

InfoScout

An Interactive Information Visualization Tool

Thomas Rieger, Wolfgang Müller

Abstract

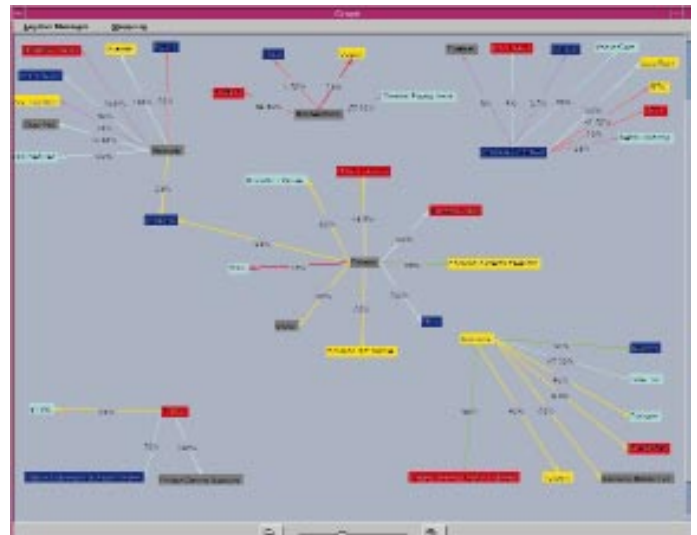
InfoScout is an interactive tool targeted at supporting the user in Data Mining and Knowledge Discovery in Databases (KDD) while making use of Information Visualization techniques.

InfoScout allows the interactive analysis of unknown data and relations. It supports the selection of relevant parts of the data and the visualization of relations using different layout strategies. Moreover, the user can select additional mappings to display additional information with the help of additional visual variables in an easy way. The user can hereby create new visualizations and new sights interactively and sift through the data to reveal discrepancies, to uncover trends, or to provide vital operational insights.

Description

As more and more corporate or institutional data is being computerized, intelligent and flexible methods for the analysis of such data to enhance decision-support processes are needed. Most of the

Figure 2:
Layout with
arrangement
function



potentially useful and valuable information in the data is usually hidden and not available to the user. This is also true for most of the structured data in the growing number of databases used in different application areas: Database Management Systems (DBMS) give access to the data stored but do not provide any means of analyzing the data.

Standard data analysis techniques (e.g., conventional numerical and statistical analysis) are not sufficient to solve the complex problem of »discovering information«. Developed for standard applications in specific scientific or business applications, they usually lack of appropriate methods of dealing with large quantities of data, complex data structures or heterogeneous data. Furthermore, these tools generally do not support the analysis of noisy, incomplete or erroneous data. Information visualization techniques, which are commonly applied to large data structures of scientific data, show how efficiently graphical interactive methods for visual data analysis can be utilized in this context.

InfoScout represents an interactive tool targeted exactly at the application domains »Data Mining« and »Knowledge Discovery in Databases« (KDD). Providing selected information visualization techniques and flexible means of combining them, InfoScout extends

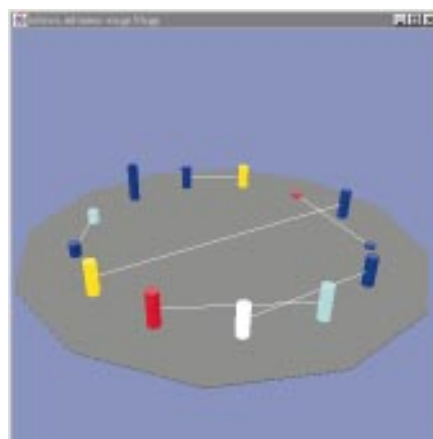


Figure 1: Differences between a 2D and a 3D representation

the repertoire of standard Data Mining tools.

Since Knowledge and Information is usually the composite of not only a large number of dates but especially of the relationships amongst themselves and between the data and a context, InfoScout allows the interactive analysis of unknown data and the relations.

