

株価時系列データを用いた市場参加者の推定

Market Participant Estimation by Using Stock Price Time Series

田島裕之^{1*} 鳥海不二夫¹ 石井健一郎¹
Hiroyuki TAJIMA¹, Fujio TORIUMI¹, Kenichiro ISHII¹

¹ 名古屋大学大学院情報科学研究科

¹ Graduate School of Information Science, Nagoya University

Abstract: In artificial market studies, agents must do the similar behavior with real market participants. In this paper, we estimate the composition of the real market participants using artificial markets. We use an inverse simulation to optimize agents who participate in artificial markets. We use stock price time series of the market to evaluate artificial markets. In the experiment for verifying the validity of the proposed method, we apply the proposed method to markets in which different participants exist. The result indicates that the proposed method is useful to estimate market participants.

1 はじめに

株式市場には多様な戦略を持った投資家たちが存在する。投資家たちは、それぞれの分析手法に基づいて株価の上下を予測し、取引を行う。株価の変動は投資家同士の取引によって生み出されるものである。したがって、バブルの発生や崩壊をはじめとする株価の変動に関する現象のメカニズムは個々の投資家の一つ一つの行動が関係しているため、非常に複雑となっている。現在の市場理論では、市場で起こる現象の全てを説明することは困難である。

人工市場研究は、市場に見られる種々の現象を解明するための新たなアプローチである。人工市場とは、エージェントと呼ばれる自動取引プログラムが取引を行う架空の市場のことである。人工市場を用いることで、ミクロなレベルの事象であるエージェント同士の取引が、マクロなレベルの事象である株価の変動に与える影響を再現、分析することができる。

しかし、人工市場を扱う上で気を付けなければならないことがある。それは現実の市場との関連付けである。設計した人工市場モデルが現実の市場を反映したものでなければ、その人工市場を用いて得られた知見は説得力を持たない。特に、エージェントが現実の投資家とは大きく異なる戦略を用いていたり、エージェントの組み合わせが現実の市場における投資家の構成から乖離した人工市場では、現実の市場で起こる現象の解明に結びつくとは考えにくい。現実の市場に即した人工市場を設計する上では、市場で取引を行う投資

家、すなわち市場参加者の戦略や組み合わせを分析することは不可欠である。

Toriumiら[1]による先行研究では、現実の市場に存在する参加者の構成を、人工市場を用いて推定している。この人工市場では、エージェントがそれぞれ異なる戦略を用いる。エージェントの組み合わせを人工市場のパラメータとし、これを最適化することで市場参加者を推定する。その際に、社会シミュレーションモデルのパラメータ推定に用いられる“逆シミュレーション”と呼ばれる手法を用いている。この手法を用いて、異なる市場における市場参加者の構成の推定を行った。その結果、異なる市場における参加者の構成に違いが見られることを示し、手法の有効性を確認した。

本論文は Toriumi らの手法を発展させ、より多様な市場における参加者の推定を目的とする。提案手法では、エージェントの取引戦略や、逆シミュレーションにおける人工市場の評価方法について拡張を行っている。その際に、株価の上昇下降といったトレンドではなく、変動の大きさや変化などに焦点を当てる。

2 人工市場

人工市場とは、コンピュータ上に作られた架空の市場のことである[2]。本論文では、人工市場に対して現実の市場のことを実市場と呼ぶこととする。人工市場では、エージェントと呼ばれる現実のトレーダーを模したプログラムが取引を行う。ただし、実市場のような連続時間における取引をコンピュータ上で再現しようとした場合、エージェントの設計が複雑になる。その

*連絡先: 名古屋大学大学院情報科学研究科
〒464-8603 愛知県名古屋市千種区不老町
E-mail: tajima@kishii.ss.is.nagoya-u.ac.jp

ため、エージェントによる注文及びエージェント同士の取引の処理は、ステップという単位で行われる。エージェントはステップ毎に、市場から情報を受け取り、買い注文又は売り注文を行う。エージェントから出された注文は、人工市場の価格決定メカニズムに従って処理され、そのステップにおける現在価格が同時に決定する。このように人工市場は、エージェント同士の取引によって価格が変動するという市場構造を再現するものである。

