technische universität
dortmund

A graph-drawing perspective to some open problems in molecular biology

Mario Albrecht
Andreas Kerren
Karsten Klein
Oliver Kohlbacher
Petra Mutzel
Wolfgang Paul
Falk Schreiber
Michael Wybrow

Algorithm Engineering Report
TR08-01-003
June 2008

Fakultät für Informatik
Algorithm Engineering (LS 11)
44221 Dortmund / Germany
http://ls11-www.cs.uni-dortmund.de/

A graph-drawing perspective to some open problems in molecular biology

Mario Albrecht[1], Andreas Kerren[2], Karsten Klein[3], Oliver Kohlbacher[4], Petra Mutzel[3], Wolfgang Paul[3⋆], Falk Schreiber[5], and Michael Wybrow[6]

[1] Max Planck Institute for Informatics, Saarbrücken, Germany
mario.albrecht@mpi-inf.mpg.de
[2] School of Mathematics and Systems Engineering (MSI), Växjö University, Sweden
andreas.kerren@vxu.se
[3] Faculty of Computer Science, Dortmund University of Technology, Germany
{karsten.klein, petra.mutzel, and wolfgang.paul}@cs.uni-dortmund.de
[4] Center for Bioinformatics, Eberhard Karls University Tübingen, Germany
oliver.kohlbacher@uni-tuebingen.de
[5] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Germany
schreibe@ipk-gatersleben.de
[6] Clayton School of Information Technology, Monash University, Australia
michael.wybrow@infotech.monash.edu.au

Abstract. Much of the biological data generated and analyzed in the life sciences can be interpreted and represented by graphs. Many general and special-purpose tools and libraries are available for laying out and drawing graphs, but they are either not adequate for handling large graphs or do not adhere to the special drawing conventions and recognized layouts of biological networks. In this paper, we describe some representative use cases that demonstrate the need for advanced algorithms for presenting, exploring, evaluating, and comparing biological network data.

1 Introduction

In recent years, the development of high-throughput experimental techniques has led to the generation of huge data sets in the life sciences. Since manual analysis of this data is costly and time-consuming, biologists are now turning towards computational methods that support data analysis. The information in many experimental data sets can be either represented as networks or interpreted in the context of networks that serve as models of the biological system under investigation. These models are used, for example, to predict the behavior of the system and to guide further experiments.

The visualization of biological networks is one of the key analysis techniques to cope with the enormous amount of data. In particular, the layout of networks should be in agreement with biological drawing conventions and draw attention to relevant system properties that might remain hidden otherwise. While the approaches and expertise of the graph-drawing community may be ideally suited for solving these problems, very little research has previously been done to solve the special layout and visualization problems arising in this area. So far, most of the available software systems for the visual analysis of biological networks (e.g., Cytoscape [4], VisAnt [7], etc.) only provide implementations of standard graph drawing algorithms like force-directed or hierarchical approaches. Nevertheless, there also some

⋆ Corresponding author

tools that offer specialist and drawing algorithms which are more suitable for application in the life sciences [8, 13].

In general, graph-drawing methods for applications in the life sciences should allow for the layout and navigation of biological networks for both their static presentation as well as their interactive exploration. Such drawing methods need to adhere to constraints that originate from recognized textbook layouts and from generally accepted drawing conventions within the life-science community.

In this paper, we aim to make graph-drawing problems originating in bioinformatics accessible to the graph drawing community. We start by presenting a characterization of common biological networks, describing their structure and semantics (Section 2.1) as well as discussing the mapping of data onto network elements (Section 2.2). In Section 3, we present a selection of use cases describing typical uses of biological networks. Section 4 will present some conclusions.

2 The nature of biological network data

Biological networks are used to communicate many different types of data. This data can be encoded in the structure of the network, the network layout, or as graphical or textual annotations. The data itself may be primary data (i.e., directly measured), secondary data (i.e., derived, inferred, or predicted), or a mixture of both. In this section, we discuss some common biological networks and the types of attributes used to annotate them.

2.1 Types of biological networks

Gene-regulatory and signal-transduction networks both use sets of directed edges to convey a flow of information. While gene regulation (regulation of the expression of genes) occurs exclusively within a cell and represents a regulatory mechanism for the creation of gene products (RNA or proteins), signal transduction refers to any process that transports external or internal stimuli via so-called signal cascades to specific cellular parts where a cell response (e.g., gene regulation) is triggered. While nodes in these networks represent molecular entities (genes, gene products, or other molecules), edges represent a flow of information (regulation or passing of a chemically encoded signal). Figure 1 gives an example for a graph representing a part of a gene regulatory network.

