エラーレートを用いた麻雀プレイヤの実力推定

久米 洋輝[1,a)]　栗田 萌[2]　保木 邦仁[1]

概要：本研究の目的は、麻雀プレイヤの実力を少ない行動記録から推定することである．実力の推定は、麻雀人工知能 (AI プレイヤ) を用いてエラーレートを評価することによりなされる．エラーレートとはバックギャモンで用いられるプレイヤの実力の指標であり，これは 1 行動あたりの推定価値 (得点期待値) の減少平均値から見積もられる．現状，麻雀プレイヤの実力は平均順位やレート値から推定される．本研究では，エラーレートと平均順位の平均レート推定性能を比較した．その結果，おおよそ 16 試合の行動記録から評価したエラーレートと 500 試合の平均順位が同程度の推定性能である可能性が示された．

1. はじめに

現実社会では，人々の熟達度を測る様々な方法が用いられる．例えば，学生の学力を測るときには試験などが用いられ，野球選手の能力を測るときには打率などが用いられる．このような熟達度の計測は，ときとして計測対象となった人にも利益をもたらし得る．自身の熟達度を知ることにより，上達への目標設定を客観的に行うことが可能となるからである．熟達度の計測はまた，ゲームのエンターテイメント性向上にも貢献し得る．自身の熟達度とかけ離れたプレイヤとプレイすることを避けることが可能となるからである．人々の熟達度を測ることは思考力を競うようなゲームにおいても重要である．すでにチェスや囲碁，将棋といった 2 人完全情報ゲームにおいてプレイヤの実力を測る指標を推定する研究がいくつかある．しかし，多人数ゲームや不完全情報ゲームにおいては，このような研究事例の報告は少ない．

本研究では，4 人不確定不完全情報ゲームである麻雀に着目する．現在，実力を測る指標としてよく用いられるものに平均順位やレート値があるが，ある程度正確に実力を測るには麻雀の性質上，数千回程度の対戦が必要である．

インターネット麻雀サイト「東風荘」の一般的なプレイヤの平均順位は，おおよそ平均 2.45，標準偏差 0.15 の正規分布に従うことが知られている [1]．ここで，実力に応じてプレイヤを上級，中級，初級の 3 つに分けることを考える．一般に明確な指標はないが，上級プレイヤを平均順位 2.4 位未満，中級プレイヤを 2.4 位以上 2.6 位未満，初級プレイヤを 2.6 位以上のプレイヤと考えることとする．この場合，上級プレイヤはプレイヤ全体の上位約 25%，初級プレイヤは下位約 25%である．大半のプレイヤの平均順位は 2.1 から 2.9 の範囲内であり，偶然の要素がゲームプレイに強い影響を与えるという麻雀の性質もあり，どのプレイヤも同じような順位確率 (各順位おおよそ 0.25) をとることから，1 戦あたりの順位の分散は 1.25 程度と考えることができる．対戦数 N のプレイヤの平均順位の標準偏差を考える．標本サイズ (対戦数) が N，分散が A であるとき，標本平均の標準誤差は $\sqrt{\frac{A}{N}}$ と見積もることができる．つまり平均順位の標準誤差は $\sqrt{\frac{1.25}{N}}$ で見積もられる [1]．従って ± 0.1 位程度の精度で平均順位を推定するためには約 500 戦必要である．500 戦行うには，通常 250 時間以上必要である．

本研究では，バックギャモンでプレイヤの熟達度をはかるエラーレートという概念に着目する．バックギャモンと麻雀は，ゲームプレイに不確実性があるという点で類似している．そのため，エラーレートは麻雀プレイヤの熟達度を測る方法としても有望ではないかと期待される．

本論文では，2 章で思考力を競うようなゲームにおける実力推定の先行研究，3 章で本研究に関する基礎知識，4 章で本研究の目的，5 章で行った実験について述べる．

2. 関連研究

この章では，本研究に関する思考力を競うようなゲームにおいてプレイヤの実力推定を行っている先行研究について概要を述べる．

2.1 囲碁プレイヤの実力推定

荒木らは，13 路 1 局の棋譜からプレイヤの棋力を推定

[1] 電気通信大学
The University of Electro-Communications
[2] HEROZ 株式会社
[a)] k1731065@edu.cc.uec.ac.jp