価格決定メカニズムとは、エージェントから受け取った注文を取引として成立させ、現在価格を決定する処理の流れを表す。本論文で扱う人工市場では、注文の処理方法にザラ場方式を用いる。ザラ場方式では、エージェントから受け取った注文は、板と呼ばれる表に記録される。板に記録された注文は、価格優先、時間優先の原則に従って、逐次的に処理される。注文は基本的に、高い買い注文、安い売り注文が優先的に処理される。同じ価格の買い注文と売り注文が入ると、その価格が現在価格として決定し、取引が成立する。ただし、同じ価格に複数の注文が入ったときは、先に入った注文が優先的に処理される。

エージェントは、人工市場が提供する情報を入力として受け取り、そのステップに出す注文を決定する。人工市場が提供する情報には、過去の価格列や出来高、エージェントが出した注文の結果などが含まれる。これらの情報を受け取ったエージェントは、自分の持つ戦略に従い、買い注文を出すのか、それとも売り注文を出すのかを決定する。また、注文価格や注文量なども、エージェントの戦略によって決定される。

3 市場参加者推定手法

3.1 逆シミュレーション

人工市場を用いた市場参加者の推定を行う場合、異なる戦略を持つエージェントの組み合わせによって人工市場を構築する。この時のエージェントの組み合わせが、人工市場のパラメータとなる。エージェントの組み合わせを最適化することで、市場参加者の構成を推定し、これを分析する。しかし、エージェントの戦略が複雑になるほど、全てのパラメータを網羅的に調べ、最適なパラメータを見つけることは困難となる。Toriumiら [1] の先行研究では、遺伝的アルゴリズムを用いた逆シミュレーションによって、人工市場パラメータの最適化を行っている。逆シミュレーションとは、エージェントベースモデルのパラメータを最適化するために用いられる手法である [3]。

逆シミュレーションによる人工市場のパラメータ推定は以下のように行う。まず、人工市場のパラメータ

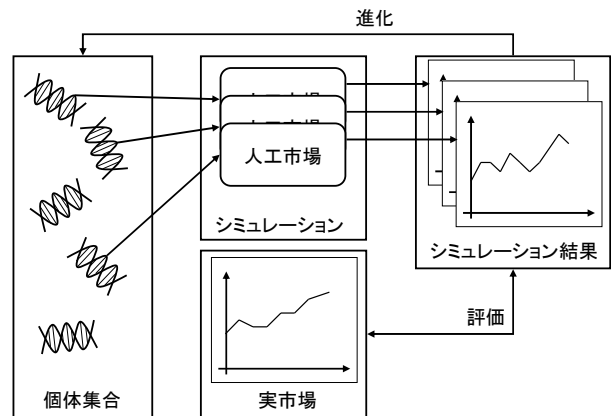


図 1: 遺伝的アルゴリズムによる逆シミュレーション

を $\{0, 1\}$ の遺伝子列として表現する。各世代における個体は、それぞれが 1 つの人工市場を表す。個体の評価には、人工市場のシミュレーション結果と実市場の類似度を用いる。シミュレーション結果には、生成された価格列、出来高、エージェントによる注文等が含まれる。シミュレーション結果に含まれるこれらの情報が、実市場から得られる情報とにどれだけ類似しているかを測り、それを人工市場の評価値とする。個体の評価値を得たら、遺伝的アルゴリズムに従い世代内の個体集合を更新し、これを繰り返すことで最終的に実市場に最も近い人工市場を得る (図 1)。

3.2 人工市場のパラメータ

人工市場は、異なる戦略を持つエージェントの集合によって構成される。したがって、逆シミュレーションを行うに当たって、人工市場を以下のように遺伝子で表現する。まず、あるエージェントの取引戦略を 10 ビットの遺伝子で表現する。取引戦略には、“注文を行うタイミング”、“1 つの注文でやり取りする株数”、“注文の頻度”などが含まれる。人工市場のパラメータは、10 ビットの遺伝子をエージェントの人数分繋ぎ合わせることで表現される (図 2)。つまり、 N 人のエージェントが参加する人工市場のパラメータは、 $10 \times N$ ビットの遺伝子となる。遺伝子によるエージェントの取引戦略の表現方法については、第 4 節で述べる。

