

# 時系列データ間の連関性と関係性理解のための ビジュアルインタラクティブティ

## Visual Interactivity for Understanding Relationships and Uncovering Correlations among Temporal Numerical Data of Thousands of Elements

中小路 久美代<sup>\*1</sup> 山本 恭裕<sup>\*1</sup> 松原伸人<sup>\*2</sup> 川嶋 稔夫<sup>\*3</sup> 羽室 行信<sup>\*4</sup> 宇野 毅明<sup>\*5</sup>  
Kumiyo Nakakoji Yasuhiro Yamamoto Nobuto Matsubara Toshio Kawashima Nobuyuki Hamuro Takeaki Uno

<sup>\*1</sup> 京都大学 <sup>\*2</sup>(株)SRA <sup>\*3</sup>はこだて未来大学 <sup>\*4</sup>関西学院大学 <sup>\*5</sup>国立情報学研究所  
Kyoto University SRA Inc Future University Hakodate Kwansai Gakuin University National Institute of Informatics

This paper describes the design of a prototyped environment and its visual interactivity for visualizing daily correlations of stock price movement among 5428 firms over the period of twenty-nine years.

### 1. はじめに

本稿では、東証上場全銘柄間の株価変動の 29 年間に亘る日々の連動性についてのデータを対象とした、データインタラクション環境のビジュアルインタラクティブティのデザインについて論じる。導出した株価の上下変動推移についての銘柄間の連関性を時間経過を踏まえながら捉え、その関係性を見出すようなビジュアルインタラクティブティを考察し、データインタラクション環境プロトタイプを説明する。

### 2. 対象とするデータ

本論で、ユーザが可視化表現を介してインタラクションを行う対象とするデータは、延べ 5248 銘柄それぞれの銘柄間の、その日から遡って 10 日間の間の株価変動(ここでは終値)の連関性の有無とその強さを、約 7000 日間分に亘って求めたものである。

これらのデータは、任意の2銘柄の類似度を定義し、ある閾値を超えた銘柄間に枝を張り銘柄間類似度グラフとすることで構築している[羽室 et al. JSAI 2016]。約 7000 日の日毎に、その日から遡って 10 日間の価格変動の距離(1-相関係数)が閾値  $\delta$  (利用データでは  $\delta = 0.3$ ) 以下の銘柄間に枝を張り、これを銘柄間の類似度とした。各ノード(銘柄)について、過去 10 日間の価格において、[1,2,...9]との相関係数が閾値  $\rho$  (ここでは  $\rho = 0.7$ ) より大きなものをアップトレンド(上昇局面)、[9,8,...1]との相関係数が閾値  $\rho$  より大きなものをダウントrend(下降局面)であると判定した。

これまでに、これら銘柄類似度グラフの時系列構造変化と TOPIX インデックスの関係についての考察を行い、得られた銘柄類似度グラフの枝密度というグラフ特徴量を用いて、株価の全体的な動きを捉えることを試みている[羽室 et al. JSAI 2016]。

我々は、本データを対象として、個々の銘柄と他の銘柄間の株価の上下変動推移に着目している。市場全体の動きを捉えつつ、着目すべき個々の銘柄に注目し、その銘柄と関係があると見受けられる他の銘柄を追いつつ、仮説やその説明を作り出すことを支援する[中小路, 山本 2004]ような、データインタラクション環境[中小路 et al. 2016]の構築を目指すものである。

### 3. データの基軸

対象となるデータの基軸となるのは、日付と銘柄とである。ユーザは、日付あるいは銘柄に着目し、その着目点を基軸点として変化を見つけ出し、変化や類似性を見て取ることを繰り返す。

#### 3.1 日付(時間)

株価変動は連続して並んだ日付という時間軸に沿って生じる。本論で扱っているデータでは、ある日から 10 日間に遡って上昇局面にあるか下降局面にあるかを求めた上で、他の銘柄との関係性がどのように変化するかを時系列に沿ってみることになる。直接的に株価変動を見ているわけではない。

時間軸上のインタラクションとしては、まず着目する日付を指定できる必要がある。日付の基軸上では、年月日を直接指定してすることもあれば、震災が起きた日や金利政策の変更が発表された日、といった特定の事象が生じた日で指定したい場合もある。また、バブル崩壊やサブプライム問題といった、ある幅を持った時期で指定したい場合もあるであろう。

時間軸には線型性がある。線形性に沿って、時間軸上で着目した日付から順に日付を追ったり、また逆に遡ったりすることがスムーズに行えることが肝要となる。また、時間の周期的な性質を踏まえれば、週初めや月初、季節といった、周期を考慮した連続性で見ることが必要となる。

