

構造を持つ時系列情報に関する3次元可視化フレームワーク

3D Visualization Framework for Temporal Changes in Data with Structure

伊藤 正彦 豊田 正史 喜連川 優*

Summary.

ソーシャルネットワーク、衛星観測データなど、構造や属性値が動的に変化するデータを多角的に分析可能とする3次元可視化フレームワークを提案する。本フレームワークでは、グラフや地形等の構造データを可視化するインタラクティブなキャンバスコンポーネントを、時間軸に沿って複数配置し、その間を属性値の変遷を表すコンポーネントで結合することで、構造の変化および属性値の変動の把握を可能とし、さらに複数の観点で作成されたデータの変遷を並べて比較することも可能とした。本稿では、様々な実データを用いたアプリケーションの構築例を挙げて、本フレームワークの有用性および多様なアプリケーションに対する適用可能性を示す。

1 はじめに

構造や値の時間変化を観測することで現象を理解する必要のある問題は数多く存在している。情報の時間変化を可視化することにより、どの時期にどのような情報が存在するかを視覚的に確認出来るようになり、その変化や傾向を観察することが可能になる。時系列情報の可視化には、連続数値の可視化を行うものや、時間に沿って変化するデータ構造を可視化するものがあり、それぞれ多くの手法が提案されている[10, 8, 11, 3, 2, 9, 12]。

本論文では、構造をもつ情報における値やその構造自体の時間変化を観測可能にする3次元アプリケーションを構築するための3次元可視化フレームワークを提案する。ここでいう構造とは、ネットワークや木構造のようなデータ構造から地図情報のような2次元の位置情報、地形情報、建築物や生命情報における部品・部位の配置情報のような立体構造など様々なものを考えている。

提案フレームワークでは、直接操作が可能な時間軸コンポーネントおよび可視化キャンバスコンポーネントを組み合わせることで3次元可視化アプリケーションを構築する。可視化キャンバス内ではネットワーク構造のような2次元の構造、および地形図のような3次元の構造を表現するための様々な種類の表現コンポーネントを利用出来る。これらのコンポーネントにより可視化された構造は、可視化キャンバスコンポーネントの時間軸上の位置の応じて変化させることが可能である。また、時間軸上に値の変化を表現するコンポーネントを追加することで、例えばネットワークの各ノードの値の変化を構造変化とともに可視化することも可能となる。

最後に本稿では、提案フレームワークを様々なデータに対し適応し、多様な可視化アプリケーションを構築可能などを示すことで、提案フレームワークの汎用性を示すと共に、具体的にどのような現象を視覚的に確認出来たのか述べ、その有用性を示す。

2 関連研究

ウェブもしくはブログのリンク構造の様なネットワーク構造に関する時間変化を可視化する仕組みとして様々な手法が提案されている。(1)アニメーションの利用[8, 10, 3]: 時間による構造変化を直感的に見ることが出来る一方、状態が移り変わることで直前の状態との直接比較が困難になる。(2)スプレッドシート風インターフェース[2, 11]: 時間変化の全体像を把握しやすく、異なる時間における構造の違いを比較することが容易になるが、隣接ネットワーク同士の時間間隔が認識しにくい。(3)3次元空間の一軸を時間軸として利用[3]: 時間変化の全体像を把握しやすくなる一方、局所的な変化の認識が難しい。(4)複数時間におけるネットワークを重畠表示する手法[10]: 異なる時間におけるネットワーク構造の詳細な比較が可能になる一方、変化の全体像の把握が難しい。我々のシステムでは、構造の時系列変化を可視化する上記4種類の手法を3次元空間上で統合し、シームレスに切り替え可能にしている。

上記の手法では頻度変化を俯瞰する仕組みは用意されていない。TIARA[12]はキーワードごとの頻度変化を俯瞰可能にしたが、構造変化の可視化には対応していない。VisualLinda[9]とDwyerらの可視化システム[1]は時間軸上の各要素の状態の変化とメッセージフローを3次元空間を利用して同時に可視化した。これらは変化と流れの俯瞰に優れる一方、複数の異なる時間における状態の違いの詳細を比較調査することが困難である。我々のシ

Copyright is held by the author(s).

