

大阪府立大学 現代システム科学域 知識
情報システム学類 森田研究室

朴 太一, 足立 匠, 名越 翔

選手の成績データを用いた エージェントベースドシミュレーションによる ボートレースの結果予測

ボートレースの競技としての特徴

- ・ 勝敗の約7割はスタートから最も近い回るポイント(第1ターンマーク)までの間に決まり、第1ターンマークを回った時点でそのレースの大まかな着順が決まる
したがって、第1ターンマークを各選手がどのように回るか(展開)を予測することが結果を予想する上で最も重要である
- ・ コースの内側を走る1号艇が最も有利な位置で、逆に6号艇が外側の最も不利な位置となる1号艇の勝率が70%ほどあり、逆転が起こりうる場合というのは
1号艇の勝率と他の選手の勝率が大きく離れている場合(1号艇の方が低い)
ある選手のスタートタイムが他の選手と大きく差が開いた場合である
- ・ 選手の能力が高いと、戦略の幅が広がることやターンマークを回るスピードが早くなることで1着になる確率が比較的高く、1着でなくてもレースタイムは短くなるため選手の能力からレースタイムを算出する

シミュレーションの目的

- ・目的

逆転が発生する予測困難な状況をシミュレーションを活用して予測すること

- ・予測条件

- ・レースの進入が1～6の順番であるレース(過去のレースの67%が該当)
- ・天候・気象条件は考慮しない(各選手が受ける影響は等しいため)
- ・レースタイムとスタートタイミングを基にシミュレーションを実行

概念モデル

・エージェント:ボートレーサー

・エージェントの想定:

1. 実際の選手は様々であるが, 入手可能なデータからは, 選手の属性と直近1年程度過去の戦績データがレースタイムに大きく影響すると考えられる.
2. そのような選手の属性データと過去の戦績は, 類似したいくつかのパターンに分類することができ, それによって類似した成績(レースタイム)となると考えられる.
3. 2による各選手のレースタイムが予測されたとしても, 6名の選手がレースを行うため6名の選手のレースタイムを予測したうえで順位を決定する必要がある.
4. そのため単に予測だけでなく, 演繹的に結果を計算する模擬レースを行うことで, よりリアルに近づくと考えたまた、展開予測のための要素としてスタートタイミングを実際の分布から確率で付与する

・選手の分類

クラスタリングのランク分け(6クラスタ)による、選手の分別

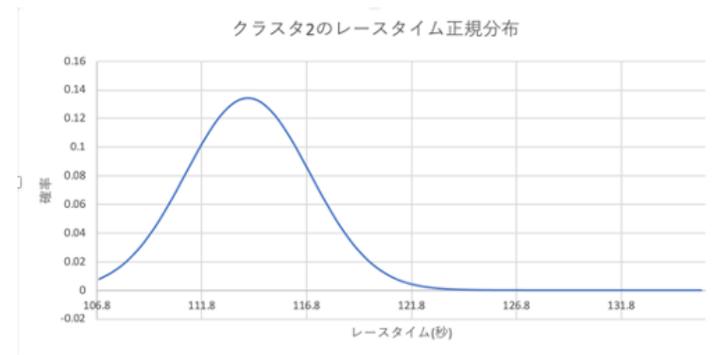
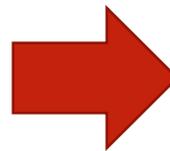
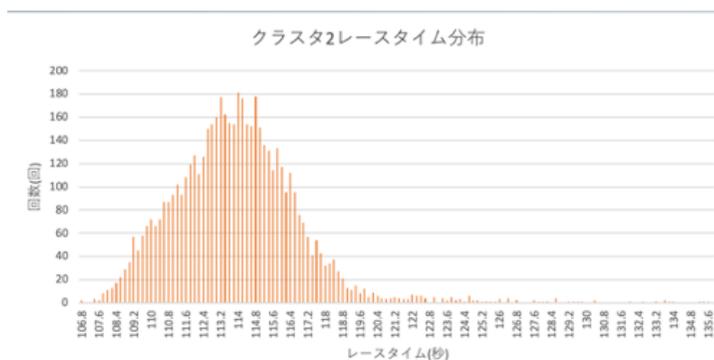
※各クラスタは性別/年齢/体重/身長/勝率/複勝率(2着以内に入る確率)/平均スタートタイミングによるものとする

