Aligning Biomolecular Networks

ARUSHI ARORA
University of Windsor
aroray@uwindsor.ca

Varied complex systems are represented using networks, however the most atypical out of them are called protein-protein interaction (PPI) networks. In the PPI networks the nodes refer to proteins, whereas the edges link to the functional interactions between them. The extensive examination and alignment of biomolecular networks can be demonstrated as a powerful approach for examining a network structure.In order to effectively apply this approach, an efficient multiple network alignment algorithms needs to be formulated. This survey discusses about various methodologies for the comparative analysis of PPI networks and aligning them.This survey summarizes research work on aligning PPI networks and thus finding an optimal alignment algorithm.The survey contains annotations of research publications describing the issue of exploring the optimal alignment and finding different approaches to maximize the complete match between the two networks, further directing to discover a correspondence between nodes and edges of the input networks.Hence, with an ever-increasing amount of available data on PPI networks the problem of aligning these networks becomes an important problem as along with aligning the PPI networks conserved patterns and complexes can also be discovered.

Categories and Subject Descriptors: A.1 [**Introductory and Survey**]: ; J.3 [**LIFE AND MEDICAL SCIENCES**]: Biology and genetics

General Terms: biomolecular alignment, Protein interaction network alignment

Additional Key Words and Phrases: pairwise alignment,multiple alignment,Protein interaction network alignment, biomolecular networks

Contents

1. INTRODUCTION

Protein-protein interaction (PPI) networks are regarded as the most commonly studied biological networks. As proteins in the network do not function alone, they interact with one another thus forming a huge network of protein-protein interactions. These interaction networks are illustrated with the help of floating graphs without self-loops. The nodes signify proteins, whereas the edges symbolize physical interactions between the proteins. A network is represented as a pair G = (V,E) where V is a set of nodes and E = V x V is a set of edges. Nodes of the network represent elements of the complex system and two nodes u and v are connected by an edge if they are somehow related. One challenge faced here is to study and compare these real world networks. The main focus of this survey is to compare the given PPI networks and align them in order to find conserved subgraphs. While it is challenging task for simplifying the network complexity and subgraph isomorphism, heuristic methods are taken into consideration. The heuristic methodology facilitates in generating a merged representation of the networks being compared, known as a network alignment graph. This helps us in finding conserved sub networks. A network alignment graph comprise of nodes that represents the sets of proteins, one from each species, whereas the edges signifies conserved protein-protein interactions across the two species. The alignment might comprise of one-to-one correspondence between proteins across the two networks; although there may be a many-to-many correspondence between proteins. This happens when a single protein from a specie is homologous to multiple proteins from the other species. PPI networks comprise of topology that provides with an understanding of the functioning of individual proteins. This enables us to find the systematic classification of protein-protein interaction (PPI) and use of algorithms for protein network alignment. This network alignment problem is important for locating network regions that are well-preserved in their sequence and interaction pattern across two or more species.

In this survey we show that network alignment problem is carried in a successful manner by various authors to search for conserved pathways and similar complexes. The first section of the survey focuses on research of various approaches used for pairwise alignment of networks. However, the need for extension to more than a few networks proved to be difficult due to the exponential growth of the alignment graph with the number of species. The second section discusses research about aligning multiple networks.

While aligning the biomolecular networks the authors face multiple problems. The main challenge is to find a model that fits PPI networks in the most appropriate way and finds the most optimal alignment.Researchers are still working on

this network alignment problem as it is difficult to compute an algorithm to show the efficiency of comparing compare large networks due to an underlying subgraph isomorphism problem. Another issue arises when it becomes hard to detect false positives and false negatives in PPI networks. There are few computational methodologies that have been adopted to address this difficulty. Another approach that authors have used is to utilize the present PPI network data sets to detect false positives in the exiting data. Also,there are also few problematic issues associated with these datasets such as; high levels of noise and incompleteness. Presently, the most well researched PPI networks are also extremely noisy and incomplete. As the noise and incompleteness are hard to evaluate therefore it is vital to discover a well-fitting graph theoretic model of PPI networks because this model will enable a better understanding of the topology and proteins of PPI networks. Such a model would allow us to better align the networks in the best possible way. These models are also required for statistical evaluations of the real world networks.

2. SURVEY OF RESEARCH

2.1 Pairwise alignment of biomolecular networks

The research papers covered in this section focus mainly on how to align two biomolecular networks. The main purpose of aligning biomolecular networks is to find functional as well as structural similarity between the two input networks.

2.1.1 *Pathway alignment of two biomolecular networks.* With the increase in availability of large protein-protein interaction network data the need to examine evolutionary changes in protein networks is gaining a lot of attention. Since we know that protein sequences are conserved Kelly et. al. [2003] addresses the problem of finding if there exists conserved interaction pathways between two protein interaction networks

The authors refer to no previous work.

The authors propose a method called PathBLAST used for aligning two protein interaction networks to extract conserved pathways. In this method an alignment graph is formed in which the proteins of the first network are paired and aligned with the similar proteins occurring in the second network. In the network alignment graph the proteins that share strong sequence homology are paired together. There are three types of protein interaction relations defined by the authors i) direct ii) gap iii) mismatch. The vertical solid line represents the direct interaction and the vertical dotted line represents an indirect interaction (gap or a mismatch) between proteins of given networks. The horizontal line between the two networks indicates the sequence similarity between the protein nodes. The path of the alignment graph is obtained by searching the highest scoring pathways between two paths one from each network. The score of the pathway alignment is given on the basis of the quality and the amount of sequence similarity of the proteins in the given networks. Because of the variation in the two networks "a gap" or "a mismatch" is introduced. A gap or a mismatch occurs when the proteins in one path are connected to proteins in the other path that are not directly linked with each other.

The authors performed a global alignment between the protein–protein interac-

tion networks of yeast (S. cerevisiae) having 14,489 interactions among 4,688 yeast proteins and bacteria (H. pylori) having 1,465 interactions among 732 proteins.The authors identified orthologous pathways between two species,paralogous pathways within a network of each species. PathBLAST is useful tool for finding conserved paths between networks and hence forming complexes by merging overlapping paths .Information about proteins similarity can be extracted with the help of paths having comparable likelihood as they are said to be biologically significant in providing knowledge about proteins.It was observed that pathway gaps and mismatches detected larger regions of the network that were generally conserved even when there were no direct interactions.

2.1.2 *Aligning networks on the basis of topology.* Considering the large amount of protein-protein interaction data available having different topologies, it is challenging to find similarities between two complex networks.

