Visual Analysis of Overlapping Biological Networks

Conference Paper · July 2009

DOI: 10.1109/IV.2009.55 · Source: DBLP

5 authors, including:

David C Y Fung
UNSW Sydney
19 PUBLICATIONS **169** CITATIONS
SEE PROFILE

Seok-Hee Hong
University of Sydney
168 PUBLICATIONS **1,456** CITATIONS
SEE PROFILE

Falk Schreiber
Universität Konstanz
214 PUBLICATIONS **3,929** CITATIONS
SEE PROFILE

Kai Xu
Middlesex University, UK
78 PUBLICATIONS **415** CITATIONS
SEE PROFILE

Visual Analysis of Overlapping Biological Networks

David C. Y. Fung, Seok-Hee Hong, Dirk Koschützki
Falk Schreiber, Kai Xu

Abstract

This investigates a new problem of visualizing a set of overlapping networks. We present two methods for constructing visualization of two and three overlapping networks in three dimensions. Our methods aim to achieve both drawing aesthetics (or conventions) for each individual network and exposing the common nodes between the overlapping networks. We evaluated our approaches using biological networks including protein interaction network, metabolic network, and gene regulatory network, from the bacterium Escherichia coli *and crop plants to demonstrate their usefulness to support biological analysis.*

1 Introduction

In this paper, we study a problem of visualizing a set of *overlapping* networks. Roughly speaking, two networks overlap if they share a subset of nodes and/or edges. Note that unlike temporal/evolution/dynamic networks, which consists of a set of similar networks arising from one single network, the overlapping networks may be *heterogeneous*, i.e., consist of different types of networks.

Overlapping networks appears frequently in many application domains. Our research problem was originally inspired by the current understanding of biological networks, which considers a biological network as a supergraph of heterogeneous but interconnected networks. Networks play a central role in the life sciences. They have been used to represent, analyze, and visualize various biological processes, and several networks are employed frequently such as metabolic networks (MN), gene regulatory networks (GRN), protein interaction networks (PIN), and signaling networks. Often these networks share common elements. For example, proteins are the result of gene expression which is regulated by the GRN. Proteins interact with one another to form the PIN. Also, special proteins known as enzymes help transforming metabolites to another, hence, they are part of the MN.

The *integration* of all these networks forms a single huge complex network. Because of the need to reduce complexity and also the limited availability of data, this large network is often divided into specific subnetworks, each dealing with a particular biological function. For example, MNs are commonly broken down into separate parts (amino acid synthesis, carbohydrate metabolism, and so forth) and then into simple pathways. Most biological studies focus only on one of these subnetworks or even pathways. However, this approach suffers from the limitation that the connections between different subnetworks or pathways are lost. As a result, the probable function of a pathway is often deduced from incomplete information.

Good visualization of overlapping networks can enable *integrated analysis*, which cannot be supported by visualizing each single network independently. Furthermore, complex high level analysis can be supported by relating two or more heterogeneous networks. More specifically, by representing a set of overlapping networks in a single visualization, we can enable visual analysis by highlighting important relationships between different networks, while emphasizing both the intersection and the difference.

In this paper, we present two methods for visualizing two and three overlapping networks in three dimensions. More specifically, we use 2.5D representation (i.e., drawing each network in a two dimensional plane) to visualize a set of overlapping networks, in order to minimize occlusion problem and ease navigation problem in three dimensions. We also use inter-plane edges for highlighting the intersection (i.e., shared nodes) between networks. Our methods aim to achieve both drawing aesthetics (or conventions) for each individual network, and an optimization criteria for minimizing the total inter-plane edge lengths in order to help the understanding the visualization.

Our approaches are evaluated with three types of biological networks to support visual analysis of *network integration* of two overlapping metabolic pathways, and three heterogeneous biological networks (MN, PIN and GRN). An integrated approach which considers the overlap between specific networks can substantially improve biological analysis.