クラスタリング結果

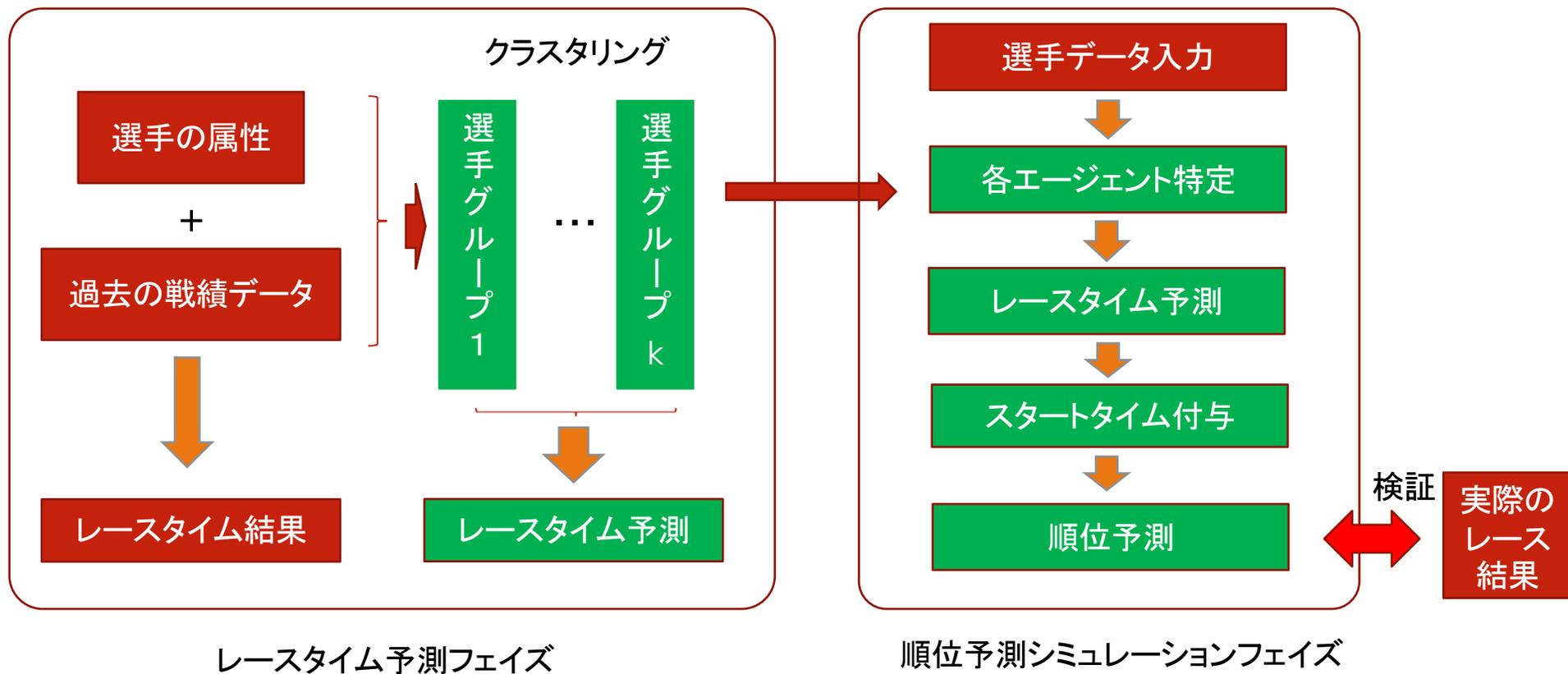
クラスタごとのレースタイム平均(秒)と不偏標準偏差

	女性クラスタ (199人)	熟年クラスタ (259人)	常勝クラスタ (307人)	熟練クラスタ (248人)	次世代クラスタ (395人)	新人クラスタ (150人)
平均	113.7	114.0	112.0	112.5	113.5	115.5
不変標準偏差	3.26	2.97	3.12	2.96	3.11	3.25

