

Systems biology

Increasing confidence of protein interactomes using network topological metrics

Jin Chen^{1,2}, Wynne Hsu¹, Mong Li Lee¹ and See-Kiong Ng^{2,*}¹School of Computing, National University of Singapore, Singapore 119260 and ²Knowledge Discovery Department, Institute for Infocomm Research, Singapore 119613

Received on February 17, 2006; revised on May 18, 2006; accepted on June 12, 2006

Advance Access publication June 20, 2006

Associate Editor: Jonathan Wren

ABSTRACT

Motivation: Experimental limitations in high-throughput protein–protein interaction detection methods have resulted in low quality interaction datasets that contained sizable fractions of false positives and false negatives. Small-scale, focused experiments are then needed to complement the high-throughput methods to extract true protein interactions. However, the naturally vast interactomes would require much more scalable approaches.

Results: We describe a novel method called IRAP* as a computational complement for repurification of the highly erroneous experimentally derived protein interactomes. Our method involves an iterative process of removing interactions that are confidently identified as false positives and adding interactions detected as false negatives into the interactomes. Identification of both false positives and false negatives are performed in IRAP* using interaction confidence measures based on network topological metrics. Potential false positives are identified amongst the detected interactions as those with very low computed confidence values, while potential false negatives are discovered as the undetected interactions with high computed confidence values. Our results from applying IRAP* on large-scale interaction datasets generated by the popular yeast-two-hybrid assays for yeast, fruit fly and worm showed that the computationally repurified interaction datasets contained potentially lower fractions of false positive and false negative errors based on functional homogeneity.

Availability: The confidence indices for PPIs in yeast, fruit fly and worm as computed by our method can be found at our website <http://www.comp.nus.edu.sg/~chenjin/fpfn>

Contact: skng@i2r.a-star.edu.sg

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

Although recent progress in high-throughput experimental techniques (Uetz *et al.*, 2000; Ito *et al.*, 2000; von Mering *et al.*, 2002) has provided us with much more protein–protein interaction (PPI) data than those accumulated using traditional detection methods from the past decades, we are still far from being able to unravel the actual interactomes completely. This is because, while the new experimental methods have allowed PPIs to be detected *en masse*, it was done at the expense of data quality.

The PPI data generated by the high-throughput methods are by no means at the same level of quality as those that were painstakingly generated by conventional small-scale, focused experimental approaches. For example, recent surveys (Legrain *et al.*, 2001; von Mering *et al.*, 2002; Sprinzak *et al.*, 2003) have revealed that the interaction data detected by the popular yeast-two-hybrid (Y2H) assay may contain as much as 50% false positives. At the same time, the false negative rate of the Y2H-constructed interaction map for *Saccharomyces cerevisiae* interaction maps has also been estimated to be as high as 70% (Deng *et al.*, 2002). Such alarmingly high levels of errors in terms of both false positives and false negatives greatly diminish the potential usefulness of the experimental data made available by technology.

As a result, further carefully-focused small-scale experiments are often needed to complement the large-scale methods to validate the detected interactions. However, the vast interactomes require much more scalable and inexpensive approaches. In this work, we explore the possibility of computationally ‘repurifying’ the highly erroneous experimentally derived interaction datasets. We propose a computational method that uses only network topological metrics to sort the PPIs according to their computed reliability. Our proposed method can then identify false positives amongst the detected interactions that have low reliability values, and false negative from the putative interactions with high computed reliability values. By iteratively removing false positives from the interactomes and replacing them with interactions that are identified as false negatives, we will show that the computationally repurified PPI datasets contain potentially lower proportions of false positive and false negative errors than the original experimentally derived interactomes.

The rest of this paper is structured as follows. In Section 2, we describe the related work and motivation for our topological approach. In Section 3, we describe IRAP*, our iterative computational repurification method, and how it uses network topological metrics to identify false positives and false negatives in the experimentally detected interactomes. We report, in Section 4, evaluation results from applying IRAP* on actual large-scale interaction datasets generated by the popular Y2H assays for yeast, fruit fly and worm that show that the computationally repurified interaction datasets contain potentially lower fractions of false positive and false negative errors. We then conclude in Section 5 with a summary and discussions about possible further work.

*To whom correspondence should be addressed.

2 BACKGROUND

The potential benefits of the current technological advances in large-scale PPI detection have been largely diminished by the abundant presence of both false positives and negatives in the resulting experimental data. Other than resorting to small-scale, focused experiments to selectively validate the PPIs of interest in the interactomes, researchers have also begun to consider various bioinformatics approaches to identify the false negatives and false positives in the PPI datasets.

The problem of false negatives was typically addressed by PPI prediction methods. Researchers have proposed numerous approaches to predict PPIs using a variety of biological information from genome sequences (Tsoka and Ouzounis, 2002; Matthews *et al.*, 2001; Wojcik and Schachter, 2001) to 3D structures (Aloy *et al.*, 2004). However, these prediction methods depended on, and are limited by, the availability and richness of biological information; few have attempted to make use of the existing interactions in a PPI network to predict new interactions for the interactome. In this paper, we focus on the latter by directly discovering the false negatives amongst the missing interactions in the experimentally derived PPI networks, using only the topological information from the PPI networks.

