EG UK Computer Graphics & Visual Computing (2015)
Rita Borgo, Cagatay Turkay (Editors)

Extending the Scaffold Hunter Visualization Toolkit with Interactive Heatmaps

W. Sturm[1,2], T. Schäfer[3], T. Schreck[1], A. Holzinger[4] and T. Ullrich[1,2]

[1]Institut für ComputerGraphik & Wissensvisualisierung, TU Graz, Austria
[2]Fraunhofer Austria Research GmbH, Visual Computing, Graz, Austria
[3]Dept. of Computer Science, Technische Universität Dortmund, Germany
[4]Institute for Medical Informatics, Statistics and Documentation, Medical University Graz, Austria

Abstract

In many application areas, large amounts of data arise, which are often hard to interpret or make use of by humans. Interactive visualization can help to overview and explore large amounts of data. An example is in the life sciences, where databases of chemical compounds need to be analyzed in terms of similarities of molecular properties. Scientists then need to explore this data in an efficient way.

The Scaffold Hunter framework is an Open Source software system for interactive visualization of high-dimensional data. In this paper, we present an extension of Scaffold Hunter with an interactive heatmap, which ties in tightly with a dendrogram visualization. We added specific interaction modalities and views tailored to the analysis of chemical compounds. Zooming capabilities allow to start from an overview of the data (showing all data elements at once) down to a detail-on-demand view which includes chemical structural views of molecules. We show how the interactive heatmap with clustered rows and columns can bring new insights into the data regarding various properties. The implementation is made available for researchers and practitioners to use.

Categories and Subject Descriptors (according to ACM CCS): H.4.3 [Information System Applications]: Communications Applications—Information browsers J.3 [Computer Applications]: Life and Medical Sciences—Medical information systems

1. Introduction

Information Visualization can provide effective tools for exploration and understanding of large data sets. Increasingly, interactive visualization tools are used in many application areas, such as financial data analysis, engineering, social media analysis, and also, the life sciences [LLC*12]. In drug design, a core problem is the analysis of similarities between chemical compounds (or molecules) based on certain properties (e.g., physical, electrical, functional, etc.) of the molecules. Recent high-throughput laboratory technologies can provide thousands of molecular properties by mass experiments, and it is a challenge to analyze, which molecules correspond to which other molecules along which properties.

A visual representation often used in this scenario is a tabular (or matrix) representation, where rows represent molecules and columns represent individual properties. Sorting algorithms can arrange the rows and columns by similarity, supporting grouping analysis in this data. Using color-coding, the cells in this matrix represent properties (or attributes) of the molecules, and visual structures like blocks, starts or lines denote patterns in the data to be further analyzed.

Such generic analysis tasks require the availability of visual analysis software. As it is not efficient for each laboratory to implement their own software, there is a need for available, versatile software frameworks, which are ready to use or can be tailored towards a specific task. Scaffold Hunter is an Open Source framework for explorative analysis of chemical compound databases. We extend this framework by an interactive heatmap representation, which supports users to examine the data from overviews to details. Our implementation is flexible in that it can incorporate different sorting and display algorithms.

In this paper, we describe the Scaffold Hunter framework, and introduce our heatmap extension. While we recognize

DOI: 10.2312/cgvc.20151247

that a number of Open Source visualization frameworks already exists, Scaffold Hunter is readily available and tailored towards chemical and molecular analysis tasks, offering efficient and effective visualization in this domain. We provide the extension to the public, hoping that it will foster research and application in chemical and molecular analysis scenarios. We also describe a use case that demonstrates how the tool is applied.

2. Related Work

Our work relates to the wider field of interactive data visualization, and specifically, open source frameworks for visualization of data, including biomedical use-cases.

Information visualization [CMS99] aims at finding visual structures for abstract, possibly large data sets. The goal is to support users by gaining an understanding of data by providing overviews, interactive exploration and drill-downs. In general, many visual mappings for many different kinds of data, including time-oriented, geospatial or high-dimensional data, have been proposed [WGK15]. Two specific visualization techniques are matrix heatmaps and dendrograms.

Matrix heatmaps show properties of data in a tabular view, where rows and columns contain data records and properties, respectively. Matrix visualization [WTC08] typically needs to provide a similarity-oriented ordering, such to allow users to see patterns and clusters of similar data. Matrix visualization usually uses color-coding to show normalized data values in the cells of the matrices. Matrix visualizations are scalable down to the pixel level, where each pixel by color-coding may represent one given value [AKK95].