3.3 実市場との類似度評価

3.3.1 類似度の評価値

本論文では、遺伝的アルゴリズムにおける個体の評価に株価時系列データを利用する。ただし、時系列中で市場参加者の構成が大きく変化した場合、その推定

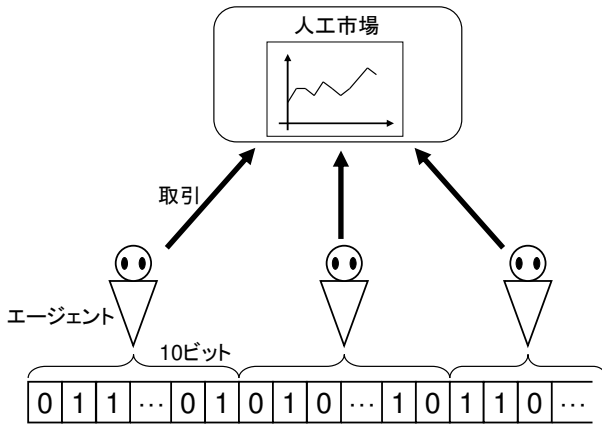


図 2: 人工市場の遺伝子表現

は困難になると考えられる。したがって、個体の評価に用いる株価時系列は、比較的短い期間に観測されたものである必要がある。ゆえに、本論文で扱う株価時系列は分足を対象とする。

市場参加者は各自の売買戦略に基づいて注文を行い、その注文を処理する過程において株価は決定する。つまり、株価時系列には市場参加者が用いる戦略に関する情報が含まれている。人工市場においても、価格列はエージェント同士の取引の結果生成されるものである。したがって、人工市場の価格列と実市場の株価時系列データが類似しているとき、双方の市場参加者の戦略及び構成は類似していると考えられる。

株価時系列データを用いて、人工市場の評価値 e を以下のようにして求める。まず、長さ T の実市場の株価時系列 $\hat{p}_1, \hat{p}_2, \dots, \hat{p}_T$ から特徴を抽出し、特徴ベクトル \hat{x} を生成する。この時用いる特徴については次節で述べる。次に、人工市場から生成される長さ T' の価格列 $p_1, p_2, \dots, p_{T'}$ から、同様に特徴ベクトル x を生成する。抽出した特徴間の距離として、人工市場の評価値 e を以下の式で表す。

$$e = D(\hat{x}, x) \quad (1)$$

ここで $D(x_1, x_2)$ は特徴ベクトル x_1, x_2 間のユークリッド距離を表す。評価値 e が小さいほど、人工市場から生成された価格列は実市場の株価時系列に類似していることを示す。

3.3.2 株価時系列の統計量

提案手法では、株価時系列から特徴を抽出し、特徴間の距離によって時系列の類似度を測る。そこで、分足の株価時系列に関する統計量を特徴として用いる。時系列の統計量は、その時系列の性質を表すものとして有用であると考えられる。

ただし、提案手法では株価の上昇下降といったトレンドは株価時系列の特徴として用いず、価格変動の大きさやその変化に関する統計量を用いる。これは、市場において買い側が強いのか売り側が強いのかではなく、あくまで市場参加者毎に異なる戦略の組み合わせを推定することを目的としているためである。そのため、トレンドの影響が強い統計量は特徴から除外している。以下に、株価時系列より特徴として抽出する統計量を示す。

変動量の分散, 尖度 価格列 p_1, p_2, \dots, p_T の変動量

$v_t = p_t - p_{t-1}$ の分散 σ^2 , 尖度 k を特徴に含める。尖度 k は、平均周りの値と、平均から離れた値の出現しやすさを表す指標である [4]。平均周りの γ 次のモーメントを μ_γ とすると、尖度 k は以下の式で表される。

$$k = \frac{\mu_4}{\mu_2^2} - 3 \quad (2)$$

価格変動確率 分足の場合、数分間価格が変動しない場合が多々ある。そこで、時間 $t, t+1$ 間で価格が変動する確率を特徴に加える。価格列 p_1, p_2, \dots, p_T において、 $p_t \neq p_{t+1}$ となる回数をカウントし、価格列の長さ T で割って正規化したものを、価格変動確率とする。

自己相関係数 自己相関係数 C_l は、時系列 y_1, y_2, \dots, y_N において、 y_t と y_{t-l} の相関係数として表される指標である [5]。値の取りうる範囲は $-1 < C_l < 1$ である。 $C_l = 0$ のとき、 y_t と y_{t-l} の間には相関がないことを示す。変動量系列 $v_t = p_t - p_{t-1}$ の自己相関係数と、変動量の 2 乗の系列 $v_t^2 = (p_t - p_{t-1})^2$ の自己相関係数を $l = 1$ について求め、特徴に加える。

