The maximum edge biclique problem is NP-complete

René Peeters

Department of Econometrics and Operations Research

Tilburg University

The Netherlands

April 5, 2005

File No. DA5734

Running head: Maximum edge biclique problem

Correspondence:

René Peeters
Department of Econometrics and Operations Research
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands

E-mail: m.j.p.peeters@uvt.nl
Phone: +31-13-4663425
Fax: +31-13-4663280

Abstract

We prove that the maximum edge biclique problem in bipartite graphs is NP-complete.

Key words: Complexity, bipartite graphs, biclique

1 Introduction

Let $G = (V, E)$ be a graph with vertex set V and edge set E. A pair of two disjoint subsets A and B of V is called a *biclique* if $\{a, b\} \in E$ for all $a \in A$ and $b \in B$. Thus the edges $\{a, b\}$ form a complete bipartite subgraph of G (which is not necessarily an induced subgraph if G is not bipartite). A biclique $\{A, B\}$ clearly has $|A| + |B|$ vertices and $|A| * |B|$ edges. In this note we restrict ourselves to case when G is bipartite. The two colour classes of G will be denoted by V_1 and V_2, so $V = V_1 \cup V_2$.

Already in the book of Garey and Johnson [2, GT24] the complexity of deciding whether or not a bipartite graph contains a biclique of a certain size is discussed. If the requirement is that $|A| = |B| = K$ for some integer K (this is called the *balanced complete bipartite subgraph problem* or *balanced biclique problem*), then the problem is NP-complete. If however the requirement is that $|A| + |B| \geq K$ (the *maximum vertex biclique problem*), the problem can be solved in polynomial time via the matching algorithm. The complexity of the maximum vertex biclique problem for general graphs depends on the precise definition of a biclique in this case. With the above definition the problem is solvable in polynomial time since there is a one to one correspondence between bicliques in the bipartite double[1] of the graph and bicliques in the graph itself (see also [4]). If one defines a biclique as an induced complete bipartite subgraph (so A and B are independent sets

[1]The bipartite double of a graph with adjacency matrix A is the bipartite graph with adjacency matrix $\begin{bmatrix} O & A \\ A & O \end{bmatrix}$.

in G), then the maximum vertex biclique problem for general graphs is NP-complete (see [8]). A natural third variant is the so-called *maximum edge biclique problem (MBP)* where the requirement is that $|A| * |B| \geq K$. Up to now the complexity of this problem was still undecided.

In various papers the complexity of MBP is mentioned and guessed to be NP-complete ([1, 4, 3, 7]. In [1] some applications of MBP are discussed and it is shown that the weighted version of MBP is NP-complete. Furthermore the authors show that four variants of MBP are NP-complete. Using different techniques Hochbaum [4], Haemers [3] and Pasechnik [7] derive upper bounds for the maximum number of edges in a biclique. Hochbaum [4] presents a 2-approximation algorithm for the minimum number of edges needed to be removed so that the remainder is a biclique based on an LP-relaxation. Inspired by the work of Lovász on the Shannon capacity of a graph ([6]), Haemers [3] and Pasechnik [7] derive similar inequalities for the maximum biclique problem. Pasechnik uses semidefinite programming techniques wereas Haemers uses eigenvalue techniques.

In the next section we prove that indeed MBP is NP-complete. The reduction used is inspired by the reduction that is used to prove the NP-completeness of the balanced biclique problem (see [5]). As a consequence MBP is also NP-complete for general graphs.

2 The reduction

We define MBP as follows:

Maximum edge biclique problem (MBP): Given a bipartite graph $G = (V_1 \cup V_2, E)$ and a positive integer K, does G contain a biclique with at least K edges?

Theorem 1 *MBP is NP-complete.*

Proof: We shall reduce CLIQUE to MBP. This reduction is a modification of the reduction from CLIQUE to BALANCED COMPLETE BIPARTITE SUBGRAPH referred to in [2, GT24] and published in [5].

Let $G = (V, E)$ and K provide an instance of CLIQUE. Without loss of generality we may assume that $K = \frac{1}{2}|V|$.