する手法を提案した [2]．この手法は畳み込みニューラルネットワーク (CNN) を用いたものであった．

棋譜には，インターネット碁会所である「囲碁クエスト」のものを用いた．CNN を用いて推定したものはプレイヤのレート値である．N 手目の盤面情報を表すビット列を入力データ，プレイヤのレート値を目標値として CNN を訓練した．盤面情報は，(1) すでに置かれている黒石の位置が 1 でそれ以外が 0 の長さ 13×13 のビット列，(2) すでに置かれている白石の位置が 1 でそれ以外が 0 の長さ 13×13 のビット列，(3) 次の 1 手が 1 でそれ以外が 0 の長さ 13×13 のビット列，の 3 つで表現される．白番もしくは黒番のみに限定して盤面情報を入力するため，1 棋譜あたり $N/2$ の盤面が 3 種類であり，CNN への入力のビット数は $13 \times 13 \times (3N/2)$ である．CNN のネットワーク構成は入力側から，入力層，畳み込み層が 3 つ，全結合層，出力層からなっていた．畳み込み層のフィルタサイズは 3×3 であり，チャネル数はそれぞれ 12，8，8 であった．畳み込み層の活性化関数には正規化線形関数を，学習方法には確率的勾配降下法を用いた．N 手目までの情報を用いる場合，ネットワークの重みの数は $162N + 1,861$ 個である．最終的な達成度として，$N = 50$ の場合ではレート推定値の平均二乗誤差の平方根が 250 程度まで減少した．

この先行研究と本研究の類似点は，ゲームの行動記録から実力を推定している点である．ニューラルネットワークを用いている点，ゲーム AI プレイヤを用いた実力推定はしていない点，2 人確定完全情報ゲームを題材としている点が異なっている．

2.2 チェスプレイヤの実力推定

Guid らはチェスの AI プレイヤを用いて棋譜からチェスプレイヤの実力を順位付けする手法を提案した [3][4]．また，プレイスタイルを数値化する可能性も示した．

2006 年の報告では，チェスの AI プレイヤ「CRAFTY」の示す評価値を用いて 3 つの指標を計算した．CRAFTY の評価値はポーン何個分勝っているかを推定するものである．1 つ目は CRAFTY の示す最善手の評価値とゲーム記録で行われた行動の評価値の差を 1 行動あたりに平均化したもの (平均損失) である．ただし，評価値が $[-2, 2]$ の範囲外にある場合では計算を行わなかった．2 つ目はある状況における最善手の評価値と次善手の評価値の差 (ゲーム状況の複雑さ) である．3 つ目は，CRAFTY の示す最善手の評価値とゲーム記録で行われた行動の評価値の差が 1 以上であった行動の数である．これらの指標を用いて，チェスプレイヤの実力の順位付けとプレイスタイルの分析を行った．

2011 年の報告では，いくつかのチェスの AI プレイヤの探索の深さを変えながら 2006 年の報告と同様にプレイヤの実力の順位付けをし，結果に差が出るのか調査した．用いた 3 つのチェスの AI プレイヤは人間のトッププレイヤよりも強い AI プレイヤである．探索深さを変化させることによっても AI プレイヤの強さは変化する．1 つの悪手が平均損失に影響しすぎないように AI プレイヤの示す最善手の評価値とゲーム記録でとられた行動の評価値の差を最大値が 3 になるように修正した．結果は，AI プレイヤの強さを変化させても推定したプレイヤの実力の順位には差が見られないというものであった．また，AI プレイヤのプレイスタイルが推定結果に影響を与える可能性は確認されたが，これはチェスチャンピオンの順位付けに大きく影響を与えるものではなかった．

AI プレイヤを用いて実力を判断する点，ゲームの行動記録から実力を判断する点が本研究との類似点である．エラーレートを用いていない点，2 人確定完全情報ゲームを題材としている点が異なっている．

2.3 将棋プレイヤの実力推定

山下は，将棋の AI プレイヤを用いて棋譜からプレイヤの実力を推定する手法を提案した [5]．推定に用いた棋譜は「将棋の棋譜でーたべーす」に 2014 年 6 月 30 日までに登録されていたものである．推定した実力はレート値である．AI プレイヤには「Bonanza 6.0」と「GPSFish」を用いた．AI プレイヤの示す最善手と評価値から，平均悪手，平均好手，AI プレイヤとの一致率などを計算した．将棋の AI プレイヤの評価値は駒の損得を推定するものである．計算時には評価値を 100 で割ったものを用いた．平均悪手は AI プレイヤの示す最善手の評価値とゲーム記録で行われた行動の評価値の差を 1 行動あたりに平均化したものである．ただし平均悪手の計算は，ゲーム記録の行動が AI プレイヤの最善手と異なる行動である場合かつ評価値が最善手よりも下がる場合で行った．また，ゲーム記録の行動が 40 手目以降かつ評価値の絶対値が 10 未満の場合のみを用いた．AI プレイヤの行う探索の深さを変えることで，AI プレイヤの強さを変化させて計算を行った．Bonanza と GPSFish の 2 つの AI プレイヤの評価値を用いて同様の計算を行った．