* Masahiko ITOH, Masashi TOYODA, and Masaru KITSUREGAWA, 東京大学生産技術研究所

ステムでは構造変化と頻度変化およびメッセージフォローを同時に可視化し探索することを可能にしている。さらに時間軸上に複数の可視化部品を配置することで変化の詳細を比較観測可能にしている。

Moving Phenomenon [7] と GeoTime [6] ではそれぞれ地図上に時間軸を上方向にとり、情報の散布や頻度変化、物流のトレース情報等を可視化している。これらの手法では地図空間の一軸を時間軸として既に利用済みのため、立体的な位置情報を持つデータの時間変化を可視化することが難しい。

3 時系列変化 3 次元可視化フレームワーク

構造および属性値の時間変化を可視化するため、(1) 様々な構造を持つデータの特定の時間のスナップショットを可視化、(2) 時間を変化させることによるアニメーション表示、(3) 複数時間のスナップショットの比較、および(4) 構成要素の値の変遷を可視化可能にする。さらに(5) 複数の観点から抽出したデータの比較をすることで対象の特徴を観測することを可能にする。

本フレームワークでは、時間軸コンポーネント (TimeLine) 上に直接操作が可能な可視化キャンバスコンポーネント (TimeSlice [5] もしくは TimeCube) を配置することで、様々な構造を持ち、かつ、その属性値もしくは構造そのものが変化するデータを可視化する 3 次元アプリケーションを構築可能にする(図 1)。

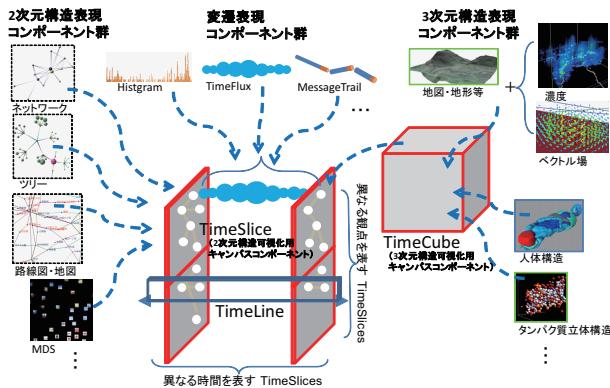


図 1. 時系列変化 3 次元可視化フレームワーク

キャンバスコンポーネントはユーザの直接操作により TimeLine 上に任意の数だけ追加可能である。また、キャンバスコンポーネントをドラッグし TimeLine 上で横方向にスライドすることで時間変化をアニメーション表示させることができる。さらに、異なる観点から抽出したデータを可視化するためのキャンバスコンポーネントを縦方向に追加し、それらの間の同期をとりながら比較することができる。複数のキャンバスコンポーネントを用いる際、時

間が近すぎてキャンバスコンポーネント同士が重なり見づらくなる場合や、それらを並べて比較したい場合がある。そのような場合、キャンバスコンポーネントを奥行き方向にスライドすることで重なり回避が可能であり、図 7 の様なタイル表示を行うことで詳細な比較を行うことも可能である。

キャンバスコンポーネントは 2 次元の構造を内部に可視化するための TimeSlice と 3 次元の構造を内部に可視化するための TimeCube からなり、目的に応じて入れ替えることが可能である。TimeSlice では、可視化する対象にあわせて 2 次元構造表現コンポーネントを選択し平面内に構造を可視化する。現在、ネットワーク表現、ツリー表現、路線図・地図、および MDS (Multidimensional Scaling) のための 2 次元構造表現コンポーネント群が利用可能である。同様に、TimeCube では、可視化する対象にあわせて 3 次元構造表現コンポーネントを選択し空間内に構造を可視化する。現在、地図・地形図、濃度およびベクトル場表現、人体立体構造、タンパク質立体構造表現のための 3 次元構造表現コンポーネント群が利用可能である。