Now construct an instance $G' = (V_1 \cup V_2, E')$, K' of MBP as follows: Let

$$\begin{aligned} V_1 &= V, \\ V_2 &= E \cup W, \\ &\quad \text{where } W \text{ is a set of } \tfrac{1}{2}K^2 - K \text{ new elements.} \\ E' &= \{\{v, e\} : v \in V; e \in E; v \notin e\} \cup \{\{v, w\} : v \in V; w \in W\} \\ K' &= K^3 - \frac{3}{2}K^2 \end{aligned}$$

This construction can clearly be performed in polynomial time. Suppose G has a clique C of size K. Take $A := V - C$ and $B := W \cup \{\{c, d\} : c, d \in C; c \neq d\}$. Then $\{A, B\}$ is a biclique in G' with $|A| * |B| = K * (\frac{1}{2}K^2 - K + \frac{1}{2}K(K-1)) = K^3 - \frac{3}{2}K^2$ edges. So if G has a clique of size K then G' has a biclique with K' edges.

Now suppose G has no clique of size K. Let $\{A, B\}$ be a biclique of G' with $A \subseteq V_1$ and $B \subseteq V_2$. We shall finish the proof by showing that $|A| * |B| < K'$ in this case. Without loss of generality $W \subseteq B$. Let $a := |A|$ and $b := |B| - |W|$.

The b elements of $B \cap E$ correspond with edges in G whose endpoints are not in A. There are $2K - a$ vertices of G that are not in A so $b \leq \frac{1}{2}(2K - a)(2K - a - 1)$, with equality if and only if $V - A$ is a clique with edge set $B \cap E$.

We consider two cases:

1. Suppose $a > K$, so $|V - A| = K - c$ with $c := a - K$ (So $0 < c \leq K$). Then $b \leq \frac{1}{2}(K - c)(K - c - 1)$, so

$$|A| * |B| \leq [K + c] * [\frac{1}{2}K^2 - K + \frac{1}{2}(K - c)(K - c - 1)]$$

   This reduces to

$$|A| * |B| - (K^3 - \frac{3}{2}K^2) \leq \frac{1}{2}c(c^2 - (K - 1)c - 2K)$$

   Now $c^2 - (K - 1)c - 2K$ is negative for $0 \leq c \leq K$, so $|A| * |B| < K'$ for $0 < c \leq K$.

2. Suppose $a \leq K$, so $|V - A| = K + c$ with $c := K - a$ (So $0 \leq c \leq K$). Since G has no cliques with K vertices, the number of edges in the subgraph of G induced by $V - A$, and consequently b, is strictly less than $\frac{1}{2}(K + c)(K + c - 1) - c$. This leads to

$$|A| * |B| < [K - c] * [\frac{1}{2}K^2 - K + \frac{1}{2}(K + c)(K + c - 1) - c]$$

which reduces to

$$|A| * |B| - (K^3 - \frac{3}{2}K^2) < \frac{1}{2}c^2(-c + 3 - K)$$

Since we may assume that $K \geq 4$, the right hand side is negative for $1 \leq c \leq K$ and zero for $c = 0$. So $|A| * |B| < K'$.

□

References

[1] M. Dawande, P. Keskinocak and S. Tayur, On the biclique problem in bipartite graphs, GSIA Working Paper 1996-04, Carnegie Mellon University, Pittsburgh, PA 15213, USA, 1996.

[2] M.R. Garey and D.S. Johnson, Computers and Intractability, A guide to the theory of NP-completeness (Freeman, San Francisco, 1979).

[3] W.H. Haemers, Bicliques and eigenvalues, J. Comb. Theory B 82 (2001) 56-66.

[4] D.S. Hochbaum, Approximating clique and biclique problems, J. Algorithms 29(1) (1998) 174-200.

[5] D.S. Johnson, The NP-completeness column: an ongoing guide, J. Algorithms 8(3) (1987) 438-448.

[6] L. Lovász, On the Shannon capacity of a graph, IEEE Trans. Inform. Theory 25 (1979) 1-7.

[7] D.V. Pasechnik, Bipartite sandwiches: bounding the size of a maximum biclique, preprint, 1999.

[8] M. Yannakakis, Node- and edge-deletion NP-complete problems, Proc. 10th Ann. ACM Symp. on Theory of Computing, Association for Computing Machinery, New York, (1978) 253-264.