以上の統計量を株価時系列について求め、5次元の特徴ベクトルを生成する。特徴ベクトルの各成分を標準偏差で割ることにより正規化した後、特徴間距離を測る。

4 エージェントの設計

4.1 概要

多くの人工市場研究では、大きく分けて 2 種類のタイプのエージェントが存在する [6, 7]。ファンダメンタリスト (fundamentalist) とチャーティスト (chartist) である。この 2 種類のエージェントは、それぞれ異なる情報を用いて将来の株価を予測し、行動を決定する。本論文においても、この 2 種類のエージェントによって構成される人工市場によって、市場参加者の推定を行う。

4.2 ファンダメンタリスト

ファンダメンタリストは、その株式を発行している企業の業績や種々の経済指標などを用いて、将来の株価を予測する。企業の業績には株式の配当や株価収益率などが含まれ、経済指標には金利や経済成長率、雇用状況などが含まれる。主に機関投資家が、ファンダメンタリストによって表される。

ファンダメンタリストは長期的な視野を持った投資をする傾向があるため、頻繁に売り買いを繰り返すことはないと思われる。特に提案手法では、分足という短期間の株価時系列を扱う。従って、ファンダメンタリストエージェントは買い側又は売り側のどちらかに分類される。例えば、買い側のエージェントは決して売ることはなく、注文をする場合は必ず買い注文のみを行う。

4.3 チャーティスト

チャーティストは、過去の株価時系列からトレンドを判定し、将来の株価を予測する。短期間において売買を繰り返す個人投資家は、チャーティストであることが多い。

チャーティストエージェントは、 n ステップ前の価格 p_{t-n} と現在価格 p_t を比較し、 $p_{t-n} < p_t$ ならば上昇トレンド、 $p_{t-n} > p_t$ ならば下降トレンドと判定する。また、チャーティストの予測には順張り、逆張りの2種類が存在する。順張りは現在のトレンドが持続するという予測、逆張りは現在のトレンドが反転するという予測である。例えば、現在のトレンドが上昇トレンドと判定された場合、順張りの場合は買い注文、逆張りの場合は売り注文を行う。

4.4 エージェントのパラメータ

第3.2節で述べたように、人工市場のパラメータは、エージェントを表す10ビットの遺伝子を繋げたものとして表現される。エージェントが注文を出す際に用いる取引戦略を、10ビットの遺伝子によって表す。遺伝子の各ビットと、パラメータとの対応を表1に示す。各パラメータの意味は以下の通りとする。

注文確率 あるステップに注文を行う確率。高いほど取引に積極的に参加することを示す。

注文量 エージェントの注文株数。大きいほど株式をまとめて購入、売却することを示し、そのエージェントの利用可能な資産が多いことを表す。

表 1: エージェントのパラメータ

ビット	パラメータ	値
共通		
1	タイプ	ファンダメンタリスト, チャーティスト
2, 3	注文確率	0.05, 0.2, 0.5, 1.0
4, 5	注文量	1, 10, 50, 100
6, 7	強気度	4 段階
ファンダメンタリスト		
8	注文側	買い側, 売り側
9, 10	目標注文量	10, 100, 500, 2000
チャーティスト		
8	トレンド戦略	順張り, 逆張り
9, 10	測定期間	1, 10, 30, 60

強気度 売買注文に対する自信を表す。4段階に分かれており、強気度が高いほど高い価格で買い注文、安い価格で売り注文を行う。

注文側 ファンダメンタリストのみが持つパラメータ。買い側、売り側の2値を取る。

目標注文量 ファンダメンタリストのみが持つパラメータ。エージェントはこの数値に達するまで株を売買し、それ以上の注文は行わない。

トレンド戦略 チャーティストのみが持つパラメータ。順張り、逆張りの2値を取る。

測定期間 チャーティストのみが持つパラメータ。トレンド判定に用いる期間を表す。例えば、トレンド測定期間が10ステップの場合、10ステップ前の価格 p_{t-10} と現在価格 p_t を比較することで、トレンドを判定する。

5 提案手法による市場参加者の推定

5.1 概要

提案手法を実市場に対して適用し、市場参加者の推定における有効性を確認する。実験の対象とする市場は、東証1部の“住友化学”、“東京電力”の2銘柄である。両銘柄の2006年7月12日前場の分足株価時系列^{*1}を用いて逆シミュレーションを行う。両銘柄の2006年7月12日前場の株価を図3に示す。グラフを見ると分かるように、この2つの市場では、株価の変動の大きさや頻度について違いが見られる。“住友化学”はほぼ毎