First example is the integration of two different metabolic pathways. Biologists who are experts in specific pathways may want to know if some pathways share common metabolites or reactions, and only by knowing the connections between pathways that one can gain an understanding on the flow of metabolites through multiple pathways. In this respect, the proposed visual analysis of the two overlapping network in Section 5.1 demonstrates its usefulness.

Second example is the integration of three different networks (MN, PIN and GRN). Each network represents one type of molecular interactions. Biologists who come from the background of complex systems or biophysics are often more interested in the functional position of a pathway in terms of systems architecture. In this respect, the proposed visual analysis of the three overlapping network in Section 5.2 demonstrates its usefulness at a much larger scale.

2 Related Work

Complex biochemical processes in living beings constitute a number of networks which attract visual analysis methods to life sciences. Several authors deal with the visual investigation of PIN and MN [3, 5, 11, 15, 19, 23]. In general, these studies on biological processes only focus on a *single* network of an organism.

The major problem of treating these networks separately is that connections between these processes, which clearly occur within living organisms, are not represented in these single networks.

Several methods have been proposed to visualize a set of *similar* biological networks in three dimensions [7, 24]. However, they are designed to visualize a set of *similar* networks that are of the same type and the difference among them is small.

Overlapping networks occur in several application area, and they have gained large interest recently. A force directed method for drawing intersecting clustered graphs which supports complex structures such as inclusion and overlap between nodes and clusters was presented [17]. Similarly, a clustered graph based approach was used to address pathway overlapping problem [6].

A *multiple alignment* approach was introduced to support biological *network comparison* [20, 21]. However, the resulting drawing in *two* dimensions have edge-edge crossings and node-edge crossings, due to the use of matching edges between the set of drawings.

Recently, a method for visualizing *two* overlapping biological networks using 2.5D representation, where one network has a given fixed layout, was introduced [8]. In this paper, we extend this approach to visualize two and *three* overlapping networks with given two fixed layouts, and conduct a detailed investigation into the biological analysis of visualization.

3 Two Overlapping Network Visualization

In order to construct 2.5D representation of two overlapping networks G_1 and G_2, we use two parallel planes P_1 and P_2 in three dimensions. More specifically, we draw G_1 and G_2 with layouts L_1 and L_2 on the top plane P_1 and the bottom plane P_2 respectively. Then inter-plane edges that represent mappings of shared nodes between G_1 and G_2 are added between P_1 and P_2.

Note that L_1 or L_2 may or may not be given based on the drawing convention of the specific network. For example, PINs are commonly drawn with force-directed method, which does not have a fixed layout, whereas metabolic pathways are usually drawn with hierarchical or KEGG layout, which is pre-defined.

When G_1 and G_2 both have fixed layouts, our main optimization criteria is to minimize the total edge lengths of inter-plane edges between P_1 and P_2. In general, minimizing the total edge length is one of a well known criteria in graph visualization [4]. More specifically, this can be achieved as follows:

1. Draw graph G_1 with layout L_1 on P_1.
2. Draw graph G_2 with layout L_2 on P_2.
3. Add inter-plane edges between P_1 and P_2.
4. Reduce the total length of the inter-plane edges by fixing one drawing, and then rotate, scale, and translate the other.

While the first three steps are straightforward, the last step is essentially an optimization problem for finding the right amount of rotation, scaling and translation in order to minimize the total edge length. Let $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$, and M_v be a mapping between a vertex $v_1 \in G_1$ and $v_2 \in G_2$ which represents the overlapping information. Without loss of generality, we fix L_1. Assume that (v_2^x,v_2^y) is the x and y coordinate of node $v_2 \in V_2$. After rotating G_2 with angle θ, the new coordinate of v_2 is:

$$(v_2^x \cos(\theta) - v_2^y \sin(\theta), v_2^x \sin(\theta) + v_2^y \cos(\theta))$$

