A Tight Analysis of the Maximal Matching Heuristic

Jean Cardinal[1], Martine Labbé[1], Stefan Langerman[1,⋆], Eythan Levy[1,⋆⋆], and Hadrien Mélot[2]

[1] Université Libre de Bruxelles, Brussels, Belgium
{jcardin,mlabbe,slanger,elevy}@ulb.ac.be
[2] Université de Mons-Hainaut, Belgium
hadrien.melot@umh.ac.be

Abstract. We study the worst-case performance of the maximal matching heuristic applied to the Minimum Vertex Cover and Minimum Maximal Matching problems, through a careful analysis of tight examples. We show that the tight worst-case approximation ratio is asymptotic to $\min\{2, 1/(1-\sqrt{1-\epsilon})\}$ for graphs with an average degree at least ϵn and to $\min\{2, 1/\epsilon\}$ for graphs with a minimum degree at least ϵn.

1 Introduction

The maximal matching heuristic is a textbook algorithm that provides a 2-approximation for the Minimum Vertex Cover and Minimum Maximal Matching problems, two classical NP-hard problems [8]. It is perhaps one of the simplest and best-known approximation algorithms. It consists in finding a collection of disjoint edges (a matching) that is maximal (with respect to edge inclusion) by iteratively removing adjacent vertices until no more edges are left in the graph. Tightness of the 2-approximation is witnessed by a number of examples, for instance by the family of complete bipartite graphs in the case of Minimum Vertex Cover. This paper addresses the question of expressing the approximation ratio in a finer way, as a function of well-chosen graph parameters. We show that density parameters are good candidates for this purpose. Actually, the approximation ratio of the maximal matching heuristic is strictly less than 2 for graphs with a sufficiently high number of edges or sufficiently high minimum degree. We characterize precisely the asymptotic approximation ratio as a function of these parameters, together with tight examples. This is, to our knowledge, the tightest analysis ever done of this algorithm. This study shows that even simple heuristics might deserve nontrivial analyses. It was initiated using GraPHedron, a newly developed software for the investigation of relations between graph invariants (see [4] and [16]).

In the Minimum Vertex Cover problem, one is asked to find a minimum cardinality set of vertices that contains at least one endpoint of each edge of

⋆ Chercheur qualifié du F.N.R.S.
⋆⋆ Corresponding author.

L. Wang (Ed.): COCOON 2005, LNCS 3595, pp. 701–709, 2005.

the graph. One can easily see that the set of endpoints of a maximal matching indeed contains at least one endpoint of each edge, and that the optimal solution – since it must contain a least one vertex of each edge of the matching – has size not less than half the size of this set, hence the 2-approximation.

The minimum vertex cover problem is thus 2-approximable using the maximal matching heuristic, but no polynomial time algorithm with constant approximation ratio better than 2 is known. The problem is further known to be APX-complete [18] and not approximable within a factor of 7/6 [11]. Monien and Speckenmeyer [17] and Bar-Yehuda and Even [1] provide algorithms that achieves a ratio of $(2 - (\ln \ln n)/\ln n)$, where n is the number of vertices in the graph. Karakostas [14] later reduced the approximation ratio to $2-\Theta(1/\sqrt{\log n})$. For graphs with maximum degree Δ, Halperin [9] provides an aproximation algorithm with a ratio of $2-(1-o(1))2\ln\ln\Delta/\ln\Delta$. The problem has further been studied under the hypothesis that the input graph is dense. We say that a graph G is *weakly ϵ-dense* if its *average* degree is at least ϵn, i.e. if $m \geq \epsilon n^2/2$, with m being the number of edges in the graph, and *strongly ϵ-dense* if its *minimum* degre is at least ϵn. It has been shown [5] that the Minimum Vertex Cover problem restricted to strongly ϵ-dense graphs is APX-complete. Eremeev [7] shows that it is NP-hard to approximate the minimum vertex cover within a ratio less than $(7+\epsilon)/(6+2\epsilon)$ in strongly ϵ-dense graphs. Nagamochi and Ibaraki ([12]) provide an approximation algorithm with a ratio of $2 - 8m/(13n^2 + 8m)$, where m is the number of edges in the graph. This algorithm can also be seen as an approximation algorithm that achieves an approximation ratio that is asymptotic to $2-4\epsilon/(13+4\epsilon)$ for weakly ϵ-dense graphs. Karpinski and Zelikovsky [15] propose an algorithm that achieves a better ratio of $2/(2-\sqrt{1-\epsilon})$ for weakly ϵ-dense graphs, and a ratio of $2/(1+\epsilon)$ for strongly ϵ-dense graphs. Finally, Imamura and Iwama [13] recently proposed a randomized approximation algorithm which, with high probability, yields an approximation factor of $2/(1+\gamma(G))$, where $\gamma(G)$ is a function of the maximum and the average degree, and runs in polynomial time if Δ, the maximum degree of the graph, is $\Omega(n \log\log n/\log n)$.