求めたいくつかの結果から，山下は平均悪手がもっとも実力を推定できると考察した．平均悪手とレート値の関係が一次関数になると仮定し，探索深さ 11 の Bonanza での平均悪手の結果から最小二乗法を用いて以下の式を得た．

$$\text{レート値} = -3148 \times \text{平均悪手} + 4620 \tag{1}$$

このときのレート値とは，インターネット将棋対局サイトの「将棋倶楽部 24」でのレート値である．また，AI プレイヤの探索深さを変えることで AI プレイヤの強さを変化させた場合でも，精度に誤差はあるものの AI プレイヤ自身よりも強いプレイヤも実力を判断できる可能性が高いことが示された．

類似点と相違点は 2.2 節のものと同じである．

3. 基礎知識

この章では，麻雀とバックギャモンのエラーレートについて述べる．

3.1 麻雀

この節では，麻雀のルール概要，行動記録を収集したインターネット麻雀サイト，AI プレイヤについて述べる．

3.1.1 ルール概要

麻雀は牌を用いて遊ぶ 4 人不確定不完全情報ゲームである．牌はマンズ，ピンズ，ソーズ，字牌の 4 つに分けられる．マンズ，ピンズ，ソーズはそれぞれ 1 から 9 までの 9 種である．字牌は東，南，西，北，白，發，中の 7 種である．全部で牌種は 34 種からなる．牌は各種 4 枚ずつあり，合わせて 136 枚である．マンズ，ピンズ，ソーズのことをまとめて数牌と呼ぶ．本研究では，数牌それぞれの 5 の牌 1 枚ずつを赤牌と呼ばれるボーナス牌に変更したルールを扱う．

各プレイヤは牌 13 枚 (手牌) を持ち，順番に牌を 1 枚引いて不要な牌を 1 枚捨てること (打牌) を繰り返す．手牌の牌種は，その手牌を所持するプレイヤにしか見えない情報である．各プレイヤは牌山と呼ばれる場所から順番に牌を引いてくる．牌山の牌は各プレイヤが引くまでどのプレイヤにもそれがどの牌種かはわからない．あるプレイヤが手牌を特定の形にそろえ，役が 1 つ以上あるとき，そのプレイヤは得点する．これを和了と呼ぶ．役とは手牌のそろえ方や状況によって成立するものである．王牌と呼ばれる 14 枚の牌以外の牌山からすべての牌を引いてもだれも和了しない場合を流局と呼ぶ．誰かが和了または流局するまでを局と呼ぶ．麻雀では複数回の局を行い，最終的な持ち点からプレイヤの順位が決まる．本研究では，基本的な局数が 8 局である東南戦を扱う．

基本的な手牌のそろえ方は 4 面子 1 雀頭を作ることである．面子とは同じ数牌または字牌の 3 枚で構成される．面子には順子と刻子の 2 種類がある．順子とはマンズの 123 やピンズの 678 のように同じ数牌が 3 つ順に並んだものである．刻子とはピンズの 222 やソーズの 555 といった同じ牌種を 3 つそろえたものである．雀頭とは同じ牌種を 2 つそろえたものである．

麻雀の行動選択肢には打牌以外にも副露と呼ばれるものがある．副露にはポン，チー，アンカン，ダイミンカン，カカンの 5 種類がある．副露とは手牌の中の面子を確定させて，全プレイヤにこれを公開する行為である．ポンとは，自分が 2 枚以上持つ牌種を誰かが捨てたときに行えて，捨てられたその牌を取得することでその牌 3 枚を刻子の形で確定し，他の全プレイヤにこれを公開する．チーとは，左隣のプレイヤから捨てられた牌を取得することで順子を確定し，その順子を他の全プレイヤに公開する．ポンまたはチーをした場合，牌を取得した後に手牌から 1 枚牌を捨てる．アンカンとは，手牌で同じ牌を 4 枚持っているときに行うことができ，その 4 枚を刻子として確定し，全プレイヤに公開する．ダイミンカンとは，同じ牌種を 3 枚持っていて，その牌種を誰かが捨てたときに行うことができる．その牌 4 枚を刻子として確定し，全プレイヤに公開する．カカンとは，自分がポンしている牌種をさらに持っているときに行うことができ，ポンしている面子にその牌種を加えた 4 枚で刻子を確定する．アンカン，ダイミンカン，カカンを行った後には嶺上と呼ばれる牌山とは異なる場所から牌を 1 枚引いて，その後 1 枚捨てる．また，アンカン，ダイミンカン，カカンを行ったあとにはドラと呼ばれるボーナス牌が増える．他のプレイヤが牌を捨てたとき、副露またはアガリできる場合でもそれらの行動をとらない選択も可能であり、その行動をパスと呼ぶ．