After scaling k and translation (Δ_x, Δ_y)—Δ_x and Δ_y are the x and y translation respectively—the new coordinate is:

$$(k(v_2^x \cos(\theta) - v_2^y \sin(\theta)) + \Delta_x, k(v_2^x \sin(\theta) + v_2^y \cos(\theta)) + \Delta_y)$$

The total inter-plane edge length of all edges connecting G_1 and G_2 can be computed as

$$S = \sum_{v_2 \in V_2,\ v_1 \in V_1,\ v_2 \leftrightarrow v_1 \in M_V} |v_2 - v_1|$$

where $|v_2 - v_1|$ is a distance between v_2 and v_1. Note that different distance metrics can be used, such as Euclidean or Manhattan distance and that the choice of the metric will have an impact on the final visualization. In the following let $\alpha = v_2^x \cos(\theta) - v_2^y \sin(\theta)$ and $\beta = v_2^x \sin(\theta) + v_2^y \cos(\theta)$. For Euclidean distance the following holds:

$$|v_2 - v_1| = \sqrt{(\alpha k + \Delta_x - v_1^x)^2 + (\beta k + \Delta_y - v_1^y)^2 + (v_2^z - v_1^z)^2}$$

Because G_1 and G_2 are drawn in parallel to the x-y plane, their distance along the z-axis is a constant d. Therefore,

$$|v_2 - v_1| = \sqrt{(\alpha k + \Delta_x - v_1^x)^2 + (\beta k + \Delta_y - v_1^y)^2 + d^2}$$

and the total inter-plane edge length according to the Euclidean distance can be computed as:

$$S = \sum_{v_2 \leftrightarrow v_1 \in M_V} \sqrt{(\alpha k + \Delta_x - v_1^x)^2 + (\beta k + \Delta_y - v_1^y)^2 + d^2}$$

To achieve a good visualization, the value of S and therefore the total inter-plane edge length should be minimized.

To solve this minimization problem we used the Nelder-Mead optimization method [16], a standard method for solving nonlinear optimization problems. An implementation of this method is available in the Apache Commons Mathematics Library[1], a Java library provided by the Apache Software Foundation under the Apache License Version 2.0. To compute the required values for θ, k, Δ_x and Δ_y, we integrated the library into GEOMI [1], a visual analysis tool for large and complex networks, and invoked the optimizer with an instance of the function S given above.

In general, 2.5D visualization has less occlusion problem and easier to navigate than the full 3D visualization. Further, GEOMI supports basic interaction and navigation methods including zooming and panning, rotation and selection.

[1] http://commons.apache.org/math/

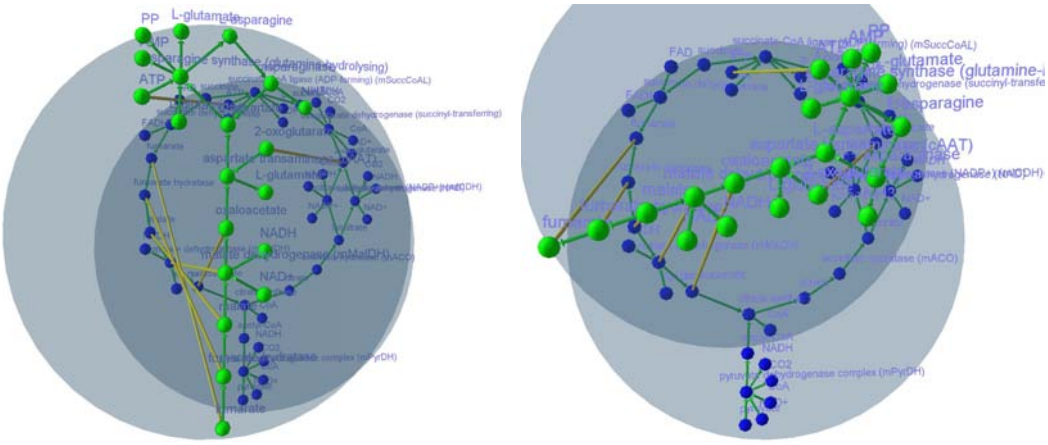

Figure 1: The pathways for Asparagine biosynthesis (green) and the tricarboxylic acid (TCA) cycle (blue) for crop plants. (left) Both pathways are drawn with the coordinates provided by the database MetaCrop. (right) The TCA cycle was fixed and the Asparagine biosynthesis pathway was scaled, rotated and translated.