In terms of false positive detection, a naive approach would be to use the intersection of PPI datasets derived from various independent methods to detect reliable interactions. However, because of the different and limited coverage of various PPI detection methods, this operation would leave only few interactions (Deane *et al.*, 2002) and it is therefore suitable only for identifying a small set of highly confident interactions. As such, some researchers have recently started to explore alternative approaches, such as the use of expected topological characteristics of PPI networks to assess the reliability of the experimentally detected interactions mathematically. Saito *et al.* pioneered such approach by developing a series of computational measures called interaction generalities (IG1 and IG2) (Saito *et al.*, 2002a,b) that used local topology and the statistics of adjacent interactions to detect false positives in a PPI network.

The IG1 measure was based on the simple observation that interacting proteins having many interacting partners that have no further interactions are likely to be false positives. This is a reasonable false positive model for Y2H data, as some so-called ‘sticky’ proteins in Y2H assays have a tendency to turn on the positive signals of the assay by themselves. In fact, we observe that this is also consistent with a real-world network model. In a real world network, neighbors are likely to also know one another—an interaction whose neighbors are all ‘lone rangers’ is therefore likely to be a false positive. Using this seemingly simplistic network-based heuristic, Saito *et al.* was able to create a confidence index for the detected interactions in a PPI network that was shown to correlate well with the probabilities of the interactions being true positives. The authors followed up their work by devising a more complex IG2 measure (Saito *et al.*, 2002a) that incorporated more complex topological properties to obtain better results.

Motivated by the success of interaction generalities, we have recently proposed an alternative topologically based quantitative measure called ‘Interaction Reliability by Alternative Path’ (IRAP) (Chen *et al.*, 2004, 2005). Instead of using small-sized, predefined network motifs such as those in the interaction

generalities, IRAP uses the observation that alternative paths are often present in many real-world networks (e.g., there is often more than one way to fly from one city to another in an airline network), and computes the reliability of a detected PPI with respect to the presence of alternative reliable interaction paths in the underlying network (see also Appendix A of Supplementary Materials for the statistics and biological examples of alternate paths in PPI networks). Our evaluation results reported in our previous works showed that IRAP outperformed the IG measures in the system-wide detection of false positives in the yeast interactome.

Both the problems of false positives and false negatives must be addressed together in order to improve the usability of experimentally derived interactomes. In this work, we extend our previous IRAP scoring method to detect both false positives and false negatives, using different topological metrics as the initial weights. We call the IRAP-based iterative approach for computational interactome repurification IRAP*, which stands for Interactome Repurification by Alternate Paths, with the asterisk indicating that it is an iterative process. We will show in this work that our approach works well in other species’ interactomes in addition to the commonly evaluated yeast interactome.

3 METHOD

Computationally, an experimentally derived protein interactome can be modeled using an undirected network $G = (V, E, w)$. Each node in the network represents a unique protein in the species’ proteome. An edge exists between two nodes v_A and v_B if there is a detected interaction between the corresponding proteins A and B . The PPI network is modelled as a weighted graph to account for the existence of experimental errors in the interactome, with $w(v_A, v_B)$ as the weight of edge (v_A, v_B) that serves also as a reliability index for the PPI.

In our previous work (Chen *et al.*, 2004, 2005), we used a quantifiable measure called IRAP to evaluate the reliability of a detected PPI with respect to the presence of a reliable alternative interaction path between the two proteins in the global PPI network. For completeness, we describe some key aspects of IRAP in detail in the next section.

3.1 IRAP: interaction reliability by alternative path

Given an edge (v_A, v_B) in network G , we use an alternative path selection strategy to nominate one of the non-reducible paths as its measure of reliability.

DEFINITION 3.1. Non-reducible Path. A path $\phi = v_1, \dots, v_n$ is a non-reducible path of edge (v_A, v_B) if we have $v_1 = v_A, v_n = v_B$ (or vice versa); and there is no shorter path ϕ' connecting node v_A and v_B that shares some common intermediate nodes with the path ϕ . That is, \nexists path $\phi' = u_1, \dots, u_m$ such that $(u_i, u_{i+1}) \in E, u_1 = v_A, u_m = v_B, u_r = v_s$ for some $r \in [2..m - 1], s \in [2..n - 1], m < n$.

To avoid selecting lengthy alternative paths with many intervening proteins that would be of relatively low value, IRAP also takes into consideration the length of the alternative paths connecting the two proteins. Whenever there is sharing of nodes, we use the shortest path to approximate the (biologically) strongest alternative path that connects the candidate interacting pair of proteins A and B in the interaction network.