表 1　プレイヤの可能な行動

打牌	手牌から牌を 1 枚捨てる
ポン	他プレイヤの捨てた牌をもらい，刻子を確定
チー	他プレイヤの捨てた牌をもらい，順子を確定
アンカン	手牌の同じ牌 4 枚を刻子として確定
ダイミンカン	他プレイヤの捨てた牌をもらい，同じ牌 4 枚を刻子として確定
カカン	ポンして確定した刻子にもう 1 枚同じ牌を加えて刻子として確定
リーチ	あと 1 枚で手牌 14 枚がそろい，手牌を入れ替えないと他プレイヤに宣言
和了	手牌 14 枚がそろい得点
パス	他プレイヤの捨てた牌に対して、副露、和了をしない

麻雀の特殊な行動選択肢として，リーチというものがある．リーチとは，1,000 点を払い，手牌が後 1 枚でそろうことと手牌をこれ以上入れ替えないことを他プレイヤに宣言する行動である．リーチと同時に打牌を行う．これによって役が 1 つ成立する．払われた 1,000 点は次に和了したプレイヤが獲得する．また，アンカンを除く副露を行ったプレイヤはリーチすることができない．

表 1 は，麻雀のゲーム進行においてプレイヤの可能な行動を表にしたものである．プレイヤの可能な行動は打牌，副露，リーチ，和了のいずれかである．表 2 はプレイヤが行動可能なタイミングとその行動を表にしたものである．副露やリーチは行動可能なタイミングかつ条件がそろったときにしか行えない．表 3 は，麻雀のゲーム状況を表す主要な要素の一覧である．

3.1.2 天鳳

天鳳とはインターネット麻雀サイトの 1 つである [6]．累計登録ユーザー数は 480 万人以上，180 日以内にプレイしているアクティブユーザー数は 34 万人以上であり，最もよ

表 2　プレイヤの意思決定点

意思決定点	可能な行動
牌を引いたとき	打牌，アンカン，カカン，リーチ，和了
正面，右プレイヤが打牌したとき	ポン，ダイミンカン，和了，パス
左プレイヤが打牌したとき	チー，ポン，ダイミンカン，和了，パス

表 3　あるプレイヤ視点でのゲーム状況の主要な構成要素

手牌	和了するためにそのプレイヤがそろえる牌
捨て牌	プレイヤ 4 人が打牌した牌
副露牌	プレイヤ 4 人が副露した確定面子
ドラ表示牌	ボーナス牌であるドラを決める牌
親番	局の最初に牌を引くプレイヤ
持ち点	プレイヤそれぞれの点数
リーチ状況	プレイヤそれぞれのリーチ有無
局数	現在のゲーム進行度
本場	親以外のプレイヤが何局和了していないか
供託	まだどのプレイヤも獲得していない，リーチによって払われた点数

く遊ばれているインターネット麻雀サイトの 1 つである．天鳳にはプレイヤの実力指標として，段位とレート値がある．段位は対戦結果に応じて上下し，対戦で得られるポイントが一定以上貯まると昇段，失うと降段する．段位は新人からスタートし，9 級，8 級，…，初級，初段，二段，…，十段，天鳳位と上昇していく．レート値は対戦結果と相手プレイヤのレート値に応じて上下する．レート値の初期値は 1,500 である．レートの変動値の計算式は以下の通りである．

$$\text{レートの変動値} = \text{対戦数補正} \times (\text{対戦結果} + \text{補正値}) \tag{2}$$

対戦数補正は 400 戦以上では 0.2 で固定値である．400 戦未満の場合では以下の式で計算する．

$$\text{対戦数補正} = 1 - \text{対戦数} \times 0.002 \tag{3}$$

対戦結果は 1 位 $+30$, 2 位 $+10$, 3 位 -10, 4 位 -30 となる．補正値は次のように計算する．

$$\text{補正値} = (\text{卓の平均レート値} - \text{自分のレート値})/40 \tag{4}$$

天鳳では，プレイヤの段位とレート値によって対戦できる場所が限定されている．段位 7 段以上かつレート値 2,000 以上のプレイヤは鳳凰卓，4 段かつレート値 1,800 以上のプレイヤは特上卓，1 級以上のプレイヤは上級卓，それ以外は一般卓で打つことができる．

天鳳では，麻雀の行動記録である牌譜を生成している．プレイヤは対戦を行うと，自分の対戦の牌譜を得られる．自分の対戦以外の牌譜は鳳凰卓の牌譜のみ公開されており，容易に手に入れることができる．しかし，鳳凰卓以外の牌譜は一般に公開されていない．有志のプレイヤが牌譜を提供するような場もないため，初級者中級者の牌譜を収集することは困難である．