特定要素の属性値の変遷を可視化するための機能は変遷表現コンポーネントとして TimeFlux [4] を TimeLine に追加することで実現出来る。現在、ヒストグラムを表示する変遷表現コンポーネントおよびデータやメッセージの流れを可視化する変遷表現コンポーネントである MessageTrail [4] が利用可能である。

提案フレームワークを用いることで時間情報、要素の位置情報、および(必要に応じて)要素間の関係情報を持つデータに対して可視化アプリケーションを開発可能である。ユーザは、データに合わせてキャンバスコンポーネントおよび時間軸コンポーネントの基本クラスを拡張することで任意のアプリケーションを開発出来る。その際、任意の既存表現コンポーネントを組み合わせて利用可能である。

4 アプリケーション例

4.1 ブログ上のインフルエンサー追跡

著者らの研究室では、過去 5 年間、3 億 5000 万エントリに及ぶ日本語ブログアーカイブの構築を進めてきている。これらのアーカイブから作成した時系列情報を持つウェブグラフのインタラクティブな 3 次元可視化を実現している [5]。この例では、ブログのエントリをノード、エントリ間のハイパーリンクをエッジとするネットワークを描画している。

図 2 にブログ空間上における「ワーキングプア」に関連する人々の興味の移り変わりを可視化した例を示す。ヒストグラム表示により、流行のピークが複数回観測出来る。また、複数 TimeSlice により、各ピーク後におけるインフルエンサーノードが移り変わっていることが確認出来る。

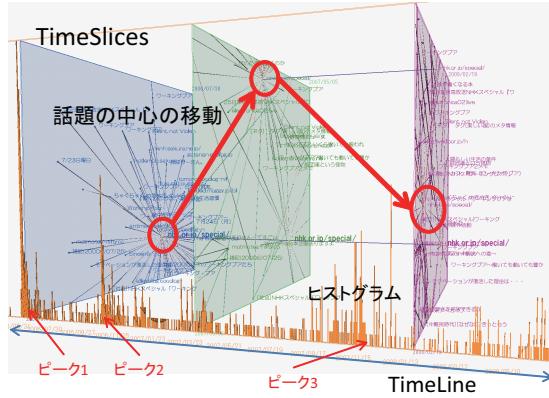


図 2. ブロギング構造の時系列変化可視化による「ワーキングプア」の話題に関するインフルエンサー追跡

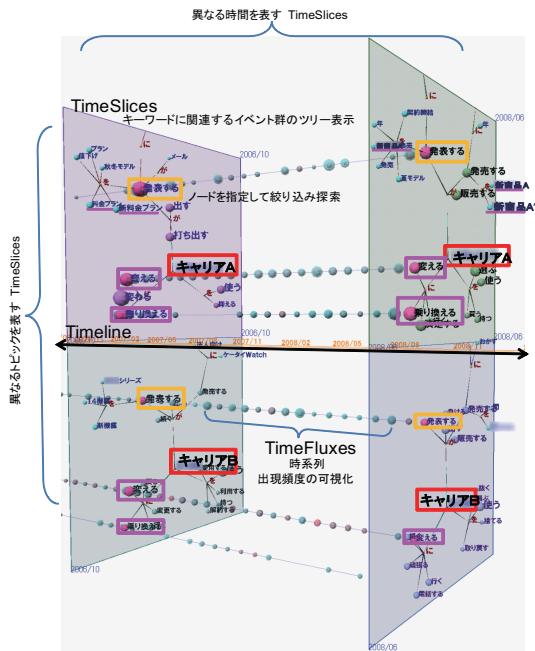


図 3. 係り受け解析を用いたブログ話題追跡システム

詳細に探索すると、ワーキングプア問題を広めるきっかけとなった NHK の番組にリンクが集中している様子が見て取れ、1回目、2回目、3回目の放送に伴って、話題の中心が移動していることが分かる。