Formally, IRAP is defined as follows:

DEFINITION 3.2. IRAP. The reliability of a candidate PPI (A, B) , IRAP (A, B) , is indicated by the collective reliability of the strongest

alternative path of interactions connecting the two proteins in the underlying PPI network.

$$IRAP(A, B) = \max_{\phi \in \Phi(A, B)} \prod_{(u, v) \in \phi} w(u, v) \quad (1)$$

where $w(u, v)$ denotes the weight value for edge (u, v) in the PPI network G ; $\Phi(A, B)$ denotes the set of non-reducible paths.

3.2 False positive detection

As mentioned earlier, some proteins in Y2H assays tend to activate transcription of the reporter gene without actually interacting with their partners, leading to an excess number of candidate partners (false positives) detected for the proteins (Serebriiskii and Golemis, 2001). The IG1 topological metric proposed by Saito et al. (2002b) was designed to detect these easily identified ‘sticky’ proteins. Since in this work, we are focusing on Y2H-derived experimental datasets (Y2H being the most popular high-throughput PPI detection method), for false positive detection, we use IG1 values as a basis for the PPI network’s initial weights:

$$IG1^G(v_A, v_B) = 1 + |\{(v'_A, v'_B) \in E \mid v'_A \in \{v_A, v_B\} \& \deg^G(v'_B) = 1\}| \quad (2)$$

$$w_+(v_A, v_B) = 1 - \left(\frac{IG1^G(v_A, v_B)}{IG1^G_{\max}} \right) \quad (3)$$

where $IG1^G_{\max}$ is the maximum IG1 value in the interaction network, and $w_+(v_A, v_B)$ is the initial weight for edge (v_A, v_B) .

Using (1) on the resulting interaction graph $G = (V, E, w_+)$, we compute an interaction reliability index for all the detected PPIs in the interactome. The IRAP scores for each of the edges in the interaction graph can be computed efficiently using the algorithm in our previous work (Chen et al., 2004). Those PPIs that are low in the IRAP reliability index can then be considered as potential false positives.

3.3 False negative detection

While IG1 was useful as initial weights for false positive detection, it turns out that a different topological metric is required for false negative detection. This is because IG1 tends to assume interaction links to be true unless there are topological evidence from the immediate neighbors to suggest otherwise. As such, while it was well-designed for detecting false positives in Y2H data, when used for false negative detection, IG1 will tend to overestimate the reliability for the missing links during the false negative detection process. For example, under IG1, all missing orphan links will be identified as false negatives since they will all have the lowest (strongest) IG1 values because there are no immediate neighbors to suggest otherwise. As such, for effective false negative detection, we would need a more stringent topological metric that assumes the missing links to be true negatives unless there are topological evidence to suggest otherwise.

Therefore, for false negative detection, we adopt a non-IG1 topological metric that is based on a different observation that interactions between proteins having a large number of common neighbors tend to be true interactions themselves. This was shown previously by Kelley and Ideker (2005) to be a useful measure for predicting interactions in genetic networks. Here, we apply such a common neighbor counting approach to compute the initial weight for edges (v_A, v_B) in a PPI network for false negative detection:

$$\text{ComNbr}^G(v_A, v_B) = |N(v_A) \cap N(v_B)| \quad (4)$$

$$w_-(v_A, v_B) = \frac{\text{ComNbr}^G(v_A, v_B)}{\text{ComNbr}^G_{\max}}, \quad (5)$$

where ComNbr^G_{\max} is the maximum number of common neighbors in the interaction network, and $w_-(v_A, v_B)$ is the initial weight for edge (v_A, v_B) .

We then use IRAP as defined in (1) on the resulting interaction graph $G = (V, E, w_-)$ to assign a refined reliability value to each candidate false negative. Those missing links having high IRAP values can then be treated as potential false negatives.

However, unlike false positive detection, a major challenge for false negative detection is the huge number of candidate false negatives. For instance, the size of the yeast proteome is ~ 6000 proteins and the current PPI network detected for it is ~ 10000 . The number of false negative candidates to be considered could be as many as $(6000 \times 5999)/2 - 10000 \approx 1.8 \times 10^7$. It is clearly impractical to evaluate the reliability of every possible link. Fortunately, as we are interested in only the top $(k/100) \times |E|$ false negatives at each iteration of computational repurification of the interactome with IRAP* (see Section 3.4 later), there is no need to consider all the false negative candidates. As such, we use the following heuristic approach to consider the most possible candidates first.

First, each node v_i in network $G = (V, E, w_-)$ is assigned a value h_i as:

$$h_i = 1 - \prod_{j \in N(v_i)} (1 - w_-(v_i, v_j)), \quad (6)$$

where $N(v_i)$ is the direct neighbor set of node v_i in network G . Then, the top Y proteins with the highest h values are pairwise joined to form a candidate set I . In a sparse network, where the number of the edges is linear to the number of the nodes, the action to add α new links is of the same scale to add new links to connect at most 2α nodes. We can thus define the Y as: $Y = 2 \times (k/100) \times |E|$. We can then compute the IRAP scores for those protein pairs that are in I but not in E .