4 Three Overlapping Network Visualization

Let G_1, G_2, G_3 be three overlapping networks. As with the case of two overlapping networks, we use three planes P_1, P_2 and P_3 in three dimensions. More specifically, we first draw each G_i with layouts L_i on the plane P_i, $i = 1, 2, 3$, respectively. Then, inter-plane edges that represent mappings of shared nodes between G_i and G_j are added between P_i and P_j, where $i, j = 1, 2, 3$ and $i \neq j$.

Based on the structure of three overlapping networks, we have designed two representations: *parallel plane layout* and *circular plane layout*. Two representations mainly differ in the method for arranging planes in three dimensions. The circular plane layout arranges planes P_i, $i = 1, 2, 3$, in a triangular shape in three dimensions. Note that if the structure of three overlapping networks forms a path (for example, G_1 and G_2 overlap, and G_2 and G_3 overlap), then parallel plane layout can be preferred. However, if the structure of three overlapping networks forms a cycle (for example, G_1 and G_2 overlap, G_2 and G_3 overlap, and G_3 and G_1 overlap), then circular plane layout can be preferred.

As with the case of two overlapping networks, each representation can have variations depending on whether G_i has a *fixed layout* $L_i, i = 1, 2, 3$. In this paper, we present one of the variation, the parallel plane layout with *Fixed-Free-Fixed Case*, where G_1 and G_3 have fixed layouts, as we have specific application with biological networks: the metabolic pathway and the GRN have fixed layouts, whereas the PIN does not.

Suppose that G_1 and G_3 have fixed layouts L_1 and L_3. In this case, we need to compute a good layout of G_2, considering the fixed layout of G_1 and G_3, in order to minimize the total inter-plane edge length. A modification of a force-directed algorithm can be used to produce a reasonably good layout of G_2 and inter-plane edges as follows:

1. Draw G_1 with a given layout L_1 on P_1.
2. Draw G_3 with a given layout L_3 on P_3.
3. Arrange planes $P_i, i = 1, 2, 3$ in parallel way.
4. Assign the initial position of node v_2 in G_2 using the barycenter of the positions of its mapped nodes v_1 in L_1 and v_3 in L_3.
5. Add inter-plane edges between planes P_1 and P_2, and P_2 and P_3 to represent mappings, and model each inter-plane edge as a zero-length natural spring (i. e. attraction force only).
6. Draw G_2 and the inter-plane edges using a force-directed layout.

At step 4, by assigning a good initial position based on L_1 and L_3, it can help the force-directed layout of G_2 at step 6 to converge quicker. At step 5, we add the zero-length natural spring for inter-plane edges, in order to reduce the total edge length of inter-plane edges. Note that at step 6, this force competes with other forces of G_2, which try to produce a good layout for G_2. As a result, the corresponding nodes may not always perfectly aligned as a straight-line. We have implemented the new layout using GEOMI [1].

5 Overlapping Biological Networks

5.1 Visual analysis of two overlapping networks

We used the database MetaCrop [10] for constructing an overlapping MN consisting of two metabolic pathways. Metabolic reactions are organized in groups called pathways. As of November 2008, 38 pathway diagrams are available in MetaCrop documenting the metabolism of major crop plants with high agronomical importance such as *Hordeum vulgare* (barley), *Triticum aestivum* (wheat) and *Oryza sativa* (rice).