(1) **Matching-neighbourhood approach.**

Singh et. al. [2007] address the problem of global alignment of two protein-protein interaction (PPI) networks in accordance with their topologies and finding maximum association between the edges and nodes of a network in order to get the best alignment of two networks. The authors have proposed a new algorithm IsoRank for pairwise alignment of two protein-protein interaction networks. For aligning the two networks the algorithm takes into consideration both PPI network data and sequence similarity data. The main idea of algorithm is that a node i in G1 is mapped to a node j in G2 if the neighborhood topologies of i and j are similar, i.e., the neighbors of i can be well-mapped to the neighbors of j.The input to the algorithm is two PPI networks and for similarity data, Blast similarity scores have been used.The output of the algorithm is a maximum common subgraph between two input networks.The the algorithm works in two stages for the given networks G1 and G2: (i) $\forall$i ϵ (G1) and $\forall$j ϵ (G2) it computes the scores Rij of matching node i with node j, (ii)it constructs a global network alignment by extracting from vector R, high-scoring pairwise mutually-consistent matches. Thus, the n1 $\times$n2 dimensional scores vector R is subject to the following constraints

$$Rij = \sum_{u \epsilon N(i)} \sum_{v \epsilon N(j)} \frac{R(u,v)}{|N(u)||N(v)|} \quad \forall i \ \epsilon V \ (G1) \text{and } \forall j \ \epsilon \ V \ (G2)$$

where N(u) is neighborhood of node u [Singh et. al. [2007],21].In the matrix form the equation can be written as R=AR where A is a schotastic matrix.This equation is further modified to include the pairwise node similarity information between the two networks: R=αAR+ (1-α)E where α is the parameter that determines the contribution of the topology in the network alignment.

The authors tested their two-way global alignment algorithm on S. cerevisiae and D. melanogaster PPI networks. The algorithm applied to this pair of networks is used to identify the common subgraph of the given two networks. Experiments were performed to find set of proteins that behave similarly in

the two PPI networks, evaluating algorithms error tolerance power, evaluating influence of α and comparing local and global alignment results.

The authors state that the algorithm, IsoRank is tolerant of noise in the input. Also as the value of α increases, the importance of network data in the alignment process also increases, for maximum weight bipartite matching strategy. The algorithm also helps in predicting functional orthologs(FO) between the fly and yeast.The author claims that IsoRank algorithm is similar to Google's Page Rank algorithm which is based on the concept of ranking web-pages in the order of their "authoritativeness". The input of the algorithm is a pair of undirected weighted graphs and the output is an alignment. The algorithm can be used for both biological and non-biological data and has also been extended for aligning multiple networks.

(2) **Seed and Extend Approach.**
Kuchaiev et al. [2010] addresses the problem of devising an algorithm for aligning two networks on the basis on topology. The advantage is that with the help of topology-based alignment we can extract both phylogeny and biological function information of the respective protein networks.

The authors refer to no previous work.

The authors propose a new algorithm GRAAL (GraphAligner) that is solely based on topological similarity of the networks which are to be aligned. The algorithm does not take into account sequence similarity of protein in the PPI networks thus it can align any two networks not just biological ones. Both local as well as global alignment of networks is possible using this algorithm. Here nodes of the PPI networks being aligned are matched and paired on the basis of signature similarity. More the signature similarity between two nodes more is the topological similarity between their neighborhoods. Cost for aligning every node pair is calculated. Lower the cost more is the probability of the nodes in the given networks to be aligned. A parameter "a" helps in calculating the cost of the pair of nodes and controls the contribution of the node signature similarity to the cost function. If there exists a situation where two pairs have similar cost function then the choice is made randomly. Hence running the algorithm multiple times can result in different solutions. The algorithm is based on a "seed and extend" approach. To begin with, a single node is chosen from each network to be aligned as a "seed". Then the alignment is extended around the seed radially.Hence, using greedy algorithm concept find the nodes with highest signature similarities around the seed. Three types of scores are defined to score GRAAL's algorithm. 1) Edge Correctness (EC). 2) Node Correctness (NC). 3) Interaction Correctness(IC).

The authors align the yeast S. cerevisiae PPI network which consists of 8,323 interactions and1004 proteins, with its 'noisy' counterparts obtained by: (1) random removal of nodes from the network; (2) random removal of edges from the network; (3) random addition of edges to the network; and (4) addition of low confidence PPIs to the network. The authors perform experiments for these four types on noise types by changing the percentages of noise types.

The authors state that if the networks to be aligned are similar to each other then the algorithm proposed gives good alignment results. For 5% random node removal, we obtain NC, EC, and IC of about 70%, 90%, and 55%.For 5% random edge removal, we obtain NC, EC, and IC of about 65%, 75%, and 45%, respectively .For 5% of random edge addition, we obtain NC, EC, and IC of about 50%, 70%, and 60%, respectively). For 5% of low-confidence edge addition, we obtain NC, EC, and IC of about 70%, 90%, and 65%.

2.1.3 *Alignment based on models of evolution.* Koyuturk et. al. [2006] discuss the problem to align two protein-protein interaction networks of different species to comprehend the conserved pathways and protein complexes etc. across different species. The main challenge is to describe a measure of similarity between the different networks that captures underlying biological phenomena accurately.

The authors refer to no previous work.

The authors propose a method called MaWish (Maximum Weight Induced Subgraph) for modeling the conservation and divergence of interactions as important design parameters, as well as the interpretation of resulting alignments. The approach behind MaWish duplication/divergence models that focus on understanding the evolution of protein interactions between two protein-protein interaction networks. It is based on a mathematical model that extends the concepts of match, mismatch, and gap in sequence alignment to that of match, mismatch, and duplication in network alignment and evaluates the similarity between graph structures through a scoring function that accounts for evolutionary events. Considering the evolutionary models, it assists in interpretation of resulting alignments in terms of not only conservation but also divergence of modularity in PPI networks.

The authors align the PPI networks of two mammalians that are available in the database; Homo sapiens (Hsapi) and Mus musculus (Mmusc).The Hsapi PPI network contains 1369 interactions among 1065 proteins while Mmusc PPI network contains 286 interactions among 329 proteins.

This paper presents a framework for local alignment of protein interaction networks. The framework is guided by theoretical models of evolution of these networks. The model is based on discovering sets of proteins that induce conserved subnets based on scoring match and mismatch of interactions, and duplication of proteins. An implementation of the proposed algorithm reveals that this framework is successful in uncovering conserved substructures in protein interaction data.

2.1.4 *Summary*

Year	Author	Title of Paper	Major Contribution
2003	Kelly *et. al.*	Conserved pathways within bacteria and yeast as revealed by global protein network alignment	Allows networks to discover conserved pathways across species. By restricting the alignment to pathways i.e. linear chains of interacting proteins, this algorithm simplies the problem and helps in preserving the biological information of proteins.

2006	Koyuturk *et. al.*	Pairwise alignment of protein interaction networks	Local pairwise alignment algorithm MaWish based on the duplication/divergence models that focus on understanding the evolution of protein interactions. It constructs a weighted global alignment graph and tries to find a maximum induced subgraph in the networks.
2007	Singh *et. al.*	Pairwise global alignment of protein interaction network by matching neighbourhood topology	The first global network alignment algorithm IsoRank which is guided by the intuition that two nodes should be matched only if their neighbours can also be matched.The algorithm is formulated as an eigenvector problem using a greedy strategy.
2010	Kuchaiev *et. al.*	Topological network alignment uncovers biological function and phylogeny	A global network alignment algorithm GRAph ALigner (GRAAL) which is based solely on network topology. It produces alignments that are topologically significant and the proteins that are aligned perform the same biological function.