3.1.3　AI プレイヤ

本研究では「Ako_Atarashi」という AI プレイヤを用いる．Ako_Atarashi は栗田が作成した AI プレイヤである [7][8][9]．ゲーム状況を与えると，可能な行動の価値を推定し，推定した価値が最も大きな行動を行う．推定する価値は，1 位 +90，2 位 +30，3 位 −30，4 位 −90 と順位点が与えられるとしたときの順位点の期待値である．天鳳では，プレイヤは通常，段位に関するポイント期待値を最大化するようにプレイするため，Ako_Atarashi と天鳳プレイヤの最大化したい価値は厳密には異なる．

Ako_Atarashi は，AI プレイヤの「爆打」と同等の実力を持つ AI プレイヤである．Ako_Atarashi が爆打と 2 対 2 の東風戦を 3000 ゲーム行った結果を表 4 に示す[*1]．爆打は天鳳において，一般的に上級者の指標の 1 つとされる 7 段レート 2,000 を達成した AI プレイヤであり，人間の上級者と同等以上の実力を持つ [10][11]．つまり Ako_Atarashi も人間の上級者と同等以上の実力を持つ十分に強い AI プレイヤである．

Ako_Atarashi はゲーム状況によっては、価値推定を行わずにルールに基づく選択をするため、推定価値を出力しない場合もあるが、経験的に 94%の割合で、推定価値を出力する。エラーレートの計算には、精度の高い推定価値を高い割合で出力する AI が必要であり、筆者の知る限り Ako_Atarashi は現在最も目的にあった AI である。

表 4　Ako_Atarashi の爆打との対局結果.

1 位率	2 位率	3 位率	4 位率	平均順位
0.260	0.255	0.241	0.243	2.469 ± 0.024

3.2　エラーレート

エラーレートとは，バックギャモンプレイヤの実力指標の 1 つである．1 行動あたりの推定価値の減少平均値を表しており，値が 0 に近いほどそのプレイヤは優秀であるといえる．エラーレートは行動記録と十分な実力をもつ AI プレイヤから計算されるものである．AI プレイヤは，得点期待値 (価値) を推定するために用いられる．

バックギャモンの商用プログラムの 1 つである eXtreme Gammon が行うエラーレートの計算方法を以下に示す [12]．バックギャモンプレイヤの意思決定過程をマルコフ決定過程として考える [13]．非終端状態の集合を S，状態 $s \in S$ における行動集合を $A(s)$ とする．複数の行動記録から状

[*1] 2 対 2 の対局を行い、1 プレイヤあたりの平均順位を求めると、その分散は 1 対 3 の対局より小さくなる．1 回の対局において 4 つの AI プレイヤは同一の計算機上で計算を行ったが、計算機の性能や状況次第で対局の結果は変化し得る．

態 s と行動 a の対 (s,a) がいくつも採取でき，対の数が N 個であったとする．n 番目の状態行動対は (s_n,a_n) と書く．n は $1 \leq n \leq N$ である．状態と行動の対の集合 D は以下のように表す．D には複数の行動記録から採取した対が属している．

$$D = \{(s_1,a_1),\ldots,(s_N,a_N)\} \tag{5}$$

状態 s において行動 a を選択したときの eXtreme Gammon が推定する価値を $Q(s,a)$ とする．エラーレート $e(D)$ の計算式は以下のように表す．

$$e(D) = \frac{500}{|D|}\sum_{(s,a)\in D}\max_{a'\in A(s)} Q(s,a') - Q(s,a) \tag{6}$$

ただし 1 ポイントマッチの場合のエラーレートはさらに 1.5 倍する．

eXtreme Gammon の推定する価値は，ゲームの設定によって異なる．アンリミットの場合では得点期待値，ポイントマッチの場合では勝率である．D に含まれる状態行動対の採取は駒の移動，ダブルの提案，ダブルのテイク選択のときに行う．状態行動対を採取する条件は，$|A(s)| \geq 2$ かつ $\max_{a\in A(s)} Q(s,a) - \min_{a\in A(s)} Q(s,a) \geq 0.001$ である．また，ダブルの提案では，提案する前の推定価値の方が相手がテイクした後より 0.2 以上高い場合，ダブルを提案する前の推定価値の方が相手がパスしたときより 0.2 以上高い場合，ダブルを提案する前も相手がテイクした後も推定価値がとても低い場合においては状態行動対を採取しない．

4. 目的

本研究の目的は，麻雀プレイヤの実力を少ない行動記録から推定することである．おおまかにプレイヤの実力を推定しようとすると 500 戦程度必要であるが，通常 250 時間程度かかってしまう．

本研究では，少ない行動記録から AI プレイヤを用いてエラーレートを計算し，プレイヤの実力推定を行う．その際，麻雀プレイヤの実力の正しい値を得ることは困難であるが，本研究で行動記録を収集した天鳳での 500 戦以上にわたって平均化されたレート値 R_p (平均レート) をそのプレイヤの最も信頼できる実力の指標とする．これは，3.1.2 節で述べた通り，プレイヤのレベルごとに対戦できる場所が分かれているため，平均順位はプレイヤの実力の指標として適切ではない可能性がある一方で，レート値は他プレイヤのレベルによってレート変動値が変化するため，プレイする場所が異なっていてもある程度実力を表していると考えられるためである．そのうえで，エラーレートの平均レート推定性能を相関関数と決定係数から評価する．また，この推定性能を平均順位から平均レートを推定する場合と比較する．

