

# ユーザの探索行為におけるグラフ型インタフェースの役割に関する分析

## A Role of Chart-based Interface for Supporting a User's Exploration

蓮井大樹<sup>1</sup>                      田中和広<sup>2</sup>                      松下光範<sup>1\*</sup>  
Daiki Hasui<sup>1</sup>                      Kazuhiro Tanaka<sup>2</sup>                      Mitsunori Matsushita<sup>1</sup>

<sup>1</sup> 関西大学 総合情報学部

<sup>1</sup> Faculty of Informatics, Kansai University

<sup>2</sup> 関西大学大学院 総合情報学研究科

<sup>2</sup> Graduate School of Informatics, Kansai University

**Abstract:** This paper presents an information retrieval system that utilizes a chart-based interface to support a user's exploratory data analysis. In such analysis, a user tends to examine data to support his/her hypothesis and to seek new perspectives about the data. To conduct this analysis, a user usually overviews data from various viewpoints first, compares multiple data to find differences and similarities among the data, and then accesses the original texts (e.g., newspaper articles) that contain the data. Our proposed system aims to support such the contiguous processes by providing a chart-based interface, which displays two charts side by side and permits a user to overlay the two charts for comparison, and then shows the original texts on demand. This paper brings the result of the user study conducted in VisEx task by analyzing the obtained data, which includes log records and post-questionnaire responses.

## 1 はじめに

現在、電子的に蓄積されるデータの量は時間の経過と共に増加している。これらの膨大なデータは単なるアーカイブとしての役割に留まらず、意思決定や問題解決のためのリソースとして期待されている [4]。これらのデータを分析に利用する際、分析者が「データをどのように利用するか」という明確な方法論や利用意図を予め持っておらず、より抽象的な目的や解決すべき問題だけを意識している場合も多い。これは、問題解決という大局から見れば目的志向的ではあるが、その過程で行われる探索自体は試行錯誤的な行為であると言える。このような分析行為は探索的データ分析 (exploratory data analysis) と呼ばれる [1]。

本研究の目的は、このような探索的データ分析の支援であり、分析者が円滑に試行錯誤できるように支援するシステムの実現である。その一環として、我々は動向情報データを対象に、異なるデータ同士の比較とそれに関する詳細な情報へのアクセスの支援を目指したグラフ型インタフェースを提案している [6]。本稿では、情報アクセス技術に関する評価型ワークショップ

NTCIR-9 のパイロットタスクである VisEx[2] で行われた被験者実験の結果を通じて、提案したインタフェースを分析者がどのように利用したかを分析し考察する。

## 2 グラフ型インタフェース

探索的データ分析は単純作業の繰り返しではなく、同じデータを様々な視点から多角的に眺める、仮説を立てながらデータを検索する、等の試行錯誤を繰り返しつつ行う高度に知的な思考活動 [1] である。このような分析行動においては、情報可視化が重要な役割を果たす。

Shneiderman は可視化表現を用いた分析者の情報探索行為を “Overview first, zoom and filter, then details on demand.” と表現している [3]。すなわち、(1) 情報の概観、(2) 情報の加工・操作 (3) 詳細情報の閲覧、の各段階が継ぎ目なく連続して行えるように支援することが、分析者の円滑な情報探索に貢献するといえる。

(2) の過程の遂行において、データ同士の「比較」は重要な行為の一つである [5]。分析者は複数のデータを比較することによって、異なるデータ間の相関や類似、相違などの関係を見出す。そして、このような発見をきっかけにして、自らの立てた仮説を確信/棄却した

