

タワーマップ： 2 部グラフ構造と量的情報を同時提示する 3 次元可視化手法

伊藤 隆朗[†] 三末 和男[†] 田中 二郎[†]

筑波大学大学院 コンピュータサイエンス専攻[†]

1. はじめに

現実世界のデータの把握を行う際、2 つの集合間の関係性に注目することがよくある。例えば、Web サイトのアクセスログの場合、Web ページと訪問者の関係に着目する。このような情報把握の手段として、データを可視化することがよく行われる。

情報把握の際には、統計グラフ等を表示することによって量的な傾向を把握することが多い。しかしながら、この方法では、個々の接続関係といった詳細な情報の把握は困難である。一方、2 集合間の関係を 2 部グラフとしてとらえ可視化することによって、接続関係は把握できるが、統計情報で得られるような大きな傾向の把握は難しい。

本研究は、この 2 つの情報を同時に提示することで、2 集合間の関係把握の支援を目的とする。そこで、2 部グラフ可視化手法のアンカーマップ [1] を拡張し、各ノードに高さの軸を持たせることで量的情報を提示する 3 次元可視化手法「タワーマップ」を開発した。本稿では、ノード形状のバリエーションの検討と、その効果について述べる。

2. タワーマップ

2.1. アンカーマップ

$G=(A \cup B, E)$ を 2 部グラフとする。集合 A と集合 B は共通要素を持たないノードの有限集合で、エッジの有限集合 E は $A \times B$ の部分集合である。アンカーマップでは、 A の要素を「アンカー」、 B の要素を「フリーノード」と呼び (A と B は交換可能)、アンカーを円周上 (正多角形の頂点上) に固定し、フリーノードをスプリングモデル [2] を利用して配置する。アンカーマップでは、関係の深いアンカーができるだけ近くに配置されるようなアンカーの並び順を求めるアルゴリズムが提案されている。

Tower maps: A 3D visualization technique representing bipartite graph structure and statistical information simultaneously

[†]Takao Ito, Kazuo Misue and Jiro Tanaka, Department of Computer Science, University of Tsukuba

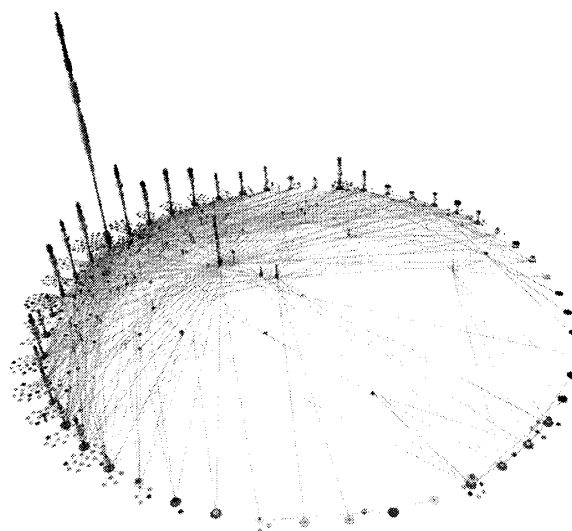


図 1 タワーマップの描画例

2.2. 量的情報の提示

アンカーマップでは、接続関係を基にした傾向の把握は可能だが、統計グラフで得られるような量的な情報の把握が困難である。そこで個々のノードに高さの軸を持たせ、それを用いることで量的な情報の提示を行った (図 1)。この表現手法を「タワーマップ」と呼ぶ。ノードごとに量的情報を提示することで、データ全体を画一的に統計処理した場合と比べ、個々の細かな情報の把握を行うことができる。また、アンカーマップのノード配置により、見たい量的情報の選択や、接続関係を考慮した比較・分析を行うことができると考えられる。

2.3. ノードの形状と配色

各ノードは接続関係以外にも属性を持つことが多い。アクセスログの場合、訪問者はアクセスした時間等の属性を持っており、この情報を時間帯等のカテゴリ毎に集計し分析することが良く行われる。本研究では、数量や属性内でのカテゴリの割合の比較の観点から 3 つのスタイルを設計した (図 2)。