Protein interaction networks represent physical interactions of proteins with each other or with other binding partners like DNA/RNA. The nodes in such networks represent proteins or sets of proteins. The time scale of protein interactions ranges from very short, transient processes (for instance, pairwise protein interactions and phosphorylation or glycosylation events) to very long lasting, permanent formation of protein assemblies (protein complexes) working as molecular machines. The interaction edges are normally undirected, but may be directed in case of heterogeneous networks (for example, protein-protein and protein-DNA/RNA interactions), resulting in mixed graphs. Each node and edge may be annotated with additional biological attributes like expression level, cellular localization, and the number of interaction partners. For an example of a protein interaction network, see Figure 2.

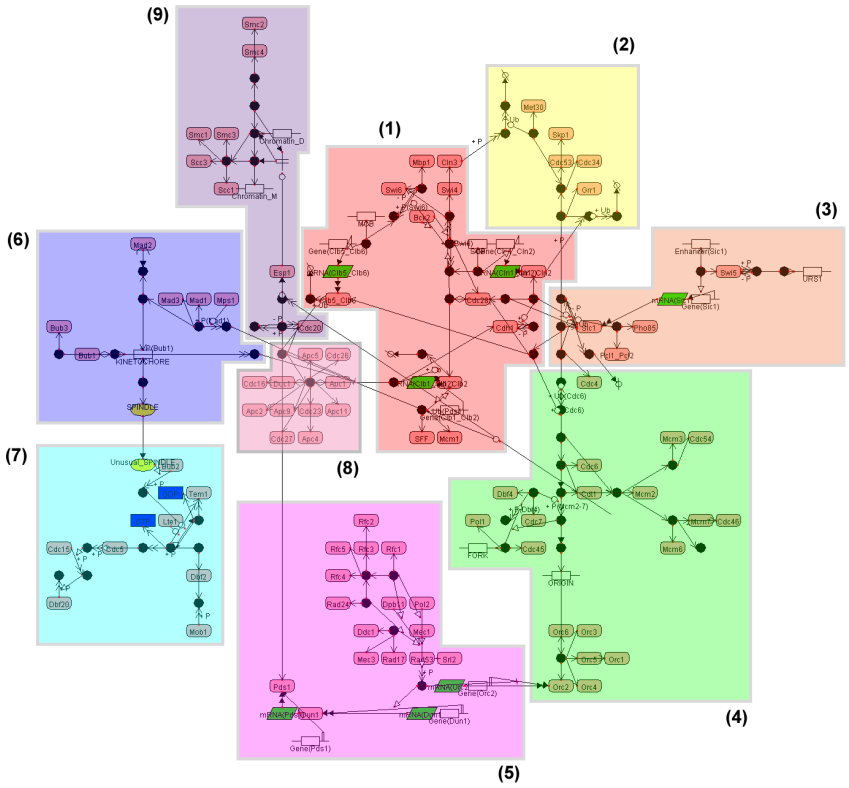

Fig. 1. A regulatory network describing yeast's cell cycle. The picture is taken from the CADLIVE homepage (`http://www.cadlive.jp/`) and was originally published in [14]. Li and Kurata used their implementation of a grid-layout algorithm.

Metabolic networks describe how metabolites (chemical compounds) are converted into other metabolites. Such a network is a hypergraph that is usually represented as a bipartite graph $G = (V_1 \cup V_2, E)$. The node set is partitioned into the set V_1 of metabolite and enzyme nodes (enzymes catalyze the chemical reactions converting metabolites) and V_2 the set of reaction nodes. There are large posters (e.g., Nicholson's [17] and Michal's [15] pathway maps) and several projects that have created graphical representations of metabolic networks and offer access to these graphs via web pages (e.g., Kyoto Encyclopedia of Genes and Genomes (KEGG) [18] or the BioCyc collection [10]). The availability of these representations has established a de facto standard for metabolic network drawings that features near-orthogonal drawings where important paths are aligned, relevant subgraphs are placed close to the center of the drawing, substances and products of a reaction are clearly separated, and co-substances are placed out of the main path close to the reaction. There are also layout algorithms that obey established drawing styles of these networks (e.g., [11, 19, 20]). Figure 3 shows an example of a metabolic network.