実際のレースタイム分布と正規分布の例



順位予測のための2段階提案モデル



エージェントの設定(概要)

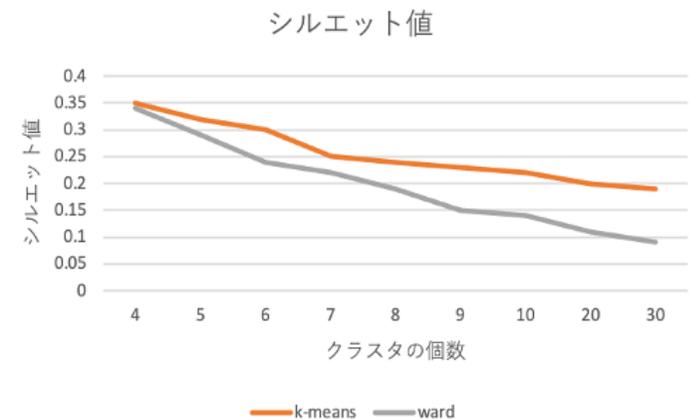
1. 各選手データ(性別/年齢/体重/身長/勝率/複勝率/平均スタートタイミング)からエージェントのクラスタを決定
2. 各エージェントのレースタイムとスタートタイムを予測
各クラスタの各コースの分布から確率的に予測
3. 予測レースタイムと予測スタートタイミングから速さを計算し、エージェントに付与

1. 選手の情報からクラスタへの分類

VAPを用いてk-means法と階層型クラスタリング(ward法)で検討した結果, シルエット係数, 各クラスタにおける人数の偏り, 各クラスタ内で項目による分散などからk-means法によるクラスタ数6のクラスタリングを採用

表:k-means法のクラスタ数6における各項目の不変標準偏差

ID	人数	性別	年齢	身長	体重	勝率	複勝率	平均ST
1	199人	0	9.94	4.53	3.12	0.14	0.14	0.027
2	259人	0	5.94	3.61	3.18	0.05	0.05	0.020
3	307人	0	6.12	3.59	2.35	0.06	0.07	0.015
4	248人	0	5.69	3.70	2.25	0.05	0.06	0.017
5	395人	0	4.89	3.81	2.62	0.06	0.06	0.019
6	150人	0	4.65	4.07	3.26	0.08	0.06	0.026



各クラスタの特徴

- 1:すべてが女性
- 2:年齢が高め。体重が重い。
- 3:勝率が最も高い
- 4:勝率が2番目に高い
- 5:比較的に若い選手が多い
- 6:新人選手が多い