2.2 Multiple alignment of biomolecular networks

The research papers covered in this section present various approaches for aligning more than two protein-protein interaction networks.

2.2.1 *Aligning on the basis of evolution.* Flannick et.al [2006] address the problem of finding an optimal alignment for multiple networks. Also,conserved functional modules in protein-protein interaction networks were found by aligning these networks.

The authors did not refer to any previous work.

The authors proposed a new network alignment algorithm Græmlin for finding conserved functional modules and conserved pairwise interactions. The conserved modules are pairs of homologous proteins and the conserved interactions are the interactions between the two proteins. This algorithm can perform both global as well as local alignment. Also, Græmlin algorithm can be used to align multiple protein-protein interaction networks and ascertain the presence of a functional module by searching a large database of modules and networks. Græmlin addresses three network alignment problems 1) Formulation of mathematical function 2) devising a scoring function which encapsulate the knowledge of module evolution. 3) A search algorithm to sort high scoring alignments in order to find conserved functional modules. Græmlin uses log scoring function and Græmlin's scoring function uses equivalence class concept which consists of proteins which are functionally orthologous. Edge Scoring Matrix(ESM) allows log-odds scoring function to apply to alignment of pathways, protein complexes and general modules.

The authors analysed the ability of the algorithm to align know functional modules by comparing the results with two alignment algorithms NetworkBLAST and MaWISh .The algorithms were tested on a set of 10 microbial protein interaction networks. The sensitivity of each method and also how many proteins were correctly aligned in the pathway were observed.

The authors state that Græmlin more sensitive than the other two method NetworkBlast and MaWish when looking for highly connected components.It is faster than NetworkBlast with respect to running time that is Græmlin can efficiently search large protein networks. While performing query to network alignment Græmlin builds index for every query and saves it in the memory for future queries thus leading to fast alignment of many queries.

2.2.2 *Using network alignment multigraph to align networks.* A merged representation of the two networks being compared is created called a network alignment graph in which nodes represent sets of proteins, one from each network and edges represent conserved protein interactions across different networks.

(1) **Avoid explicit information of the network.**
Kalaev et.al [2008] address the problem of aligning biomolecular networks and find areas that are conserved in their interaction pattern and sequence among various biomolecular networks using network alignment graph.

The authors refer to no specific previous work.

The authors introduce a new algorithm to align multiple biomolecular networks which helps in avoiding explicit representation of the network which in turn helps in reducing time complexity and memory utilization. The authors designed NetworkBlast-M which can be used for multiple network alignment. The algorithm is based on the concept of k-layer alignment graph where k is any constant. Each layer in thus graph represents single specie and contains corresponding network. The authors define the term k-spine as a subgraph of size k containing all the vertices obtained from the different layers that are functionally orthologous. All the vertices in the k-spine are connected and hence form a spanning tree. The search algorithms use the greedy approach of finding conserved subgraphs between networks. The two main tasks performed by the algorithm is i) finding the seeds having maximum weight ii) extending the seed greedily. The former task of finding the seed with maximum weight is done in the following two ways 1) Obtaining a set of "d" k-spines with identical topologies and maximum score which is based on the assumption that k-spines of the same seed have the same topology of inter-connections. 2) Obtaining set of "d" k-spines based on the corresponding phylogeny having maximum score.

The authors tested their algorithm for eukaryotic and microbial PPI networks and compared the results with the NetworkBlast algorithm. The scoring function and the scoring parameters were the same for both algorithms. The results of the new algorithm were also compared with Graemlin on a set of 10 microbial networks.

The authors state that the new algorithm has higher sensitivity than the NetworkBlast algorithm and its sensitivity further improves when using relaxed order variant. While running NetworkBlast-M on a given data, the time utilized to compute the result was less than thirty seconds in comparison to NetworkBlast which took about six hours to compute the result. Also NetworkBlast-M was better in performance when compared with Graemlin as the latter had lower specificity and sensitivity. The authors claim that the framework discussed in this paper is better than the previous alignment algorithms. NetworkBlast can only align up to 3 networks whereas NeworkBlast-M can align up to 10 microbial networks with improved time efficiency and less memory requirement. The authors observed that NetworkBlast-M works much faster in the case of restricted-order as compared to relaxed order.

(2) **Connected component approach.**
Denielou Y.P. [2009] address the problem of finding similarity among biomolecular network in terms of similarity which is applied on aligned sub networks. While aligning multiple biomolecular networks the complexity of how to avoid explicit construction of network graph is faced. Various methods have been proposed to face this problem.

The authors refer to previous work by Kalaev, M., Bafna, V., Sharan, R.[2008].

The authors introduced a new approach towards solving the problem of avoiding explicit construction of network graph and hence achieving an appropriate alignment among large number of networks. This approach is based on connected components conserved in multiple networks. The author introduces a new algorithm C3Part-M which is an improvement over previous algorithm C3Part. The main purpose of this algorithm is to avoid the initial construction of the multigraph.The basic idea is to build a multigraph on the given biomolecular network. The multigraph is constructed by starting with the connected component on the primary graph and the expanding it to the second graph by splitting it into two colors and the further splitting on the third graph and so on. The previous algorithm followed the process of finding connectons in the network alignment multigraph by considering one single class of all vertices and hence partitioning at every step. Whereas the author enhanced the same process by first starting with one single class of vertices and then calculating the intersection of all connected components on all colors and continuing this process until partition doesn't change. The author introduced two new operations.1) SPLIT1-i – which is used to split a class on colors within range 1-i. 2).EXPANDi+1 – adds (i+1)th color to the current network alignment multigraph.

Denielou Y.P. [2009] compared their algorithm results with NetworkBlast algorithm on 10 microbial PPI networks. Both the algorithms were conducted on relaxed mode and the tree guided path mode.The authors claim that the C3Part-M algorithm proposed can be used for multiple networks. It is observed that the results obtained from both the C3Part-M and NetworkBlast-M algorithm are quite similar despite the difference in their input data.

2.2.3 *Globally aligning multiple networks.* Aligning networks alobally provides a unique alignment from every node in the smaller network to exactly one node in the larger network.It may ay lead to inoptimal machings in some local regions.various methods have been discussed in the following section for global network alignment.

(1) **k-partite matching method.**
Singh et.al [2008] address the problem of global alignment of multiple protein-protein interaction (PPI) networks and finding maximum matching in the given networks.

The authors refer to no previous work.

The authors introduce a new algorithm which works in two parts . First is to compute a k-partite graph of the given k input networks in which each of the k-parts contains node from each input network. The edges mentioned are the intra-network edges i.e. edge eij denotes node i of network G1 is connected to node j of network G2.each edge is assigned an edge weight Rij that indicate the similarity between proteins. In the second stage of the algorithm, nodes are combined on the basis of k-partite matching to form clusters. The k-partite matching has transitive property i.e if i is matched to j and j is matched to l then it implies that i must be matched to l. All the nodes in cluster are mapped to each other in the corresponding global network alignment. The authors also introduce a new concept called functional coherence which is an programmed approach of predicting similarity between various functions of proteins It is used to calculate an aggregated score of proteins which is indicative of how well our algorithm works in comparison to other algorithm.