#### 3.2 銘柄

それぞれの銘柄には、会社名や創業年数、資本金や所在地といった基本的な会社情報と銘柄コードが付されている。また、(上場期間中の)各日付ごとに株価(始値, 終値, 高値, 安値)を数値データとして有している。銘柄コードはある程度業種によって番号帯が分かれており、銘柄コード順に並べるとほぼ類似の業種が並ぶことになるが、必ずしも厳密ではない。時間軸と異なり、銘柄間に一意に定まる線形順序は存在しない。

### 4. ビジュアルインタラクティブティの要件

本データを可視化した環境とのインタラクションを介してユーザは、個別の銘柄間の類似性を探し出しその要因を考え出したり、多くの銘柄が同じ株価変動を呈する時期(日付)に至るまでやその日以降の株価連動性の変遷や経過を見て、株価に影響を及ぼした事象を考察したり、といったことが行えると考えている。

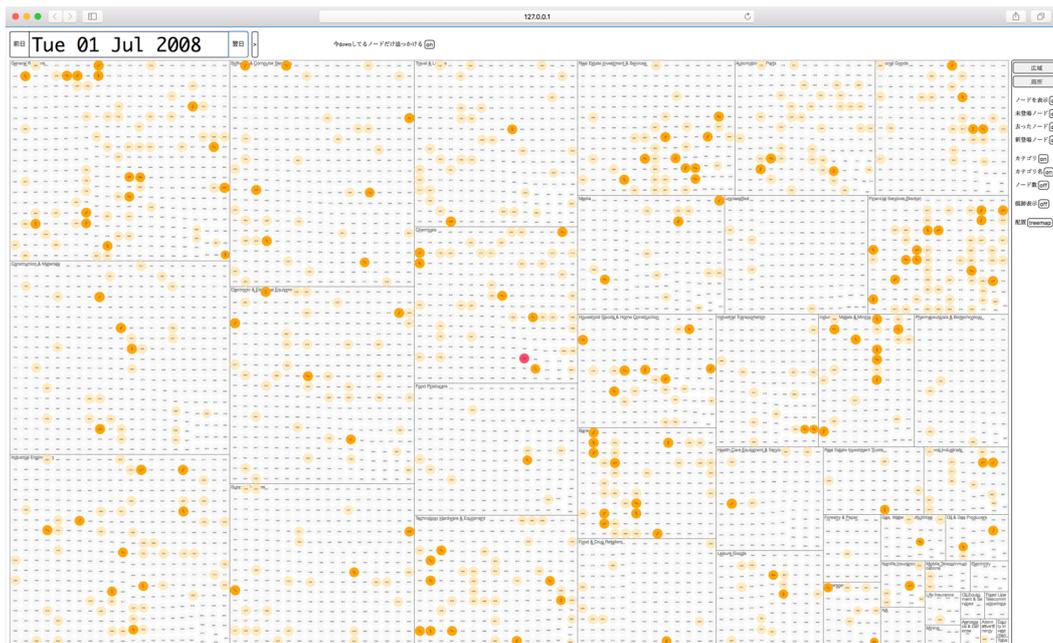


図 1: データインタラクション環境 ViNL のスクリーンイメージ

データインタラクション環境においてユーザが見てとることができるべきと考えるインタラクションの要件を、下記に列挙する。

- それぞれの銘柄を区別できること
- 指定した日付の類似した株価変動を呈する銘柄群を特定できること
- 時間経過に従って類似した株価変動を呈する銘柄群の変化を見て取れること
- 時間経過を遡って類似した株価変動を呈する銘柄群の変化を見て取れること
- 類似した株価変動を呈する銘柄群を構成する銘柄数がわかること
- 類似した株価変動を呈する銘柄群の特性がわかること (業種, 地域, 創立時期, 社長の年齢等)
- ある銘柄と類似した株価変動を呈する銘柄群を構成する銘柄数がわかること
- ある銘柄と類似した株価変動を呈する銘柄を特定できること
- ある銘柄と類似した株価変動を呈する銘柄群の中で, それぞれの銘柄との類似性の強弱を見て取れること
- 時間経過に伴いある銘柄と株価変動の類似性が新たに生じた銘柄群を特定できること
- 時間経過に伴いある銘柄と株価変動の類似性が無くなった銘柄群を特定できること
- 時間経過に伴いある銘柄と株価変動の類似性が新たに生じた (無くなった) 銘柄群を構成する銘柄数がわかること
- 時間経過に伴いある銘柄と株価変動の類似性が新たに生じた (無くなった) 銘柄群の特性がわかること
- 着目してきた銘柄を辿れること
- 着目してきた日付を辿れること