From MetaCrop we extracted two pathways, the TCA cycle and the Asparagine biosynthesis using the tool VANTED [12]. In

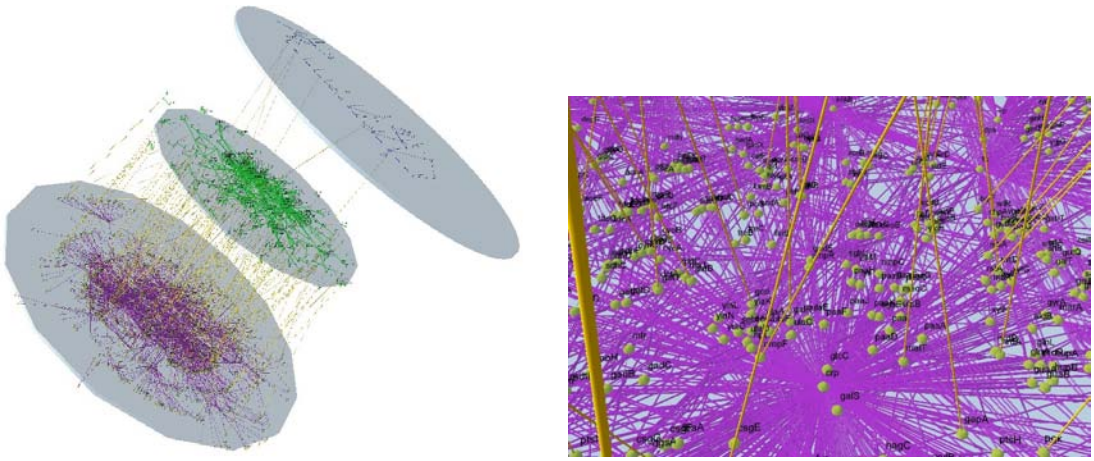

Figure 2: (left) Three overlapping network in the parallel plane layout (side view). (right) Network topology of the central GRN protein crp. The yellow edges are the inter-plane edges between GRN and PIN.

total the networks consist of 43 nodes and 44 edges (TCA cycle) and 24 nodes and 26 edges (Asparagine biosynthesis). Metabolic pathways consist of two kinds of nodes: enzymes and metabolites. Both node types can overlap between different pathways.

To construct an integrated network, both pathways were combined. Every metabolite in one pathway having at least one corresponding metabolite in the other pathway was connected to each other by inter-plane edges. For enzymes, the same process was applied. In total an overlap of 7 metabolites (fumarate, malate, oxaloacetate, NAD+, NADH, 2-oxoglutarate, ATP) and 2 enzymes (fumarate hydratase, malate dehydrogenase) were detected.

Based on the fixed layouts of two metabolic pathways, we computed a visualization of the integrated network by applying the method described in Section 3. The result is shown in Fig. 1 (right) with a viewpoint from the top. For a comparison, see Fig. 1 (left) shows the same pathways without performing the optimization.

From the top viewpoint, it is clearly comparable, that the optimization resulted in a rotation ($\theta \approx 270$ degrees) and scaling ($k \approx 0.73$) of the Asparagine biosynthesis pathway. Additionally, the optimization resulted in a minor translation of the x and y coordinates ($\Delta_x \approx -1.22$ and $\Delta_y \approx 6.67$). By this optimization the total edge length of the edges for overlapping nodes was reduced from ≈ 25.24 to ≈ 22.43.

Using our method the overlapping parts between the two pathways are easily recognizable and it becomes clear, that Asparagine biosynthesis starts with some reactions of the TCA cycle and is therefore dependent of the activity of this pathway.

5.2 Visual analysis of three overlapping networks

To understand how the GRN in *E. coli* regulates the MN by influencing the physical organization of the PIN, we used three overlapping network visualization. This should provide an overview on the different modes of interaction within *E. coli*.

We used data sets downloaded from three public databases to construct the overlapping networks. The GRN data was downloaded from the RegulonDB database version 6 (accessed April 2008) [9] from which the largest connected component was extracted. This subset contains 1371 operons and 3030 interactions. Each operon is a DNA sequence encoding a protein. The MN data was downloaded from the pathway section of the KEGG databases (accessed June 2008) [14] from which the glycolytic pathway was extracted. This subset contains 57 nodes and 62 edges. Of which, 29 nodes represent proteins (also known as enzymes) and the rest represent metabolites. The PIN data was downloaded from the DIP database (accessed April 2008) [18] from which the largest connected component was extracted. This subset contains 1434 proteins and 6567 interactions.

