Approximating the Minimum Maximal Independence Number

Magnús M. Halldórsson

School of Information Science
Japan Advanced Institute of Science and Technology
Tatsunokuchi, Ishikawa 923-12, Japan
magnus@jaist-east.ac.jp

Abstract

We consider the problem of approximating the size of a minimum non-extendible independent set of a graph, also known as the minimum dominating independence number. We strengthen a result of Irving [2] to show that there is no constant $\epsilon > 0$ for which this problem can be approximated within a factor of $n^{1-\epsilon}$ in polynomial time, unless $\mathcal{P} = \mathcal{NP}$. This is the strongest lower bound we are aware of for polynomial-time approximation of an unweighted $\mathcal{NP}$-complete graph problem.

Keywords: Combinatorial problems, Approximation algorithms.

1 Introduction

A *maximal independent set* in an undirected graph is a set of mutually non-adjacent vertices such that the introduction of an additional vertex destroys the non-adjacency property. The *minimum maximal independence number* (MMIN) of a graph is the size of the smallest such set. It is also commonly known as the size of a minimum independent dominating set. Determining this number for arbitrary graphs is known to be $\mathcal{NP}$-complete.

In this paper, we show that assuming $\mathcal{P} \neq \mathcal{NP}$, MMIN can not be approximated in polynomial time within a factor of $n^{1-\epsilon}$ for any $\epsilon > 0$, where n is the number of vertices. In fact, we obtain a spectrum of results, with tradeoffs between assumptions on the solvability of the satisfiability problem and lower bounds on the approximation of MMIN.

This work is a followup to a paper of Irving [2], who shown that MMIN could not be approximated within any *constant* factor (assuming $\mathcal{P} \neq \mathcal{NP}$). Our basic result is a simple generalization of his construction. On a related note, Kann [3] has since showed that the MMIN problem is complete for the class of polynomial-bounded minimization problems, implying, among other things, a somewhat weaker lower bound on the approximability of MMIN of n^{δ}, for *some* constant $\delta > 0$.

1.1 Notation

For an algorithm (or a function) **A**, let $A(G)$ denote the number output on input G, representing the approximation to $MMIN(G)$, the minimum maximal independence number of G. We say algorithm **A** approximates the MMIN function within a factor of $s(n)$, if, for all n and every graph G on n vertices,

$$1 \leq \frac{\mathbf{A}(G)}{MMIN(G)} \leq s(n)$$

Let ϕ denote a Boolean formula in conjunctive normal form (CNF), and let p and m be its number of clauses and variables, respectively. SAT denotes the problem of deciding if a CNF formula has a satisfying assignment. Let n denote the number of vertices of the graph under consideration.

2 Hardness of approximation

In this section we prove strong hardness results for approximating the Minimum Maximum Independence Number problem. We first extract and generalize the core of Irving's [2] argument, and then state a general theorem about the solvability of the SAT problem in terms of the approximability of MMIN. One corollary is that approximating MMIN within a factor of $n^{1-\epsilon}$ is $\mathcal{NP}$-complete; another is that obtaining even an $n/\log^3 n$ approximation is unlikely. Finally we note that these hardness results extend to the approximation of MMIN for bipartite graphs.

First, the essence of Irving's construction.

Lemma 1 *For any CNF formula ϕ with p clauses and m variables, and any integer t, there is a graph $G_{\phi,t}$ on $2m + tmp$ vertices with the property that*

$$MMIN(G_{\phi,t}) \begin{cases} \leq & m \quad \text{if } \phi \text{ is satisfiable} \\ > & tm \quad \text{if } \phi \text{ is not satisfiable} \end{cases}$$

Proof. Given ϕ, the graph $G_{\phi,t}$ has two vertices, labeled x_i and $\overline{x}_i$, for each variable x_i, and tm vertices, labeled $C_{j,1}, \ldots, C_{j,tm}$, for each clause C_j. The edges of $G_{\phi,t}$ are $\{x_i, \overline{x}_i\}$ for each i, $\{x_i, C_{j,k}\}$, for all k whenever literal x_i is in clause C_j, and $\{\overline{x}_i, C_{j,k}\}$, for all k whenever literal $\overline{x}_i$ is in clause C_j.

Now, suppose ϕ is satisfiable, and consider a particular satisfying assignment. Then the vertex set

$$\bigcup_i \{x_i : \text{variable } x_i \text{ is true}\} \cup \bigcup_i \{\overline{x}_i : \text{variable } x_i \text{ is false}\}$$

is a maximal independent set of size m, since every clause contains a true literal.

On the other hand, any maximal independent set must include either all the $C_{j,k}$ vertices associated with a clause C_j, or some x_i (or $\overline{x}_i$) representing a variable in the clause. However, an independent set may not contain both x_i and $\overline{x}_i$. Therefore, if ϕ is not satisfiable, no matter what combination of the x and $\overline{x}$ vertices is chosen, there must be at least one clause C_j that contains none of the corresponding literals. So a maximal independent set that contains the chosen x and $\overline{x}$ vertices must also contain the vertices $C_{j,1}, \ldots, C_{j,tm}$, and therefore be of size greater than tm. ∎

From this we obtain a general result about tradeoffs between the solvability of SAT and the non-approximability of MMIN.