4.2 係り受け解析を用いた話題追跡システム

4.1 節で用いたブログアーカイブに対して、係り受け解析を行い、ブログユーザの行動・興味に関する記述（イベントと呼ぶ）を月ごとに収集し可視化した [13]。図 3 は、複数の携帯キャリアに関するイベント群を可視化、比較した例である。ノードサイズはイベントの出現頻度を表す。図中、上部の TimeSlices はキャリア A、下部はキャリア B に関する話題を可視化している。

時間軸に沿って上下のイベント群をアニメーショ

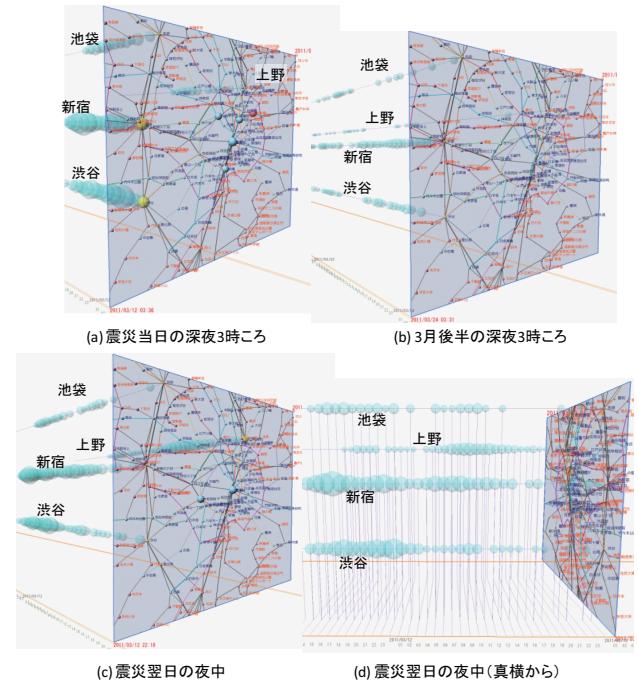


図 4. 2011 年 3 月 11 日震災後における東京周辺各駅に関するツイート頻度推移

ンさせ比較することで、両者の傾向の違いを観測出来る。これにより、キャリア B に比べてキャリア A では「乗り換える」の記述が目立つことが分かった。TimeFluxes を表示することで「乗り換える」ピークが来たタイミングを直感的に確認出来る。このタイミングにおける詳細を観測すると、キャリア A では各ピーク時に「新料金プラン」や「新商品 A」を発表もしくは発売していることが分かり、キャリア A のマーケティング戦略は社会に大きなインパクトを与えていたことが伺える。

4.3 鉄道・地下鉄路線図上の状況観測システム

東京近郊の地下鉄、JR、私鉄の路線図上の各駅に関連するツイートを収集し、時間毎に集約し、その頻度とツイートを各駅上に可視化することによる状況観測システムを構築した。選択された時間周辺におけるツイート出現頻度に応じて駅を表すノードのサイズが変化する。駅ノードを選択することで TimeFlux を表示し、指定駅に関するツイート数の変遷を一覧出来できる。図 4 では 2011 年 3 月 11 日に発生した東日本大震災後の状況を可視化した。

図 4(a-b) では震災当日と数週間後の深夜 3 時頃における各駅に関するツイート数を可視化・比較した。通常、図 4(b) の様に深夜に駅に関するツイートがされることはないが、図 4(a) から震災直後は多くの人が徹夜で駅の状況をつぶやき続けたことが分かる。