The edges in I are good false negative candidates since their associated alternative paths are likely to have high IRAP values given the high neighbor edge weights of the nodes. This heuristic is effective in reducing the candidate space—in such a sparse network as the PPI network, the h value for a large proportion of nodes are 0.

3.4 IRAP*: iterative refinement of interactome

Our previous works (Chen et al., 2004, 2005) focused on using IRAP for false positive detection. In this work, we extend IRAP to also detect false negatives by using a different network topological metric as the initial scoring scheme. Next, we will describe IRAP*, an iterative application of IRAP for increasing the confidence of protein interactomes by computationally ‘repurifying’ the interactomes from its experimental errors. This is achieved by iteratively removing false positive and false negative interactions from the interactomes using IRAP. In each iteration, we remove $k\%$ of the interactome that are identified as very likely to be false positive interactions. Then, we add an equal number of new interactions that have been identified as potential false negatives into the interactome. In other words, each iteration consists of the following steps:

- *Step 1: False positive detection.* Compute the reliability index for all the interactions in $G = (V, E, w_+)$ using IRAP. Replace E with E' , the set of $[1 - (k/100)]|E|$ interactions that have the highest IRAP values.
- *Step 2: False negative detection.* Compute the IRAP-based reliability index for all the putative interactions in I derived from $G' = (V, E', w_-)$ as described above (Section 3.3), and determine the new edge set E'' , the set of $(k/100)|E|$ new interactions that have the highest IRAP values in G' .
- *Step 3:* Set $E = E' \cup E''$ for the next iteration.

A step-by-step example for IRAP* can be found in Appendix B of Supplementary Materials.

4 EVALUATION

We perform various evaluation experiments to ascertain that the application of IRAP* can improve the confidence of experimentally derived protein interactomes. First, we validate that IRAP is

effective in identifying false positives and false negatives. After that, we show that the iterative refinement process in IRAP* leads to increasingly better interactomes in terms of functional homogeneity of the protein partners in the interactions.

4.1 Datasets

We perform our evaluation on experimentally derived interaction data for three different species, namely *S.cerevisiae* (yeast), *Caenorhabditis elegans* (worm) and *Drosophila melanogaster* (fruit fly). We focus here only on interactomes that are derived by the popular high-throughput assays such as Y2H, which are well-known for their high experimental error rates. For *S.cerevisiae*, the experimental Y2H-derived interactome was downloaded from the BIND database GD *et al.* (2003) that comprises 7686 non-redundant Y2H interactions between 4141 of the yeast proteins. For *C.elegans*, we obtained an interaction network of 5025 non-redundant high-throughput PPIs between 2911 of the worm proteins from the BIND database. For *D.melanogaster*, we obtained a much larger interaction network of 22437 non-redundant high-throughput PPIs between 7621 of the fly proteins from the BIND database. The three PPI datasets used in this work can also be found from our website.

We evaluate the quality of the interactomes by the degree of cellular functional homogeneity amongst the interacting protein pairs. Under the oft-quoted ‘guilt-by-association’ principle (Oliver, 2000), we would expect that as the rate of true positive interactions increases in the resulting interactomes processed by IRAP*, the proportion of interacting proteins with a common functional roles should also increase correspondingly. We use the molecular functional annotations in Gene Ontology (GO) (<http://www.geneontology.org/>) for the functional annotation for the three genomes. The GO is one of the most important ontologies within the bioinformatics community. A total of 3554 out of the 4141 *S.cerevisiae* proteins, 2175 out of the 2911 *C.elegans* proteins and 6132 out of the 7621 *D.melanogaster* proteins were annotated under GO for our evaluation experiments.

4.2 GO term similarity measure

To compute the degree of functional homogeneity for any two proteins according to their GO molecular function annotations, we adopt an enriched GO similarity measure from (Lord *et al.*, 2003), which gives different weights to GO terms based on the GO term frequency in the target species, and the similarity values of different GO terms are determined by their shared parents, as follows:

$$\text{Similarity}(t_a, t_b) = \frac{2 \times \ln p(t_{ab})}{\ln p(t_a) + \ln p(t_b)} \quad (7)$$

Here t_a and t_b are any two GO terms, and t_{ab} is their common parent. The probabilities $p(t_a)$, $p(t_b)$ and $p(t_{ab})$ refer to the probabilities of the respective GO terms occurring in the target species based on their term frequencies.