\*連絡先： 関西大学総合情報学部  
大阪府高槻市豊仙寺町 2-1-1  
E-mail: mat@res.kutc.kansai-u.ac.jp

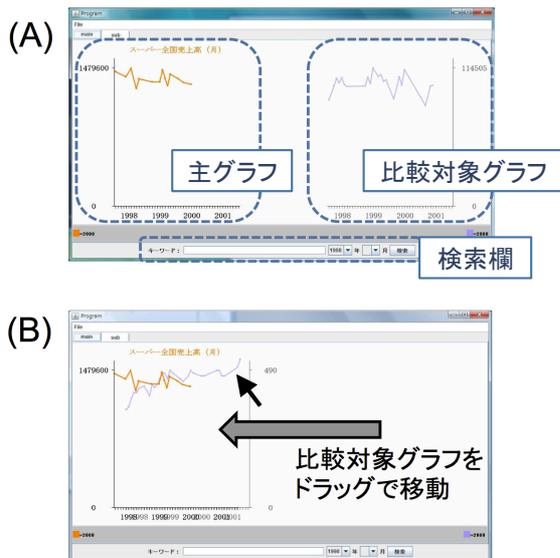


図 1: 提案するグラフ型インターフェース

り、必要なデータを取捨選択したりしていく。我々はこのような分析者の比較行為に着目し、データ全体の概観、データ間の比較、データの詳細の閲覧、を簡便に行うためのグラフ型インターフェースを持つシステムを実装した [6]。このシステムは、(1) 情報を概観するために折れ線グラフを用いて動向情報の推移を表示する、(2) 情報の比較をするために同時にふたつのグラフを並べて表示し、それらを直接操作により部分的に重ね合わせられるようにする、(3) グラフ上のデータをクリックすることで、そのデータの時点に関連する新聞記事にアクセスできるようにする、という機能を通じて、分析者の探索行為の包括的な支援を試みている。

図 1 に作成したプロトタイプシステムのインターフェースを示す。図 1-(A) はデータの比較を行うためにふたつのグラフを並べて表示した状態である。ふたつのグラフのうち、左側のグラフは分析者の主たる関心となるデータを表示したものである。本稿ではこれを主グラフと呼ぶ。また、右側のグラフは主グラフと比較したいデータを表示したものである。これを比較対象グラフと呼ぶ。図 1-(B) に示すように、比較対象グラフをマウスでドラッグすることにより、これらふたつのグラフを任意の割合で重ねることができる。

グラフとして表示させるデータは、予め用意された 70 個のデータの中から分析者自身が選択する。システムは、分析者が関心のあるデータを探す際に、他のデータを一覧として周辺知覚的に提示する。そのため、分析者がその時点で関心を持っていないデータにも気づきを得られるように設計されている。

また、分析者がグラフ上にプロットされている点 (data point) をクリックすることで、その点に対応する新聞



図 2: 検索結果の例

記事データがブラウザ上に表示される (図 2)。加えて、システム下部に設けられた検索欄 (図 1-(A) 参照) に分析者がキーワードと粒度 (年もしくは月) を指定することで、それに即した詳細情報が表示される。

### 3 実験

提案システムの被験者実験は、NTCIR-9 VisEx のトレンド要約課題 [2] を用いて行われた。被験者は実験を実施した企業によって募集で集められた 20 代の男女 10 名であった。被験者はランダムに 2 群に分けられ、一方の群が我々の提案システムを、もう一方の群がオーガナイザによって提供されたベースラインシステムを使用して課題を遂行するように指示された。実験によって、各被験者のシステム操作ログ、与えられた課題に対するレポート、各課題の実施直後に行った事後アンケート、全ての課題終了後に行った最終アンケートが収集された。

被験者には、課題の結果にかかわらず実験終了後に一定額の報酬が支払われた。被験者に与えられた課題は、携帯電話等の通信機器の加入状況 (T1)、ガソリン価格の状況 (T2)、雇用の変化の状況 (T3)、人口構成の変化の状況 (T4) の 4 つであった。なお、実験デザインの詳細については NTCIR-9 成果報告会の VisEx タスク実施報告 [2] を参照されたい。

被験者実験によって得られた結果から各課題における提案システムの各機能の使用回数を表 1 に示す<sup>1</sup>。表 1 において“キーワード検索”はキーワード検索を行った回数、“点のクリック”はグラフ上にプロットされた各点をクリックして新聞記事の検索を行った回数、“グラフ選択”は表示する統計グラフを選択するためにメ

<sup>1</sup>被験者 F07 の課題 T4 遂行時のログに関しては、当該被験者が統制条件を満たしていなかった (使用が許可されていない別のシステムを使用していた) ため、分析対象から除外した。

表 1: 各機能を利用した回数 (A: キーワード検索、B: 点のクリック、C: グラフ選択、D: グラフ移動、E は年月の変更)