Ontologies have become widely used in the life-science community. Collaborative efforts such as the Gene Ontology (GO) [1] or the KEGG Ontology (KO) [9] have developed

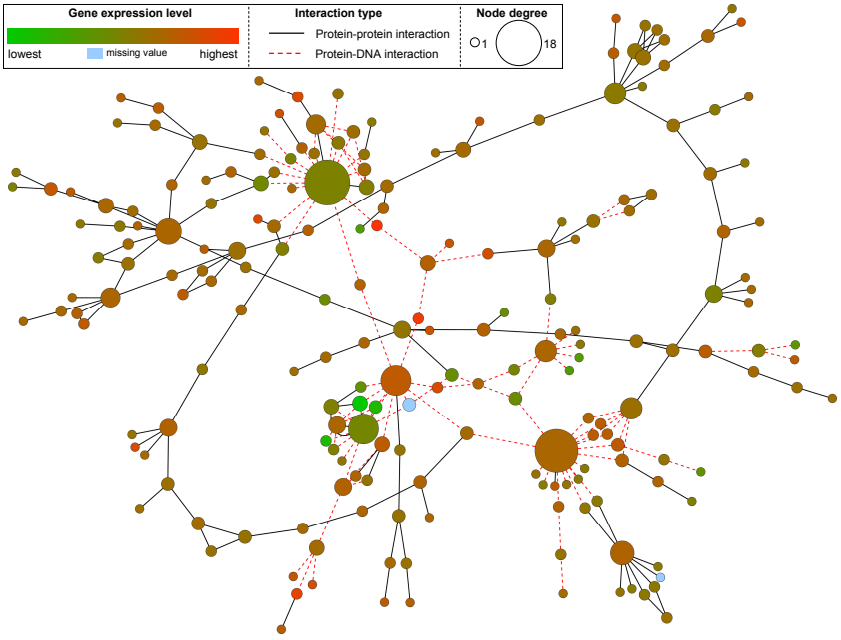

Fig. 2. Part of a human protein interaction network. The protein nodes are given a shade gradient according to their expression value; light grey represents the lowest, dark grey the highest value. The node size corresponds to the number of interactions. The shades and styles of the edges represent different interaction types; solid lines indicate protein-protein, and dashed lines protein-DNA interactions. The graph was drawn with Cytoscape [4] using its implementation of the spring-embedder algorithm.

controlled vocabularies to describe biological processes on different levels (e.g., participation of a gene product in signaling networks) and help biologists to identify available knowledge concerning a specific area of research. The network structure of such ontologies is best described through directed acyclic graphs in which nodes represent specific terms from the predefined controlled vocabularies and relationships between different terms (e.g., "is part of" or "is a type of") are represented by edges. Gene Ontology categories have also been used to convey functional aspects in gene regulatory networks [21].

Phylogenetic networks originally started out as rooted phylogenetic trees, but have recently received a great deal of attention [12]. They represent models for the evolutionary development of and the relationships between existing and extinct species. The structure of these networks is best described by graphs in which nodes represent different species and hereditary relationships between two species are represented by edges. The evolutionary time axis (and hence the direction of the edges) is usually implied through the layout of the graphs.

2.2 The attributes of network elements

The representation of primary and secondary data that has been mapped onto the elements of a molecular network is an important research field. This is mainly due to the fact that

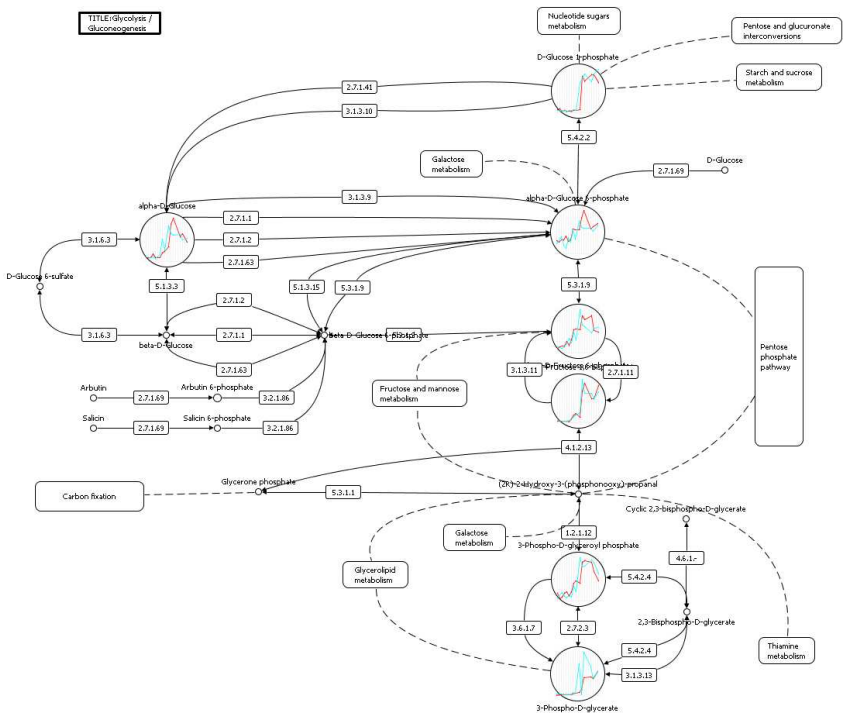

Fig. 3. Part of the glycolysis/gluconeogenesis pathway with additional data mapped onto some nodes. Circles encode metabolites, rectangles represent enzymes catalyzing the reaction, and rectangles with rounded corners denote other pathways. Solid and dashed lines represent reactions and connections to other pathways, respectively. The pathway data was derived from KEGG [18], and the graph was drawn with VANTED [8] in a style similar to the KEGG pathway picture.