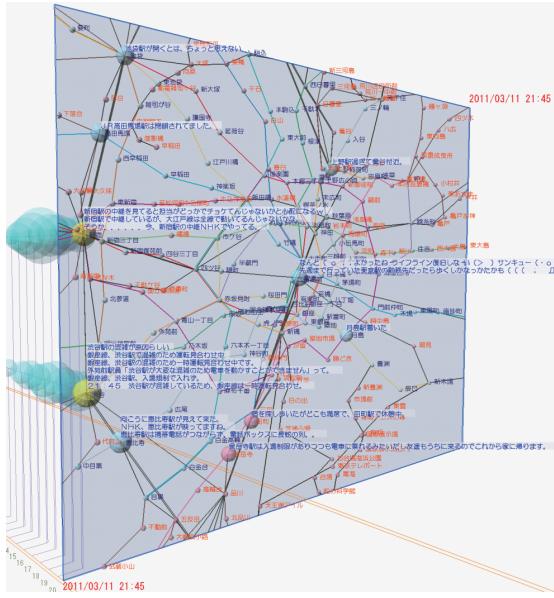


図 5. 2011 年 3 月 11 日震災後における東京周辺各駅の状況に関するツイートの時間変化

図 4(c) および図 4(d) では震災翌日の夜中までの渋谷、新宿、上野、および池袋各駅に関するツイート数推移を TimeFlux により可視化・比較した。これにより、各駅の時間帯による傾向の違いとピーク点を直感的に理解可能になる。渋谷駅では震災当日深夜になるとツイート数が急激に減少している。一方、新宿駅では翌日の朝まで多くのツイートがあることが分かる。図 5 のようにツイートを表示し詳細を見ることで、多くの人が新宿駅周辺で野宿をするなど夜を明かしたことなどが確認出来た。また、TimeFlux から上野駅では朝方に向けてツイート数が上昇していることが確認出来る。ツイートを見ることで、JR を目当てに集まった人々が JR に乗れずに大混乱が起きていることが分かった。

4.4 ツイッター上のコミュニケーション可視化

図 6 の例では、Twitter のフォローネットワークと、そのツイート数、リプライとリツイートの流れを可視化した。フォローネットワークは、まず入力したユーザに関して指定した段階数のフォロワーを取得し TimeSlice 上にネットワークを可視化する。ユーザを選択して展開することでネットワークを広げていくことが可能である。ノードサイズには、ツイート数、被リプライ、被リツイート数等を割り当て可能である。また、MessageTrail コンポーネントを用いてリプライとリツイートの流れと実際のツイートを可視化すること（図ではツイートは省略している）で、ユーザ間でどのような会話がなされているか追いかけることや、頻繁なやりとりなのか長期的なやりとりなのかといった傾向、話題による会話の相手に違い等を可視化可能である。これらによ

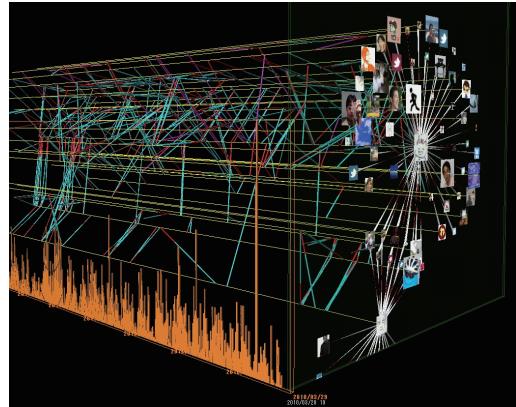


図 6. ツイッターのフォローネットワークとコミュニケーションの可視化

り情報がどのように伝搬しているか、ユーザの役割等を観測することが出来る [4]。

4.5 史料データにおける人間関係構造変化可視化

歴史史料から抽出された人物関係ネットワークに対する 3 次元可視化システムを実現した [14]。東京大学・史料編纂所にて公開されている大日本史料データベース¹から、ノードを人物、エッジをその関係とした人物関係ネットワークを抽出した。ノードのサイズは人物の重要度を求め割り当てている。

図 7 では左から順に 1568 年、1569 年、1570 年の“戦”という観点からの織田信長周辺人物ネットワークを表示している。1568 年以降、信長が大きな中心ノードに成長していることが確認できる。1568 年は信長が足利義昭を奉じて上洛する年であり、この年から急激に信長が様々な人物と関係し始めていることが伺える。

