Motivation
With the recent increase in availability of biological data and improvements to biological models, biological network analysis has become a promising research area. An emerging technique to analyse biological networks is through network alignment. Network alignment has been used to calculate genetic distance [1,2], similarities between regulatory structures [3,4], the effect of external forces on gene expression [5], to depict conditional activity of expression modules in cancer [6], and to outline the defensive systems of rice to Xanthomonas oryzae pv. Oryzae [7]. Network alignment can be viewed as an extension of traditional sequence alignment, however performing network alignment is much more algorithmically complex than aligning linear sequences, therefore we must rely on heuristics, ideally as efficient and accurate as possible. The majority of current techniques [8-15] for network alignment rely on pre-computed information such as protein sequence alignment, or on tunable network alignment parameters, which may introduce an increased computational overhead.

Aims
• To define an alignment strategy that is able to quickly and accurately align two or more networks.
• To provide a simple method for comparing networks.

Significance
Current network analysis methods have been used to identify gene expression signatures of cancer, hypothesized to identify and validate drug targets, finding evolutionary paths, and identify condition-specific topological changes in organisms.

Method
Multiple experiments are performed to compare the performance of this algorithm (NF) with established alignment methodologies such as IsoRankN and MI-GRAAL as well as the latest alignment strategy (Ghost).

Experiment 1 - Saccharomyces cerevisiae
This experiment involves performing a pairwise alignment between a high confidence protein-protein interaction network of S. cerevisiae and networks that include interactions with progressively more interactions with lower experimental confidence.

Results

Experiment 1 - Saccharomyces cerevisiae

Node Correctness

Edge Correctness

Experiment 2 - Simulated networks

This experiment involves performing a pairwise alignment between a simulated biological network using GNLab. Twenty networks are generated with the default parameters. These networks are then modified by removing interactions randomly. The original networks are then aligned against their modified versions.

Fig 2. Saccharomyces cerevisiae – 20% Noise

Fig 3. Simulated network – Original

Fig 4. Simulated network – 20% Interactions removed

Challenges
• Large datasets pose a problem for most alignment strategies as the computational complexity of current algorithms is too high.
• Datasets contain errors.

Conclusion
I have described a novel approach to the network alignment problem. This approach is able to quickly and accurately align networks. We believe that this alignment strategy is the first practical method to align large biological networks.

Further work
• Use GPGPU techniques to further increase performance as this alignment strategy is embarrassingly parallel.
• This method is immediately extendable to multiple alignment.
• Devising a clear visualization of network alignments.
• Application on non-biological datasets such as social networks.

References