4.3 False positive detection

Our recent work (Chen *et al.*, 2005) reported on the performance of IRAP for false positive detection on a combined set of *S.cerevisiae* interaction data that included interactions detected via both high-throughput and low-throughput methods. To better illustrate how IRAP can be a useful computational complement to

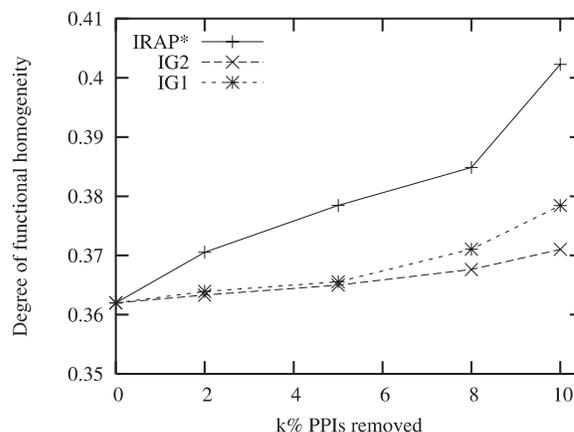


Fig. 1. Degree of functional homogeneity increases at different rates as potential false positives are removed from the yeast interactome under different interaction reliability measures.

current high-throughput experimental assays for PPIs, we report here the performance on the *S.cerevisiae* interaction data that were derived only using high-throughput Y2H experimental assays.

Recall that in each iteration of IRAP*, we seek to identify the most likely false positives amongst the worst $k\%$ of the interactions in terms of their IRAP values. We verify here that the interactions in the lower-end spectrum of the IRAP-indexed interactome indeed contained a larger proportion of biologically unlikely (i.e. functionally non-homogeneous) interactions than those indexed by other such topological metrics as IG1 and IG2. Figure 1 shows that IRAP is the best in detecting false positives—as more interactions that were detected as potential false positives were removed from the interactome, the degree of functional homogeneity in the resulting interactome increases at a faster rate than using other topological filtering metrics¹.

4.4 False negative detection

Similarly, for false negative detection, we verify whether the top new interactions proposed with IRAP exhibit a higher degree of functional homogeneity than those detected using other topological measures. Figure 2 shows that the new interactions proposed by IRAP were indeed of better quality than the corresponding sets proposed by IG1 and IG2².

4.5 Iterative refinement by IRAP*

Finally, we evaluate whether the iterative refinement process of IRAP* leads to incrementally improved interactomes. We perform two sets of experiments here. First, we verify whether true ‘gold standard’ interactions are retained in the interactome over the iterations, while the hidden true gold standard interactions

¹A detailed comparison of IRAP with the recently proposed PathRatio measure is also provided in Appendix C of Supplementary Materials, with results showing that IRAP performed comparably, if not slightly better than PathRatio. In the key spectrum of PPIs for false positive detection.

²Comparison with PathRatio was not feasible as the intensive computation required by PathRatio was prohibitive for false negative detection.

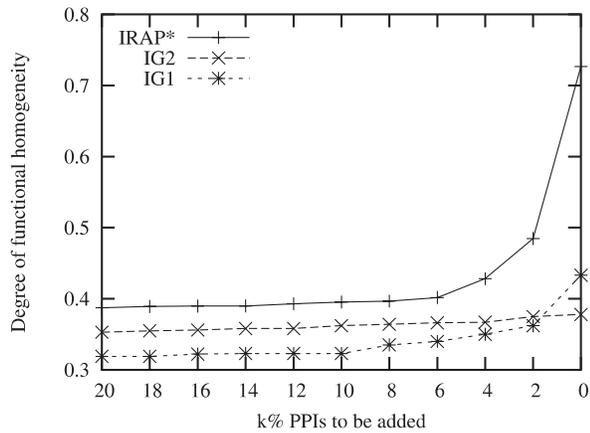


Fig. 2. Different degrees of functional homogeneity in the various proportions of potential false negative PPIs to be added to the yeast interactome under different interaction reliability measures.

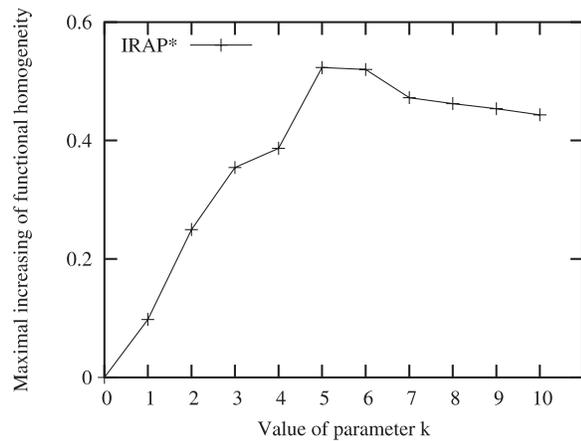


Fig. 3. Maximal increasing of functional homology in 15 iterations on the *S.cerevisiae* interactome varies with the parameter *k*.

are rediscovered. Then, we verify whether the degree of functional homogeneity of the repurified interactomes consistently improves over iterations. For this, we will present the results for interactomes of different species.

For evaluation, we ran IRAP* with parameter *k* for sufficient iterations to repurify the PPI networks while keeping the number of interactions unchanged in the procedure. To determine the value for parameter *k*, we tested with different values of *k* on the *S.cerevisiae* PPI network. Figure 3 shows the maximal similarity scores for GO molecular function annotations within 15 iterations for *k* = 1–15. We use *k* = 5 since it gives the best overall similarity score.