In the Minimum Maximal Matching problem, one is asked to find a maximal matching of minimum cardinality, i.e. a minimum-cardinality set of disjoint edges that cannot be augmented. It is fairly easy to see that any maximal matching has a size that is at most twice the size of the minimum maximal matching. Much less is known about the minimum maximal matching problem than about the minimum vertex cover problem. Chlebìk and Chlebìková [3] do nevertheless show that it is NP-hard to approximate the problem within a constant factor better than 7/6.

Our Results. We study the worst-case approximation ratio of the maximal matching heuristic for the Minimum Vertex Cover and Minimum Maximal Matching problems in weakly and strongly ϵ-dense graphs. For both problems in weakly ϵ-dense graphs, we characterize the exact worst-case approximation ratio as a function of ϵ and obtain a function that is asymptotic to 2 when $\epsilon \leq 3/4$ and to $1/(1-\sqrt{1-\epsilon})$ otherwise. In the case of strongly ϵ-dense graphs, again we characterize the exact worst-case approximation ratio as a function of ϵ and obtain a function that is asymptotic to 2 when $\epsilon \leq 1/2$ and to $1/\epsilon$ otherwise. It is interesting to compare the approximation ratios we obtain for MINIMUM VERTEX COVER with the ones obtained by Zelikovsky and Karpinski: we note that in both the weakly and the strongly ϵ-dense cases the ratios differ only by one unit in the numerator and one in the denominator. The tight bounds we obtain for the MINIMUM VERTEX COVER problem are greater than those that were obtained using more sophisticated approximation algorithms [12, 15]. We nevertheless believe that a tight worst-case study of the classical heuristic is interesting as a point of comparison to the ratios obtained by other algorithms for the same problem, or by the same heuristic applied to other problems. On the other hand, the approximation ratios obtained for MINIMUM MAXIMAL MATCHING are, to the best of our knowledge, the best ones known for this problem since it does not seem to have been studied under density constraints yet. Finally, the results obtained for this problem are also valid for a variant problem, namely the MINIMUM EDGE DOMINATING SET problem since, as noted by by Yannakakis and Gavril [19], both problems always admit optimal solutions of the same size, and an optimal solution to one can always be transformed into an optimal solution to the other in polynomial time.

Section 2 is devoted to graph-theoretic preliminaries. In section 3 we study the worst-case aproximation ratio for the MINIMUM VERTEX COVER problem in weakly and strongly ϵ-dense graphs. The same kind of analysis is performed for MINIMUM MAXIMAL MATCHING in section 4. Full proofs are omitted through lack of space and can be found in [2].

2 Preliminaries

In the sequel we shall use the classical definition of a simple, loopless, undirected graph $G = (V, E)$, with vertex set V and edge set E. We denote by $\mathcal{G}_{n,m}$ the set of all non isomorphic graphs having n vertices and m edges. We use $n(G)$ to denote $|V|$, $m(G)$ to denote $|E|$ and $\delta(G)$ for the minimum degree of G. We also use the classical notions of *complete graph*, *empty graph*, *independent set*, *clique*, *complete bipartite graph* $K_{a,b}$, *matching*, *perfect matching* and *augmenting path*. Readers that are not familiar with these are referred to standard graph theory texts such as Diestel [6]. The *join* of two graphs G_1 and G_2 with vertex sets respectively V_1 and V_2 is the graph having $V_1 \cup V_2$ as vertex set and containing all edges of G_1, G_2, and all edges between vertices in V_1 and V_2. We denote by $\tau(G)$ the size of a minimum cardinality vertex cover of G, by $\nu(G)$ the size of a maximum cardinality matching of G, and by $\mu(G)$ the size of a minimum maximal matching of G.