primary and secondary data are quite complex in nature. The various types of primary data are defined by the various types of biochemical entities and experiments (e.g., time-series experiments, differential studies, etc.) and entities that are the subjects of analysis, namely, gene sequences, transcripts, expression levels, proteins, protein concentrations, metabolites, metabolite concentrations, or fluxes (of mass or information). The structure of the class of secondary data is, however, even more complex, and different categories of inferred, derived, or predicted information can be distinguished such as results of correlation analysis or comparisons of different biological states (e.g., healthy vs. diseased, before vs. after treatment with a drug, different organisms). The data belongs to different data types, namely:

- nominal data: sequence names, categories, etc.
- ordinal data: ontologies, rankings, partly ordered information, etc.
- scalar data: comparisons, ratios, etc.
- categorized spatial data: data points that refer to biological entities from various parts or substructures of a cell, etc.

The list of secondary data is not necessarily complete and the distinction between the categories may not always be clear. It should also be mentioned that primary as well as

secondary data are subject to uncertainties (measurement errors, prediction confidences, etc.). This is a general problem that has to be taken into account when trying to draw biological networks, and often the visualization of this uncertainty is also desired.

3 Use cases and related graph-drawing problems

This section contains a list of typical use cases that arise in constructing and viewing networks representing life-science data. Along with each use case we present a formalized description of the graph-drawing, information-visualization, or visual-analytics problem behind the use case. The list should not be interpreted as a complete collection of all use cases relevant to the graph drawing community. Rather, it should stimulate research and demonstrate that there are many interesting and important graph drawing problems within the life sciences.

3.1 Visual analysis of data correlation

Biologists frequently use correlation graphs as a means for visually expressing and exploring complex forms of correlation within their data. Normally, the information contained in the data is mapped as annotation onto a graph that represents the pathways characterized by experimental data. Biologists are then interested in a graphical representation that highlights the interrelation between the connectivity structure of pairs/subsets of nodes in the original network and their correlation. An example for an interesting correlation pattern would be a set of nodes that is closely connected within the underlying graph but exhibits only weak correlation in the data or vice versa. The represented connectivity structure should include only statistically significant correlations, for instance, significant up- or down-regulation of co-expressed genes or proteins. In particular, two or more nodes representing biological entities with multiple annotations may be considered correlated if a minimum number of node annotations corresponds with each other, for example, regarding genotype, time value, number of the biological replicates, etc.

One possible way to attack this problem would be to model the correlation data as a weighted graph. Then we have a given graph $G_1 = (V, E_1)$ (called *network* in order to distinguish it from the correlation graph) with correlation data that induces a second graph $G_2 = (V, E_2)$ with edge weights on the same set of vertices V. This possibly dense correlation graph may, for example, contain two types of edges representing positive and negative correlation. This gives us a simultaneous embedding problem [6] in which the two given graphs typically do not have too many edges in common. We search either a layout of the union graph $G = (V, E_1 \cup E_2)$, in which the given network G_1 and the correlations are clearly displayed, or two disjoint layouts, in which the coordinates of the vertices in both layouts are identical. In the first case, a challenging task is to provide a layout which clearly emphasizes the two different edge sets E_1 and E_2. Often, a layout π_1 of the graph G_1 is given, which has to be preserved as closely as possible. In this case, there is a trade-off between mental map preservation and emphasizing the correlation structure. Possible solutions may either fix the layout given in π_1, or try to preserve the mental map by keeping the orthogonal relations, the topological embedding, or the layout of a backbone.

3.2 Visual comparison of biological networks

Conservation of biochemical function during evolution results in structurally similar molecular subnetworks across different organisms and species. Uncovering relevant similarities and

differences or comparing networks in different states (e.g., diseased vs. healthy), at different time points, or under various environmental conditions (temperature, pressure, substrate concentrations, etc.) supports the biologists' knowledge-discovery process, for example, by identifying disease-specific patterns (biomarker discovery).