As there are currently no comparable measures that can detect both false positives and false negatives, we compared IRAP* with a similar iterative refinement process, which we will call IG1+ComNbr, that uses only those measures that were used in IRAP* as initial weights, namely, IG1 for false positive detection³

³We use IG1 instead of IG2 since it has comparable performance as IG2 for the Y2H dataset in the lowest 10% interactions (see Figure refigure:low) and IG2 is much slower to compute.

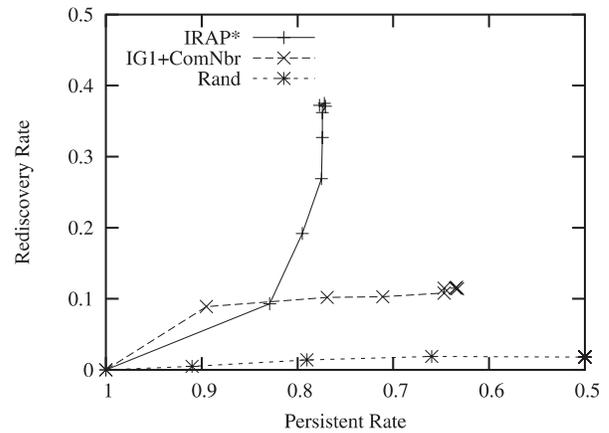


Fig. 4. Persistent and rediscovered rates for IRAP*, IG1+ComNbr and the baseline random process.

and ComNbr for false negative detection in each iteration. In other words, in each iteration of IG1+ComNbr, the initial weights of IG1 and ComNbr are directly used for filtering the false positives and false negatives without being further refined with IRAP. In this way, we show the effect of IRAP in the repurification process.

4.5.1 Persistence and rediscovery First, we are interested in finding out how well IRAP* retains true interactions and also how well it can rediscover those true interactions that were hidden from the experimentally detected interactome. We use the yeast PPIs in BIND that were indicated as having been obtained using low-throughput experimental methods as the set of ‘gold standard’ interactions here. Out of 1313 such low-throughput interactions found in BIND, 410 interactions were also found in our experimentally derived yeast interactome (using high-throughput Y2H assays). We randomly hide 50% of the 410 gold standard PPIs from the network, to see whether IRAP* could rediscover these lost interactions and retaining the other 50% in its refined interactome. This hiding and re-discovery process was repeated randomly for 100 times. Figure 4 shows the average results of the persistent and rediscovery rates for each iteration of IRAP*. In the final IRAP*-repurified interactome, an average of 139 out of 205 gold standard interactions were retained in the PPI network, where as 77 out of 205 hidden gold standard interactions were rediscovered. IG1+ComNbr performed significantly less well. After iterating for 10 times, only 102 gold standard interactions were still kept in the repurified interactomes, and a mere 21 hidden interactions were rediscovered. For reference, Figure 4 also shows the results of a baseline process that randomly added and removed PPIs from the network.

4.5.2 Functional homogeneity Finally, we check whether the quality in terms of functional homogeneity of the repurified interactomes improves over each iteration of IRAP*. To demonstrate IRAP*'s consistent performance across different kinds of interactomes, we show the results of IRAP* on three different species: *S.cerevisiae*, *C.elegans* and *D.melanogaster* in Figures 5–7, respectively. For reference, we also show the degree of functional

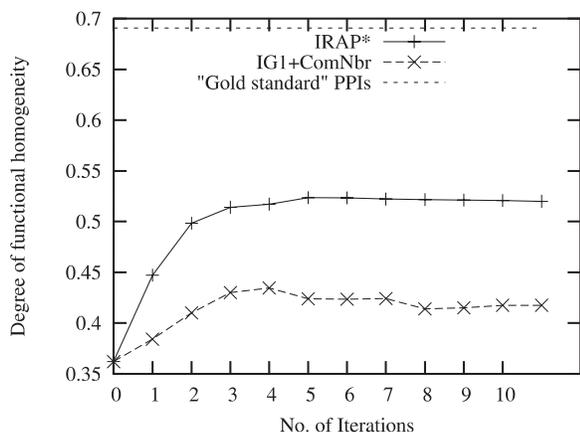


Fig. 5. PPI similarity score based on enriched GO terms increases at different rates with IRAP* and IG1+ComNbr on the *S.cerevisiae* interactome.

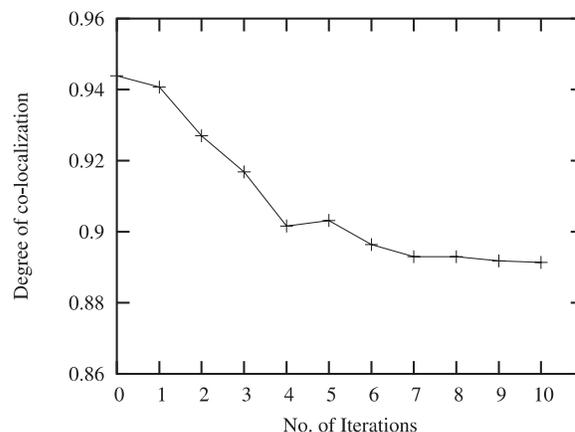


Fig. 8. Degree of co-localization decreases in each iteration.