We shall make extensive use of the following family of graphs, that arise as extremal graphs for several graph invariants (see [10] and [4]). A *complete split graph* $\Psi_{n,\alpha}$ with $1 \leq \alpha \leq n-1$, is a graph that can be decomposed in an independent set of size α and a clique of size $n - \alpha$, with each vertex of the independent set being adjacent to each vertex in the clique.

Our proofs make use of the following basic results on the values of invariants of complete split graphs.

Lemma 2.1. $m(\Psi_{n,\alpha}) = \binom{n}{2} - \binom{\alpha}{2} = (n-\alpha)(n+\alpha-1)/2.$

Lemma 2.2. $\nu(\Psi_{n,\alpha}) = \begin{cases} \lfloor \frac{n}{2} \rfloor & \text{if } \alpha \leq n/2, \\ n-\alpha & \text{otherwise.} \end{cases}$

Lemma 2.3. $\tau(\Psi_{n,\alpha}) = n - \alpha.$

Note that $\tau(\Psi_{n,\alpha}) = n - \alpha$ is obtained only by choosing all vertices in the clique as a vertex cover.

Lemma 2.4. $\mu(\Psi_{n,\alpha}) = \lceil \frac{n-\alpha}{2} \rceil.$

The following two simple properties shall also be useful.

Lemma 2.5. *For any graph* G*, we have* $\tau(G) \geq \delta(G)$*.*

Lemma 2.6. *For any graph* G*, we have* $\mu(G) \geq \left\lceil \frac{\delta(G)}{2} \right\rceil$*.*

3 Minimum Vertex Cover

We analyze the worst-case behavior of the maximal matching heuristic when applied to the Minimum Vertex Cover problem. We first consider weakly ϵ-dense graphs, which amounts to express the approximation ratio as a function of the number of edges.

3.1 Approximation Ratio vs Number of Edges

Lemma 3.1. *Let* n *and* m *be positive integers such that* $m(\Psi_{n,\alpha+1}) < m \leq m(\Psi_{n,\alpha})$ *for some* α*. The minimum value of* $\tau(G)$ *attained by a graph* G *in* $\mathcal{G}_{n,m}$ *is* $\tau(\Psi_{n,\alpha}) = n - \alpha$*.*

Sketch of the proof. The proof is in two steps: we first show by contradiction that a graph G in $\mathcal{G}_{n,m}$ cannot have $\tau(G) < \tau(\Psi_{n,\alpha})$, and second, by construction, that there exists a graph G in $\mathcal{G}_{n,m}$ having $\tau(G) = \tau(\Psi_{n,\alpha})$. □

Lemma 3.2. *Let* n *and* m *be positive integers such that* $m(\Psi_{n,\alpha+1}) < m \leq m(\Psi_{n,\alpha})$ *for some* α*. There exists a graph* G *in* $\mathcal{G}_{n,m}$ *such that* $\tau(G) = \tau(\Psi_{n,\alpha})$ *and* $\nu(G) = \nu(\Psi_{n,\alpha})$*.*

Sketch of the proof. A graph satisfying the required conditions is given by the removal of $m(\Psi_{n,\alpha}) - m$ edges from the edges joining the clique and the independent set of $\Psi_{n,\alpha}$. This graph can be shown to satisfy the required conditions by the use of classical tools, among wich Hall's condition on perfect matchings (see [6]). □

Using Lemmata 3.1 and 3.2 to maximize the numerator and minimize the denominator of the ratio, we obtain Theorem 3.1:

Theorem 3.1. *Let $\beta(G)$ be the worst-case approximation ratio for graph G. Let $\beta(m,n)$ be the worst approximation ratio attained by a graph in $\mathcal{G}_{n,m}$. We have:*

$$\beta(m,n) = \beta(\Psi_{n,\alpha^*(m,n)}) = \begin{cases} 2 & \text{if } \alpha^*(m,n) > n/2, \\ \frac{2\lfloor \frac{n}{2} \rfloor}{n-\alpha^*(m,n)} & \text{otherwise,} \end{cases}$$

where $\alpha^*(m,n) = \left\lfloor 1/2 + \sqrt{n(n-1)+1/4-2m} \right\rfloor$ is the integer value α such that $m(\Psi_{n,\alpha+1}) < m \leq m(\Psi_{n,\alpha})$.

