On Enumerating All Maximal Bicliques of Bipartite Graphs

Enver Kayaaslan

enver@cs.bilkent.edu.tr

Key words: bipartite graph, biclique, maximal biclique

1 Introduction

Enumerating all maximal bicliques of a bipartite graph has two main contributions to literature: First, we can use all maximal bicliques to find minimum number of bicliques covering a subset of edges or vertices for given bipartite graph. This problem is referred as Minimum Biclique Cover (MBC) Problem and dates back to the study of Orlin [5]. A recent work by Cornaz and Fonlupt deals with MBC problem for general graphs and useful references can be found therein [3]. Secondly, Agarwal et al. show that extracting all maximal bicliques can be used for compression [1].

In this work, we focus on finding all maximal bicliques in general bipartite graphs. In this point of view, this work can be considered as the specified version of study by Alexe et. al. [2] where the authors try to solve maximal biclique generation problem (MBGP) which is the problem of enumerating all maximal bicliques in simple graphs. Here, we present some theoretical results including a sufficient and necessary condition for a maximal biclique in a bipartite graph as well as a practical algorithm for generating all maximal bicliques of a given bipartite graph.

2 Theoritical Results

A bipartite graph $\mathcal{B} = (X, Y, \mathcal{E})$ is a graph, where the vertices can be divided into two disjoint sets X and Y such that every edge $e_{ij} \in \mathcal{E}$ connectes a vertex in X to one in Y. A biclique (S_x, S_y) of a bipartite graph $\mathcal{B}$ is a complete bipartite subgraph of $\mathcal{B}$ induced by vertex set $S_x \cup S_y$. The neighbor set $N_y(x_i)$ of a vertex $x_i \in X$ is defined as the set of vertices y_j such that there is an edge $e_{ij} \in \mathcal{E}$, i.e., $N_y(x_i) = \{y_j \in Y : (x_i, y_j) \in \mathcal{E}\}$. Similarly, $N_x(y_j) = \{x_i \in X : (x_i, y_j) \in \mathcal{E}\}$. The empty biclique $(\emptyset, \emptyset)$ is simply denoted as $\emptyset$.

Definition 1 (Maximal Biclique) *Given a bipartite graph $\mathcal{B} = (X, Y, \mathcal{E})$, a biclique (S_x, S_y) is a maximal biclique of $\mathcal{B}$ if no proper superset of (S_x, S_y) is a biclique, i.e., there exists no biclique $(S'_x, S'_y) \neq (S_x, S_y)$ such that $S_x \subseteq S'_x$ and $S_y \subseteq S'_y$.*

For the sake of simplicity, pair of sets $(S_x, \{y_j\})$ and $(\{x_i\}, S_y)$ are denoted as (S_x, y_j) and (x_i, S_y), respectively.

Definition 2 (Consensus Set) *Given a bipartite graph $\mathcal{B} = (X, Y, \mathcal{E})$, for a subset $S_x \subseteq X$ ($S_y \subseteq Y$) the consensus set $P_y(S_x)$ ($P_x(S_y)$) is defined as the intersection of neighbor set of each vertex $x_i \in S_x$ ($y_j \in S_y$),i.e,*

$$P_y(S_x) = \bigcap_{x_i \in S_x} N_y(x_i) \tag{1}$$

By definition, $P_x(\emptyset) = P_y(\emptyset) = \emptyset$.

Note that this is equaivalent to $P_y(S_x) = \{y_j \in Y : (S_x, y_j) \text{ is biclique}\}$ and similarly $P_x(S_y) = \{x_i \in X : (x_i, S_y) \text{ is biclique }\}$.

Theorem 1 *$(S_x, S_y) \neq \emptyset$ is a maximal biclique $\Leftrightarrow$ $S_y = P_y(S_x)$ and $S_x = P_x(S_y)$.*

Proof: Let $(S_x, S_y) \neq \emptyset$ be a biclique and so $S_x \subseteq P_x(S_y)$ and $S_y \subseteq P_y(S_x)$. By defintion of maximality, S_x, S_y is a maximal biclique if and only if there exists no $y_j \in Y/S_y$ such that (S_x, y_j) is a biclique and similarly there exists no $x_i \in X/S_x$ such that (x_i, S_y) is biclique. This is equivalent to that there exists no $y_j \in Y/S_y$ such that $y_j \in P_y(S_x)$ and there exists no $x_i \in X/S_x$ such that $x_i \in P_x(S_y)$. Equivalently $S_y = P_y(S_x)$ and $S_x = P_x(S_y)$. $\square$.

Lemma 1 *For any $S_y \subseteq Y$, $S_y \subseteq P_y(P_x(S_y))$. Similarly, for any $S_x \subseteq X$, $S_x \subseteq P_x(P_y(S_x))$.*

Proof: Consider a vertex $y_j \in S_y$. Since $(P_x(S_y), S_y)$ is biclique, $(P_x(S_y), y_j)$ is also biclique. Thus, $y_j \in P_y(P_x(S_y))$ which concludes that $S_y \subseteq P_y(P_x(S_y))$. Similar proof can be conducted for $S_x \subseteq P_x(P_y(S_x))$. $\square$

Theorem 2 *For any $S_y \subseteq Y$ such that $P_x(S_y) \neq \emptyset$, $(P_x(S_y), P_y(P_x(S_y)))$ is a maximal biclique. Similarly, for any $S_x \subseteq X$ such that $P_y(S_x) \neq \emptyset$, $(P_x(P_y(S_x)), P_y(S_x))$ is a maximal biclique.*