The algorithm is tested using PPI networks for five species: S. cerevisiae, D. melanogaster, C. elegans, M. musculus, and H. sapiens.the authors also evaluate functional orthologs (FO) across five species.

The authors state that the functional orthologs predicted by the algorithm proposed have certain limitations.The algorithm lacks accuracy when dealing with noisy and incomplete data and hence may not be able to define accurate functionally related set of proteins. The correctness of these functional orthologs is similar to the current sequence-only functional orthologs. The author claims that the algorithm finds conserved functional components across the various species. It also helps us to predict functional orthologs between these five species.

(2) **Spectral partitioning method.** Liao et. al. [2009] addressed the problem of global alignment of two protein-protein interaction (PPI) networks and hence finding dense and clique-like clusters of proteins.

The authors refer to Flannick et al. [2008].

The authors propose a new algorithm IsoRankN (IsoRank-Nibble) which is an extension of IsoRank algorithm defined by Singh et al. [2007].The first step of IsoRankN algorithm is similar to IsoRank but finds pairwise alignment scores between every pair of networks hence forming a k-partite graph. The second part is to find optimal k-partite matching. On the basis of scores clusters are

formed using spectral partitioning methods. Since the information of protein networks is large and incomplete the algorithm uses star spread method to find similar cliques with maximum weight in order to align the multiple networks. The central idea of IsoRankN is to build a multiple network alignment by local partitioning of the graph of pairwise functional similarity scores.

The IsoRankN algorithm is compared with IsoRank and Græmlin 2.0.The algorithm is applied on five known eukaryotic PPI networks i.e. human, mouse, fly, worm, and yeast. The algorithm is tested on two factors 1) Consistency between proteins within a cluster 2) Coverage i.e. number of proteins assigned to a cluster.

The authors state that IsoRankN produce more exact and higher number of clusters than previous existing techniques.It is consistent with respect to local alignment algorithms which are ambiguous, inconsistent and predict overlapping clusters.For pairwise alignment of networks IsoRank produce for accurate clusters but in case of multiple network alignment IsoRankN predicts most accurate clusters. The author claims that IsoRankN algorithm has higher coverage and consistency compared to other existing approaches, and thus helps in predicting accurate functional orthologs. The clusters formed by IsoRankN algorithm have much lower entropy. It handles noisy data as well as incomplete data.

(3) **Gradient ascent method.**
Zaslavskiy et.al [2009] also focused on how to align PPI networks using state-of-the-art graph matching methods. The main focus of authors here is Global Network Alignment of PPI networks.

The authors refer to no specific previous work.

The authors have proposed two new methods to measure the performance of two types of Global Network Alignment(GNA) i.e. Balanced and Constrained. The authors proposed new and simple Gradient Ascent(GA) Method for Balanced GNA and for the Constrained GNA problem the authors introduced an efficient algorithm based on the MP method. The authors tested their algorithms to align PPI networks of yeast and fly and hence comparing their performance with other GNA methods .The experiments were conducted to check how well different algorithms can align the given PPI networks with maximum interactions and whether more number of interactions help in finding more functional orthologs.

The authors state that the two new methods proposed for Global Network Alignment give better results than the popular method like IsoRank algorithm by giving 78 % more interactions for specified level of sequence similarity. The authors claim that the two new methods proposed provide better alignment of PPI in terms of getting more conserved interactions. Since Balance GNA problem can also be used to solve constrained GNA problem, the GA method proposed gives better results in both the cases. The MP method provides most favorable solution for constrained GNA.

2.2.4 *Summary*

Year	Author	Title of Paper	Major Contributions
2006	Flannick *et.al.*	Graemlin: General and robust alignment of multiple large interaction networks	Graemlin algorithm scores a possibly conserved module between different networks by computing the log-ratio of the probability that the module is subject to evolutionary constraints.It takes into account phylogenetic relationships between species whose networks are being aligned.
2008	Kalaev *et.al.*	Fast and accurate alignment of multiple protein networks	NetworkBlast-M algorithm uses a novel representation of multiple protein-protein interaction networks as a layered network alignment graph. It builds a set of n-tuple seeds with maximum score and then extends them greedily.
2008	Singh *et.al.*	Global alignment of multiple protein interaction networks	Focuses on the global network alignment problem which helps in controlling the relative weights of the protein sequences and network data in the alignment. Introduces the first known global alignment of protein-protein interaction networks of five species: yeast, fly, worm, mouse and human.
2009	Denielou *et.al.*	Multiple alignment of biological networks: A flexible approach	Uses a correspondence multigraph to extract connected components conserved in multiple networks. By avoiding the explicit construction of the network alignment multigraph, the algorithm deals with a large number of networks.
2009	Liao *et.al.*	IsoRankN: Spectral methods for global alignment of multiple protein networks	IsoRankN performs local and global alignments between multiple networks.Based on spectral clustering on the induced graph of pairwise alignment scores, the algorithm is both error-tolerant and computationally efficient.
2009	Zaslavskiy *et.al.*	Global alignment of protein-protein interaction networks by graph matching methods	Introduces a message-passing (MP) algorithm where matching pairs belong to clusters of proteins produced by the Inparanoid algorithm and the PPI networks of both species are not too dense.When matched pairs are not limited to the same clusters, use gradient-ascent(GA) method which computes the similarities between proteins.

3. CONCLUDING COMMENTS

Kelley et al. [2003] was the first one to develop a local network alignment algorithm called PathBlast. In this paper authors focused on exploring pathways that lead to high-scoring alignments of the two bimolecular networks. The scoring function is devised assuming that proteins in the pathways are homologous and that their interactions are not false positives. A few years later Singh et al. [2007] and Kuchaiev et al. [2010] also proposed network alignment algorithm addressing the same problem of aligning the two biomolecular networks. The advantage of the algorithms proposed in these research papers is that they globally align the two given networks. For aligning the networks the authors focus on the topological information. Singh et al. [2007] developed the first global alignment algorithm called the IsoRank. The authors proposed a method which aligns the two networks if and only if their neighbors can also be matched with each other. This algorithm is devised as an eigen vector problem and is based on greedy strategy. In scoring functions the authors used Blast similarity scores and used a user-defined parameter which has the command over the contribution of topology in the aligning two networks. The algorithm GRAAL(GraphAligner) discussed by Kuchaiev et al. [2010] relies solely on topology of the networks. It uses graphlet degrees which highly contributes in providing topological information of nodes. networks. GRAAL is a seed-and-extend approach and can align any type of network.

