Antonina Mitrofanova, Martin Farach-Colton and Bud Mishra.

1 The Algorithm

In this section we give a more formal description of the algorithm.

Let $G = (P, E)$ be a *protein-protein interaction network*, where P is a set of proteins and there is an edge $e_{\alpha,\beta} \in E$ iff there exists Y2H interaction between proteins p_α and p_β.

Our algorithm will proceed by removing edges from G in phases, as described below. Let G^i be the graph obtained after i phases of our algorithm. We designate the jth connected component of G^i by $G^{i,j}$. Consider any weighted forest $T^i = (P, E_T, W_T)$ spanning the protein set P, where W_T represents max-flow values on the edges produced by the Gomory-Hu calculations on each connected component. Then T^i_k is the forest obtained from T^i by eliminating all edges of weight at most k. As before, $T^{i,j}_k$ is the jth connected component of T^i_k.

Here we summarize the algorithm. We denote the original Y2H PPI network as G^0. Let us define the procedure that takes X as an input and produces Y as an output as $X \longrightarrow Y$.

Let $i = 0$ and $k = min(w_{\alpha,\beta} : e_{\alpha,\beta} \in E_{T^i})$.

1. $\{G^{i,1}, \ldots G^{i,x}\} \longrightarrow \{T^{i,1}, \ldots, T^{i,x}\}$. /* During the ith phase, a Gomory-Hu tree is computed for each connected component in G^i. */
2. Let k_i be the minimum weight of an edge in T^i. $\{T^{i,1}, \ldots, T^{i,x}\} \longrightarrow \{T^{i,1}_k, \ldots, T^{i,x'}_k\}$. /* We eliminate all minimum-weight edges from the Gomory-Hu trees. */
3. *Output:* $P_{T^{i,j}}$ for $j = 1, \ldots, x'$.
4. $\{T^{i,1}_k, \ldots, T^{i,x'}_k\} \longrightarrow \{G^{i+1,1}, \ldots, G^{i+1,x'}\}$. /* Eliminating an edge from a Gomory-Hu tree corresponds to eliminating all edges of some cut in the graph. */
5. $i = i + 1$.
6. go to step 1, unless $E_{G^i} = \emptyset$.