Dendrograms are a well-known visualization of clusters of a hierarchical clustering. An orthogonal node-link diagram shows the similarity relationships between elements in a hierarchical fashion. The similarity relationships can be computed by algorithms, e.g., hierarchical clustering algorithms [HKP11]. Dendrograms and matrix visualizations are often combined, with dendrograms showing the similarity relationships within rows and columns on each side of sorted matrix displays.

There exist many interactive data visualization systems today, of which we can only reference a few ones here. Tableau (http://www.tableau.com) is a commercial visualization system based on a spreadsheet-like user interface, allowing to construct views by intelligent diagram selection and drag and drop operations. D3 [BOH11] is a widely-used JavaScript-based framework for developing interactive data visualization for the Web. Also, some data mining frameworks like SAS JMP (http://www.jmp.com/en_us/home.html) or KNIME [BCD*08] include visualization modules to show the outcome of applied data analysis algorithms.

In previous work, we have proposed interactive visualization and comparison of matrix displays [BSB*10] and sets of dendrograms [BLH*11] for analysis of molecular properties and phylogenetic relationships. Finally, we reference the area of molecule visualization, where the goal is to show spatial and chemical-molecular properties of compounds [HOF04]. To this end, molecule visualization including visual abstraction methods [vdZLBI11] have been proposed before.

3. Extending Scaffold Hunter with dendrogram and Matrix Visualization

In the following, we will describe an implementation of a configurable clustered heatmap for the Open Source project Scaffold Hunter (http://scaffoldhunter.sourceforge.net) [KKM13, KKK*13, WKR*09].

3.1. Scaffold Hunter as Basis for our Implementation

Scaffold Hunter is an Open Source visual analytics tool, which is freely available under GNU GPL v3 and it is implemented in Java. It enables chemists to perform drug discovery by providing visual and interactive methods to explore the chemical space of molecular databases (e.g., PubChem or ChemDB). The current version (Scaffold Hunter 2.4.1) supports the following visualization methods:

- scaffold tree
- table
- dendrogram
- scatter plot (2D and 3D)
- tree map
- heatmap

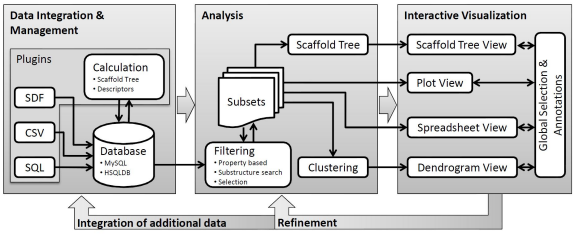

Figure 1: *The workflow of Scaffold Hunter for visual analysis of chemical space. Several data formats can be used to import data into the central database. On the molecules data, further calculations can be performed. An integrated subset and filter management supports users to refine the currently investigated data set. To analyze the data visually, Scaffold Hunter offers a broad set of interactive visualizations to analyze the current data set.* (Image source: Kriege [Kri13])

As depicted in Figure 1, the workflow of the Scaffold Hunter framework can be divided into three main parts,

namely *Data Integration & Management*, *Analysis* and *Interactive Visualization*. These parts are discussed in the following paragraphs.

Data Integration & Management The layer *data integration & management* provides all functions needed to import and to access the data (see Figure 1). The central database (eg., MySQL, HSQLDB) is accessed via Hibernate, which abstracts the database access by an object-relational mapping. Besides importing new data from various file formats (SDF, CSV and SQL), several calculation plugins can be used to compute additional properties. Such calculations might be the computation of molecular fingerprints or a scaffold tree. For several internal data structures and calculations, the library *Chemistry Development Kit* [SHK*03] is used. Moreover, the support of further data formats and calculations can be easily extended by a flexible plugin system.

Analysis The central part manages subsets of the data. Such sets are created by applying filters, performing a substructure search or by using the current selection defined by the user. In Scaffold Hunter, scaffolds and molecules are selectable elements. A special type of sets is the result of a clustering. As the whole data set is often too large, subsets are important to support the user to refine the explorable data space.

Interactive Visualization Scaffold Hunter provides a framework to manage and integrate data views including appropriate controls and integration methods seamlessly to the current system. For large and zoomable user interfaces, the 2D scene graph library *Piccolo2D* [BJM04] is used. The library *Batik* (https://xmlgraphics.apache.org/batik) provides functions to render and store SVG. The framework provides ready-to-use functions for subset management and current selections (see right sidebar in Figure 2). Each data visualization has to implement corresponding callbacks to support such global functions. As selection and annotation management is supported by the framework, selections are synchronized between all active data views. Therefore, Scaffold Hunter provides a clean realization of the technique called *linking and brushing*.

As the sophisticated workflow of Scaffold Hunter is well designed and greatly open for extensions, it was an accommodating opportunity to extend it with an additional clustered heatmap view combined with dendrograms (see Figure 2).