^{*1}NEEDS (日経の総合経済データバンク) を使用した。

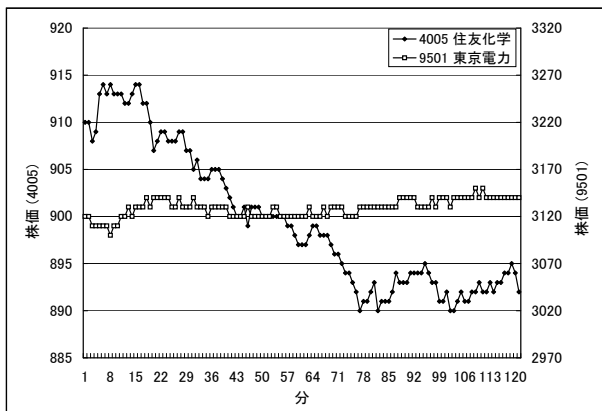


図 3: 対象市場の株価

分変動しており、大きく変動する箇所もあるが、“東京電力”には大きな価格変動は見られず、変動幅も小さい。この2つの市場に対して提案手法を適用することで、双方の市場参加者の構成が異なっていることを明らかにする。

5.2 遺伝的アルゴリズムの設定

逆シミュレーションの際に用いた遺伝的アルゴリズムの設定を表2に示す。なお、今回用いた遺伝的アルゴリズムでは、選択方法にルーレット選択を用いた。ルーレット選択を行う際、適応度が大きいほど優秀な個体であり、なおかつ適応度が非負値でなければならない。しかし、個体の評価に用いるのは特徴間距離であるため、数値が小さいほど優秀な個体である。そのため、適応度 f を以下の式によって求める。

$$f = \begin{cases} \alpha - e & e \leq \alpha \\ 0 & e > \alpha \end{cases} \quad (3)$$

α は定数であり、本実験では $\alpha = 10$ とした。評価値 e は式 (1) で示した特徴間の距離である。人工市場のシミュレーションを 20 回を行い、適応度の平均値をその個体の適応度とした。

5.3 逆シミュレーションの結果

逆シミュレーションによる評価値の収束の様子を図4に示す。ここで評価値は、実市場と人工市場から抽出した特徴間の距離を表す。評価値が小さいほど、人工市場が実市場に類似していることを示す。グラフより、対象とした2つの市場に対して行った逆シミュレーションでは、100 世代より前で収束している。

表 2: 遺伝的アルゴリズムの設定

個体数	30
世代数	200
遺伝子長	500
選択方法	最良個体を残し、 残りをルーレット選択
交叉方法	二点交叉
突然変異確率	5%

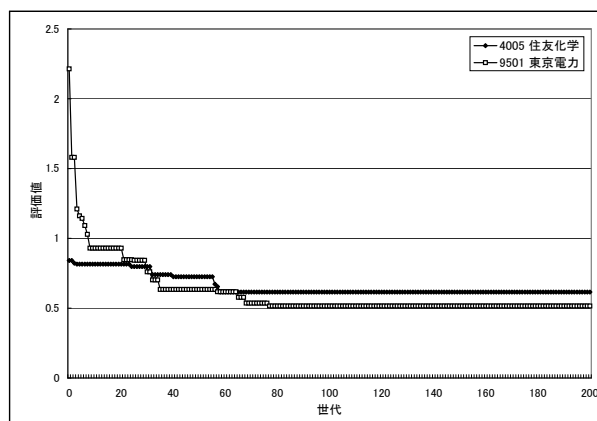


図 4: 逆シミュレーションによる評価値の収束

続いて、対象とした2つの市場それぞれについて、最適化した人工市場から生成される価格列と実際の株価を比較した。“住友化学”、“東京電力”それぞれの結果を図5、図6に示す。ここで注意されたいのは、提案手法による最適化は、株価のトレンドを類似させることが目的ではないということである。そのため、人工市場の価格列が実市場に追従しているかどうかについては考慮せず、あくまで価格変動の特性が類似しているかどうかを最適化の評価基準としている。図5、図6より、逆シミュレーションした価格列を比較すると、人工市場の価格列の動きは実市場の株価の変動に類似していることが分かる。