## 5. データインタラクションの基本的なフロー

ユーザが行うデータとのインタラクションの基本的な流れは, 下記ようになる。

1. データ基軸(日付, 銘柄)に沿ったブラウジングを行う
2. データ基軸上の1点(日付, 銘柄)に着目する(ベースポイント)
3. ベースポイントに基づく関連データ群がハイライトされる
4. ハイライト内のデータ要素を(データ基軸に沿って)トラックする

データインタラクション環境において, ユーザは, ニュートラルに表示された全体を概観しながら興味を持ったり気になったりした点(データ要素やデータ基軸上の1点)が見つければそれに着目する。本研究では, この着目点を, それに続くインタラクションを行う上での「ベースポイント」と呼ぶ。ベースポイントは, 何らかの仮説を立てる足がかりになりそうだったり, 予測を立てるヒントになりそうな可能性を有するものとなる。

データインタラクション環境は, ユーザが指定したベースポイントに応じたデータの表現を提示する。ベースポイントという視点に基づいて, データ群がハイライトされることになる。ハイライトされた表現を見ながら, ユーザはブラウジングを進める。さらに興味を引いた点や, 思い至った観点がでてくると, ベースポイントを変更したり, 先のベースポイントと比較を行ったりしながら, 漸次的に仮説や予測を立てていくことができると考えている。

## 6. データインタラクション環境のプロトタイプ

ViNL (Visual Numerical Landscape) は, データインタラクション環境として現在構築中のウェブブラウザベースのプロトタイプシリーズである。以下に, ViNL シリーズの現状の実装について概要を述べる(図 1)。

図 1 に示す ViNL では、二次元空間上に全銘柄を銘柄コードで表示している。左上の日付表示部に示されている日付より前 10 日間の株価変動の連動性の有無が示されている。このスクリーンショットでは、銘柄は、業態毎に分かれるようなツリーマップとして配置、表示されている。配置の方式としては他に、名称順や地域毎など、異なる配置方式が考えられる。銘柄間の株価変動の連動性の有無は、ノードの色で表示している。ノードを右クリックすると、その銘柄の会社情報及び株価の変動をグラフで表示する。

右端に並ぶボタンで、日付や銘柄をベースポイントとして指定し、データをハイライトすることができる。「未登場ノード」「去ったノード」「新登場ノード」といったトグルで、現在見ている日付をベースポイントとして指定し、その状態と比較して株価変動の連動性の有無の状態が変化した場合に、それぞれノードに色をつけて強調表示することができる。

表示されているノードをクリックして1個選択することで、銘柄をベースポイントとして指定することになる。そのノード(銘柄)と株価変動の連動性があるノードに色がつけられて強調表示される。その際、その銘柄との連動性の強弱は、強調表示されたノードの回転によって示している。連動性の強いノードは角速度が速く回転している。人間は、周辺視野では色や形よりも動きの解像度が高いことが知られており、変化を見て取りやすいと考えている。

## 7. 終わりに

本稿では、各銘柄間の株価変化の連動性の時系列変化を、ブラジニングと日付や銘柄への着目、ハイライト表示を見ることでの着目点の移動、というインタラクションの繰り返しの介して体験することで、ユーザの仮説や推論の生成を支援するようなナレッジインタラクションデザインについて論じた。ブラウザベースの ViNL のプロトタイピングを通して、データインタラクション環境に必要なビジュアルインタラクティブ性の要件について考察した。

今後は、各銘柄間の株価変化の連動性でクラスタリングした [Uno et al. 2015] データを取り入れる予定である。これにより、ユーザが着目点を見つけやすくなったり、より効果的なハイライト表示が行えるのではないかと考えている。限られたディスプレイ空間上での表示と、人間の視覚的認知特性を踏まえながら、ビジュアルインタラクティブ性の在り方を探っていきたい。

## 謝辞

本研究の一部は、JST CREST の支援によるものである。

## 参考文献(論文誌と同じスタイルを推奨)

- 羽室行信, 岡田克彦, Stephane Cheung, 銘柄類似度グラフの時系列構造変化に基づく株価予測, 人工知能学会全国大会, 3L4-OS-16b-2, June, 2016. pp.1-4.
- 中小路久美代, 山本恭裕, 創造的情報創出のためのナレッジインタラクションデザイン, 人工知能学会論文誌, Vol.19, No.2, pp.154-165, March, 2004.
- 中小路久美代, 山本恭裕, 松原伸人, 北雄介, 時間情報を有するテキストデータの視覚的インタラクティブ性とデータ体験に関する考察, 2016年度人工知能学会全国大会, 12E-1in2, 北九州市, Japan, pp.1-4, June, 6th, 2016.
- Uno, T., Maegawa, H., Nakahara, T., Hamuro, Y., Yoshinaka, R., Tatsuta, M, Micro-Clustering: Finding Small Clusters in Large Diversity, CORR abs, 1507.03067, 2015.