Given a set of graphs $G_1, \ldots, G_k$ with a high degree of similarity between each other, the task is to layout them in a way so that the differences (or the similarities) are highlighted. This problem can be attacked via simultaneous embedding, which requires to obtain either one layout of the union graph $G = G_1 \cup \ldots \cup G_k$ or k disjoint layouts of the graphs G_i $(i = 1, \ldots, k)$ such that the coordinates of all vertices common to two or more subgraphs are the same. An alternative presentation has been given in [3] where the third dimension has been used to stack the k layouts above each other. In the layouts, crossings between edges belonging to different graphs $G_i \neq G_j$ are either completely ignored or counted as less important than "real" crossings. In the biological context, the stronger simultaneous embedding problem with fixed edges occurs, which forces not only the vertices but also the edges occurring in two or more graphs to be drawn identically. This guarantees that identical subnetworks have an identical layout. Sometimes, it may be important to keep a mental map of already given layouts of some of the graphs or their backbone structure. In any case, the layouts must obey the given biological constraints concerning the specific network type. Sometimes, the networks may be large, and then it would be desirable to hide some parts of the network and to highlight only the specific points of interest. Points of interest may be differences between the networks, but could also be important network structures such as the main pathways in a metabolic network. Here, one possibility would be to generate layouts in which the differences are all concentrated within only a few layout areas.

3.3 Integrated representation of multiple overlapping networks

The different types of biological networks describe different functional aspects of the whole cell, tissue, or organism in question. To get a deeper, system-wide understanding, these networks need to be combined. The enzymes acting in metabolic networks for example are regulated and this regulation is described by a regulatory network. It is thus becoming increasingly common to integrate these different types of networks into joint networks. Figure 4 shows an example of integrating a gene-regulatory network and a metabolic network (see [23]). A good joint layout of these networks should reveal the *interaction* between these networks, for example, how specific nodes of the gene regulatory network activate or inactivate whole subnetworks of the metabolic network. In order to simplify the identification of these subnetworks, mental map preservation on the level of the metabolic network is helpful.

We need a representation of combined networks in which the conventional layouts (there may be several different ones) of each of these networks need to be respected. Moreover, some groups of vertices in one network may belong to groups of vertices in another network. This mapping (which may be a $1:1$, $1:n$, or $n:m$ mapping) needs to be displayed in the layout. We consider the case of integrating two networks, in which the involved mapping partners can be viewed as a cluster in a cluster graph $C = (G, T)$ of cluster depth 2. Then the problem may be attacked via the following formalized graph drawing problem. We are given two cluster graphs $C_1 = (G_1, T_1)$ and $C_2 = (G_2, T_2)$ with $G_i = (V_i, E_i)$ $(i = 1, 2)$ and cluster depth 2, and a mapping function $\Phi : \mathcal{C}_1 \to \mathcal{C}_2$, where $\mathcal{C}_i$ denote the clusters in G_i. Generate a layout $\pi(G)$ of the union graph $G = G_1 \cup G_2 \cup G[F]$, where F denotes the edges induced by the mapping, respecting the clusters as well as the conventional layouts of

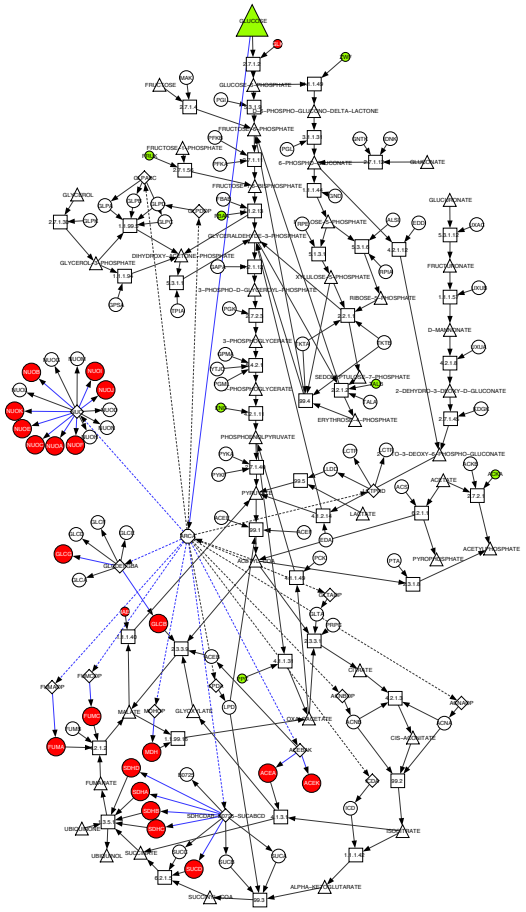

Fig. 4. An example of an integrated network consisting of a part of the glycolysis network and a gene regulatory network. The picture is taken from [23].

each of these networks. In the simpler case when the graph G_2 is highly disconnected (e.g., $E_2 = \emptyset$), we may prefer a solution in which the connected components of G_2 are integrated into a layout of G_1 (see, e.g., Figure 4). In this case, we do not require to respect the root clusters in the graph drawing problem mentioned above. Sometimes, the layout of G_1 may be given. If this is the case, there is a trade-off between sticking to the given layout as closely as possible (or mental map, backbone, etc.) and ignoring it.