The above theorem gives a tight upper bound on the approximation ratio of the maximal matching heuristic to the minimum vertex cover problem in graphs of n vertices and m edges, in the form of a discrete step function of m. The function equals 2 when $\alpha > n/2$ and begins to decrease afterwards.

Corollary 3.1. *Let $\tilde{\beta}(\epsilon, n)$ be the worst approximation ratio attained by a graph with n vertices and an average degree at least ϵn. We have:*

$$\lim_{n\to\infty} \tilde{\beta}(\epsilon, n) = \begin{cases} 2 & \text{if } \epsilon \leq 3/4, \\ \frac{1}{1-\sqrt{1-\epsilon}} & \text{otherwise .} \end{cases}$$

Sketch of the proof. The corollary follows from expressing $\tilde{\beta}(\epsilon, n)$ as $\beta(\lceil \epsilon n^2/2 \rceil, n)$ and studying the asymptotics of this expression. □

This asymptotic result is to be compared with the results of [12] and [15] quoted in the introduction (see figure 1).

3.2 Approximation Ratio vs Minimum Degree

Let $A_{n,\alpha}$ be the set of all graphs of minimum degree $n-\alpha$ that can be expressed as the join of an independent set of order α and a graph of order $n-\alpha$. Note that $A_{n,\alpha}$ contains $\Psi_{n,\alpha}$.

Lemma 3.3. *For all $G \in A_{n,\alpha}$ we have* $\nu(G) = \begin{cases} n-\alpha & \text{if } \alpha \geq n/2, \\ \lfloor \frac{n}{2} \rfloor & \text{otherwise.} \end{cases}$

Sketch of the proof. When $\alpha \geq n/2$, the result is straightforward. When $\alpha < n/2$, through a careful analysis of the configurations and degrees of unmatched vertices, we show that any non-perfect matching can be augmented. □

Lemma 3.4. *For all $G \in A_{n,\alpha}$ we have $\tau(G) = n - \alpha$. Furthermore, among all graphs with n vertices and minimum degree $n-\alpha$, this value of τ is minimal, and is attained only by graphs in $A_{n,\alpha}$.*

Sketch of the proof. The proof follows from a direct application of Lemma 2.5 and simple graph-theoretic arguments. □

Theorem 3.2 follows from Lemmata 3.3 and 3.4:

Theorem 3.2. *Let $\gamma(\delta, n)$ be the worst approximation ratio attained by a graph with n vertices and minimum degree δ. We have:*

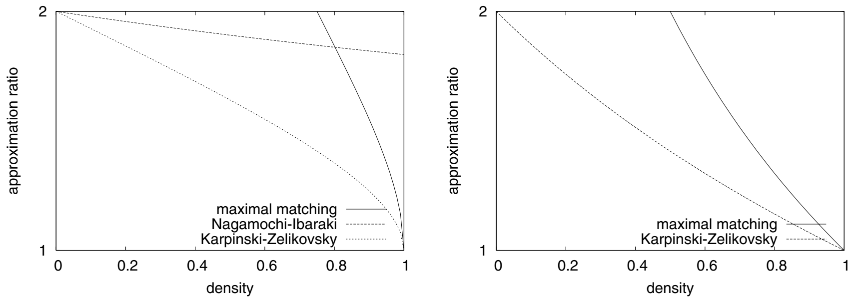
$$\gamma(\delta, n) = \begin{cases} 2 & \text{if } \delta \leq \frac{n}{2}, \\ \frac{2}{\delta}\lfloor \frac{n}{2} \rfloor & \text{otherwise.} \end{cases}$$

Furthermore, the only graphs that maximize the approximation ratio among all graphs with n vertices and a minimum degree of $n - \alpha$ when $\delta > n/2$ are in $A_{n,\alpha}$.

Corollary 3.2. *Let $\tilde{\gamma}(\epsilon, n)$ be the worst approximation ratio attained by a graph with n vertices and minimum degree at least ϵn. We have:*

$$\lim_{n \to \infty} \tilde{\gamma}(\epsilon, n) = \begin{cases} 2 & \text{if } \epsilon \leq 1/2, \\ \frac{1}{\epsilon} & \text{otherwise .} \end{cases}$$

This asymptotic result is again to be compared with the result of [15] quoted in the introduction (see figure 1).



