

論文要旨

学位論文題目「詳細度制御を適用した複合型時系列データの可視化」

八木 佐也香

ビッグデータ時代の到来に伴い、多種多様なデータベースに時系列性を有する情報が蓄積されている。その中には数値情報にあわせて、タグや位置情報といったメタ情報、あるいはリンクや階層といった構造などが付与されているものが多い。例として、気温や降水量などの気象情報と測定地の位置情報や天気などのメタ情報、株価と政治的・経済的事象を表すメタ情報あるいはウェブアクセスページ間のリンク構造などが考えられる。本論文ではこのような時系列データを複合型時系列データと定義し、時系列数値情報と付随情報の双方を関連性を捉えやすく可視化する手法を提案する。この情報可視化からのアプローチは、時系列データの観察と理解、またそれを扱う業務での能動的な意思決定を支援する道具として有効である。

本論文における複合型時系列データ可視化における着眼点は、別々の事象として蓄積された元の時系列数値情報と付随情報の双方を、関連性を捉えやすく可視化すること、画面の中で直接データの特徴を理解しながら、事象を観察・分析することである。しかし、近年のデータ量の急激な増加に伴い、観察対象となる時系列データも大規模・複雑化しており、画面上に描画した場合、観察が困難になることが多い。そこで、本論文では時系列性自体の描画を重視しながら、可視化手法の操作手順における"**Overview, Zoom, Filter, and Detail on Demand**"という標準的な考え方にもとづき、時系列データのマッピングを工夫することに加え、詳細度制御を実現することで、上記の2点を満たす可視化手法を検討する。

本論文では、まず実務面での意思決定支援を目的とした事例研究として、水防災データを対象にした複合型時系列データ可視化手法を提案する。局地的集中豪雨などの被害対策として、水防の専門家が水防災データを表示した結果を観察し、過去の危険時の数値情報(降雨量、河川水位など)の分析を行うことが重要である。それを踏まえ、単純な降雨量と水位の依存関係だけではなく、地理的分布や地形の特徴などの影響による知見を得ることで、より複合的な警告を行うことを目標としたものである。この事例では時系列数値変化の詳細を示す長方形を、重なりを回避しつつ、地図上に選択的に配置することで、地理的特徴と関連付けながら多数の数値情報を表示する。また、時系列数値変化の表現として、全体の俯瞰に適した表示と詳細な数値の読み取りに適した表示の2種類を用意し、俯瞰の度合いに応じて選択・切替を行うという詳細度制御を実現する。

続いて、より大規模なデータを対象にした汎用的な複合型時系列データの可視化手法の研究成果として、詳細な数値の読み取りに適した折れ線グラフベースの可視化手法とクラスタ構造の変遷を観察するのに適した **storyline** ベースの可視化手法の2種類を提案する。折れ線グラフベースの手法においては、付随情報としてメタ情報を持つ複合型時系列データを対象に、数値を折れ線の形状、メタ情報を折れ線の

色にマッピングする。この手法では、要素同士の比較を容易にするため、複数の折れ線グラフを一画面上に表示するというアプローチを取る。そのため、データが大規模になると、折れ線同士の絡まりによって視認性が低下するという問題が起こる。この問題に対し、数値とメタ情報の両方を考慮したクラスタリングによって詳細度制御を行うことで、グラフの表示本数を低減して読みやすさを向上し、かつ数値と付加情報の関係性も逃さない表示を実現する。また、クリックおよびスケッチ入力による要素抽出機能を実現することで、興味深い要素の時間変化を対話的に観察することを可能にする。storyline ベースの可視化手法においては、付随情報としてクラスタ構造に着目し、時系列データ要素の短時間単位の特徴と全体のクラスタ変遷の双方を同時に観察することを目標とする。この事例では、局所時間ごとに時系列実数値情報のクラスタリングを行い、隣接する局所時間帯間のクラスタ群にマッチングを適用することで画面上のクラスタ配置を決定する。この手順により類似した要素同士が画面内で近接するような storyline を描くことで、複合型時系列データのクラスタ構造の変遷を俯瞰することが可能になる。また、storyline と折れ線グラフによる数値表示の連携可視化を行うことで、興味深い局所の時系列数値情報の変化を直感的に理解することを可能にする。

将来的には、本論文で提案した storyline による可視化と水防災データ可視化手法のような地図上に時系列データをマッピングする手法を連携可視化することで、"Analyze First - Show the Important - Zoom, Filter and Analyze Further - Details on Demand"という Visual Analytics による分析の工程を実現できると考える。この一連の処理を繰り返すことで、複雑かつ大規模な複合型時系列データの効率的な分析が可能になり、その知見を知識化することで将来の予測を高精度で行うことが可能になると考えられる。