Later, Flannick et al. [2008] extended the alignment algorithms in aligning more than two biomolecular networks. The authors in this paper introduced a global network alignment algorithm called Graemlin for multiple networks. It can align both local as well as global networks. The global alignment is an equivalence relation over all nodes in a network whereas local alignment is a relation over a subset of nodes of a network. The equivalence relation is a transitive in nature and it partitions nodes into disjoint classes of orthologous proteins. The scoring function uses the features of networks representing several evolutionary events. Kalaev et.al [2008] modified PathBlast into NetworkBlast-M which is a multiple alignment algorithm. The NetworkBlast-M algorithm is used to find conserved protein complexes in multiple networks. It is based on forming a network alignment graph by finding the maximum weight node and then greedily finding adjacent maximum weight nodes.A year later based on the similar concept of devising a network alignment graph Denielou Y.P. [2009] proposed a new algorithm called C3-PartM. . The basic idea of creating a network graph is by finding connected components in a network then splitting into two networks of two different colors and further splitting the graphs until possible. Both NetworkBlast-M and C3-PartM give similar results. IsoRank has been extended to perform local and global alignments between multiple networks by Liao et al. [2009]. The authors of this paper present an extension of IsoRank called IsoRankN. The alignment is done by partitioning of the graph locally by computing pairwise functional similarity scores.MaWISh is a local alignment algorithm is based on the duplication/divergence models that focus on understanding the evolution of protein interactions; it constructs a weighted global alignment graph and tries to find a maximum induced subgraph in it (Koyuturk et al. [2006]). GA and MP algorithms given by Zaslavskiy et al. [2009] use the objective function which balances between matching similar pairs and increasing the

number of aligned interactions. They are based on relaxations of the cost function over the set of doubly stochastic matrices. Hence. this survey discusses various methods and approaches for aligning both pairwise and multiple biomolecular networks.

The future work identified by the researchers in this survey is the need for more extensive development of network alignment algorithm taking into account the evolution of protein interaction network which is becoming increasingly important as protein-protein interaction databases are continuously growing in size and species coverage.The researchers intend to explore more deeply the differences and similarities between the predicted protein functional orthologs and currently used ortholog lists.

4. ANNOTATIONS

4.1 Denielou Y.P. et. al. 2009

Citation. DENIELOU.Y.P. AND BOYER, F. AND VIARI, A. AND SAGOT, M.F.2009. Multiple alignment of biological networks: A flexible approach. *Lecture Notes in Computer Science, 2009,263-273.*

Problem. The problem that authors discuss in this paper is to find similarity among biomolecular network is in terms of similarity which is applied on aligned sub networks. While aligning multiple biomolecular networks the complexity of how to avoid explicit construction of network graph is faced. Various methods have been proposed to face this problem.

Previous Work. The authors refer to previous work by Kalaev. et al.[2008].

Shortcomings of Previous Work. The authors criticize Kalaev. et al's [2008] method for aligning networks which builds a set of n-tuple seeds with maximum score and then extends them greedily.They also criticize this approach for not being able to align more than three networks as when the correspondence relation S is not one-to-one, the size of the network alignment graph grows exponentially with both the number of vertices and of the edges.

New Idea /Algorithm. Denielou Y.P et al. [2009] introduce a new approach towards solving the problem of avoiding explicit construction of network graph and hence achieving an appropriate alignment among large number of networks. This approach is based on connected components conserved in multiple networks. The authors introduce a new algorithm C3Part-M which is an improvement over previous algorithm C3Part. The main purpose of this algorithm is to avoid the initial construction of the multigraph.The basic idea is to build a multigraph on the given biomolecular network. The multigraph is constructed by starting with the connected component on the primary graph and the expanding it to the second graph by splitting it into two colors and the further splitting on the third graph and so on. The previous algorithm followed the process of finding connectons in the network alignment multigraph by considering one single class of all vertices and hence partitioning at every step. Whereas the author enhanced the same process by first starting with one single class of vertices and then calculating the intersection of all connected components on all colors and continuing this process until partition

doesn't change.The authors introduce two new operations.1) SPLIT1-i – which is used to split a class on colors within range 1-i. 2).EXPAND i+1 – adds (i+1)th color to the current network alignment multigraph.

*Experiments and analysis conducted.*Denielou Y.P et al.[2009] compared their algorithm results with NetworkBlast algorithm on 10 microbial PPI networks. Both the algorithms were conducted on relaxed mode and the tree guided path mode.

Results. Deni´elou et al.[2009] claim that the C3Part-M algorithm proposed can be used for multiple networks. The authors observed that the results obtained from both the C3Part-M and NetworkBlast-M algorithm are quite similar.

4.2 Flannick et. al. 2006

Citation: FLANNICK, J. AND NOVAK, A. AND SRINIVASAN, B.S. AND MCADAMS, H.H. AND BATZOGLOU, S. 2006. Graemlin: general and robust alignment of multiple large interaction networks. Genome research, 16,9, 1169–1181

Problem. Flannick et.al [2006] addressed the problem of finding an optimal alignment for multiple networks. Also,conserved functional modules in protein-protein interaction networks were found by aligning these networks.

Previous Work. The authors refer to previous work by Kelly et.al [2003].

Shortcomings of Previous Work. The authors criticize the work by Kelly et.al [2003] as it does not address the problem of finding conserved modules of arbitrary topology within an arbitrary number of network.

New Idea /Algorithm. The authors proposed a new network alignment algorithm Græmlin for finding conserved functional modules and conserved pairwise interactions. The conserved modules are pairs of homologous proteins and the conserved interactions are the interactions between the two proteins. This algorithm can perform both global as well as local alignment. Also, Græmlin algorithm can be used to align multiple protein-protein interaction networks and ascertain the presence of a functional module by searching a large database of modules and networks. Græmlin addresses three network alignment problems 1) Formulation of mathematical function 2) devising a scoring function which encapsulates the knowledge of module evolution. 3) A search algorithm for identifying high scoring alignments in order to find conserved functional modules. Græmlin uses log odds scoring function and Græmlin's scoring function uses equivalence class concept which consists of proteins which are functionally orthologous and Edge Scoring Matrix(ESM) which allows log-odds scoring function to apply to alignment of pathways, protein complexes and general modules.

Experiments and analysis conducted: The authors analysed the ability of the algorithm to align know functional modules by comparing the results with two alignment algorithms NetworkBLAST and MaWISh .The algorithms were tested on a set of 10 microbial protein interaction networks. The sensitivity of each method and also how many proteins were correctly aligned in the pathway were observed.

Results.The authors state that Græmlin more sensitive than the other two method NetworkBlast and MaWish when looking for highly connected components.It is

faster than NetworkBlast with respect to running time that is Græmlin can efficiently search large protein networks. While performing query to network alignment Græmlin builds index for every query and saves it in the memory for future queries thus leading to fast alignment of many queries.

4.3 Kalaev et. al. 2008

Citation: KALAEV, M. AND BAFNA, V. AND SHARAN, R. 2008. Fast and accurate alignment of multiple protein networks. *Lecture Notes in Computer Science, 2008, 4955/2008, 246-256.*

Problem. According to the authors the main problem in aligning biomolecular networks is to find areas that are conserved in their interaction pattern and sequence among various biomolecular networks. Hence the authors in this paper design a model of aligning multiple networks solving the problem of finding conserved complexes and pathways.