A central element of InfoScout is the interactive data selection panel, which allows for picking relevant parts of the data for the visualization and for interactively determining the objects and relations of observation. For the visualization of relations, different layout strategies are available, such as a 2D hierarchical layout, a 2D circular layout, or a 2D and 3D constrained layout considering further visualization goals, such as limited edge crossings and context-based neighborhood relations.

Additional visualization techniques for qualitative and quantitative data extend the visualization spectrum in InfoScout. In case

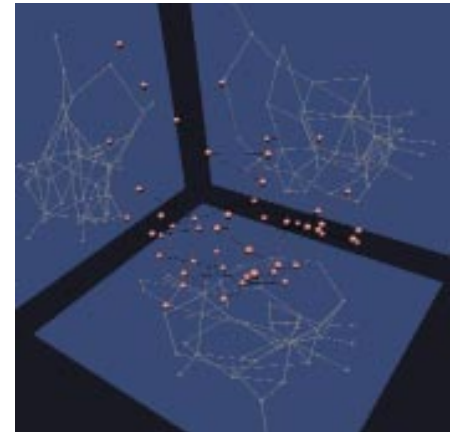
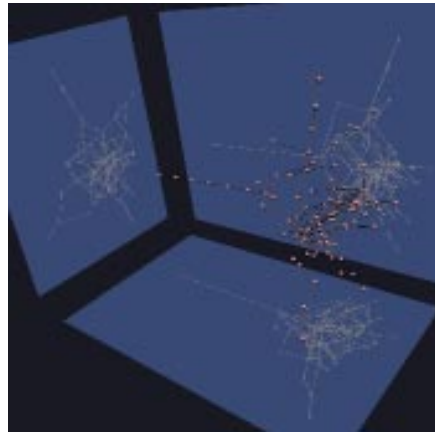


Figure 4: 3D representation with dependencies shown on the proximity to each other

of multi-parameter data, extra information can be displayed by mapping corresponding information on free visual variables such as color, transparency, size. Furthermore, conclusions and hypotheses regarding the structure of the data can be added to the data directly by way of defining new relations. The user can hereby create new visualizations

and new views of the data interactively and sift through the data to reveal discrepancies, to uncover trends, or to provide vital operational insights.

InfoScout is completely implemented in Java and therefore available on a large number of workstation systems.

Point of contact

Thomas Rieger
Technische Universität Darmstadt
Department of Computer Science
Interactive Graphics Systems Group
email: rieger@gris.informatik.tu-darmstadt.de

Dr. Wolfgang Müller
Technische Universität Darmstadt
Department of Computer Science
Interactive Graphics Systems Group
email: mueller@gris.informatik.tu-darmstadt.de

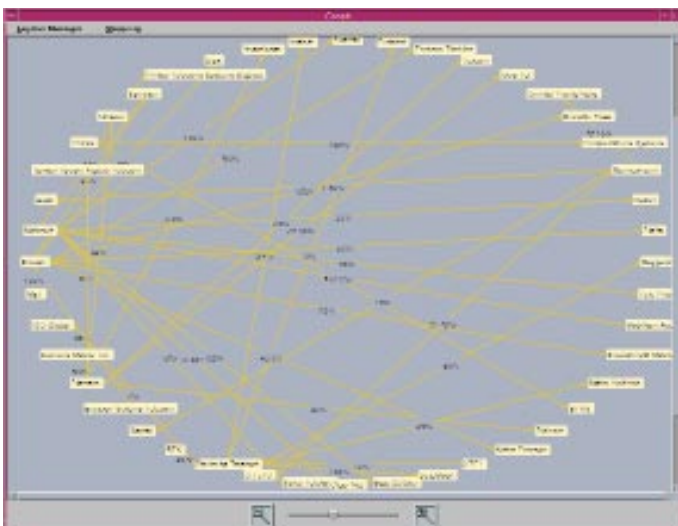


Figure 3: Clear representation of dependencies