3.4 Visualization of sub-cellular localization

Cells consist of distinct compartments, subcellular locations, separated from each other by membranes. Examples for these are the cytosol (the inner space of cell), the nucleus, the mitochondria, or chloroplasts in plants. The membranes enclosing a compartment separate parts of the biological networks as well. Different partitions of the network will be localized in different subcellular locations and hence cannot interact with each other directly. It is thus essential for an understanding of the network's function to integrate that spatial information

into the layout of the network. The required localization data is either already contained in data sets derived from experiments, can be extracted from external sources, or can be predicted.

Given a network $G = (V, E)$ and additional localization annotation for the nodes in V, we search for a layout that reflects the topographical information of G and that conforms to the drawing conventions for that type of network. It should, of course, be at the same time aesthetically pleasing. Note that the subcellular localization does not just give a clustering of the nodes, because, for example, a specific relative position of the cellular compartments may be implicitly given by the biological morphology. If G reflects a flow of mass or information, the direction of the flow also needs to be displayed, e.g., by hierarchical layering of the nodes. A simple representation of a cross-sectional cut through a cell would be a stacking of layers as in [2]. In a layout that is inspired by biological morphology, subgraphs of the network should be arranged according to their subcellular localization in such a fashion that the physical structure of the compartments delimiting these subgraphs becomes evident (e.g., [5, 16]). In order to increase the acceptance of such layouts by biologists, it may be necessary to resemble the (manually generated) layouts from established projects like BioCarta (`http://www.biocarta.com/`).

3.5 Visualization of multiple attributes

Often multiple attributes have to be considered when analyzing biological data. One example is time-series data which is frequently collected in order to better understand the dynamic behavior of a biological system. The combined representation of such time-series data and a corresponding network should allow biologists to gain new insights concerning the underlying system, for example, co-regulated sets and their connection within the network. Such a combination can, for example, be achieved by mapping the data onto the nodes of a network, see Figure 3 where time series data (20 days) from two series (day: red, night: blue) were mapped on some nodes which have been enlarged. A complex use case of multiple attributes occurs in cancer treatment, when antibody data and class labels (clinical data) of up to 200 patients should be mapped to a protein interaction network. In addition to standard statistical clustering methods the underlying network is essential in order to identify structural patterns that are significant among groups of patients with the same or similar class labels.

Mapping the given data onto nodes and/or edges is one possibility of solving the problem. In this case, we have to solve information-visualization problems so that the information can be quickly perceived. This is a challenging task because we must find an efficient interactive visualization for these attributes. The simple use of small visualizations that replace, for example, the node representations is most of the time not sufficient, because a visual comparison of such small graphics in a large graph is impossible. To address such problems, we must develop new interaction metaphors that support the user to filter out uninteresting attributes, to use so-called preattentive features [22] for discovering patterns or correlations between the attributes, etc.

Sometimes, it may be helpful to generate a drawing of the network in which the co-regulated subgraphs (e.g., maximal connected subgraphs of up- and down-regulated sets or sets having the same dynamic behavior) are grouped together, which will lead to cluster drawing problems. In the case that patterns have been identified in the network, we search a layout in which these patterns are displayed prominently (see, e.g. the following use case).

3.6 Visualization of flows and paths in networks

The qualitative and quantitative distribution of mass and signal flows (fluxes) within a network has to be analyzed under uncertainties of the data. The flow of certain paths may change over time (time-series of measurements) and the number of paths through the network is so big that not all of them can be displayed. Biologists are therefore only interested in the main paths through the network, i.e., those paths which possess a statistically significant flow and that transport a considerable percentage of the overall flow through the entire network. The scientist is then interested in the metabolites and reactions that are involved in these paths.

Here, a potentially large network is given together with quantitative and qualitative information about the flow of mass or information given as edge weights. For directed graphs like metabolic networks the layout must reflect the hierarchical nature of the flow, preserve layouts for subnetworks from textbook representations as closely as possibly, adhere to drawing conventions and at the same time focus on relevant parts of the network, e.g., paths that at a certain point in time transport a large part of the flow. These main paths therefore have to be visually emphasized (e.g., placed at the center of the layout and drawn as straight lines) and the distribution of the flow within the network has to be depicted, e.g., by using edge width or color. If the dynamic change of the flow over time also needs to be visualized, smooth animations are required to preserve the users' mental map.