Previous Work. The authors refer to no specific previous work.

Shortcomings of Previous Work. Kalaev et. al [2008] criticizes the various network alignment paradigms that have been applied successfully by a number of authors to search for conserved pathways and protein complexes, as their extension to more than three networks proved difficult due to the exponential growth of the alignment graph with the number of species.

New Idea /Algorithm. The authors introduce a new algorithm to align multiple biomolecular networks which helps in avoiding explicit representation of the network which in turn helps in reducing time complexity and memory utilization. The authors designed NetworkBlast-M which can be used for multiple network alignment. The algorithm is based on k-layer alignment graph concept where k is any constant and each layer represents a specie and contains corresponding network. The authors discuss two new concepts 1) Obtaining a set of d k-spines with identical topologies and maximum score which is based on the assumption that k-spines of the same seed share the same topology of inter-connections. 2) Obtaining set of d k-spines guided by underlying phylogeny with maximum score.

Experiments and analysis conducted. The authors tested their algorithm for eukaryotic and microbial PPI networks and compared the results with the NetworkBlast algorithm. The scoring function and the scoring parameters were the same for both algorithms. The results of the new algorithm were also compared with Graemlin on a set of 10 microbial networks.

Results. The authors state that the new algorithm has higher sensitivity than the NetworkBlast algorithm and its sensitivity further improves when using relaxed order variant. While running NetworkBlast-M on a given data, the time utilized to compute the result was less than thirty seconds in comparison to NetworkBlast which took about six hours to compute the result. Also NetworkBlast-M was better in performance when compared with Graemlin as the latter had lower specificity and sensitivity.The authors claim that the framework provided for multiple alignment of networks is fast and accurate and is better than previous progressive alignment

algorithms. NetworkBlast can only align up to 3 networks whereas NeworkBlast-M can align up to 10 microbial networks with improved time efficiency and less memory requirement. The authors observed that NetworkBlast-M works much faster in the case of restricted-order as compared to relaxed order.

4.4 Kelly et. al. 2003

Citation. KELLEY, B.P. AND SHARAN, R. AND KARP, R.M. AND SITTLER, T. AND ROOT, D.E. AND STOCKWELL, B.R. AND IDEKER, T 2003.Conserved pathways within bacteria and yeast as revealed by global protein network alignment. *Proceedings of the the National Academy of Sciences of the United States of America,100,20,11394-11399.*

Problem. The problem stated by the authors is to find a complexes and conserved pathways in protein-protein interaction network. Also, problem areas like finding conserved interaction pathway between bacteria v/s yeast , yeast v/s yeast and yeast v/s specific queries have been discussed.

Previous Work. The authors refer to no previous work.

Shortcomings of Previous Work. No shortcomings of previous work were mentioned.

New Idea /Algorithm. The authors propose a method called PathBLAST used for aligning protein networks.In this method the proteins of the first network are paired and aligned with the similar proteins called putative homologs occuring in the second network path where the path is obtained by searching for the highest scoring pathways alignment between two paths one from each network.The score of the pathway alignment is given on the basis of the quality and the extent of sequence similarity of the proteins in the given networks.Because of the variation in the two networks "gap" concept has been introduced in this method which is used when the proteins in one path are connected to proteins in the other path that are not directly linked with each other.

Experiments and analysis conducted. The authors performed a global alignment between the protein–protein interaction networks of yeast (S. cerevisiae) having 14,489 interactions among 4,688 yeast proteins and bacteria (H. pylori) having 1,465 interactions among 732 proteins.Th authors identified orthologous pathways between two species,paralogous pathways within a network of each specie.

*Results.*The authors state that PathBlast is a useful tool for predicting protein functions, revealing signaling crosstalk, and distinguishing true orthologs from among multiple proteins between species. Also, we paths having similar likelihood are more inclined to be more biologically significant and hence provide more detailed information about the proteins.It was observed that pathway gaps and mismatches detected larger regions of the network that were generally conserved even when there were no direct interactions.

4.5 Koyuturk et. al. 2006

Citation. KOYUTURK, M. AND KIM, Y. AND TOPKARA, U. AND SUBRAMANIAM, S. AND SZPANKOWSKI, W. AND GRAMA, A. 2006. Pairwise alignment

of protein- protein interaction networks. *Journal of Computational Biology, 13, 2,182-199.*

Problem. The problem stated by the authors is to to align protein-protein interaction networks of different species to comprehend the conserved pathways and protein complexes etc. across different species. The main challenge is to describe a measure of similarity between the different networks that captures underlying biological phenomena accurately.

Previous Work. The authors refer to no previous work.

Shortcomings of Previous Work. No shortcomings of previous work were mentioned.

New Idea /Algorithm. The authors propose a method called MaWish (Maximum Weight Induced Subgraph) for modeling the conservation and divergence of interactions as important design parameters, as well as the interpretation of resulting alignments, are important design parameters. The approach behind MaWish duplication/divergence models that focus on understanding the evolution of protein interactions between two protein-protein interaction networks. It is based on a mathematical model that extends the concepts of match, mismatch, and gap in sequence alignment to that of match, mismatch, and duplication in network alignment and evaluates the similarity between graph structures through a scoring function that accounts for evolutionary events. Considering the evolutionary models, it assists in interpretation of resulting alignments in terms of not only conservation but also divergence of modularity in PPI networks.

Experiments and analysis conducted. The authors align the PPI networks of two mammalians that are available in the database; Homo sapiens (Hsapi) and Mus musculus (Mmusc).The Hsapi PPI network contains 1369 interactions among 1065 proteins while Mmusc PPI network contains 286 interactions among 329 proteins.

Results. This paper presents a framework for local alignment of protein interaction networks. The framework is guided by theoretical models of evolution of these networks. The model is based on discovering sets of proteins that induce conserved subnets based on scoring match and mismatch of interactions, and duplication of proteins. An implementation of the proposed algorithm reveals that this framework is successful in uncovering conserved substructures in protein interaction data.The authors state that are currently working on a comprehensive implementation of the proposed framework which will enable more reliable assessment of statistical significance. Once these enhancements are completed, the proposed framework will be established as a tool for pairwise alignment of PPI networks, which will be publicly available through a web interface.

4.6 Kuchaiev et. al. 2010

Citation. KUCHAIEV, O., MILENKOVI_C, T., MEMI_SEVI_C, V., HAYES, W., AND PR_ZULJ, N. 2010. Topological network alignment uncovers biological function and phylogeny. *Journal of the Royal Society Interface 7, 50, 1341- 1354.*

Problem. The problem discussed by authors is devising an algorithm for aligning two networks on the basis on topology. The advantage is that with the help of

topology-based alignment we can extract both phylogeny and biological function information of the respective protein networks.

Previous Work. The authors refer to no previous work.

Shortcomings of Previous Work. No shortcomings of previous work were mentioned.