図 8 は“戦”および“贈答”といった複数の観点からなる時系列人物ネットワークを同時に可視化し比較した例である。双方のネットワークに共通して現れる人物ノードは緑色でハイライト表示されている。図から信長の時代には、戦に関わる人間関係の活発さが見て取れるが、贈答に関しては、小さな人間関係のみが観測出来る。一方、大阪夏の陣のころは、贈答と戦に関する人間関係の活発さが同様に示され、二つの観点間で多数の共通人物によりネットワークが構成されていることが観測出来る。ここから武力のみではない駆け引きが徳川家を中心に行われていたことが伺える。

4.6 地球観測データ・インタラクティブ 3 次元可視化システム

図 9 は、地球観測データに対するインタラクティブ 3 次元可視化環境の構築例である。地球観測デー

¹ <http://wwwwap.hi.u-tokyo.ac.jp/ships/db.html>

3D Visualization Framework for Temporal Changes in Data with Structure

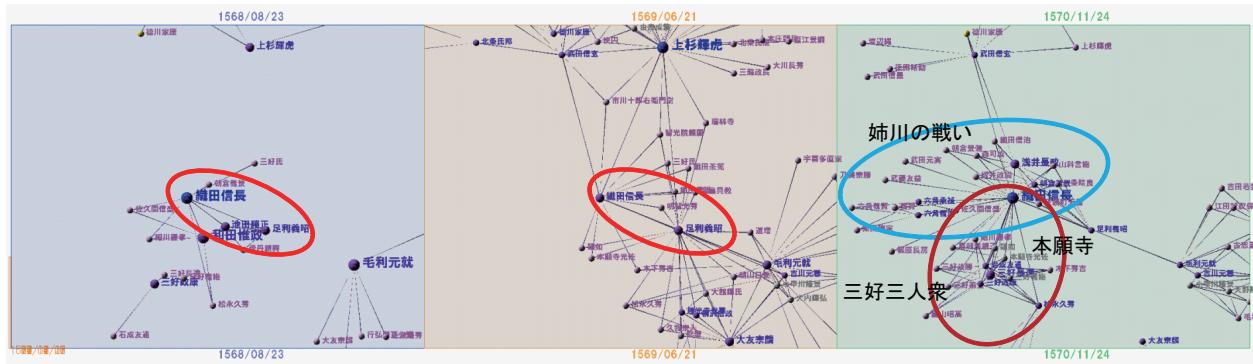


図 7. 史料データから抽出された 1568-1570 年における“戦”に関する信長周辺ネットワーク（タイル表示）

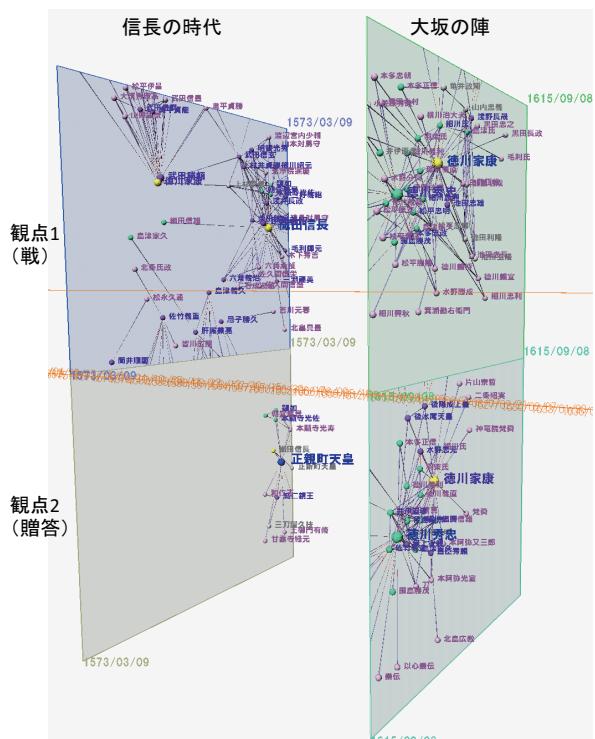


図 8. 史料データから抽出された戦国期 - 江戸時代初期における“戦”および“贈答”に関する人物ネットワークの比較

タとしては、著者らの研究室が中心となり構築している、海洋地球観測探査システム：データ統合・解析システム (DIAS)²において収集された TRMM PR の降雨強度データを用いている。これは、時間、緯度経度、高度、および濃度からなる 3 次元分布データで、TimeCube コンポーネントと濃度表示コンポーネントにより可視化している。青を低濃度、赤を高濃度にマップしている。

