

# 「左京と右京：2つの平安京ビューによるマトリクス型データの可視化」

橘春帆 伊藤貴之

お茶の水女子大学 理学部情報科学科

## 1. 概要

情報可視化は、世の中にある一般的な情報を可視化する研究分野であり、最近になって活発に研究が進められている。情報可視化が有効に活用できると考えられるデータの一つに、アンケートの集計結果があげられる。

アンケート集計結果は回答者数だけで列を持ち、質問に対する選択肢数だけ行を持つマトリクス型データの一つと考えることができる。このようなマトリクス型データを可視化する手法として、本研究では情報可視化手法である「平安京ビュー」を二つ用いて一画面に表示する手法「左京と右京」を提案する。本手法では、右京は回答者数だけのノードを有する階層型データ、左京は選択肢数だけのノードを有する階層型データを表示する。

アンケート集計結果の可視化には、単純な帯グラフや棒グラフなどがよく使用される。しかし、このようなグラフでは、回答者の全体の傾向などがわかるだけで、一人一人の特徴を見たいときには、そのグラフに表現されていない詳細情報を引き出すような操作や調査が必要だった。本研究の提案手法では例えば、ある特定の選択肢を選んだ回答者群の分布や、特定の回答者が選んだ選択肢の分布、などを可視化することが可能になる。これにより、少数だけど興味深い局所的傾向の発見ができると考えられる。

## 2. 関連研究

本研究に用いる情報可視化手法「平安京ビュー」[1]は大規模な階層型データの可視化に向いている。平安京ビューは階層型データ全体を一画面に配置する情報可視化手法である。葉ノード、枝ノードの直交配置を意識したアルゴリズムで、平安京の地図に似ていることから命名された。

階層型データを可視化する手法には、木構造を表示する手法[2][3]や、個々のノードが持つ占有率に比例した面積で画面を分割した手法[4]がある。

これらの手法に比べて、平安京ビューは階層型データを構成するすべての葉ノードを、一画面に、同じ大きさで、重なることなく表現できるという特徴がある。

この特徴は、アンケート情報を構成するすべての回答者と選択肢を一画面に表示したい、という本研究の目的に向いているといえる。また葉ノードの個数に比例した面積で画面を分割し、葉ノードを平等に一画面に表示する手法[5]もある。この手法は平安京ビューに近い特徴を持つといえる。

## 3. 提案内容

本研究では、平安京ビューを用いた新しい試みとして、マトリクス型データに対する階層型データクラスタリングの結果として得られる階層型データの可視化を考える。

$n$  行  $m$  列のマトリクス型データに対して、 $m$  次元ベクトルを持つ  $n$  個のデータ要素と考え、ベクトル間距離の近いものをボトムアップ的にクラスタリングし、結果として数段階の階層を持つ階層型データを形成する。

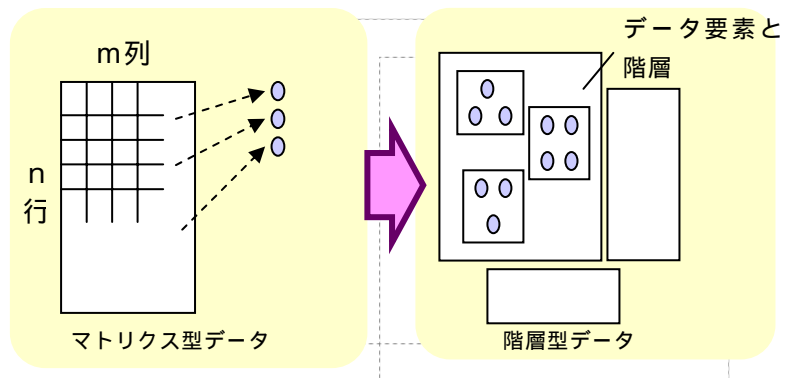


図1 マトリクス型データに対する階層型クラスタリング

$n$  行に対して1つの平安京ビュー、 $m$  列に対しても1つの平安京ビューを出し、ひとつのディスプレイ上に2つの平安京ビューを立ち上げることが本研究「左京と右京」の新しい点である。

本稿にて「左京と右京」が対象とするアンケート集計データは、選択肢、回答者を行と列にするマトリクス型データである。「右京」は回答者をノードとする階層型データを表示する。「左京」は選択肢をノードとする階層型データを表示する。「右京」および「左京」で表示する階層型データの生成にはクラスタリングを適用する。クラスタリングには Cluster3.0[6]を使用する。