また、最適化された人工市場のシミュレーションを 200 回を行い、評価値の平均と標準偏差を調べた。結果を表3に示す。表より、最適化された人工市場の評価値の標準偏差は小さい。人工市場から生成される価格列は毎回異なるが、価格列から特徴を抽出して求めた評価値に関しては、ばらつきが小さいと言える。

5.4 エージェント構成の分析

最適化された人工市場を構成するエージェントについて分析を行う。図7、図8は、人工市場を構成する

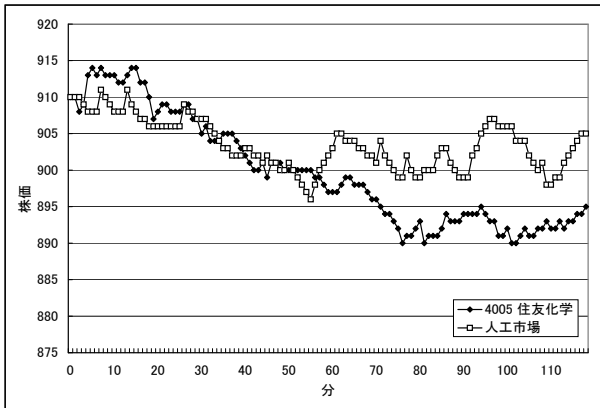


図 5: 住友化学と人工市場の価格列

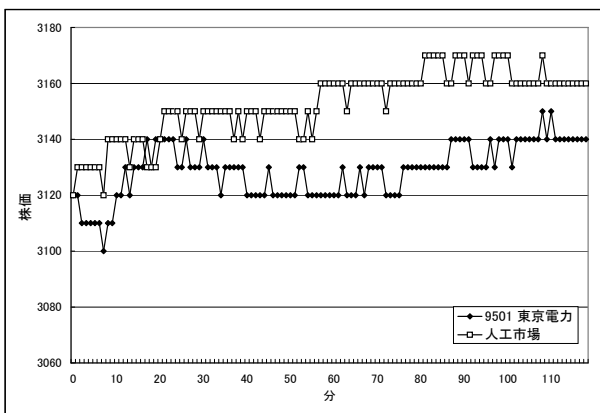


図 6: 東京電力と人工市場の価格列

エージェントの割合を，エージェントのタイプ及び注文量について調べたものである．グラフの買い側，売り側はファンダメンタリストを，順張り，逆張りはチャートリストを表す．

図 7 より，“住友化学”はチャートリストが多く，“東京電力”はファンダメンタリストが多い．“東京電力”のような電力株は企業の業績が景気に左右されにくく，安定していると言われることが多い．“東京電力”にファンダメンタリストが多いのは，機関投資家のような安定志向の投資家が多いためであると考えられる．一方，“住友化学”のように株価の変動が激しい銘柄には，短期で売買を繰り返すチャートリストが多い，このような投機的な変動をする銘柄は，短期間で大きな収益を上げやすく，個人投資家の注目を集めやすい．

次に図 8，注文量別エージェント構成を見る．グラフより，“住友化学”よりも“東京電力”の方がエージェントが一度に行う注文量が多いことが分かる．一度に行う注文量が多いということは，その投資家の利用可能な資産が多いことを示している．“東京電力”