To construct the three overlapping networks, all networks are integrated as follows. For every protein that has a corresponding node in either the glycolytic pathway or the GRN, an inter-plane edge is being added. The intersection between PIN and GRN contains 250 proteins, and the intersection between PIN and the glycolytic pathway contains 7 proteins.

To better expose the highly connected proteins (also known as hubs) that also have corresponding nodes in the other two networks, the largest connected component in the PIN is further reduced to the subnetwork consisting of neighbors of the overlapping proteins. The final PIN being visualized contains 514

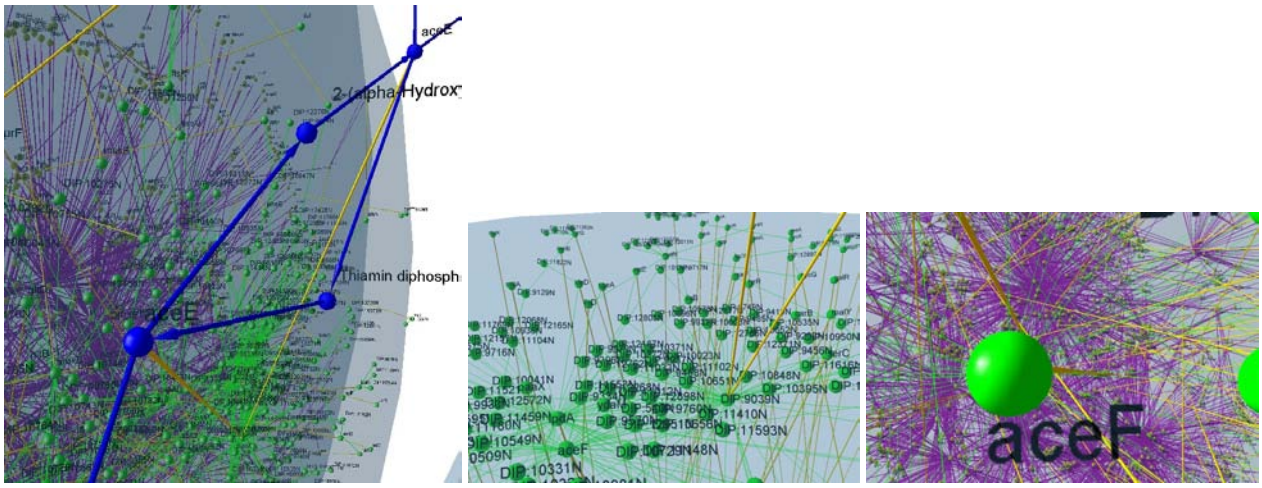

Figure 3: Connectivity of the metabolic protein *aceE* to GRN and PIN. (left) *aceE* in G_1 (blue) corresponds to a highly connected hub in the PIN (green in G_2). (middle) *aceE* (DIP:9039N) and its neighbour *aceF* in the PIN. (right) *aceF* in the PIN has its corresponding node in the GRN.

proteins and 807 interactions. The largest connected component of the GRN is visualized as it is in order to expose any protein-coding operons that may have dual functionalities, i.e. proteins that can catalyze metabolic reactions but can also function in gene regulation. At the same time, the full gene regulatory hierarchy for controlling the organization of the PIN is also preserved in the visualization. This enables us to deduce the control points that regulate the flow of metabolites through the glycolytic pathway and beyond.

Figure 2(left) shows the side view of the parallel plane layout. Here, G_1 represents the MN (blue), the glycolytic pathway. G_2 represents the largest connected component of the PIN (green) and G_3 represents the GRN (yellow nodes; magenta edges). The layers are arranged in this order to capture the current biological model that the formation of PIN is being controlled by GRN and the MN is being operated by PIN.