3.2. Scaffold Hunter Extensions

This section explains our implementation which realizes an interactive clustered heatmap for Scaffold Hunter. Subsection 3.1 describes that the architecture of the framework consists of three main parts, namely *data integration & management*, *analysis* and *interactive visualization*. Besides several small code extensions, the whole implementation of the heatmap is located in the latter layer of the framework. Scaffold Hunter provides a generic template to implement a new data view. Every view needs to inherit from the generic view class which manages all essential interfaces and callbacks for data access and global features in order to preserve consistency. Two essential global features are selection management (for linking and brushing) and subset management (to create subsets by using filters or manual selection). *Linking and brushing* is realized by using an observable set. This set contains all currently selected molecules and when the set is being changed, all observers (views) will be notified. When an individual view is notified, it is supposed to highlight all selected elements in its representation space. In case of a heatmap, each column of the map represents one molecule and therefore, the corresponding column will be highlighted.

All corresponding controls of selection management are shown in the right sidebar of every data view (Figure 2). New views do not have to implement or adjust it. In contrast to the right sidebar, the left one is a placeholder for individual user interface elements to control the data view. Therefore, all required controls have to be implemented or ported from other views. The property list in the left sidebar and additional elements, which are needed to configure the heatmap are newly implemented components (controls to configure color coding, clustering and sorting of properties).

We ported the clustering configuration, clustering execution management and the rendering component (canvas) of the dendrogram view to render the dendrogram above the heatmap. In addition to that, we extended the molecule clustering mechanism to cluster general vectors (heat map rows). For that, several generalizations and dependency resolutions were necessary.

Considering that, all visual components placed between the left and right sidebar are parts of our extension. These parts include the actual heatmap, a rendering of a color legend for each row on the left side of the heatmap and a canvas to render property names (including an optional vertical dendrogram tree) on the right side. The dendrogram canvas above the heatmap has been modularized to realize viewport synchronizations between the dendrogram canvas and the heatmap canvas.

3.3. Clustered Heatmap

Our implementation of a clustered heatmap is completely integrated into Scaffold Hunter supporting all its interaction paradigms.

3.3.1. Interactions

Including all user interactions mentioned in Figure 3, the clustered heatmap supports the following user interactions:

- panning
- zooming

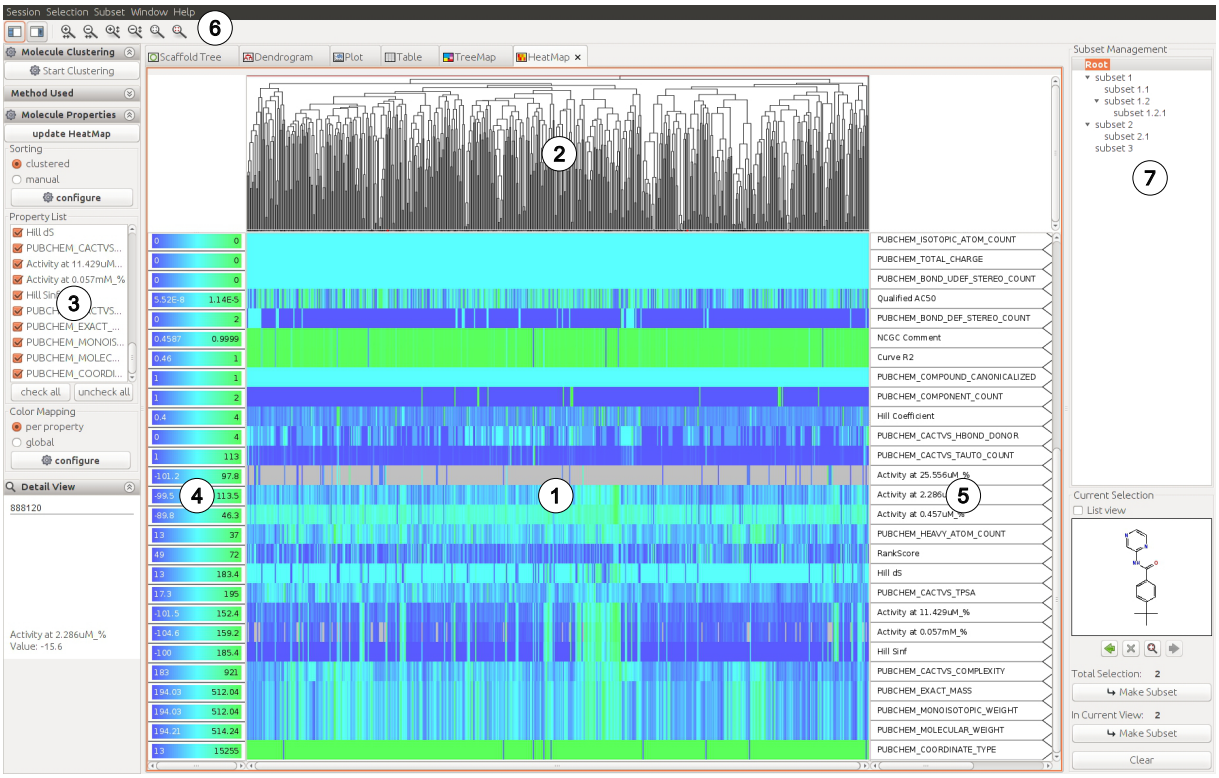

Figure 2: *An overview of the heatmap view and its user interface implemented for the Open Source project Scaffold Hunter: In the center, the actual heatmap visualizes a data array (molecules vs. molecule properties) by using a three-colored color mapping (1). Above the heatmap, a dendrogram is shown to depict the hierarchy of clustered molecules (2). The sidebar on the left provides controls to configure the clustering and the heatmap (3). Between the left sidebar and the actual heatmap, a legend for each row is shown to depict the rows value range (4). To the right of the heatmap, the name of a molecule property is shown in each appropriate row (5). User interactions such as zoom and panning are supported to explore the data set and they can be performed by using keyboard and mouse or by using the toolbar in the upper left corner (6). The selection of a data subset can easily be done using the tree structure (7).*

- resizing
- selection
- zoom-to-fit-selection
- tagging molecules
- detail view and tooltip

Panning Panning is important when the heatmap is larger then the actual display. Horizontal and vertical panning can be performed by dragging the heatmap or using the corresponding scroll bar. Vertical panning can also be carried out by using the mouse wheel on both adjacent legends on the left and right side.

Figure 3: *The heatmap view supports several user interactions. From left to right: hide left and right sidebar, horizontal zoom in and zoom out, vertical zoom in and zoom out, zoom to overview, zoom to fit selection.*

Zoom A horizontal zoom and vertical zoom can be performed by using the mouse wheel above the heatmap or dendrogram. To improve usability, the center of the zoom is defined by the current position of the mouse pointer. Moreover, Figure 4 shows that the heatmap uses *semantic zoom* to adjust the level of detail according to the zoom level.

Resizing Besides toggling sidebars, it is possible to resize the heatmap canvas by moving the borders between heatmap and dendrogram or legends. This enables the user to distribute the available space to individual segments of the heatmap view. While resizing, an internal zoom level is being updated to synchronize the size of the heatmap (when completely zoomed-out) and the size of the canvas. When

creating a new view, the proportions are automatically adjusted to utilize the available space as much as possible.

Selection In the heatmap view, a molecule can be selected by clicking on a column or on a dendrogram node. The dendrogram also enables users to select a whole subtree of the dendrogram. As shown in Figure 5, selected molecules are highlighted as red dendrogram leaves and corresponding columns have additional red borders. Each selected column is additionally highlighted with a red overlay.

Zoom-to-Fit-Selection In Scaffold Hunter, all data views are supposed to offer controls to visually navigate through the current selection. The feature *zoom-to-fit-selection* supports the user to adjust the view-port in such a way that all currently selected elements are visible at once. For that, the heatmap pans and zooms to the appropriate level to fit all selected columns. Furthermore, the selection management offers controls to switch visually to the next selected item. When switching to the next one, the view automatically pans the heatmap matrix to show the selected column (molecule) in the center of the view.

Tagging Molecules The framework of Scaffold Hunter provides a feature to tag molecules to support the user to recognize tagged molecules easily within other visualizations (linking And brushing). As shown in Figure 6, a molecule can be tagged via the dendrogram which is located above the heatmap.

Detail View and Tooltip A detail view of the molecule, at which the cursor is currently pointing at, is shown in the left sidebar. It shows the value of the hovered property and the molecules structure as a SVG. When pointing at a column of the heatmap or a dendrograms node for 3 seconds, a tooltip window will be shown. This global feature of Scaffold Hunter shows all properties of the current molecule. In addition to that, it is possible to attach comments.