New Idea /Algorithm. The authors propose a new algorithm GRAAL (GraphAligner) that is solely based on topological similarity of the networks which are to be aligned. The algorithm does not take into account sequence similarity of protein in the PPI networks thus it can align any two networks not just biological ones. Both local as well as global alignment of networks is possible using this algorithm. Here nodes of the PPI networks being aligned are matched and paired on the basis of signature similarity. More the signature similarity between two nodes more is the topological similarity between their neighborhoods. Cost for aligning every node pair is calculated. Lower the cost more is the probability of the nodes in the given networks to be aligned. A parameter "a" helps in calculating the cost of the pair of nodes and controls the contribution of the node signature similarity to the cost function. If there exists a situation where two pairs have similar cost function then the choice is made randomly. Hence running the algorithm multiple times can result in different solutions. The algorithm is based on "seed and extend" approach. To begin with, a single node is chosen from each network to be aligned as a "seed". Then the alignment is extended around the seed radially.Hence, using greedy algorithm concept find the nodes with highest signature similarities around the seed. Three types of scores are defined to score GRAAL's algorithm. 1) Edge Correctness (EC). 2) Node Correctness (NC). 3) Interaction Correctness(IC).

Experiments and analysis conducted. The authors align the yeast S. cerevisiae PPI network which consists of 8,323 interactions and1004 proteins, with its 'noisy' counterparts obtained by: (1) random removal of nodes from the network; (2) random removal of edges from the network; (3) random addition of edges to the network; and (4) addition of low confidence PPIs to the network. The authors perform experiments for these four types on noise types by changing the percentages of noise types.

Results. The authors state that if the networks to be aligned are similar to each other then the algorithm proposed gives good alignment results. For 5% random node removal, we obtain NC, EC, and IC of about 70%, 90%, and 55%.For 5% random edge removal, we obtain NC, EC, and IC of about 65%, 75%, and 45%, respectively .For 5% of random edge addition, we obtain NC, EC, and IC of about 50%, 70%, and 60%, respectively). For 5% of low-confidence edge addition, we obtain NC, EC, and IC of about 70%, 90%, and 65%.

4.7 Liao C.S et. al. 2009

Citation. LIAO, C.S. AND LU, K. AND BAYM, M. AND SINGH, R. AND BERGER, B. 2009. IsoRankN: Spectral methods for global alignment of multiple protein networks. *Bioinformatics 25, 12,253-258.*

Problem. The problem addressed by authors is global alignment of two protein-protein interaction (PPI) networks and hence finding dense and clique-like clusters of proteins.

Previous Work. The authors refer to Flannick et al. (2008).

Shortcomings of Previous Work. No shortcomings of previous work were mentioned.

New Idea /Algorithm. The authors propose a new algorithm IsoRankN (IsoRank-Nibble) which is an extension of IsoRank algorithm defined by Singh et al. 2007.The first step of IsoRankN algorithm is similar to IsoRank but finds pairwise alignment scores between every pair of networks hence forming a k-partite graph. The second part is to find optimal k-partite matching. On the basis of scores clusters are formed using spectral partitioning methods. Since the information of protein networks is large and incomplete the algorithm uses star spread method to find similar cliques with maximum weight in order to align the multiple networks. The central idea of IsoRankN is to build a multiple network alignment by local partitioning of the graph of pairwise functional similarity scores.

Experiments and analysis conducted. The IsoRankN algorithm is compared with IsoRank and Græmlin 2.0. d .The algorithm is applied on five known eukaryotic PPI networks i.e. human, mouse, fly, worm, and yeast. The algorithm is tested on two factors 1) Consistency between proteins within a cluster 2) Coverage i.e. number of proteins assigned to a cluster.

*Results.*The authors state that IsoRankN produce more exact and higher number of clusters than previous existing techniques.It is consistent with respect to local alignment algorithms which are ambiguous, inconsistent and predict overlapping clusters.For pairwise alignment of networks IsoRank produce for accurate clusters but in case of multiple network alignment IsoRankN predicts most accurate clusters. The author claims that IsoRankN algorithm has higher coverage and consistency compared to other existing approaches, and thus helps in predicting accurate functional orthologs. The clusters formed by IsoRankN algorithm have much lower entropy.It handles noisy data as well as incomplete data.

4.8 Singh et. al. 2007

Citation. SINGH, R. AND XU, J. AND BERGER, B. 2007. Pairwise global alignment of protein interaction networks by matching neighborhood topology. *Research in computational molecular biology, Lecture Notes in Computer Science, 4453/2007, 16-31.*

Problem. The problem addressed by authors is global alignment of two protein-protein interaction (PPI) networks and finding maximum association between the edges and nodes of a network in order to get the best alignment of two networks.

Previous Work. The authors refer to no specific previous work.

Shortcomings of Previous Work. No shortcomings of previous work were mentioned.

New Idea /Algorithm. The authors have proposed a new algorithm IsoRank for pairwise alignment of two protein-protein interaction networks. For aligning the two networks the algorithm takes into consideration both PPI network data and sequence similarity data. The main idea of algorithm is that a node i in G1 is mapped to a node j in G2 if the neighborhood topologies of i and j are similar, i.e., the neighbors of i can be well-mapped to the neighbors of j.The input to the algorithm is two PPI networks and for similarity data, Blast similarity scores have been used.The output of the algorithm is a maximum common subgraph between two input networks.The algorithm consists of two main steps 1) Associate score r with each pair in both thye networks 2) Finding out high scoring pairs and mapping for alignment.

Experiments and analysis conducted. The authors tested their two-way global alignment algorithm on S. cerevisiae and D. melanogaster PPI networks. The algorithm applied to this pair of networks is used to identify the common subgraph of the given two networks. Experiments were performed to find set of proteins that behave similarly in the two PPI networks, evaluating algorithms error tolerance power, evaluating influence of α and comparing local and global alignment results.

Results. The authors state that the algorithm, IsoRank is tolerant of noise in the input. Also as the value of α increases, the importance of network data in the alignment process also increases, for maximum weight bipartite matching strategy. The algorithm also helps in predicting functional orthologs(FO) between the fly and yeast. The author claims that IsoRank algorithm is similar to Google's Page Rank algorithm which is based on the concept of ranking web-pages in the order of their "authoritativeness". The input of the algorithm is a pair of undirected weighted graphs and the output is an alignment. The algorithm can be used for both biological and non-biological data and has also been extended for aligning multiple networks.

4.9 Singh et.al 2008

Citation. SINGH, R. AND XU, J. AND BERGER, B. 2008. Global alignment of multiple protein interaction networks with application to functional orthology detection. *Pacific Symposium on Biocomputing 2008 ,303-314.*

Problem. The problem addressed by authors is global alignment of multiple protein-protein interaction (PPI) networks.For aligning networks it is important to find maximum matching in the given networks.Thus the authors in this paper discuss about the algorithm that tries to find an optimal aligment by maximizing the matching of nodes in the given networks.

Previous Work. The authors refer to no previous work.