5. 実験

この章では，麻雀の行動記録を収集し 500 戦以上の対戦を行っていたプレイヤ p に対して R_p を計算し，$n = \{1,2,4,8,16,32,64,128,500\}$ に対しプレイヤ p の n 戦分のエラーレート，n 戦分の平均順位の 2 つの指標と R_p の相関を計算した．性能の評価は相関係数，決定係数により行った．エラーレートの計算には Ako_Atarashi を用いた．収集した行動記録は，天鳳での東南戦喰いタンあり赤ありのルールの行動記録に対応する．

5.1 実験方法

この節では，エラーレート，相関係数，決定係数の計算方法について説明する．

本実験では，Ako_Atarashi の推定する価値を用いてエラーレート $e(D)$ を計算する．計算方法は 3.2 節で紹介した eXtreme Gammon での計算方法とほぼ同じである．異なる点は，推定価値 $Q(s,a)$ が Ako_Atarashi が推定する価値である点，状態行動対の採取の仕方が Ako_Atarashi が推定価値を出力する場合には $|A(s)| \geq 2$ であるときに状態行動対を採取し，推定価値を出力しない場合では採取を行わない点，AI プレイヤの推定価値に関わらず採取を行う点，計算式が以下の式となる点である．

$$e(D) = \frac{1}{|D|}\sum_{(s,a)\in D}\max_{a'\in A(s)} Q(s,a') - Q(s,a) \tag{7}$$

エラーレートと平均レートの相関係数と決定係数の計算について述べる．採取した状態行動対の全体集合を $D_A = \{D_1,\ldots,D_P\}$ と書く．ただし $D_p = \{D_{p1},\ldots,D_{pM_p}\}$, D_{pi} はプレイヤ p の行動記録から得られた i 番目の状態行動対の集合で，n 戦分の対からなる．さらに $M = \sum_{p=1}^{P} M_p$, M_p はプレイヤ p のデータ数とする．

$e(D_{pi})$ のデータ全体にわたって計算された平均値を e_{ave} とする．式は以下の通りである．

$$e_{\mathrm{ave}} = \frac{1}{M}\sum_{p=1}^{P}\sum_{i=1}^{M_p} e(D_{pi}) \tag{8}$$

M_p を考慮した R_p の重み付き平均値 R_{ave} は以下のように計算される．

$$R_{\mathrm{ave}} = \frac{1}{M}\sum_{p=1}^{P} M_p R_p \tag{9}$$

エラーレート $e(D_{pi})$ と平均レート R_p の相関係数 r は以下の式で計算する．

$$
\begin{aligned}
r &= \frac{S_{xy}}{S_x S_y} \\
S_{xy} &= \frac{1}{M}\sum_{p=1}^{P}\sum_{i=1}^{M_p}(e(D_{pi}) - e_{\mathrm{ave}})(R_p - R_{\mathrm{ave}}) \\
S_x &= \sqrt{\frac{1}{M}\sum_{p=1}^{P} M_p(e(D_{pi}) - e_{\mathrm{ave}})^2} \\
S_y &= \sqrt{\frac{1}{M}\sum_{p=1}^{P} M_p(R_p - R_{\mathrm{ave}})^2} \qquad (10)
\end{aligned}
$$

相関係数の値は，推定値と目標値の正の相関が大きいほど 1 に近づき，負の相関が大きいほど -1 に近くなる．無作為に R_p を推定すると，相関係数は 0 である．相関係数 r の標準誤差は，r の標準偏差を $(1-r^2)/\sqrt{M}$ により近似的に計算して見積もった [14].

決定係数は以下の式で計算する [15].

$$
\text{決定係数} = 1 - \frac{\sum_{p=1}^{P}\sum_{i=1}^{M_p}\{R_p - f(e(D_{pi}))\}^2}{\sum_{p=1}^{P} M_p(R_p - R_{\mathrm{ave}})^2} \qquad (11)
$$

ただし関数 f はエラーレートから平均レートを推定する関数である．決定係数の値は，推定が理想的な場合 1 となる．値が 0 の場合では R_{ave} を出力する推定と推定精度が同程度であることを表す．

平均順位と平均レートの相関係数，決定係数を計算する場合には，式 (10)，(11) のエラーレートを平均順位に，エラーレートの平均値 e_{ave} を平均順位の平均値に，関数 f を平均順位から平均レートを推定する関数に置き換えて相関係数，決定係数の計算を行う．