提案手法では、注目している特定の回答者、特定の選択肢に対する対話的な詳細表示をする。右京の回答者ノードをクリックすると、左京のその人が回答した選択肢ノードが光って示される。逆に、左京の選択肢ノードをクリックすると、誰が答えているか右京によって一目でわかるようにする。

右京のノードはユーザーが選んだ項目値を形と色で表示。左京のノードはアンケート選択肢に該当する人数を高さで表示。

この可視化方法では、クラスタリングした結果として得られるグループ内の相関性の真偽、強弱などを知るのに有効と考えられる。またアンケート集計結果をクラスタリングした結果の大人数のグループだけでなく少人数のグループについてもその回答の詳細を可視化するのが容易になる。

#### 4. 実行結果

前節で述べた提案手法を用いて、本大学情報科学科の B1～M2 の学生を対象として、自作したアンケート「女子大生の就職に対する意識調査」を配布し、回答を回収した。アンケートの選択肢数は 179、回答者は 135 であった。



図2 「右京による回答者分布の可視化結果」

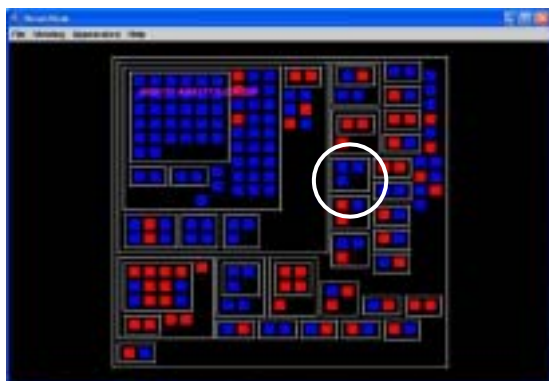


図3 「左京による選択肢分布の可視化結果」

図2は「右京」の結果である。回答者がノードである。色は学年で色分けしている。ひし形のノードは就職活動をしたことがある人。正方形のノードは就職活動したことがない人を表している。人目で色のついた学年のうち何人が就職活動しているかがわかる。

図3は「左京」の結果である。選択肢がノードで、色はある回答者が選んだものだけ別の色になっている。図の部分は一つのクラスタであり、このクラスタは以下の3つの選択肢

- 学年が「B1」という選択肢
- 就職活動に対するイメージで、「知人が増えると思うか?」「得るものが多い?」という2つの質問に対して、「どちらともいえない」という選択肢で構成されている。このクラスタから、B1の学生には、まだ就職活動イメージがしっかりわからない、という傾向がわかる。

#### 5. 応用例

本稿では、「左京と右京」でアンケート集計結果の可視化をした事例を報告した。

今後の課題として、以下のような「左京と右京」の応用例を考えている。応用例の1つ目は「オンラインショッピングの可視化」である。右京に購入者、左京に商品名を表示して、商品購入の動向を可視化できると考えられる。2つ目は、「文献とキーワードの可視化」である。右京にキーワード、左京に論文名を表示して、どのキーワードがどの論文に使用されているかという傾向を可視化できると考えられる。

#### 謝辞

生活科学系アンケートの採取方法に関する情報提供をしてくださった本大学生活科学部の御船美智子教授、永瀬伸子助教授、畑江敬子教授、香西みどり助教授、Cluster3.0を開発された Michael de Hoon 氏、アンケート作成時にアドバイスしていただきました本大学人間社会科学科小園絢子氏、アンケートの回答にご協力していただいた本大学情報科学科の皆様へ感謝の意を表します。

#### 参考文献

- [1]伊藤, 小山田, 平安京ビュー ~ 階層型データを基盤状に配置する視覚化手法, 可視化情報学会第9回ビジュアライゼーションカンファレンス, 2003.
- [2] Carriere J., et al., Research Paper: Interacting with Huge Hierarchies beyond Cone Trees, IEEE Information Visualization 95, pp.74-81 (1995).
- [3]Lamping J., Rao R., The Hyperbolic Browser: A Focus+context Technique for Visualizing Large Hierarchies, Journal of Visual Languages and Computing, Vol. 7, No. 1, pp. 33-55 (1996).
- [4]Johnson B et al., Tree-Maps: A Space Filling Approach to the Visualization of Hierarchical Information Space, IEEE Visualization '91, pp. 275-282 (1991).
- [5]Bederson B., Schneiderman B., Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies, ACM Transactions on Graphics, Vol. 21, No. 4, pp. 833-854 (2002).
- [6]Cluster3.0 ソフト  
<http://bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster/software.htm#ctv>