(a) weakly ϵ-dense (b) strongly ϵ-dense

Fig. 1. A comparison of the approximation ratios for Minimum Vertex Cover.

4 Minimum Maximal Matching

In this section we show that the analysis proposed for Minimum Vertex Cover can be performed for Minimum Maximal Matching as well.

4.1 Approximation Ratio vs Number of Edges

The proof of the following lemma is straightforward and omitted.

Lemma 4.1. *For any fixed k, the only graph G with n vertices that maximizes the number of edges among all graphs having $\mu(G) = k$ is $\Psi_{n,n-2k}$.*

Lemma 4.2. *Let n and m be positive integers such that $m(\Psi_{n,\alpha+1}) < m \leq m(\Psi_{n,\alpha})$ for some α. The minimum value for $\mu(G)$ attained by a graph G in $\mathcal{G}_{n,m}$ is $\mu(\Psi_{n,\alpha}) = \lceil \frac{n-\alpha}{2} \rceil$.*

Sketch of the proof. Using arguments similar to those used in the proofs of Lemmata 3.1 and 3.2, we first show, using Lemma 4.1, that a graph in $\mathcal{G}_{n,m}$ cannot have a maximal matching of size less than $\mu(\Psi_{n,\alpha})$, and then that the value $\mu(\Psi_{n,\alpha})$ is indeed attained by some graph in $\mathcal{G}_{n,m}$. □

Lemma 4.3. *Let n and m be positive integers such that $m(\Psi_{n,\alpha+1}) < m \leq m(\Psi_{n,\alpha})$ for some α. There exists a graph G in $\mathcal{G}_{n,m}$ such that $\mu(G) = \mu(\Psi_{n,\alpha})$ and $\nu(G) = \nu(\Psi_{n,\alpha})$.*

Sketch of the proof. The graphs described in the proof of Lemma 3.2 can easily be shown to satisfy $\mu(G) = \mu(\Psi_{n,\alpha})$ and $\nu(G) = \nu(\Psi_{n,\alpha})$. □

Using Lemmata 4.2 and 4.3 to maximize the numerator and minimize the denominator of the ratio, we obtain Theorem 4.1:

Theorem 4.1. *Let $\rho(G)$ be the worst approximation ratio for graph G. Let $\rho(m,n)$ be the worst approximation ratio attained by a graph in $\mathcal{G}_{n,m}$. For each n, m we have:*

$$\rho(m,n) = \rho(\Psi_{n,\alpha^*(m,n)}) = \begin{cases} 2 & \text{if } \alpha^*(m,n) > n/2+1, \\ \frac{\lfloor \frac{n}{2} \rfloor}{\left\lceil \frac{n-\alpha^*(m,n)}{2} \right\rceil} & \text{otherwise.} \end{cases}$$

Corollary 4.1. *Let $\tilde{\rho}(\epsilon, n)$ be the worst approximation ratio attained by a graph with n vertices and an average degree at least ϵn. We have:*

$$\lim_{n\to\infty} \tilde{\rho}(\epsilon, n) = \begin{cases} 2 & \text{if } \epsilon \leq 3/4, \\ \frac{1}{1-\sqrt{1-\epsilon}} & \text{otherwise.} \end{cases}$$

4.2 Approximation Ratio vs Minimum Degree

Let $B_{n,\delta}$ be the set of graphs of order n having minimum degree δ and a maximal matching of size $\lceil \delta/2 \rceil$. Note that $B_{n,\delta} = A_{n,n-\delta}$ when δ is even.

Lemma 4.4. *Each graph in $B_{n,\delta}$ with $\lceil \delta/2 \rceil > n/4$ has a perfect matching.*

Sketch of the proof. As in the proof of Lemma 3.3, we show that any non-perfect matching in our graph can be augmented, using a careful analysis of the configuration and degrees of the vertices. □

Theorem 4.2 follows directly from Lemmata 2.6 and 4.4:

Theorem 4.2. *Let $\sigma(\delta, n)$ be the worst approximation ratio attained by a graph with n vertices and minimum degree δ. We have:*

$$\sigma(\delta, n) = \begin{cases} 2 & \text{if } \lceil \delta/2 \rceil \leq n/4 \\ \frac{\lfloor n/2 \rfloor}{\lceil \delta/2 \rceil} & \text{otherwise.} \end{cases}$$

Furthermore, when $\lceil \delta/2 \rceil > n/4$, $B_{n,\delta}$ is the exact set of graphs that maximize the ratio among all graphs with n vertices and minimum degree δ.