The side view showed the direct connections between these three networks. The fixed layout for G_1 were computed by the KEGG coordinates whereas those of G_3 were pre-computed using the Kamada-Kawai layout [13]. The parallel plane layout exposed the control hierarchy as a series of layers. G_1 contains the enzyme-mediated metabolite flow in glycolysis. G_2 contains the physical interactions between proteins. G_3 contains the flow of control in gene expression from operon to operon.

By zooming into the visualization of G_3, six highly interconnected hubs (*arcA, crp, fis, hns, ihfAB*, and *lrp*) can be identified. This is suggesting that they are the master switches controlling the global state of the GRN.

Visually, *crp* appears to have the highest node degree which suggests that it is the central GRN protein in *E. coli* (Figure 2(right)) and could exert the most extensive influence on the organization of the PIN. This observation agrees with the current consensus that *crp* is the key hub in the *E. coli* GRN [2]. So far, *crp* has been known to regulate more than 200 proteins involved in a variety of biological processes such as carbohydrate metabolism, amino acid metabolism, ion transportation, energy production, and gene regulation [26]. The on/off state of *crp* should therefore exert the highest impact on not just the glycolytic pathway but probably the entire MN.

Figure 2(right) also shows that *crp* does not have a corresponding node in G_2. Thus it does not require any protein cofactors to facilitate gene regulation. Such a design will allow the bacterium to fine tune its PIN organization and MN dynamics rapidly in response to any external environmental challenges.

Figure 2 (left) shows that there are seven proteins (enzymes) from the glycolytic pathway that correspond to the PIN as indicated by the inter-plane edges between G_1 and G_2. These proteins are *aceE, frmA, agp, pgmA, ptsG, pykF*, and *tpiA*.

Of these, *aceE* (DIP:9039N) appears as a high degree hub of over 20 in G_2 (Figure 3(left)). One of its neighbor *aceF* is also a high degree hub of almost comparable size (Figure 3(middle)). Both *aceE* and *aceF* are protein subunits of pyruvate dehydrogenase multi-enzyme complex which is a known junction point between glycolysis and alanine, aspartate, valine, leucine, and isoleucine biosynthesis [25]. Figure 3(right) shows that only *aceF* has a corresponding node in the GRN suggesting that it is the point for regulating the abundance of functional pyruvate dehydrogenase complexes within *E. coli*. This regulatory mechanism effectively reduces the need for the just-in-time production of every protein that make up a metabolic protein complex allowing the bacterium to shutdown or switch on part of the MN rapidly [22].

In conclusion, the three overlapping network is visually more complex than the two overlapping network, but the former can provide a more complete overview of the organism's molecular network. For this reason, the three overlapping network allows more in-depth analysis than its two overlapping counterpart. The two overlapping network should be more suitable for investigating biological problems with a specific focus on a particular biological process.

Biologists can use the two overlapping and three overlapping networks as a two step visual analysis process. One possible use case scenario could be the biologist uses the two overlapping network to investigate how a signal transduction network is being regulated by a GRN, followed by the use of the three overlapping network to investigate if the same GRN regulates proteins within the PIN but are outside the signal transduction network. This should allow the biologist to deduce hypotheses on how the signal transduction network controls cellular responses in relation to cellular sensing.

6 Conclusions

In this paper, we study the problem of visualizing overlapping networks, and present three methods for visualizing two and three overlapping networks using 2.5D representation. Our methods can achieve both drawing aesthetics and conventions for each individual network, and simultaneously highlight the overlapping part between them. The usability of our visualization approaches has been studied on two real world applications from biology. It showed that our visualization can support integrated visual analysis of three heterogeneous biological networks (MN, GRN and PIN), and two overlapping metabolic pathways.

Our methods can be used for large networks, by reducing the size of the whole networks to a subset of relevant overlapping networks, as in our examples. Our current work is to design a new method for visualizing a set of k overlapping networks, where $k \geq 3$.