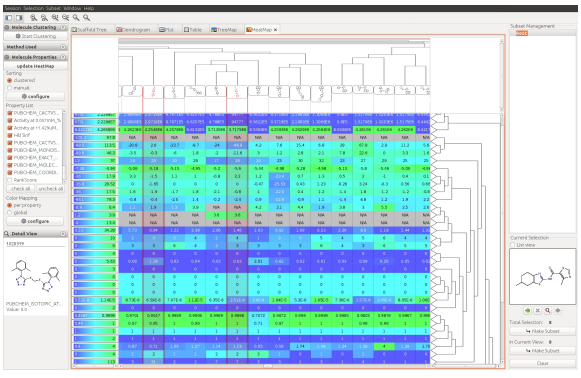

Figure 4: *Semantic zoom: When zoomed in, values are rendered in each cell to show additional information.*

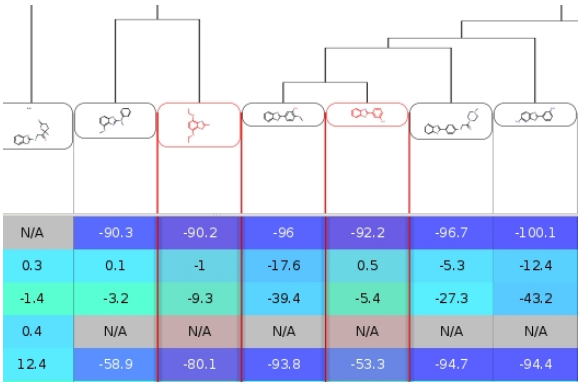

Figure 5: *Selections are highlighted as red dendrogram leaves and columns with red borders and a transparent red overlay.* (Image source: Scaffold Hunter manual, http://scaffoldhunter.sourceforge.net/)

3.3.2. Configuration

The clusted heatmap can be configured in various ways:

Clustering Before a heatmap can be rendered, a clustering of all molecules within the selected subset has to be performed. For that, a clustering configuration is shown. The Scaffold Hunter framework provides several parameters to configure a clustering. Basic parameters are the type of clustering, (normal *exact clustering* or *heuristic clustering*), the linkage criteria (*complete linkage*, *group average linkage*, *single linkage*) and the distance function (*Euclidean*, *Tanimoto*, *Jaccard*). Besides these parameters, a subset of properties, used for distance calculation, can be defined.

In addition to that, it is possible to configure the ordering of rows manually while columns have to be clustered any-

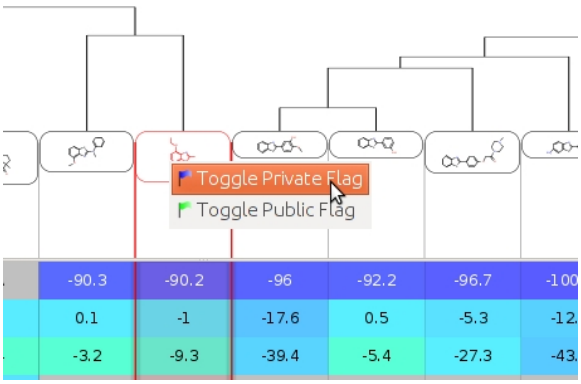

Figure 6: *Scaffold Hunter supports users to tag single molecules with a flag to find it in other data views.* (Image source: Scaffold Hunter manual, http://scaffoldhunter.sourceforge.net/)