5.2 実験結果

この節では，エラーレートと平均順位の平均レート推定性能評価の実験結果を示す．本節の表に示す不確かさは標準誤差から見積もられた 95%信頼区間である．

天鳳の一般卓から鳳凰卓まですべてのレベルでの牌譜を集めた．その中で，500 戦以上対戦しているプレイヤは 13 人であった．13 人の試合数は合計で 7,934 であった．本実験のデータ全体は，この 13 人の誰かが対戦した牌譜から構成した (表 5 参照)．Ako_Atarashi のレート値は 2000 以上であることが予想されるので，Ako_Atarashi は表 (5) の 13 人の誰よりも実力が高いと考えられる．

エラーレートと平均レートの関係が一次関数になると仮定し，500 戦以上の行動記録 D_p から得られたエラーレートと平均レートから最小二乗法を用いて式 (12) を得た．

$$
\text{レート値} = -1648.2 \times \text{エラーレート} + 2296.0 \qquad (12)
$$

同様に，平均順位と平均レートの関係が一次関数になると仮定し，500 戦以上の行動記録 D_p から得られた平均順位と平均レートから最小二乗法を用いて式 (13) を得た．

表 5 データ全体の統計情報

p	対戦数	平均値とエラーレートを計算した標本の大きさ	平均レート	平均順位	エラーレート
1	1177	533	1969.9	2.49	0.247
2	2802	580	1936.2	2.49	0.277
3	680	674	1896.7	2.39	0.231
4	529	529	1863.1	2.45	0.282
5	715	536	1833.5	2.40	0.275
6	1065	736	1732.9	2.47	0.292
7	3128	549	1726.3	2.45	0.315
8	1820	673	1710.1	2.46	0.314
9	6024	803	1700.3	2.52	0.336
10	1033	505	1695.3	2.48	0.377
11	1114	562	1571.6	2.53	0.466
12	12085	658	1553.3	2.50	0.475
13	626	596	1453.6	2.67	0.486

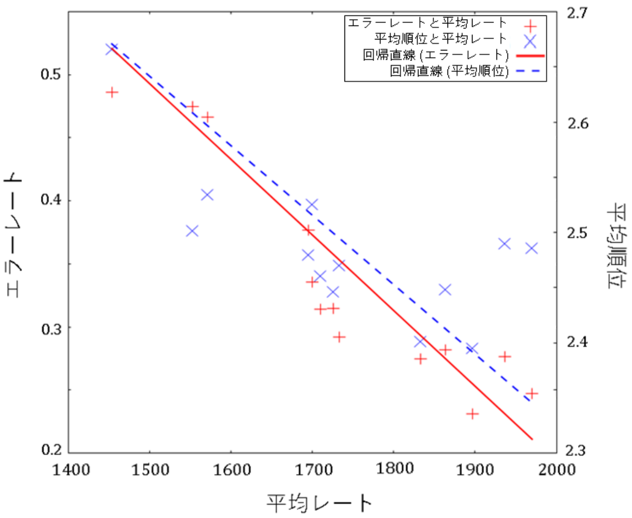

図 1 表 5 の統計情報の散布図．+の赤点がエラーレートと平均レートのデータ点を表し，×の青点が平均順位と平均レートのデータ点を表す．実線（赤）と破線（青）はそれぞれ平均レートの、エラーレートと平均順位に関する回帰直線である．

$$
\text{レート値} = -1580.8 \times \text{平均順位} + 5669.7 \qquad (13)
$$

500 戦以上のエラーレートと平均レートの散布図，500 戦以上の平均順位と平均レートの散布図を図 1 に示す．図の実線（赤）はエラーレートと平均レートの式 (12) であり，破線（青）は平均順位と平均レートの式 (13) である．エラーレート，平均順位ともに平均レートが高くなると小さくなる傾向が確認できる．

表 6 は，n 戦分の行動記録 D_{pi} から得られたエラーレートと平均レートから相関係数と決定係数を求めた結果である．決定係数は式 (12) を用いて計算を行った．対戦数 n が大きくなるに従って相関が大きくなることが確認された．対戦数 1 でもエラーレートと平均レートの間に弱い相関があり，対戦数 16 では強い相関があると言える．対戦数が 4 以下の場合では，決定係数が負の値になっていることから式 (12) による推定はうまくいっていない様子がうかがえる．対戦数 n が大きくなるに従い，決定係数が大き

表 6 エラーレートの平均レート推定性能

対戦数 n	相関係数	決定係数
1	-0.374 ± 0.023	-11.44
2	-0.49 ± 0.03	-1.42
4	-0.59 ± 0.03	-0.40
8	-0.69 ± 0.04	0.15
16	-0.76 ± 0.05	0.41
32	-0.78 ± 0.06	0.51
64	-0.82 ± 0.07	0.62
128	-0.84 ± 0.09	0.66
500	-0.93 ± 0.07	0.87

