

ALIGNMENTS OF A PAIR OF GRAPHS

Cesim Erten

Kaidr Has University, Turkey
cesimerten@gmail.com

The graph alignment problem has important applications in biological network alignment. In the case of protein-protein interaction networks for instance, undirected graphs G_1, G_2 correspond to PPI networks from a pair of species, where each of the vertex sets V_1, V_2 represent the sets of proteins, and E_1, E_2 represent respectively the sets of known protein interactions pertaining to the networks of species under consideration. The informal goal is to find a one-to-one mapping between V_1, V_2 that maximizes the conservation of interactions between the pairs of mappings. A specific version of the problem reduces the size of the problem by restricting the output alignment mappings to those chosen among certain subsets of protein mappings. The "constrained graph alignment" is considered a graph theoretical generalization of this biological network alignment problem.

Let $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$ be undirected graphs and S be a bipartite graph defined on (V_1, V_2) as the partition. For S , assume the degree of each vertex from the part V_i is at most m_i , for $i = 1, 2$. A *legal alignment* A is a matching of S . Let u_1u_2, v_1v_2 be a pair of edges of A such that $u_1, v_1 \in V_1$ and $u_2, v_2 \in V_2$. This pair of edges gives rise to a *conserved edge* if and only if $u_1v_1 \in E_1$ and $u_2v_2 \in E_2$. The constrained alignment problem is that of finding a legal alignment that maximizes the number of conserved edges. A 4-cycle x denoted with $c_4(x) = abcd$ is a cycle $a - b - c - d - a$ where $ab \in E_1, cd \in E_2$ and $ad, bc \in S$. We say that two c_4 s "conflict" if their edges from S cannot coexist in any legal alignment; the c_4 s contain at least one pair of S edges which cannot coexist in a matching of S . For a given $\langle G_1, G_2, S \rangle$ instance, we construct a "conflict graph", \mathcal{C} , as follows: For each c_4 create a vertex in \mathcal{C} and for each pair of conflicting c_4 s create an edge between their respective vertices in \mathcal{C} . With this construction, the constrained alignment problem obviously reduces to the maximum independent set problem. Let \mathcal{M} be an independent set of the conflict graph \mathcal{C} . The set of S edges included in the c_4 s corresponding to the vertices in \mathcal{M} constitute an optimum solution of the constrained alignment instance $\langle G_1, G_2, S \rangle$, that is a legal alignment with the maximum number of conserved edges, if and only if \mathcal{M} is a maximum independent set of \mathcal{C} . The following results apply to the case where $m_2 = 1$, that is each vertex of G_2 can be mapped to a single vertex of G_1 . Fertin et al. studied the constrained alignment problem under the same setting, that is $m_2 = 1$. Although they also define a conflict graph and provide a fixed-parameter tractability result based on a simple argument regarding the degrees of the conflict graphs, they do not provide any further structural properties. In what follows we extend their results and provide several graph-theoretic properties of conflict graphs arising from possible constrained alignment instances under various restrictions. Such properties are useful in applying relevant maximum independent set results. We skip all the proofs due to space restrictions.

Denote a k -cycle with C_k and denote the complement of G with \overline{G} . A "weakly triangulated" graph contains neither a C_k nor \overline{C}_k , for $k \geq 5$. The following theorem in connection with the "strong perfect graph theorem" of Chudnovsky et al. implies that the conflict graphs under the considered setting are perfect.

Theorem: If G_1 is acyclic and $m_2 = 1$ then \mathcal{C} is weakly triangulated.

Below we present graph theoretic properties of conflict graphs removing any restriction on G_1 or G_2 , but imposing a further constraint on m_1 .

Theorem: If $m_1 = 2$ and $m_2 = 1$, the wheel $W_k, k \geq 5$ is not an induced subgraph of a conflict graph.

We now generalize the previous results by providing further structural properties of conflict graphs for the more general case where m_1 can be any positive integer constant. We first present our result analogous to the previous theorem.

Theorem: If $m_2 = 1$, W_k is not an induced subgraph of \mathcal{C} , for $k \geq 7$.

Next we present our results regarding the existence of cliques as subgraphs of conflict graphs for any m_1 .

Theorem: If $m_2 = 1$, the maximum size of any clique in \mathcal{C} is m_1^2 .

The following result characterizes the conflict graphs that do not contain certain claws as induced subgraphs. A d -claw is an induced subgraph of an undirected graph, that consists of an independent set of d vertices, called *talons*, and the *center* vertex that is adjacent to all vertices in this set. Let $\Delta_{min} = \min(\Delta_1, \Delta_2)$, where Δ_1, Δ_2 are the maximum degrees of G_1 and G_2 respectively.

Theorem: If $m_2 = 1$, $(2\Delta_{min} + 2)$ -claw is not an induced subgraph of \mathcal{C} .

Joint work with Ferhat Alkan (University of Copenhagen, Denmark), Turker Biyikoglu (Izmir Institute of Technology, Turkey) and Marc Demange (RMIT University, Australia).