

グラフ構造を用いたデータマイニング結果の可視化方法について

氏名† 井又 圭太 氏名† 宇田川 佳久

所属† 東京工芸大学工学部コンピュータ応用学科

1. はじめに

今日の情報化社会において、大規模なデータを扱う機会が増えてきたが、大量のデータの中から必要な情報を得ることは容易ではない。必要なデータを自動的に抽出するアルゴリズムは様々あるが、データを自分の目で見て確かめつつ情報を取り出す方法は、直感的であるだけでなく、自動的な手法よりも有効な場合も多い。また、系列データマイニングでは一般に大量のデータが出力され、全体像や頻出するパターンの把握が難しい。本研究の目的は、ノードとエッジから構成されるネットワークデータを効果的に可視化する方法を構築し、パターンの把握を容易にする手法として、グラフを使った表示方法を開発する。

2. Vis.js について

Vis.js[1]は、米国の MIT などが開発したブラウザベースの可視化ツールである。Vis.js には、タイムライン、ネットワーク、2D グラフや 3D グラフ等を可視化する機能がある。データを動的に可視化する機能があり、表示したデータを部分的に移動したり、特定のオブジェクトに繋がっているデータだけを色を変えて表示したりするといった機能がある。

3. ネットワークデータについて

ネットワークとは、複数の要素が互いに接続された網状の構造体のことである。ネットワークを構成する各要素のことをノード(Node)、ノード間の繋がりのことをエッジ(Edge)と呼ぶ。多くの情報がネットワークで表現することができる。数学的にはグラフと呼ばれ、行列によって表現される。グラフは、相互に関連のある多数の要素の結びつきを表現するために広く使われている。紙の上に描かれた自由形式のマインドマップであれ、多数の要素間の結びつきの一般的なデータモデルを視覚的に表現したものであれ、データのネットワークを視覚表現することによって、それに精通している人にとって意味のある関連性が示されているからである。

4. Vis.js のコーディング方法

Vis.js のネットワーク可視化機能では、数千ものノードとエッジをブラウザ上で滑らかに動作することが可能である。Vis.js では、ノード、エッジのプロパティ（属性）を指定することで、様々な表示ができる。しかし、ブラウザ上に表示するためには、Vis.js が解釈できるよう、元のデータを HTML ファイルに変換する必要がある。

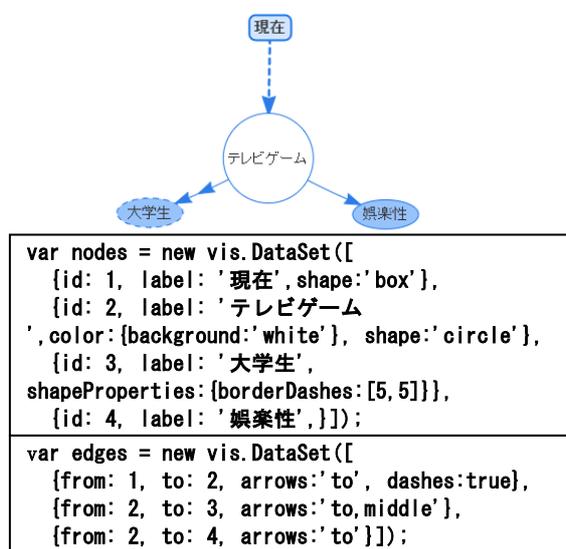


図 1. Vis.js のネットワークデータの例

図 1 は Vis.js が処理できるネットワークデータの例である。ノードを設定する際には、ID、Label 等の指定を行う。図 1 のノードでは、長方形や楕円形に形を変え、破線状になっている。エッジを設定する際には、指定されたノード間の ID で繋がりを設定し、矢印のプロパティを加えることができる。図 1 のエッジでは、矢印の数を増やし、ノードと同じように破線状に変更している。

5. データマイニングからの変換

Apriori アルゴリズムに代表されるデータマイニングの抽出結果は、大量に生成されることが知られている。また、抽出されるデータのパターンも、類似するもの、類似しないものが混在している。図 2 は、系列データマイニングの出