3.7 Exploration of hierarchical networks

Biological networks often comprise several thousand nodes and edges. To help exploring such large and complex structures the entire network is usually broken down in a hierarchical manner into pathways and subpathways. Biologists commonly focus on (sub)pathways in a region of interest and explore their relation to other pathways. However, due to the many connections between different pathways often an abstract overview-like picture of all pathways and their connections as well as an interactive navigation from a set of pathways to other connected or related pathways is desired. An example of some pathways and the derived overview graph is shown in Figure 5.

Given a huge biological network, methods for the biologically meaningful visualization of selected subsets $G_1, \ldots, G_k$ (e.g., pathways) and their interrelations, as well as techniques for the navigation within the network are needed. In order to allow the user to keep his orientation during exploration of the network, the layout changes resulting from a user interaction (e.g., selection of an additional pathway) should be small and also context information needs to be represented in an appropriate way. Expand-and-collapse mechanisms therefore need to be incorporated into layout algorithms such that drawing conventions and the mental map are preserved. These operations could be restricted to certain levels of abstraction, for example by only collapsing/expanding semantically meaningful substructures like pathways. One of the main challenges is that layouts for such subnetworks as well as their relative position to each other, may be given. This layout information needs to be preserved as closely as possible. As these networks are too large to be laid out nicely as a whole, some overview graph or backbone could be defined by reduction or abstraction that covers the topologically or semantically relevant features of the network, thus helping the biologist in navigating through the network (see Figure 5 for an example). The subsets G_i do not need to be disjoint but may partially overlap. This poses an additional challenge for the visualization problem: Either the duplicates are merged, which complicates the task of

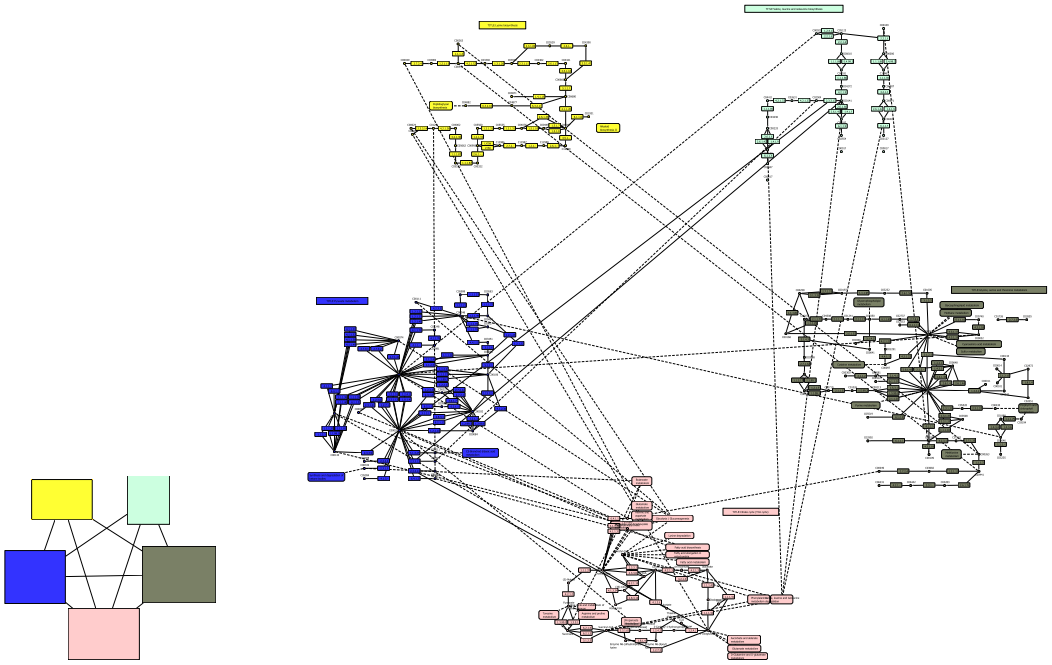

Fig. 5. The combination of a specific overview graph of five pathways (here done with a circular layout method, left part of the figure) and subsequent replacement of graph nodes with complete pathways (right part). The graph was drawn with KGML-ED [13].

mental map preservation, or it has to be clearly emphasized somehow that they represent the same biological entity.

4 Conclusions

In this paper, we have presented some common use cases describing the visualization of biological networks and their applications. These examples have revealed graph-drawing and information-visualization problems, predominantly tackled in the past by experts from the life sciences. Developing improved solutions to these problems will require custom state-of-the-art graph-drawing approaches, and more importantly, collaboration between graph drawing, information visualization, visual analytics, and the life-sciences.