表:k-means法でクラスタ数6での各項目の平均値

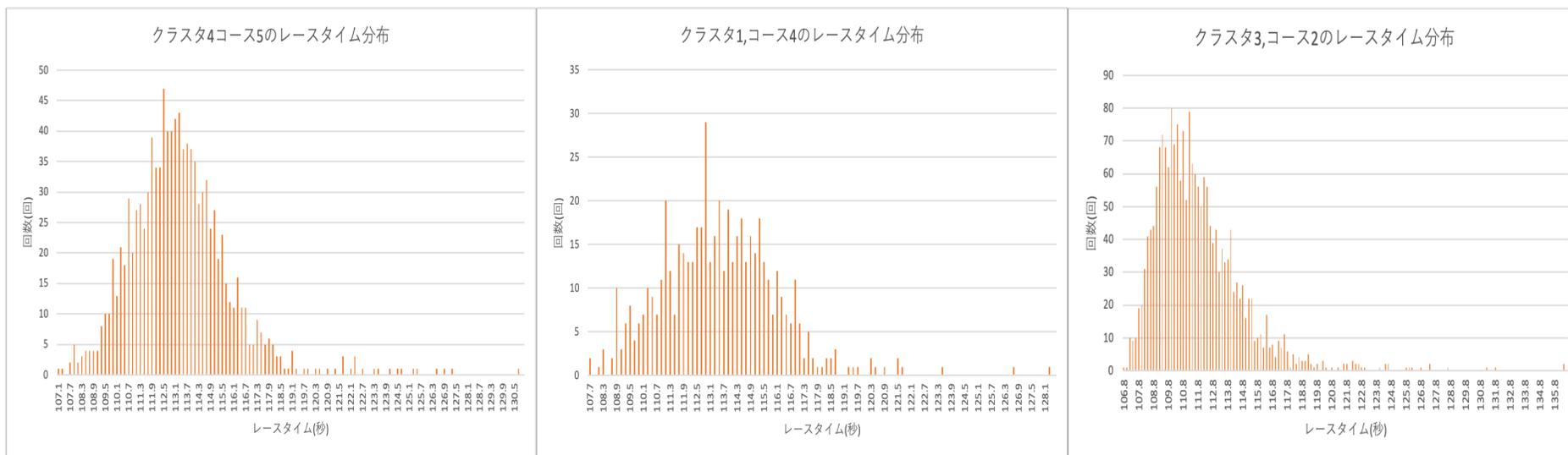
ID	クラスタ	人数	性別	年齢	身長	体重	勝率	複勝率	平均ST
1	女性クラスタ	199人	女性	35.1歳	157.2cm	47.9kg	0.453	0.262	0.181
2	熟年クラスタ	259人	男性	49.1歳	163.8cm	55.1kg	0.417	0.201	0.188
3	常勝クラスタ	307人	男性	36.0歳	166.1cm	53.4kg	0.669	0.494	0.149
4	熟練クラスタ	248人	男性	47.0歳	163.0cm	53.5kg	0.575	0.388	0.167
5	次世代クラスタ	395人	男性	32.8歳	166.7cm	54.2kg	0.482	0.280	0.169
6	新人クラスタ	150人	男性	25.3歳	165.6cm	54.6kg	0.265	0.074	0.186