表 7 平均順位の平均レート推定性能

対戦数 n	相関係数	決定係数
1	-0.041 ± 0.022	-142.89
2	-0.06 ± 0.03	-71.77
4	-0.08 ± 0.04	-35.77
8	-0.12 ± 0.06	-16.71
16	-0.17 ± 0.09	-8.08
32	-0.23 ± 0.12	-3.85
64	-0.30 ± 0.16	-1.67
128	-0.41 ± 0.21	-0.47
500	-0.7 ± 0.3	0.48

くなっていることから式 (12) による推定がある程度はうまくいっていると考えられる．

平均順位と平均レートの関係は表 7 の通りである．決定係数の計算は式 (13) を用いて行った．表 7 から，対戦数を 32 まで増やしても平均順位は平均レートとの相関がほとんど確認できなかった．そのため，行動記録が少ない場合では，平均順位は実力を表す指標として適切ではないと考えられる．対戦数 500 の場合では，誤差が大きいものの平均順位はある程度実力を推定できていると考えられる．

また，表 6 の 16 戦での相関係数と表 7 の 500 戦での相関係数を比較すると，16 戦で計算したエラーレートの方が有意な差ではないが，大きな相関係数の値を持つ．また，決定係数の値も同等程度である．これらの結果から，16 戦の行動記録から得られたエラーレートは，500 戦の行動記録から得られた平均順位と同程度の性能であると考えられる．

6. まとめ

本研究ではエラーレートを用いた麻雀プレイヤの実力推定を行い，平均レート推定性能を評価した．その結果，16 戦のエラーレートは 500 戦の平均順位と同等の推定性能を持つ可能性があることが明らかとなった．麻雀において，大半のプレイヤの実力は長期の平均順位を基準に 2.5 ± 0.4 程度であること，1 人のプレイヤの 500 戦の平均順位は ± 0.1 程度の範囲に収まることを考えると，本研究の結果からエラーレートを用いることでプレイヤの実力を 8 時間程度のゲームプレイからおおまかに上級，中級，初級の 3 つ程のカテゴリに分類できると予想される．

参考文献

[1] とつげき東北．科学する麻雀．講談社，2014．
[2] 荒木伸夫，保木邦仁，村松正和．畳み込みニューラルネットワークを用いた囲碁における 1 局の棋譜からの推定．情報処理学会論文誌，Vol. 57，No. 11，pp. 2365-2373，2016．
[3] Matej Guid，Ivan Bratko．Conputer Analysis of World Chess Champions．ICGA Journal，Vol. 2，No. 29，pp. 65-73，2006．
[4] Matej Guid，Ivan Bratko．Using Heuristic-Search Based Engins for Estimating Human Skill at Chess．ICGA Journal，Vol. 2，No. 34，pp. 71-81，2011．
[5] 山下宏．将棋名人のレーティングと棋譜分析．The 19th Game Programming Workshop 2014，pp. 9-16，2014．
[6] オンライン対戦麻雀 天鳳．url: http://tenhou.net．最終アクセス 2019．
[7] 栗田萌，保木邦仁．有効非巡回グラフで表現された 1 人麻雀の探索アルゴリズム．The 22nd Game Programming Workshop 2017，pp. 42-49，2017．
[8] 栗田萌，保木邦仁．麻雀 1 局の目的に応じた抽象化と価値推定からなるプレイヤの開発．The 22th Game Programming Workshop 2017，pp. 72-79，2017．
[9] 栗田萌，保木邦仁．麻雀における他家の手牌と待ちの予測に基づく放銃率推定．情報処理学会研究報告，Vol. 2017-GI-38，No. 5，2017．
[10] 水上直樹，中張遼太郎，浦晃，三輪誠，鶴岡慶雅，近山隆．多人数性を分割した教師付き学習による 4 人麻雀プログラムの実現．情報処理学会論文誌，Vol. 55，No. 11，pp. 2410-2420，2014．
[11] 水上直樹，鶴岡慶雅．期待最終順位の推定に基づくコンピュータ麻雀プレイヤの構築．The 20nd Game Programming Workshop 2015，pp. 179-186，2015．
[12] eXtreme gammon．url: http://www.extremegammon.com/extremegammon2.pdf．最終アクセス 2019．
[13] Csaba Szepesvári，小山田創哲，前田新一，小山雅典．速習 強化学習 基礎理論とアルゴリズム．共立出版，2017．
[14] A. L. Bowley, Journal of the American Statistical Association, Vol. 23, No. 161, pp. 31-34, 1928.
[15] Tarald O. Kvalseth．Cautionary Note about R^2．The American Statistician，Vol. 39，No. 4，pp. 279-285，1985．