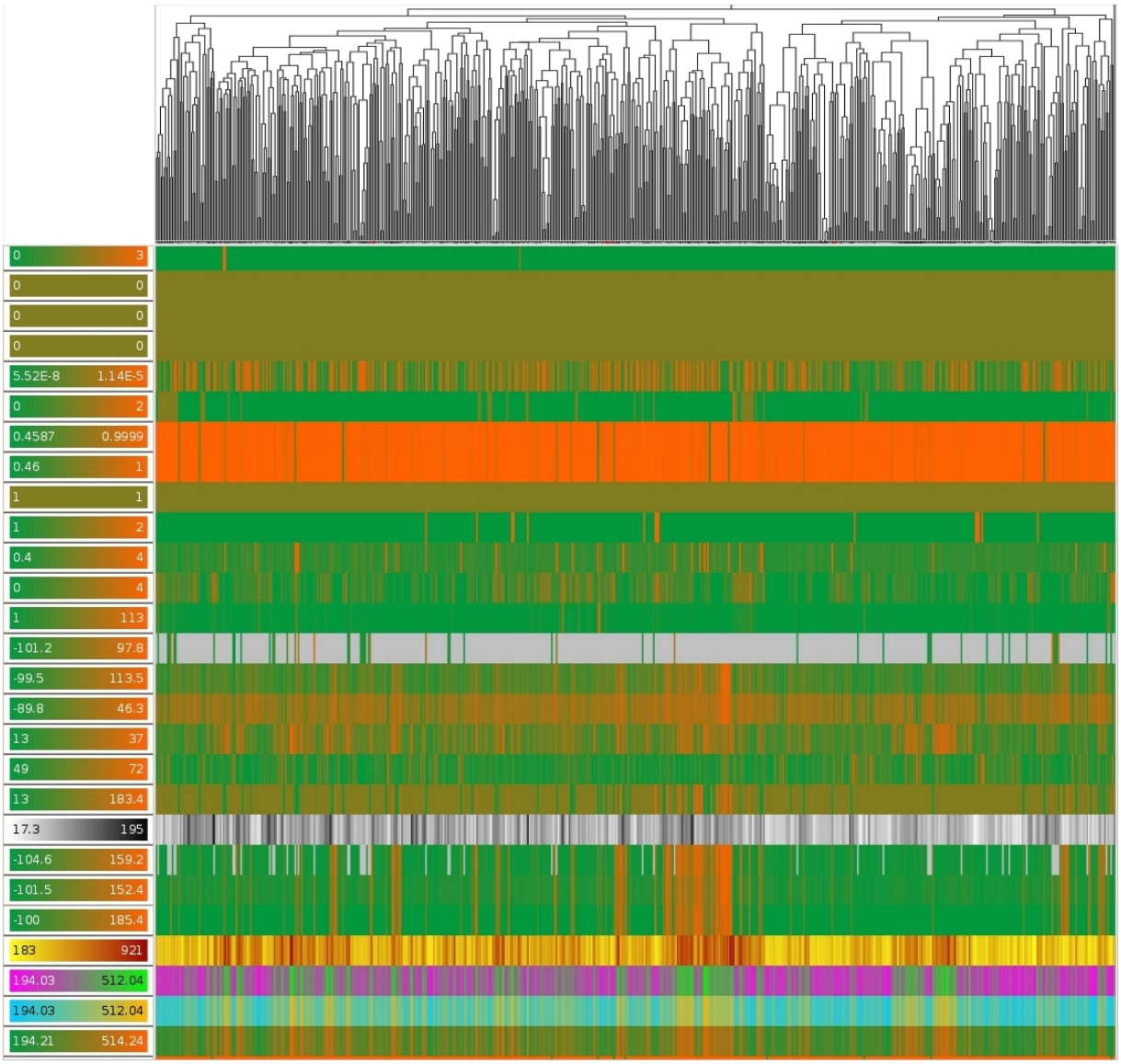

Figure 7: *Individual color mapping functions are used to highlight specific properties in the heatmap.*

way. All properties are shown as list in the left side bar. It is possible to exclude individual properties to be shown as a row in the rendered heatmap by unchecking it in the list. For manual ordering, items in the list can be moved and reordered by drag-and-drop.

Color Mapping The heatmap view provides several possibilities to configure the color mapping function for each row individually (per property) or as a global function applied to each row. As shown in Figure 7, the assignment of an individual color mapping function to each row enables the user to configure individual color ranges to highlight specific properties. The heatmap supports three different mapping functions: a two-colored gradient mapping, a three-colored gradient mapping and an interval mapping. An interval mapping supports the user to highlight individual intervals of values.

4. Conceptual and Practical Application of Scaffold Hunter

Our implementation of Scaffold Hunter provides useful facilities for interactive visual analysis in compound data. We next describe its benefits in conceptual and practical terms.

4.1. Relevance of Visual-Interactive Data Analysis

This implementation can be used to continue experiments in interactive machine learning approaches particularly for applications in personalized medicine [Hol14]. However, biomedical data sets are full of probability, uncertainty, incompleteness, vagueness, noise, etc., which makes the application of automated machine learning approaches often impossible [HDJ14]. Moreover, the complexity of current machine learning algorithms have demotivated non-experts from the application of them, but for machine learning and data mining in the biomedical domain to be effective, it is important to include a human domain expert into the data exploration process, and combine the flexibility, creativity, and general knowledge of the human with the enormous computational capacity and analytical power of algorithms and systems. One such approach is Visual Data Mining (VDM) [OPHJ14], which integrates the human into the data exploration process and aims to effectively represent data visually to benefit from sophisticated human perceptual abilities, allowing the expert to get insight into the data by direct interaction with the data. Such an approach can be particularly helpful in the biomedicine, where we have little previous knowledge, have to deal with complex data sets, or when the exploration goals are unclear or may evolve gradually over time [TJHH14]. Such processes are hypothesis generation processes and pose a lot of challenges for further research: The visualizations of the data enable the user to gain novel insight into the data, and generate new hypotheses to support data mining and interpretation [WXH11].