Proof: Let S_y be a subset of Y such that $P_x(S_y)$ and let $\tilde{S}_y$ denote $P_y(P_x(S_y))$. Consider a vertex $x_i \in P_x(\tilde{S}_y)$. Then, $(x_i, \tilde{S}_y)$ is biclique. Since $S_y \subseteq \tilde{S}_y$,

(x_i, S_y) is also biclique. Thus $x_i \in P_x(S_y)$ which concludes that $P_x(S_y) \supseteq P_x(\tilde{S}_y)$. Since $P_x(S_y) \subseteq P_x(\tilde{S}_y)$, $P_x(S_y) = P_x(\tilde{S}_y)$. Since both $P_y(P_x(S_y)) = \tilde{S}_y$ and $P_x(S_y) = P_x(\tilde{S}_y)$, $(P_x(S_y), \tilde{S}_y)$ is a maximal biclique. Similar proof can be conducted for maximality of biclique $(P_x(P_y(S_x)), P_y(S_x))$. □

3 Algorithm

Theoretical results suggest that it is sufficient to enumerate on the consensus sets, in order to find all maximal bicliques. Therefore we find all consensus X subsets of bipartite graph $\mathcal{B} = (X, Y, \mathcal{E})$. The algorithm requires only the bipartite graph $\mathcal{B}$. Then we initialize a set of consensus X subsets $\mathcal{S}$ with neighbor sets of every vertex $y_j \in Y$. Concurrently, we hold a priority queue Q which works in a FIFO manner, and it is also initialized by the same set of consensus X subsets. We iteratively grow the set $\mathcal{S}$ as follows. At each iteration, we select an unselected consensus set S_x from queue Q. For each vertex $y_j \in Y$ which is not in the consensus set of S_x, we construct a set S_{new} by the intersection of S_x and neighbor set $N(y_j)$. S_{new} corresponds to the consensus set of $P_y(S_x) \cup \{y_j\}$. We do not consider a vertex in $P_y(S_x)$, because for such a vertex, the intersection will result again S_x which wouldn't be a new consensus set in $\mathcal{S}$. If S_{new} is a new consensus set in $\mathcal{S}$, we insert S_{new} to $\mathcal{S}$. We also enqueue it to priority queue Q in order to expand new consensus sets based on S_{new}. The iterations terminate whenever there remains no consensus set to generate new ones. By the termination, we compute the maximal bicliques by taking pairs $(S_x, P_y(S_x))$ for each subset $S_x \in \mathcal{S}$.

Algorithm 1 FIND-ALL-MAXIMAL Algorithm

Require: Bipartite graph $\mathcal{B} = (X, Y, \mathcal{E})$
  $\mathcal{S} \leftarrow \{N_x(y_j) : y_j \in Y\}$
  $Q \leftarrow \mathcal{S}$
  while $Q \neq \emptyset$ **do**
    $S_x \leftarrow$ DEQUEUE(Q)
    for each $y_j \in Y / P_y(S_x)$ **do**
      $S_{new} \leftarrow S_x \cap N(y_j)$
      if $S_{new} \notin \mathcal{S}$ **then**
        $\mathcal{S} \leftarrow \mathcal{S} \cup \{S_{new}\}$
        ENQUEUE(Q, S_{new})
      end if
    end for
  end while
  $\mathcal{C}_{max} = \{(S_x, P_y(S_x)) : S_x \in \mathcal{S}\}$
  return $\mathcal{C}_{max}$

For each maximal biclique we check for at most $|Y|$ new bicliques. With a naive implementation, the checking procedure can be done in polynomial time on number of total maximal bicliques and number of vertices. Thus, the whole procedure runs in polynomial-time in total of input and output size with a naive implementation which conludes that FIND-ALL-MAXIMAL is a *total polynomial* algorithm [4] for the problem of enumerating all maximal bicliques of a bipartite graph.

4 Conclusion

In this work, we study the problem of enumerating all maximal bicliques in a given bipartite graphs. We devised a necessary and sufficient condition for a maximal biclique in bipartite graph which happened to be useful for enumeration. As well as, we constructed an algorithm for generating all maximal bicliques of a bipartite graph which runs in *total polynomial* time.

References

[1] P. K. Agarwal, N. Alon, B. Aranov, and S. Suri, *Can visibility graphs be represented compactly?*, Discrete and Computational Geometry, 12 (1994), pp. 347 – 365.

[2] G. Alexe, S. Alexe, Y. Crama, S. Foldes, P. L. Hammer, and B. Simeone, *Consensus algorithms for the generation of all maximal bicliques*, Discrete Applied Mathematics, 145 (2004), pp. 11 – 21. Graph Optimization IV.

[3] D. Cornaz and J. Fonlupt, *Chromatic characterization of biclique covers*, Discrete Mathematics, 306 (2006), pp. 495 – 507.

[4] D. S. Johnson, M. Yannakakis, and C. H. Papadimitriou, *On generating all maximal independent sets*, Information Processing Letters, 27 (1988), pp. 119 – 123.

[5] J. Orlin, *Contentment in graph theory: Covering graphs with cliques*, Proceedings of the Koninklijke Nederlandse, (1977), pp. 406–424.