図の例では台風発生シーズンにおける 2001 年 9 月の台北周辺上空の降雨量の時間変化をユーザの直

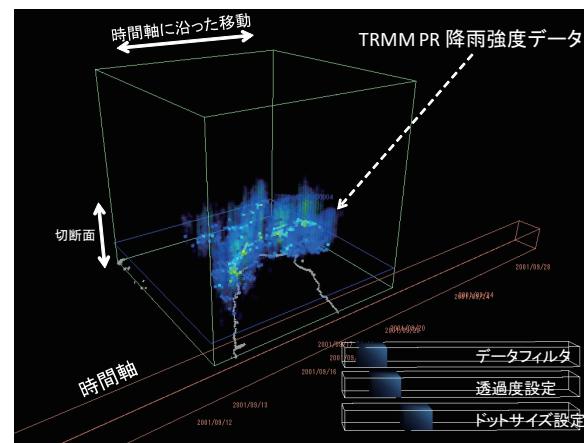


図 9. 台風発生シーズンにおける台湾上空の降雨量の時間変化可視化

接操作により探索している。台風シーズンに台湾の切り立った海岸線上空に向かって雨が集中している様子が読み取れる。

5まとめ

本稿では、構造および値の時間変化を可視化するための 3 次元可視化フレームワークを提案した。単一のフレームワークにより、3 次元データをも対象にした様々なトレンド探索アプリケーションを構築可能にする点を提案フレームワークの最大の特徴とする。提案フレームワークは、今後ますます時間とともに累積していく様々な種類の膨大な量のデータやそれらに対する解析結果に対して、多角的に分析を行うための基盤となることを目指す。

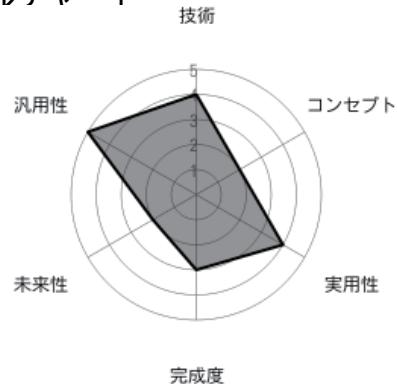
3 次元可視化には情報重なり問題など様々な課題が残されているが、本稿で挙げたような、構造と複数要素の頻度変化の同時可視化や、立体構造を持つデータの可視化など 3 次元を用いることが有用な場合は数多くある。今後の課題としては、情報重なり問題、およびパースによる認知の妨げ問題への回避手法の導入などが挙げられる。