Theorem 2 *Let A be an algorithm that approximates MMIN within a factor of $s(n)(= o(n))$ in time $h(n)$. Define $s'(x)$ as the smallest n for which $n/s(n) \geq x$. Then A can be used to decide SAT in time $h(s'(mp))$.*

Proof. Given a CNF formula ϕ with m variables and p clauses, let N be $s'(mp)$ and t be $s(N)$. Observe that $N \geq s(N)mp = tmp$, where the latter equals n or the number of vertices of $G_{\phi,t}$, if we ignore the lower order term. When applied to $G_{\phi,t}$, A will output a number at most $s(n) \cdot m \leq tm$ if ϕ is satisfiable, and a number greater than tm if ϕ is not satisfiable. Hence, A decides SAT in time $h(n) \leq h(s'(mp))$. ∎

Strong $\mathcal{NP}$-hardness results now follow easily.

Corollary 3 *For every constant $\epsilon > 0$, the problem of approximating MMIN within a factor of $n^{1-\epsilon}$ is $\mathcal{NP}$-complete.*

Proof. Assume the existence of a $n^{1-\epsilon}$ approximate algorithm A running in time $h(n)$, for some polynomial h and a fixed $\epsilon > 0$. Then by thm. 2, SAT can be decided in time $h((mp)^{1/\epsilon})$, which is polynomial in m and p. ∎

If we allow for superpolynomial time we obtain stronger lower bounds, the strongest being when we allow for time slightly less than the known $2^{O(p)}$ upper bound.

Corollary 4 *If MMIN can be approximated within a factor of $n/\log^{2+\epsilon} n$ in time $O(n^{\log^{\epsilon/3} n})$, for some constant $\epsilon > 0$, then 3-SAT is solvable in time $2^{p^{1-\delta}}$, for some constant $\delta(\epsilon) > 0$.*

Proof. Consider the notation of the statement of thm. 2, with $h(n) = n^{(\log n)^{\epsilon/3}}$ and $s(n) = n/(\log n)^{2+\epsilon}$. Here, $s'(x) = 2^{x^{1/(2+\epsilon)}}$, and $h(s'(x)) = 2^{x^{1/2(1-\epsilon/6(2+\epsilon))}}$. It then follows from thm. 2 that given the above hypothesis, 3-SAT is solvable in time $h(s'(3p^2)) \leq 2^{p^{1-\delta}}$, for some $\delta = \delta(\epsilon) > 0$. ∎

A monotone formula is a boolean formula where each clause contains either only literals or only negated literals. As observed by Irving [2], the graphs $G_{\phi,t}$ are bipartite whenever ϕ is monotone.

Corollary 5 *An algorithm that approximates MMIN on bipartite graphs within a factor of $s(n)$ in time $h(n)$ can be used to decide Monotone SAT in time $h(s'(mp))$.*

Since satisfying monotone formulas is also $\mathcal{NP}$-complete, cor. 3 holds also when the input is restricted to bipartite graphs.

3 Further evidence of hardness

We further observe that the MMIN problem is still hard to solve when the MMIN of the graph is some fraction of the number of vertices. That closes one option for obtaining a non-trivial approximation. To show this we need a recent result of Arora, Lund, Motwani, Sudan, and Szegedy [1] about the hardness of approximating Maximum Satisfiability.

Lemma 6 (ALMSS) *There is a constant α for which the following problem is $\mathcal{NP}$-complete:*
Given a CNF formula with the property that either
a) all the clauses are satisfiable, or
b) at most $(1-\alpha)$ fraction of the clauses are simultaneously satisfiable,
determine which of the two holds.

Theorem 7 *There exists a constant $\alpha > 0$ such that, deciding if $MMIN(G) \leq dn$ is $\mathcal{NP}$-complete, for all $d \leq \alpha$.*

Proof. Observe from the proof of lemma 1, that the property of $MMIN(G_{\phi,t})$ can be restated as

$$MMIN(G_{\phi,t}) \begin{cases} = & m & \text{if } \phi \text{ is satisfiable} \\ > & \alpha ptm & \text{if at most } (1-\alpha)p \text{ clauses are satisfiable} \end{cases}$$

Thus, deciding if $MMIN(G_{\phi,t}) \leq dn \leq \alpha n = \alpha ptm$ would decide the $\mathcal{NP}$-complete SAT approximation problem of lemma 6. ∎

4 Discussion

It would be interesting to find out whether MMIN can be approximated within anything less than the trivial factor of $O(n)$. The main difficulty appears to lie in the fact that the problem is non-monotonic: no proper subset or superset of a feasible solution is also feasible.

For its cousin, the monotone Maximum Independent Set problem, the existence of a large independent set in the graph implies the existence of a multitude of feasible solutions of a significant size which are statistically likely to be found. In contrast, we have shown in thm. 7 that solving MMIN when it is known to be large is still hard.

We must conclude that the hopes for efficient non-trivial approximations of MMIN are less than promising.

Acknowledgement

I thank Jaikumar Radhakrishnan for careful readings and comments on drafts of this paper.

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and hardness of approximation problems. In *Proc. 33nd Ann. IEEE Symp. on Found. of Comp. Sci.*, October 1992.

[2] R. W. Irving. On approximating the minimum independent dominating set. *Inform. Process. Lett.*, 37:197–200, February 1991.

[3] V. Kann. *Polynomially bounded minimization problems which are hard to approximate.* Manuscript, October 1992.