Acknowledgements

We would like to thank Schloss Dagstuhl, Leibniz Center for Informatics, for providing the congenial atmosphere for the seminar 08191 "Graph Drawing with Applications to Bioinformatics and Social Sciences", which led to the conception of this paper. We would also like to thank Hagen Blankenburg for providing us with the drawings of the protein interaction networks and Stephan Diehl, Aaron Quigley, and Idan Zohar for helpful and stimulating discussions during the Dagstuhl seminar.

References

1. M. Ashburner, C.A. Ball, J.A. Blake, and D. Botstein et al. Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25:25–29, 2000.
2. A. Barsky, J.L. Gardy, R.E.W. Hancock, and T. Munzner. Cerebral: a cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation. *Bioinformatics*, 23(8):1040–1042, 2007.
3. U. Brandes, T. Dwyer, and F. Schreiber. Visualizing related metabolic pathways in two and a half dimensions. In *Graph Drawing*, volume 2912 of *LNCS*, pages 111–122, 2003.
4. M.S. Cline, M. Smoot, E. Cerami, and A. Kuchinsky et al. Integration of biological networks and gene expression data using cytoscape. *Nature Protocols*, 2(10):2366–2382, 2007.
5. E. Demir, O. Babur, U. Dogrusoz, and A. Gursoy et al. Patika: An integrated visual environment for collaborative construction and analysis of cellular pathways. *Bioinformatics*, 18(7):996–1003, 2002.
6. C. Erten and S.G. Kobourov. Simultaneous embedding of planar graphs with few bends. In *Graph Drawing*, volume 3383 of *LNCS*, pages 195–205. Springer Berlin, 2005.
7. Z. Hu, J. Mellor, J. Wu, and T. Yamada et al. VisANT: data-integrating visual framework for biological networks and modules. *Nucleid Acids Research*, 33:W352–W357, 2005.
8. B.H. Junker, C. Klukas, and F. Schreiber. VANTED: A system for advanced data analysis and visualization in the context of biological networks. *BMC Bioinformatics*, 7:109, 2006.
9. M. Kanehisa, M. Araki, S. Goto, and M. Hattori et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Research*, 36(Database issue):D480–D484, 2008.
10. P.D. Karp, C.A. Ouzounis, C. Moore-Kochlacs, and L. Goldovsky et al. Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. *Nucleic Acids Research*, 33:6083–6089, 2005.
11. P.D. Karp and S.M. Paley. Automated drawing of metabolic pathways. In H. Lim, C. Cantor, and R. Bobbins, editors, *Proc. International Conference on Bioinformatics and Genome Research*, pages 225–238, 1994.
12. T.H. Kloepper and D.H. Huson. Drawing explicit phylogenetic networks and their integration into splitstree. *BMC Evolutionary Biology*, 8:22, 2008.
13. C. Klukas and F. Schreiber. Dynamic exploration and editing of KEGG pathway diagrams. *Bioinformatics*, 23(3):344–350, 2007.
14. W. Li and H. Kurata. A grid layout algorithm for automatic drawing of biochemical networks. *Bioinformatics*, 21(9):2036–2042, 2005.
15. G. Michal. *Biochemical Pathways, 4th edition (Poster)*. Roche, 2005.
16. M. Nagasaki, A. Doi, H. Matsuno, and S. Miyano. Genomic Object Net: a platform for modeling and simulating biopathways. *Applied Bioinformatics*, 2:181–184, 2004.
17. D. E. Nicholson. *Metabolic Pathways Map (Poster)*. Sigma Chemical Co., 1997.
18. H. Ogata, S. Goto, K. Sato, and W. Fujibuchi et al. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 27:29–34, 1999.
19. F. Schreiber. High quality visualization of biochemical pathways in BioPath. *In Silico Biology*, 2(2):59–73, 2002.
20. M. Sirava, T. Schäfer, M. Eiglsperger, and M. Kaufmann et al. BioMiner - modeling, analyzing, and visualizing biochemical pathways and networks. *Bioinformatics*, 18(Suppl. 2):219–230, 2002.
21. A. Symeonidis, I.G. Tollis, and M. Reczko. Visualization of functional aspects of microrna regulatory networks using the gene ontology. In *Biological and Medical Data Analysis*, volume 4345 of *LNCS*, pages 13–24. Springer, 2006.
22. Colin Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann, 2nd edition, 2004.
23. C.H. Yeang and M. Vingron. A joint model of regulatory and metabolic networks. *BMC Bioinformatics*, 7:322, 2006.