※ST:スタートタイミング

2. レースタイム・スタートタイミングの予測

1. レースタイム予測

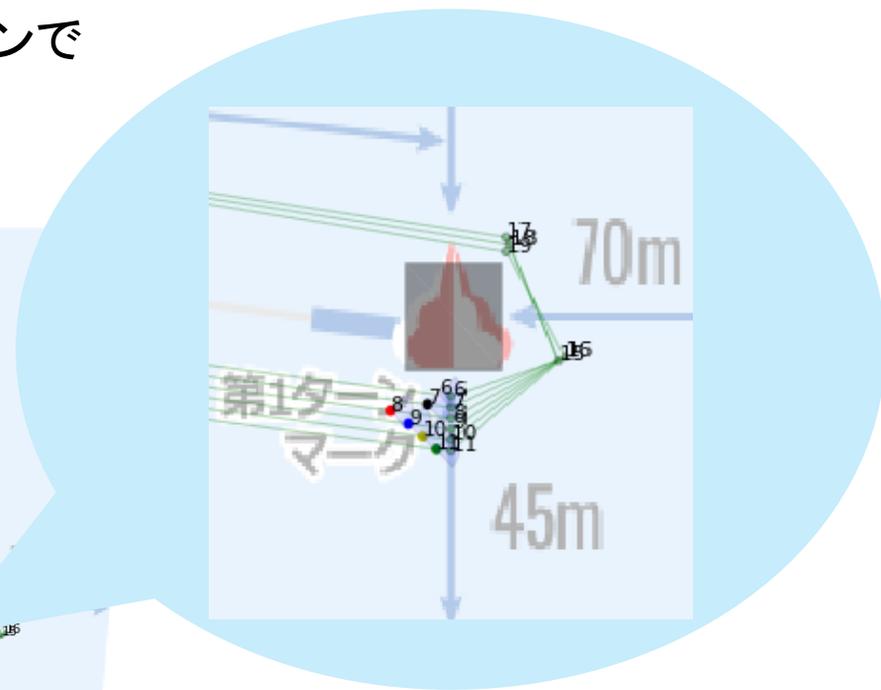
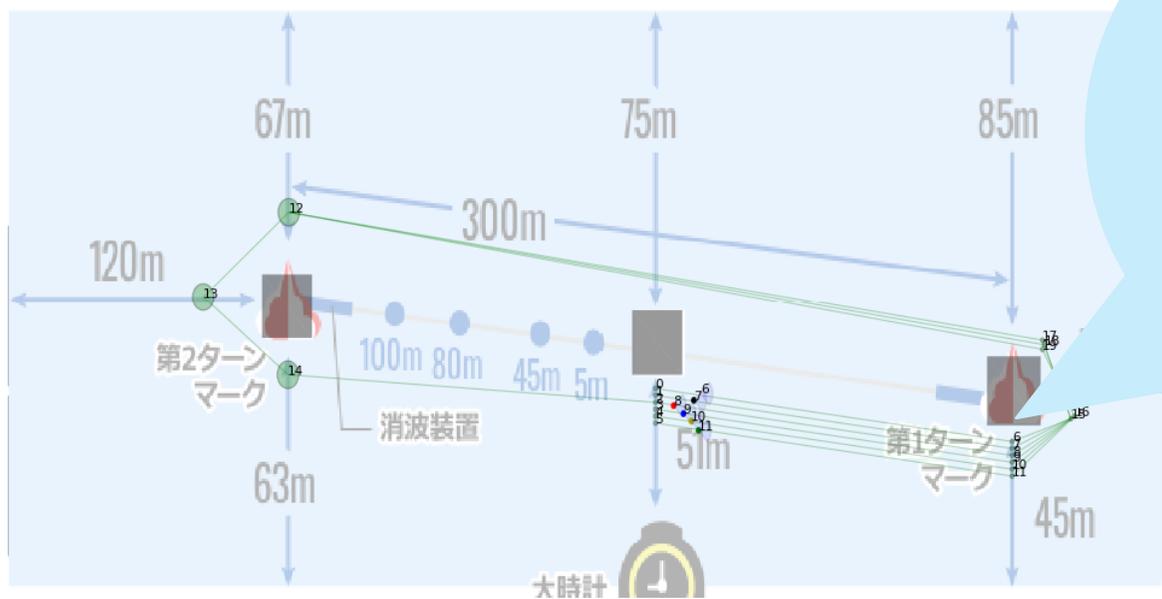
クラスタのコースごとのレースタイムのヒストグラムを確認した結果例は、以下のようになった。



→正規分布に従うと考えこの分布から、レースタイムを
毎レース・コースごとに予測する

レースのシミュレーション

実際にはボートが周回する状況をシミュレーションで再現している



実証 → 結果

実証 レース会場のグラフは住之江競艇場のレース会場を参考にした、会場のサイズや水の質(淡水・海水)などが異なるため、使用する過去のデータはすべて住之江で行われた対象のレース(527レース)に統一した
その中からランダムに50レースを選択、それらのレースを10回シミュレーションし、1着を的中させた確率を調べた

結果 実際のレースでは、1号艇が1着になったレースは50レース中30レースだった。

1号艇が1着になる確率が高いので

1号艇以外が1着になったレースの場合と分けて、的中率を計算した

1着の的中率	1,2着の的中率	1,2,3着の的中率	1号艇が1着のレースでの1着の的中率	1号艇以外が1着のレースでの1着の的中率
31%	5.8%	1.4%	45.3%	9.0%

結論 → 課題

結論 予想の精度としてはそれほど高い結果にはならなかった
しかし、勝率の高い選手を予想の上位に持っていくという、基本的な予想の再現に加え、1号艇が1着になるという予想は的中率を見ると予測できたと感じる
しかし、1号艇以外が1着になるレースでの予想の的中率は低く、課題が残る結果となった

課題

- ・場面設定での道中でのエージェントの動きの自由度が少ない
- ・水面の波などの他のエージェントへの影響を考慮できていない
- ・ボートの加速に影響するモーターの勝率を加味できていない
- ・一定のスピードで計算しているので、ターンマークを回る速さも一定になっている