References

[1] Adel Ahmed and et al. GEOMI: GEOmetry for Maximum Insight. In *Graph Drawing*, pages 468–479, 2005.

[2] C. L. Barrett and et al. The global transcriptional regulatory network for metabolism in Escherichia coli exhibits few dominant functional states. *National Academy of Science U S A.*, 102:19103–19108, 2005.

[3] W. Basalaj and K. Eilbeck. Straight-line drawings of protein interactions. In *Graph Drawing*, pages 259–266, 1999.

[4] G. Di Battista, P. Eades, R. Tamassia, and I. G. Tollis. Graph drawing: Algorithms for the visualization of graphs. *Prentice*, 1999.

[5] M. Y. Becker and I. Rojas. A graph layout algorithm for drawing metabolic pathways. *Bioinformatics*, 17(5):461–467, 2001.

[6] R. Bourqui and et al. Metabolic network visualization eliminating node redundance and preserving metabolic pathways. *BMC Systems Biology*, 3:1–29, 2007.

[7] U. Brandes, T. Dwyer, and F. Schreiber. Visual understanding of metabolic pathways across organisms using layout in two and a half dimensions. *Journal of Integrative Bioinformatics*, 1:e2, 2004.

[8] D. C. Y. Fung and et al. 2.5D visualisation of overlapping biological networks. *Journal of Integrative Bioinformatics*, 5(1):e90, 2008.

[9] S. Gama-Castro and et al. RegulonDB version 6.0: gene regulation model of Escherichia coli K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. *Nuclei Acids Research*, 36:D120–D124, 2008.

[10] E. Grafahrend-Belau and et al. MetaCrop - a detailed database of crop plant metabolism. *Nucleic Acids Research*, 36:D954–D958, 2008.

[11] K. Han and B. H. Ju. A fast layout algorithm for protein interaction networks. *Bioinformatics*, 19(15):1882–1888, 2003.

[12] 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:e109, 2006.

[13] T. Kamada and S. Kawai. An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31:7–15, 1988.

[14] M. Kanehisa and et al. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Research*, 34:D354–D357, 2006.

[15] P. D. Karp and S. M. Paley. Automated drawing of metabolic pathways. In *International Conference on Bioinformatics and Genome Research*, pages 225–238, 1994.

[16] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7:308–313, 1965.

[17] H. Omote and K. Sugiyama. Method for drawing intersecting clustered graphs and its application to web ontology language. In *APVIS 2006*, pages 89–92. Australian Computer Society, Inc., 2006.

[18] L. Salwinski and et al. The Database of Interacting Proteins: 2004 update. *Nucleic Acids Research*, 32:D449–D451, 2004.

[19] F. Schreiber. High quality visualization of biochemical pathways in BioPath. *In Silico Biology*, 2(2):59–73, 2002.

[20] R. Sharan and et al. Conserved patterns of protein interaction in multiple species. *PNAS*, 102(6):1974–1979, 2005.

[21] R. Sharan and T. Ideker. Modeling cellular machinery through biological network comparison. *Nature Biotechnology*, 24(4):427–433, 2006. doi:10.1038/nbt1196.

[22] T. Shlomi, Y. Eisenberg, R. Sharan, and E. Ruppin. Interplay between transcriptional regulation and metabolism. *Molecular Systems Biology*, 3:101, 2007.

[23] Wegner and Kummer. A new dynamical layout algoritmh for complex biochemical reaction networks. *BMC Bioinformatics*, 6:212, 2005.

[24] K. Wegner. SimWiz3D - visualising biochemical simulation results. In *Proc. International Conference BioMedical Visualization (MediViz)*, pages 77–82, 2005.

[25] C.H. Yeang and M. Vingron. A joint model of regulatory and metabolic network. *BMC Bioinformatics*, 7:332, 2006.

[26] D. Zheng and et al. Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nuclei Acids Research*, 32:5874–5893, 2004.