Shortcomings of Previous Work. No shortcomings of previous work were mentioned.

New Idea /Algorithm. The authors introduce a new algorithm which works in two parts . First is to compute a k-partite graph of the given k input networks in which each of the k-parts contains node from each input network. The edges

mentioned are the intra-network edges i.e. edge eij denotes node i of network G1 is connected to node j of network G2.each edge is assigned an edge weight Rij that indicate the similarity between proteins. In the second stage of the algorithm, nodes are combined on the basis of k-partite matching to form clusters. The k-partite matching has transitive property i.e if i is matched to j and j is matched to l then it implies that i must be matched to l. All the nodes in cluster are mapped to each other in the corresponding global network alignment. The authors also introduce a new concept called functional coherence which is an programmed approach of predicting similarity between various functions of proteins It is used to calculate an aggregated score of proteins which is indicative of how well our algorithm works in comparison to other algorithm.

Experiments and analysis conducted. The algorithm is tested using PPI networks for five species: S. cerevisiae, D. melanogaster, C. elegans, M. musculus, and H. sapiens.the authors also evaluate functional orthologs (FO) across five species.

*Results.*The authors state that the functional orthologs predicted by the algorithm proposed have certain limitations.The algorithm lacks accuracy when dealing with noisy and incomplete data and hence may not be able to define accurate functionally related set of proteins. The correctness of these functional orthologs is similar to the current sequence-only functional orthologs. The author claims that the algorithm finds conserved functional components across the various species. It also helps us to predict functional orthologs between these five species.

4.10 Zaslavskiy et.al 2009

Citation. ZASLAVSKIY, M. AND BACH, F. AND VERT, J.P. 2009.Global alignment of protein–protein interaction networks by graph matching methods. *Oxford University Press , 25 , 12 ,259-267.*

Problem. The problem addressed by authors is how to align PPI networks using state-of-the-art graph matching methods. The main focus of authors here is Global Network Alignment of PPI networks.

Previous Work. The authors refer to no specific previous work.

Shortcomings of Previous Work. No shortcomings of previous work were mentioned.

New Idea /Algorithm. The authors propose two new methods to measure the performance of two types of Global Network Alignment(GNA) i.e. Balanced and Constrained. The authors proposed new and simple Gradient Ascent(GA) method for Balanced GNA and for the Constrained GNA problem the authors introduced an efficient algorithm based on the MP method.

Experiments and analysis conducted. The authors tested their algorithms to align PPI networks of yeast and fly and hence compared their performance with other GNA methods.The experiments were conducted to check how well different algorithms can align the given PPI networks with maximum interactions and whether more number of interactions help in finding more functional orthologs.

Results. The authors state that the two new methods proposed for Global Network Alignment give better results than the popular method IsoRank algorithm by giving 78 % more interactions for specified level of sequence similarity. The authors claim that the two new methods proposed provide better alignment of PPI in terms of getting more conserved interactions. Since Balance GNA problem can also be used to solve constrained GNA problem , the GA method proposed gives better results in both the cases.

REFERENCES

Chen, L., Wang, R., and Zhang, X. 2009. *Biomolecular Networks: Methods and Applications in Systems Biology.* John Wiley and Sons.

Deniélou, Y., Boyer, F., Viari, A., and Sagot, M. 2009. Multiple alignment of biological networks: A flexible approach. In *Combinatorial Pattern Matching.* Springer, 263–273.

Flannick, J., Novak, A., Srinivasan, B., McAdams, H., and Batzoglou, S. 2006. Graemlin: General and robust alignment of multiple large interaction networks. *Genome Research 16,* 9, 1169–1181.

Hsieh, M. and Sze, S. 2010. Graphlet alignment in protein interaction networks. In *Genomic Signal Processing and Statistics (GENSIPS), 2010 IEEE International Workshop.* IEEE, 1–4.

Kalaev, M. 2008. Algorithms and tools for the alignment of multiple protein networks. Ph.D. thesis, Tel-Aviv University.

Kalaev, M., Bafna, V., and Sharan, R. 2008. Fast and accurate alignment of multiple protein networks. In *Research in Computational Molecular Biology.* Springer-Verlag Berlin, Heidelberg 2008, 246–256.

Kalaev, M., Smoot, M., Ideker, T., and Sharan, R. 2008. Networkblast: Comparative analysis of protein networks. *Bioinformatics 24,* 4, 594–596.

Kelley, B., Sharan, R., Karp, R., Sittler, T., Root, D., Stockwell, B., and Ideker, T. 2003. Conserved pathways within bacteria and yeast as revealed by global protein network alignment. *Proceedings of the the National Academy of Sciences of the United States of America 100,* 20, 11394–11399.

Koyutürk, M., Grama, A., and Szpankowski, W. 2005. Pairwise local alignment of protein interaction networks guided by models of evolution. In *Research in Computational Molecular Biology.* Springer-Verlag Berlin, Heidelberg 2005, 48–65.

Koyutürk, M., Kim, Y., Topkara, U., Subramaniam, S., Szpankowski, W., and Grama, A. 2006. Pairwise alignment of protein interaction networks. *Journal of Computational Biology 13,* 2, 182–199.

Kuchaiev, O. 2010. Modeling and alignment of biological networks. Ph.D. thesis, University of California.

Kuchaiev, O., Milenković, T., Memišević, V., Hayes, W., and Pržulj, N. 2010. Topological network alignment uncovers biological function and phylogeny. *Journal of the Royal Society Interface 7,* 50, 1341–1354.

Liao, C., Lu, K., Baym, M., Singh, R., and Berger, B. 2009. Isorankn: Spectral methods for global alignment of multiple protein networks. *Bioinformatics 25,* 12, 253–258.

Milenković, T., Ng, W., Hayes, W., and Przulj, N. 2010. Optimal network alignment with graphlet degree vectors. *Cancer Informatics 9,* 121–137.

Singh, R., Xu, J., and Berger, B. 2007. Pairwise global alignment of protein interaction networks by matching neighborhood topology. In *Research in Computational Molecular Biology.* Lecture Notes in Computer Science, 2007, 16–31.

Singh, R., Xu, J., and Berger, B. 2008a. Global alignment of multiple protein interaction networks. In *Proc. Pacific Symp. Biocomputing.* Vol. 13. Citeseer, 303–314.

Singh, R., Xu, J., and Berger, B. 2008b. Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proceedings of the National Academy of Sciences 105,* 35, 12763–12768.

Tianut, W. and Samatova12, N. 2009. Pairwise alignment of interaction networks by fast identification of maximal conserved patterns. In *Pacific Symposium on Biocomputing 2009*. World Scientific Pub Co Inc, 99–110.

Towfic, F., Greenlee, M., and Honavar, V. 2009. Aligning biomolecular networks using modular graph kernels. *Algorithms in Bioinformatics*, 345–361.

Zaslavskiy, M., Bach, F., and Vert, J. 2009. Global alignment of protein–protein interaction networks by graph matching methods. *Bioinformatics 25*, 12, 259–267.