² <http://www.editoria.u-tokyo.ac.jp/dias/index.html>

参考文献

- [1] T. Dwyer and P. Eades. Visualising a Fund Manager Flow Graph with Columns and Worms. In *Proc. IV 2002*, pp. 147–152, 2002.
- [2] E. H.Chi and S. K.Card. Sensemaking of Evolving Web Sites Using Visualization Spreadsheets. In *Proc. INFOVIS 1999*, pp. 18–25, 1999.
- [3] E. H.Chi, J. Pitkow, J. Mackinlay, P. Pirolli, R. Gossweiler, and S. K.Card. Visualizing the evolution of Web ecologies. In *Proc. CHI 1998*, pp. 400–407, 1998.
- [4] M. Itoh. 3D Techniques for Visualizing Users' Activities on Microblogs. In *Proc. the IET International Conference on Frontier Computing 2010*, pp. 384–389, 2010.
- [5] M. Itoh, M. Toyoda, and M. Kitsuregawa. An Interactive Visualization Framework for Time-Series of Web Graphs in a 3D Environment. In *Proc. IV 2010*, pp. 54–60, 2010.
- [6] T. Kapler and W. Wright. GeoTime Information Visualization. In *Proc. INFOVIS 2004*, pp. 25–32, 2004.
- [7] K.-S. Kim, K. Zettsu, Y. Kidawara, and Y. Kiyoki. Moving Phenomenon: Aggregation and Analysis of Geotime-Tagged Contents on the Web. In *Proc. W2GIS 2009*, pp. 7–24, 2009.
- [8] M. Kitsuregawa, T. Tamura, M. Toyoda, and N. Kaji. Socio-Sense: A System for Analysing the Societal Behavior from Long Term Web Archive. In *Proc. APWeb 2008*, pp. 1–8, 2008.
- [9] H. Koike, T. Takada, and T. Masui. Visu-aLinda: A Framework for Visualizing Parallel Linda Programs. In *Proc. VL 1997*, pp. 174–182, 1997.
- [10] N. Nakazono, K. Misue, and J. Tanaka. NeL2: Network Drawing Tool for Handling Layered Structured Network Diagram. In *Proc. APVis 2006*, pp. 109–115, 2006.
- [11] M. Toyoda and M. Kitsuregawa. A System for Visualizing and Analyzing the Evolution of the Web with a Time Series of Graphs. In *Proc. HYPERTEXT 2005*, pp. 151–160, 2005.
- [12] F. Wei, S. Liu, Y. Song, S. Pan, M. X. Zhou, W. Qian, L. Shi, L. Tan, and Q. Zhang. TIARA: A Visual Exploratory Text Analytic System. In *Proc. KDD 2010*, pp. 153–162, 2010.
- [13] 伊藤正彦, 吉永直樹, 豊田正史, 喜連川 優. プログューザの行動・興味に関する時系列推移 3 次元可視化システム. 第 3 回データ工学と情報マネジメントに関するフォーラム, 2011.
- [14] 伊藤正彦, 赤石美奈. 3 次元可視化による史料データにおける人間関係構造変化の俯瞰. 第 82 回人工知能基本問題研究会, pp. 31–36, 2011.

アピールチャート



未来ビジョン

時系列データは今後ますます爆発的に増えている、より生活に密着したものになっていくことが予想される。例えば、ソーシャルネットワークのデータのようなウェブ上のデータはもちろんのこと、今後はセンサー情報等の多種多様な非ウェブデータがますます重要なとなると考えられる。非ウェブデータは非常に身近な生体情報や GPS による移動情報から、スマートグリッドや農業の土壤センサリングのような地域レベルのデータ、さらには論文中でも用いた地球観測データのようなグローバルレベルのデータまで様々である。

ここであげたデータの特徴として、まず「超巨大」であるということがあげられる。超巨大で構造を持つデータの可視化手法は既に様々なものがあるが、時間変化の可視化に対応したものや、ドリルダウンなどの探索操作をインタラクティブに行えるものは少なく、超巨大時系列構造データのインタラクティブ可視化環境

の構築はひとつのチャレンジである。

センサデータやツイッターのようなソーシャルメディアのデータにおける高い「リアルタイム性」も特徴の一つである。今現在から過去へのログを可視化し続けることにより、構造異常や値の異常を視覚的に認識することが可能になる。また、例えば論文中で用いた路線状況のツイッター表示などは、現状（路線が止まっているとか）の可視化を行うことでユーザが行動指針を決める助けになり得る。

データの様々なレイヤ間（生体情報、GPS 情報とスマートグリッドなど）における「連携統合」が情報探索には重要になってくる。そのために構造的な連携および時間的な連携の仕組みが必要になる。

提案フレームワークは、様々な種類の「超巨大」でかつ「リアルタイム」なデータやそれらに対する解析結果に対して、「連携統合」を行い多角的に分析を行うための基盤となることを目指す。