表 3: 最良個体の評価値のばらつき

	評価値の平均	標準偏差
4005 住友化学	0.655	0.275
9501 東京電力	0.636	0.269

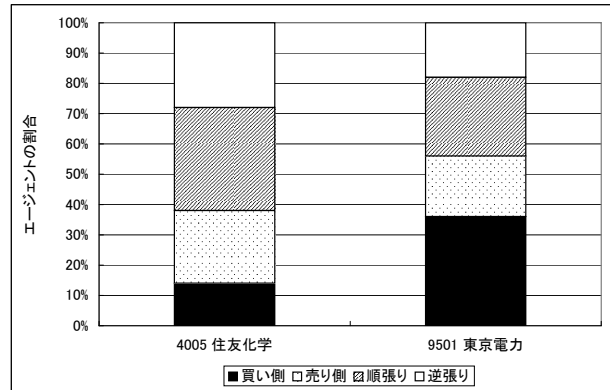


図 7: タイプ別エージェント構成

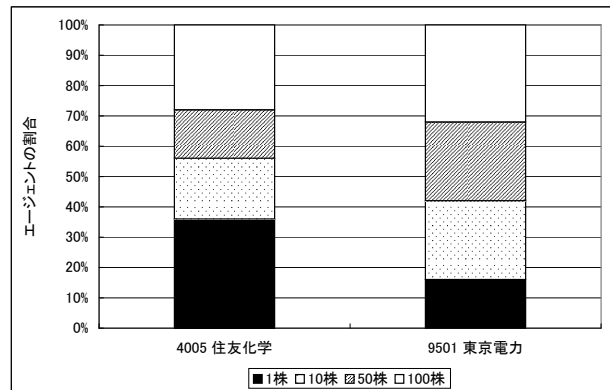


図 8: 注文量別エージェント構成

のエージェントの注文量が多い傾向にあることは，先ほど述べたように，電力株に機関投資家が多いことを裏付けている．それに対し“住友化学”では，注文量の小さいエージェントが多い．これは“住友化学”には個人投資家が多いことを示している．

以上より，最適化された人工市場のエージェントは，実市場の参加者構成について共通点が確認でき，実市場の市場参加者を再現できていると分かる．ゆえに，市場参加者の構成を推定するに当たって，提案手法は有効であると考えられる．

6 おわりに

本論文では、市場参加者の戦略や組み合わせを分析することを目的とし、市場参加者推定手法を提案した。提案手法では、エージェント毎に異なる戦略を用いる人工市場を構築し、エージェントの組み合わせを逆シミュレーションによって最適化することによって、市場参加者の構成を人工市場上に再現した。なお、最適化の際の評価値には、実市場と人工市場の株価時系列からそれぞれ特徴を抽出し両者の特徴間距離を用いた。

提案手法の有効性を評価するために、価格変動の傾向が異なる2銘柄について提案手法を適用した。その結果、“東京電力”のように安定していると言われる銘柄と“住友化学”のように株価が大きく上下する銘柄では、エージェントの組み合わせに違いが見られた。前者は安定志向な機関投資家、後者は投機的な個人投資家が多いという結果となった。これは実市場における参加者の構成に共通していると思われ、提案手法の有効性が示された。

ただし、提案手法にはいくつかの改善点が挙げられる。一つ目は、エージェントの戦略に関することである。提案手法では10ビットの遺伝子によってエージェントの戦略を表現した。そのため、エージェントの戦略に関するパラメータの数をある程度に絞った。しかし、実市場の複雑な価格変動は、より複雑な戦略によって生み出されるものである可能性がある。

二つ目の改善点は、人工市場の評価値の算出方法である。今回は株価時系列のみを考慮し、これに関する統計量の特徴として、人工市場の評価を行った。しかし、株価と並んで出来高もまた市場参加者の取引に関する重要な情報である。出来高も同時に利用することで、より実市場に近い人工市場を構築可能であると考えられる。

また、有効性評価実験におけるエージェント構成の分析では、エージェントのタイプと注文量に注目し分析を行った。しかし、“注文確率が高く、注文量が少ないファンダメンタリスト”のように、特定の戦略を持つエージェントが人工市場に大きな影響を与える可能性もある。このようなエージェントを調べるには、一人ひとりのエージェントに着目した分析が必要となる。しかし、エージェントのパラメータを増やすほどその組み合わせは膨大になり、エージェントを個別に分析することは困難になると考えられる。これは先に述べたエージェントの戦略に関する改善点にも関係しており、提案手法をさらに発展させる上で重要な課題である。

参考文献

- [1] F. Toriumi, K. Izumi, and H. Matsui. Market Participant Estimation by Using Artificial Market. *Advances in Practical Multi-Agent Systems*, p. 201, 2010.
- [2] 和泉潔. 人工市場: 市場分析の複雑系アプローチ. 森北出版, 2003.
- [3] S. Kurahashi, U. Minami, and T. Terano. Why not multiple solutions: agent-based social interaction analysis via inverse simulation. In *Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on*, Vol. 2, pp. 522–527. IEEE, 2002.
- [4] DN Joanes and CA Gill. Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, Vol. 47, No. 1, pp. 183–189, 1998.
- [5] 北川源四郎. 時系列解析入門. 岩波書店, 2005.
- [6] T. Lux and M. Marchesi. Volatility clustering in financial markets: A microsimulation of interacting agents. *International Journal of Theoretical and Applied Finance*, Vol. 3, No. 4, pp. 675–702, 2000.
- [7] H.P. Boswijk, C.H. Hommes, and S. Manzan. Behavioral heterogeneity in stock prices. *Journal of Economic Dynamics and Control*, Vol. 31, No. 6, pp. 1938–1970, 2007.