課題	被験者	A	B	C	D	E
T1	F06	0	3	2	1	12
	F07	1	25	2	1	1
	F08	3	12	2	0	35
	M06	0	16	8	4	62
	M07	3	11	7	3	3
	平均	1.4	13.4	4.2	1.8	22.6
T2	F06	8	2	3	2	5
	F07	0	26	1	0	0
	F08	0	15	1	0	17
	M06	0	10	3	0	35
	M07	0	25	1	0	13
	平均	1.6	15.6	1.8	0.4	14.0
T3	F06	2	6	9	8	7
	F07	0	36	2	7	0
	F08	0	10	3	0	12
	M06	1	5	6	5	52
	M07	0	21	6	8	4
	平均	0.6	15.6	5.2	5.6	15.0
T4	F06	14	0	2	1	2
	F07	—	—	—	—	—
	F08	8	7	3	0	43
	M06	4	1	2	0	33
	M07	18	6	2	1	64
	平均	11.0	3.5	2.3	0.5	35.5

ニュー画面を開いた回数、“グラフ移動”はグラフの重ね合わせなどのためにグラフを移動した回数、“年月の変更”は年月を前後に遷移させたり粒度を変更したりした回数、を各々表している。また、図3は課題毎のそれらの平均値である。これらを手がかりにして、2章で述べた情報探索行為の各段階、すなわち概観フェーズ、比較フェーズ、詳細閲覧フェーズの3つのフェーズにおけるユーザ行為の特徴について観察する。

概観フェーズは、表示する統計グラフを選択する行為とグラフへ可視化する行為から観察できる。表1によると、グラフの選択回数はT3で平均5.2と最も多く、T2では1.8、T4では2.25と少なくなっている。ログを観察すると、課題T3では、被験者らは共通して“完全失業率”ないし“完全失業者数”のグラフを選択していた。それに加えて“事由別完全失業者数”や“有効求人倍率”などのグラフを選択する被験者も観察された。したがって、T3でグラフの選択回数が増加しているのは、事前に準備されたグラフの中に課題T3

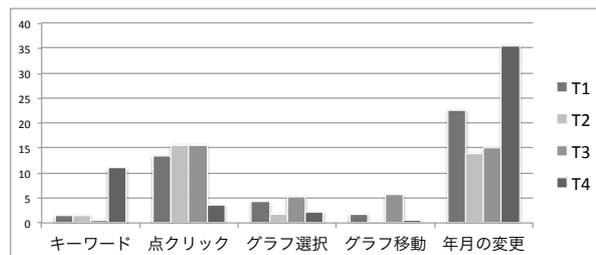


図 3: 各行為の課題ごとの平均回数

に適したグラフ種が豊富に存在していたためと推察される。

比較フェーズは、2つのグラフを並べて表示したり、重ね合わせたりする操作から観察できる。表1によると、グラフを移動させる回数はT3では5.6と多いのに対し、T2とT4ではそれぞれ0.4、0.5と少なかった。これは、概観フェーズにおけるグラフを選択する回数と同じ傾向である。また全ての実験を終えた後の最終アンケートによると、T1においては日本のレギュラーガソリン価格と原油価格との間に逆相関があることに気づいた被験者がいた。このような気づきは、ベースラインシステムを用いた被験者では観察されなかった。この結果は、グラフの比較・重ね合わせ機能が相関や逆相関を発見する一助となりえることを示唆している。

詳細閲覧フェーズは、グラフ上の点をクリックしたり、キーワードによる検索を行ったりする操作から観察できる。表1によると、T1、T2、T3ではグラフ上の点をクリックする回数がそれぞれ13.4、15.6、15.6と多く、キーワード検索の回数は1.4、1.6、0.6と少なかった。一方T4においては、グラフ上の点をクリックする回数が3.5と最も少なく、キーワード検索の回数は11.0と最も多かった。実験前の予想では、システムが事前に用意したグラフの中に課題遂行に適切なものが見つからなかった場合に、キーワード検索が用いられると考えていた。T2では、被験者のほとんどが“内閣支持率”のグラフしか選択してしなかったにもかかわらず、キーワード検索の回数が少なかった。その理由としては、グラフ上の点のクリックで検索される記事だけでもレポートを纏めるのに十分な量であったためではないかと推察される。一方課題T4も課題T2同様、事前に用意されたグラフの中に適切なものが少なかったが、この課題ではキーワード検索が多用されていた。その理由としては、T4では被験者に提示されたレポート作成指示文書に“年少人口”や“高齢人口”といった具体性の高い用語が含まれていたために「何を検索すべきか」がより明確になり、結果としてキーワード検索が誘発されたのではないかと考えている。