4.2. Application in Practice and Extension Possibilities

Scaffold Hunter is developed in collaboration with Chemists and Molecular Biologists from the pharmaceutical industry. As part of their routine work, these experts aim to find groups in chemical compounds and relate them to chemical and pharmaceutical properties, i.e., the effect the compounds gave on certain processes. The ultimate goal is to advance drug development and understand better the reactions. As discussed in "Scaffold hunter: visual analysis of biological activity data" [KKK*14], a typical workflow starts by loading all compound data (rows) and apply hierarchical clustering on the rows based on the attributes (i.e., the columns). The hierarchical clustering is mapped to a 1-dimensional sort order. Then, the expert by visual inspection of the dendrogram identifies groups of compounds which are similar to each other and have some interesting behavior for certain attributes of interest (e.g., high reaction indicators). Then, for each relevant cluster, a partitioning is attempted to identify alternative compounds which exhibit a desired reaction, but offer alternative compound structure as relevant for pharmaceutical production. While the originally described workflow [KKK*14] was based on the dendrogram and compound structure views, Scaffold Hunter as the result of our implementation effort includes a heatmap representation, which is expected to enhances the subgrouping and correlation search substantially for our collaborators.

An open problem that needs further research attention is a specific evaluation of the effect of the heatmap inclusion as compared to using only dendrograms alone. Depending on the data normalization and color scales chosen, expert findings in the clusters and relationships may vary. While we are interested in providing the best possible color-mappings, more research is needed here. Just as one effect, there may be differences between different users relating to contrast or color perception, which may eventually affect the distinction of groups. To that end, it would be desirable to include recent approaches for perceptional compensation [MSK14] into the color-mapping process.

Other extension possibilities include improving the dimension selection process. Our data usually includes many dimensions, and currently all user-selected dimensions are used to create the hierarchical clustering. However, this may lead to unstable results due to the potential inclusion of noisy or irrelevant dimensions. Methods from feature selection [LM07] or subspace search [BPR*04] could provide means to arrive at smaller dimension sets and improve the overall exploration process. Also, we may aim at interactively supporting the feature selection process by glyph-based approaches as proposed recently by KRAUSE ET AL. [KPB14].

5. Conclusions

We motivated the problem of visual interactive data analysis in biomedical applications. A key problem here is to group data for similarities and explore for relationships (e.g., similarities and differences) within and between groups of compounds and their attributes. Visualization can help analysts to efficiently search for such relationships, making use of interactive navigation and activation of background knowledge. We extended the Scaffold Hunter Open Source visualization system with a sortable heatmap and dendrogram as basic and powerful facilities to search for relationships in compound data. Semantic zoom interactions allow a user to go from the overview of all compound data down to individual compound structures. Our implementation is made available and represents a step forward to provide visual-analysis capabilities ready for use in biomedical research. Tools like Scaffold Hunter contribute to the tool landscape, eventually making it affordable to test visual-interactive data analysis in many domains. Our contribution is an extension of Scaffold Hunter; i.e. we integrated our implementation into an existing, well-known framework and did not present "yet-another-solution", which merely integrates into an existing working pipeline.

References

[AKK95] ANKERST M., KEIM D., KRIEGEL H.-P.: Recursive Pattern: A Technique for Visualizing Very Large Amounts of Data. *Proceedings of Visualization 1* (1995), 279–286. 2

[BCD*08] BERTHOLD M. R., CEBRON N., DILL F., GABRIEL T. R., KÖTTER T., MEINL T., OHL P., SIEB C., THIEL K., WISWEDEL B.: *KNIME: The Konstanz Information Miner*. Springer, 2008. 2

[BJM04] BEDERSON B. B., JESSE G., MEYER J.: Toolkit Design for Interactive Structured Graphics. *IEEE Transactions on Software Engineering 30* (2004), 535–546. 3

[BLH*11] BREMM S., LANDESBERGER T. V., HESS M., SCHRECK T., WEIL P., HAMACHER K.: Interactive visual comparison of multiple trees. *IEEE Conference on Visual Analytics Science and Technology 9* (2011), 31–40. 2

[BOH11] BOSTOCK M., OGIEVETSKY V., HEER J.: D^3 Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics 17* (2011), 2301–2309. 2

[BPR*04] BAUMGARTNER C., PLANT C., RAILING K., KRIEGEL H.-P., KROGER P.: Subspace selection for clustering high-dimensional data. *IEEE International Conference on Data Mining 4* (2004), 11–18. 7

[BSB*10] BREMM S., SCHRECK T., BOBA P., HELD S., HAMACHER K.: Computing and Visually Analyzing Mutual Information in Molecular Co-evolution. *BMC Bioinformatics 11* (2010), 330ff. 2

