繁夫准教授に厚く御礼申し上げます。また、日頃より多くの御助言を与えてくださりました石田亨教授に心より感謝致します。そして、日頃からさまざまなご協力と議論をしてくださいました石田・松原研究室の皆様には感謝致します。

参考文献

- [1] Gale, D. and Shapley, L.: College admissions and the stability of marriage, *American Mathematical Monthly*, Vol. 69, pp. 9–15 (1962).
- [2] 久保幹雄, 松井知己, 田村明久: 応用数理計画ハンドブック, 朝倉書店.
- [3] A.E.Roth.: Misrepresentation and Stability in the Marriage Problem, *Journal of Economic Theory*, Vol. 34, pp. 383–387 (1984).
- [4] A.E.Roth.: Conflict and Coincidence of Interest in Job Matching: Some New Results and Open Questions, *Mathematics of Operations Research*, Vol. 10, pp. 379–389 (1985).
- [5] A.E.Roth.: The College Admissions Problem is not Equivalent to the Marriage Problem, *Journal of Economic Theory*, Vol. 36, pp. 277–288 (1985).
- [6] A.E.Roth. and E.Peranson.: The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design, *National Bureau of Economic Research Cambridge, Mass.*