Corollary 4.2. *Let $\tilde{\sigma}(\epsilon, n)$ be the worst approximation ratio attained by a graph with n vertices and minimum degree at least ϵn. We have:*

$$\lim_{n\to\infty} \tilde{\sigma}(\epsilon, n) = \begin{cases} 2 & \text{if } \epsilon \leq 1/2, \\ \frac{1}{\epsilon} & \text{otherwise.} \end{cases}$$

References

1. R. Bar-Yehuda and S.Even. A local-ratio theorem for approximating the weighted vertex cover problem. *Ann. Discrete Math.*, 25:27–45, 1985.
2. J. Cardinal, M. Labbé, S. Langerman, E. Levy, and H. Mélot. A tight analysis of the maximal matching heuristic. Technical Report 545, Université Libre de Bruxelles, 2005. Available at http://www.ulb.ac.be/di/publications.
3. M. Chlebìk and J. Chlebìková. Approximation hardness of minimum edge dominating set and minimum maximal matching. In *Proceedings of 14th International Symposium on Algorithms and Computation (ISAAC)*, Lecture Notes in Computer Science. Springer–Verlag, 2003.
4. J. Christophe, S. Dewez, J. Doignon, S. Elloumi, G. Fasbender, P. Grégoire, D. Huygens, M. Labbé, H. Mélot, and H. Yaman. Linear inequalities among graph invariants: using GraPHedron to uncover optimal relationships. Submitted. Available at http://www.optimization-online.org/DB_HTML/2004/09/964.html, 2004.
5. A. Clementi and L. Trevisan. Improved non-approximability results for vertex cover problems with density constraints. *Theoretical Computer Science*, 225(1–2):113–128, 1999.
6. R. Diestel. *Graph Theory, Second Edition*. Springer-Verlag, 2000.
7. A.V. Eremeev. On some approximation algorithms for dense vertex cover problem. In *Proceedings of SOR*. Springer–Verlag, 1999.
8. M. R. Garey and D. S. Johnson. *Computers and intractability. A guide to the theory of NP-completeness*. Freeman and Company, 1979.
9. E. Halperin. Improved approximation algorithms for the vertex cover problem in graphs and hypergraphs. *Siam Journal on Computing*, 31:1608–1623, 2002.
10. P. Hansen and H. Mélot. Variable neighborhood search for extremal graphs 9. Bounding the irregularity of a graph. To appear in S. Fajtlowicz *et al.* (Eds.), Graphs and Discovery, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, (forthcoming), Providence, American Mathematical Society.
11. J. Hastad. Some optimal inapproximability results. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 1–10, 1997.
12. T. Ibaraki and H. Nagamochi. An approximation of the minimum vertex cover in a graph. *Japan J. Indust. Appl. Math.*, 16:369–375, 1999.
13. T. Imamura and K. Iwama. Approximating vertex cover on dense graphs. In *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2005.
14. G. Karakostas. A better approximation ratio for the vertex cover problem. *ECCC Report TR04-084*, 2004.
15. M. Karpinski and A. Zelikovsky. Approximating dense cases of covering problems. In P. Pardalos and D. Du, editors, *Proceedings of the DIMACS Workshop on Network Design: Connectivity and Facilites Location*, volume 40 of *DIMACS series in Disc. Math. and Theor. Comp. Sci.*, pages 169–178, 1997.
16. H. Mélot. Facets defining inequalities among graph invariants: the system GraPHedron. In preparation.

17. B. Monien and E. Speckenmeyer. Ramsey numbers and an approximation algorithm for the vertex cover problem. *Acta Inf.*, 22:115–123, 1985.
18. C. H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *J. Comput. System. Sci.*, 43(3):425–440, 1991.
19. M. Yannakakis and F. Gavril. Edge dominating sets in graphs. *SIAM J. Appl. Math.*, 38(3):364–372, 1980.