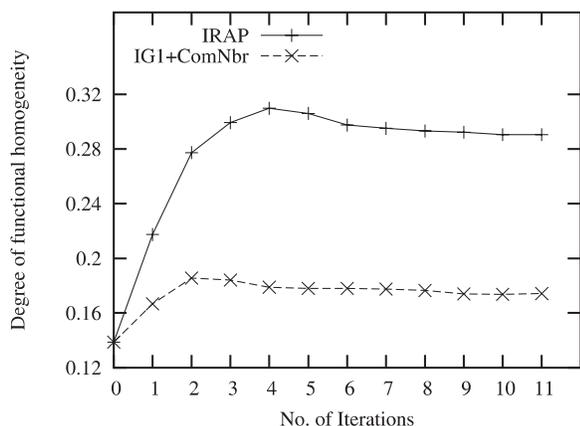


Fig. 6. PPI similarity score based on enriched GO terms increases at different rates with IRAP* and IG1+ComNbr on the *C.elegans* interactome.

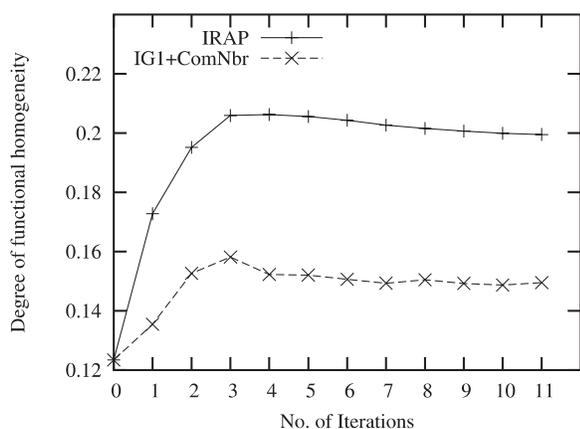


Fig. 7. PPI similarity score based on enriched GO terms increases at different rates with IRAP* and IG1+ComNbr on the *D.melanogaster* interactome.

homogeneity of the ‘gold standard’ yeast interactions used in Section 4.5 in Figure 5.

Figure 5 shows that with IRAP*, the average functional homogeneity score of the interacting *S.cerevisiae* proteins increases from the original 0.362–0.523 whereas with IG1+ComNbr, the similarity score in its repurified yeast interactome increases to only 0.435. For the sparser *C.elegans* interactome, the improvement with IRAP* was 0.139–0.310 while IG1+ComNbr obtained an increase to only 0.189. As for the much larger *D.melanogaster* interactome, IRAP* improved the degree of functional homogeneity quality from 0.123 to 0.206 whereas IG1+ComNbr could only increase it to 0.158. These results show that IRAP* is effective as a computational complement for the repurification of various experimentally derived interactomes with different densities and sizes. Note also that in all cases, only a small number of iterations was required since the quality of the repurified interactomes tend to converge in <10 iterations.

4.6 Cross-talkers

Thus far, we have evaluated the repurified interactomes based on the degree of functional homogeneity of the interactions. In the other related works (Saito *et al.*, 2002a,b; Pei and Zhang, 2005), the degree of cellular co-localization in the interactions was also used as a quality measure. However, as there are many important biological interactions (such as those in signaling pathways) that occur across cellular localizations, we have chosen not to use cellular co-localization in our evaluation.

In fact, we found that a number of the PPIs in our repurified interactomes that have been assigned high confidence values actually involved protein partners from different cellular localizations. Figure 8 shows that the degree of co-localization in the protein pairs detected by IRAP* decreases in each iteration until about 7–8 iterations. Under the cellular co-localization evaluation measure, these interactions would have been considered as errors. However, on inspecting the corresponding cellular functional roles of the non co-localizing protein interaction partners, we found that many of them actually shared the same function. For example, 32% of the non co-localized pairs discovered by IRAP* involved functionally homogeneous protein partners.

Protein A	Protein B	Com. Function	Localization A	Localization B
MCM1	CDC43	mitotic cell cycle and cell cycle control	nucleus	cytoplasm
FIG1	GPA1	pheromone response, mating-type determination, sex-specific proteins	integral membrane / endomembranes	plasma membrane
STE24	SLA2	incompatibility	mitochondrial matrix	actin cytoskeleton
END3	YCK2	cytokinesis (cell division) / septum formation	cytoskeleton	plasma membrane
MCM10	PRT1	mitotic cell cycle and cell cycle control	nucleus	cytoplasm
TPS2	SED1	stress response	cytoplasm	plasma membrane

Fig. 9. Examples of interactions between non co-localized proteins ('cross-talkers') that are involved in the same cellular pathways as discovered by IRAP*.