次に、被験者の操作ログから、被験者が検索した新

聞記事の発行日を抽出して調査した。その結果、日付の古い記事から新しい記事へと順番に検索する傾向が観察された。この傾向は、T1 から T4 へと実験が進むにつれてより明確に観察できた。実験前の予想では、被験者がグラフ上の点をクリックする際は、特徴的な点（極大/極小値や変曲点など）や変化（急騰/緩化した時点）に注意が向けられると考えていた。これは、今回の実験ではレポート作成が課題であり、4年間分のデータを時系列に沿って順番に見ていく方が手早く纏められたからではないかと考えている。

## 4 考察

3章で観察された結果を踏まえ、システムの利用傾向と改善点について考察する。

概観フェーズと比較フェーズの結果から、取り組んでいる課題に関連したグラフが豊富に利用できる場合、そのインタフェースを介した情報アクセスが増加すること、グラフの特徴を生かした特徴発見が促されることが示唆された。今回の実験では、VisEx のタスク説明<sup>2</sup>を参考に、グラフ描画に必要なデータを予め手で用意したため、課題によって利用できるデータの数にばらつきがあった。しかし、本来の探索的データ分析の過程では、事前にデータを用意しておくことはできない。そのため、ユーザの検索意図に基づいて、グラフ描画に必要なデータを収集し動的にグラフを描画する機能が必要となる。加えて、探索的データ分析では、最初から明確な目的が決まっていないため、ユーザの要求に応じてグラフの一覧を提示するだけでなく、システム側からの働きかけも必要になると考えている。

全課題終了後に実施した最終アンケートによると、グラフから新聞記事へアクセスする機能が使いやすかったという意見があった反面、新聞記事の検索結果がブラウザに表示された後の操作に関しては、使いづらいという意見が得られた。また、グラフ上にプロットされた点が小さいためにクリックがしづらい、検索精度が悪いといった指摘もあった。これらの指摘は、インタフェースがユーザの行為の支援に有望ではあるものの、現時点では十分なユーザビリティを有していないことを意味している。

検索精度の問題に関しては、今回の実装で用いたグラフの点の「意味づけ」が影響していると考えている。今回の実装ではクリックした点に応じて、描画しているグラフのトピック名から作成したキーワードとクリックされた時点を用いて、新たに新聞記事の検索を行う実装であった。これは、統計グラフをその時点の状況にアクセスするための時間的手がかりとみなしていた

ためである。一方で、統計グラフはあるトピックに関する統計量の視覚的要約とみなすこともできる。被験者が「精度が悪い」と感じたのは、点をクリックした意図がその時点の値を含む新聞記事テキストへのアクセス、すなわち「元記事」へのアクセスであるにもかかわらず、記事検索が実行されてその結果が提示されたためだと推察される。これら二つの実装方式については、どちらも一長一短があるため、今後解決策の検討をすすめていく。

## 5 終わりに

本稿では、NTCIR-9 VisEx タスクで実施された被験者実験の結果を通じて、グラフ型インタフェースを各被験者がどのように利用したかを分析し考察した。今後、実験によって観察された被験者の利用傾向を精査し、システムへの改良につなげたい。

本研究の遂行にあたり、文部科学省科学研究費（課題番号：22300048）の助成を受けた。記して謝意を表す。

## 参考文献

- [1] Hartwig, F. and Dearing, B.: *Exploratory Data Analysis*, SAGE Publications (1979).
- [2] Kato, T., Matsushita, M. and Joho, H.: Overview of the VisEx Task at NTCIR-9, In *Proc. 9th NTCIR Workshop Meeting on Evaluation of Information Access Technologies* (2011).
- [3] Shneiderman, B.: The Eyes Have It: a Task by Data Type Taxonomy for Information Visualizations, In *Proc. IEEE Symposium on Visual Languages*, pp. 336–343 (1996).
- [4] 加藤恒昭, 松下光範, 神門典子: 動向情報の要約と可視化: 言葉と図で情報をまとめる, *情報処理*, Vol. 47, No. 9, pp. 1013–1020 (2006).
- [5] 高間康史, 狩野真次: 比較分析に着目した時空間的動向情報の探索的分析支援, *人工知能学会論文誌*, Vol. 26, No. 4, pp. 494–503 (2011).
- [6] 蓮井大樹, 田中和広, 松下光範: 時系列データの探索的分析を支援するグラフ型インタフェース, 第21回WEBインテリジェンスとインタラクション研究会公演予稿集, pp. 35–36 (2011).

<sup>2</sup><http://must.c.u-tokyo.ac.jp/visex/>