[A] ノードを積み上げグラフのように提示する形式。各カテゴリの数量をそのまま高さに割り当て色分けする。

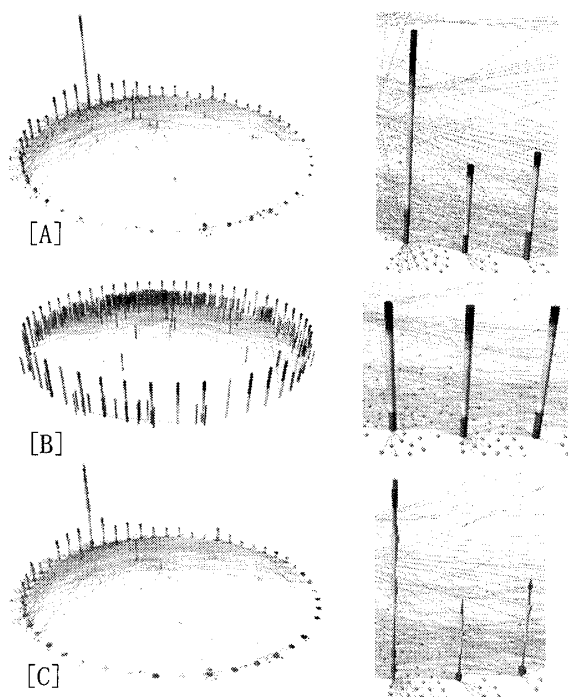


図 2 ノード形状と配色

- [B] ノードの高さを統一し、各カテゴリの占める割合を帯グラフのように提示する形式。カテゴリ毎に色分けをする。
- [C] 各カテゴリの高さを統一して割合を太さで表し、ノードの高さをを用いて数量情報を提示する形式。ただし、占める割合が 0 のカテゴリは彩度の低い色で表す。

[A]形式は、各カテゴリの数量がそのまま高さで反映されるため、ノード同士の数の比較を行うことができる。それに対し、[B]ではノードの高さが一定なため、割合の比較を行うことができる。[C]形式は、ノードの太さによって割合を表すことで、数量と割合の両方を提示することができる。

3. 可視化例：アクセスログ

図 1 は、第一著者の Web サイトの 1 週間分のアクセスログをタワーマップによって可視化した例である。この図では、Web ページをアンカー、訪問者 (IP アドレス) をフリーノードとしてレイアウトしている。ノードの形状は[C]スタイルを用いており、1 時間毎のアクセス頻度を表している。高さは 1 日のアクセス数、色は時間帯、太さは各時間におけるアクセス頻度を示している。アンカーを見渡すことで、アクセス数の多いページや、ページごとにアクセスされている時間帯の傾向が読み取れる。また、最もアクセス数の高いページのアンカーには、放射状にエ

ッジが広がって接続されている。このページは、サイト中でトップページであり、エッジの接続関係から、多くの訪問者はこのページを中心として閲覧している傾向が読み取れる。このように 1 つの可視化結果から数量的な傾向と接続関係の傾向の両方を読み取ることができる。

図 3 は、中央付近のノードにズームしたものである。これらの訪問者はフリーノードの接続関係から比較的多くのページを見て回っていることが分かる。しかし、(a)のページは他の訪問者と同様同一の時間帯に閲覧しているのに対し、(b)の訪問者群は閲覧の時間帯が分散している。これらの訪問者のログを調べたところ、(a)は google の検索で訪れたユーザであったのに対し、(b)は yahoo 等のクローラであった。このように量的情報をノードに付加することによって、詳細な情報の比較・分析を行うことができる。

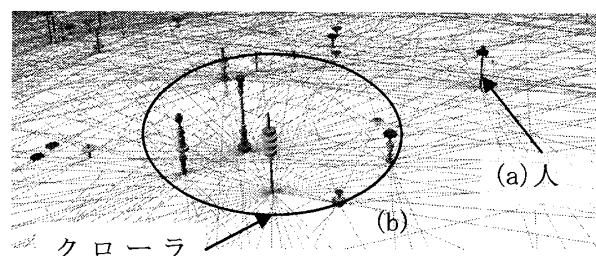


図 3 中央付近のズーム

4. まとめ

本稿では、2 部グラフ可視化手法アンカーマップを拡張し、ノードを 3 次元表現することで個々の量的情報を提示する手法「タワーマップ」について述べた。本手法により、2 部グラフの接続関係を考慮しながらの量的情報の比較・分析が行えるようになる。また、本手法は、ノードのクラスタリング手法と組み合わせ、クラスタを縮退・展開した際に量的情報を統合・分解することで、大規模な情報解析の支援として応用できると考えている。

参考文献

- [1] K. Misue. Drawing Bipartite Graphs as Anchored Maps. In *Proceedings of Asia-Pacific Symposium on Information Visualization (APVIS2006)*, pp. 169-177, 2006
- [2] P. Eades. A heuristic for graph drawing. *Congressus Numerantium*, Vol. 42, pp. 149-160, 1984.