These 'cross-talkers' are thus likely to be actual biological interactions in pathways that occur across cellular sublocalizations. Figure 9 shows some examples of cross-talkers discovered by IRAP* in *S.cerevisiae* PPIs.

5 CONCLUSIONS

Much of the recent technological advance has been focused on the high-throughput detection of PPIs in order to map the tremendously vast protein interactome. Unfortunately, the potential benefits of these technological advances cannot be fully realized as the PPI data generated using the high-throughput technologies have very high error rates. In this paper, we have proposed a novel computational complement for the repurification of the experimentally derived interactomes. We iteratively refine an interactome by removing interactions that are identified as false positives and adding interactions detected as false negatives into the interactome. The computationally repurified interaction datasets were shown to contain potentially lower fractions of false positive and false negative errors. In addition, biologically interesting interactions such as cross-talkers may also be discovered using our method.

Note that in this work, the detection of the potential experimental errors was intentionally done using only the topological information that were mathematically derived from the underlying interaction graphs. This is to allow us to clearly illustrate the potential usefulness of such a topological approach. In practice, other biological information (e.g. gene expression correlation and even the functional and co-localization information that were used for evaluation here) should be incorporated to improve the quality of the repurified interactomes. As future work, it will be thus interesting to investigate how other topological measures, as well as additional biological information, can be incorporated for interactome repurification in practice.

ACKNOWLEDGEMENTS

The authors would like to thank Limsoon Wong for his advice on their work and Ho Nian Chua for letting them use his program for computing the PathRatio scores.

Conflict of Interest: none declared.

REFERENCES

- Aloy,P. et al. (2004) Structure-based assembly of protein complexes in yeast. *Science*, **303**, 2026–2029.
- Chen,J. et al. (2004) Systematic assessment of high-throughput experimental data for reliable protein interactions using network topology. In *Proceedings of International Conference on Tools with Artificial Intelligence*, Boca Raton, Florida pp. 368–372.
- Chen,J. et al. (2005) Towards discovering reliable protein interactions from high-throughput experimental data using network topology. *Artif. Intell. medi.*, **35**, 37–47.
- Deane,C.M. et al. (2002) Protein interactions: two methods for assessment of the reliability of high throughput observations. *Mol. Cell Proteomics*, **1**, 349–356.
- Deng,M. et al. (2002) Inferring domain-domain interactions from protein–protein interactions. *Genome Res.*, **12**, 1540–1548.
- GD,B. et al. (2003) Bind: the biomolecular interaction network database. *Nucleic Acids Res.*, **31**, 248–250.
- Ito,T. et al. (2000) Toward a protein–protein interaction map of the budding yeast: a comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc. Natl Acad. Sci. USA*, **97**, 1143–1147.
- Kelley,R. and Ideker,T. (2005) Systematic interpretation of genetic interactions using protein networks. *Nat. Biotechnol.*, **23**, 561–566.
- Legrain,P. et al. (2001) Protein–protein interaction maps: a lead towards cellular functions. *Trends Genet.*, **17**, 346–352.
- Lord,P. et al. (2003) Semantic similarity measures as tools for exploring the gene ontology. In *Proceedings of the Pacific Symposium on Biocomputing*, Lihue, Hawaii, pp. 601–612.
- Matthews,L.R. et al. (2001) Identification of potential interaction networks using sequence-based searches for conserved protein–protein interactions or 'interologs'. *Genome Res.*, **11**, 2120–6.
- Oliver,S. (2000) Proteomics: guilt-by-association goes global. *Nature*, **403**, 601–603.
- Pei,P. and Zhang,A. (2005) A topological measurement for weighted protein interaction network. In *Proceedings of Computational Systems Bioinformatics Conference*, Stanford, CA, pp. 268–278.
- Saito,R. et al. (2002) Construction of reliable protein–protein interaction networks with a new interaction generality measure. *Bioinformatics*, **19**, 756–763.
- Saito,R. et al. (2002) Interaction generality, a measurement to assess the reliability of a protein–protein interaction. *Nucleic Acids Res.*, **30**, 1163–1168.
- Serebriiskii,I. and Golemis,E. (2001) Two-hybrid system and false positives. approaches to detection and elimination. *Methods Mol. Biol.*, **177**, 123–134.
- Sprinzak,E. et al. (2003) How reliable are experimental protein–protein interaction data? *J. Mol. Biol.*, **327**, 919–923.
- Tsoka,S. and Ouzounis,C.A. (2000) Prediction of protein interactions: metabolic enzymes are frequently involved in gene fusion. *Nat. Genet.*, **26**, 141–142.
- Uetz,P. et al. (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature*, **403**, 623–627.
- von Mering,C. et al. (2002) Comparative assessment of largescale data sets of protein–protein interactions. *Nature*, **417**, 399–403.
- Wojcik,J. and Schachter,V. (2001) Protein–protein interaction map inference using interacting domain profile pairs. *Bioinformatics*, **17** (Suppl. 1), S296–305.