[CMS99] CARD S. K., MACKINLAY J. D., SHNEIDERMAN B.: *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999. 2

[HDJ14] HOLZINGER A., DEHMER M., JURISICA I.: Knowledge Discovery and interactive Data Mining in Bioinformatics - State-of-the-Art, future challenges and research directions. *BMC Bioinformatics 15*, S6 (2014), I1. 7

[HKP11] HAN J., KAMBER M., PEI J.: *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers Inc., 2011. 2

[HOF04] HALM A., OFFEN L., FELLNER D.: Visualization of Complex Molecular Ribbon Structures at Interactive Rates. *International Conference on Information Visualisation 8* (2004), 737–744. 2

[Hol14] HOLZINGER A.: Trends in Interactive Knowledge Discovery for Personalized Medicine: Cognitive Science meets Machine Learning. *IEEE Intelligent Informatics Bulletin 15*, 1 (2014), 6–14. 7

[KKK*13] KLEIN K., KOCH O., KRIEGE N., MUTZEL P., SCHÄFER T.: Visual analysis of biological activity data with scaffold hunter. *Molecular Informatics 32* (2013), 964–975. 2

[KKK*14] KLEIN K., KOCH O., KRIEGE N., MUTZEL P., SCHÄFER T.: Scaffold hunter: visual analysis of biological activity data. *Journal of Cheminformatics 6* (2014), 33. 7

[KKM13] KLEIN K., KRIEGE N., MUTZEL P.: Scaffold hunter: facilitating drug discovery by visual analysis of chemical space. *Computer Vision, Imaging and Computer Graphics. Theory and Application 1* (2013), 176–192. 2

[KPB14] KRAUSE J., PERER A., BERTINI E.: INFUSE: Interactive Feature Selection for Predictive Modeling of High Dimensional Data. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1614–1623. 7

[Kri13] KRIEGE N.: Visual Analysis of Chemical Space with Scaffold Hunter. http://www.opentox.org/data/documents/development/meeting/opentoxeuro2013_submission_8.pdf, 2013. Accessed: 2015-04-15. 2

[LLC*12] LIPSA D. R., LARAMEE R. S., COX S. J., ROBERTS J. C., WALKER R., BORKIN M. A., PFISTER H.: Visualization for the Physical Sciences. *Computer Graphics Forum 31* (2012), 2317–2347. 1

[LM07] LIU H., MOTODA H.: *Computational Methods of Feature Selection (Chapman & Hall/Crc Data Mining and Knowledge Discovery Series)*. Chapman & Hall/CRC, 2007. 7

[MSK14] MITTELSTÄDT S., STOFFEL A., KEIM D. A.: Methods for Compensating Contrast Effects in Information Visualization. *Computer Graphics Forum 33* (2014), 231–240. 7

[OPHJ14] OTASEK D., PASTRELLO C., HOLZINGER A., JURISICA I.: Visual Data Mining: Effective Exploration of the Biological Universe. *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics: State-of-the-Art and Future Challenges. Lecture Notes in Computer Science 8401* (2014), 19ff. 7

[SHK*03] STEINBECK C., HAN Y., KUHN S., HORLACHER O., LUTTMANN E., WILLIGHAGEN E.: The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics. *Journal of Chemical Information and Computer Sciences 43* (2003), 493–500. 3

[TJHH14] TURKAY C., JEANQUARTIER F., HOLZINGER A., HAUSER H.: On Computationally-Enhanced Visual Analysis of Heterogeneous Data and Its Application in Biomedical Informatics. *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics - State-of-the-Art and Future Challenges 8401* (2014), 117–140. 7

[vdZLBI11] VAN DER ZWAN M., LUEKS W., BEKKER H., ISENBERG T.: Illustrative Molecular Visualization with Continuous Abstraction. *Computer Graphics Forum 30* (2011), 683–690. 2

[WGK15] WARD M., GRINSTEIN G., KEIM D. A.: *Interactive Data Visualization: Foundations, Techniques, and Application*. Apple Academic Press Inc., 2015. 2

[WKR*09] WETZEL S., KLEIN K., RENNER S., RAUH D., OPREA T. I., MUTZEL P., WALDMANN H.: Interactive exploration of chemical space with Scaffold Hunter. *Nature Chemical Biology 5* (2009), 581–583. 2

[WTC08] WU H.-M., TZENG S., CHEN C.-H.: *Handbook of Data Visualization*. Springer, 2008. 2

[WXH11] WONG B. L. W., XU K., HOLZINGER A.: Interactive Visualization for Information Analysis in Medical Diagnosis. *Information Quality in e-Health, Lecture Notes in Computer Science 7058* (2011), 109–120. 7