数理解析研究所講究録
1093巻 1999年 45-50

On greedy algorithms for maximum weighted independent set problem

Shuichi Sakai, Mitsunori Togasaki, Koichi Yamazaki [*]

Department of Computer Science Gunma University
1-5-1 Tenjin-cho, Kiryu zip:376-8515, Gunma, Japan
shuichi,togasaki,koichi@comp.cs.gunma-u.ac.jp

1 Introduction

The optimization problem Maximum (Weighted) Independent Set (MIS(MWIS)) is one of most important problems in computer science [16, 14]. It is well known that the decision problem for MIS is NP-complete. In the eighties, many natural graph classes for which MIS allows to have a polynomial time algorithm were found, and many other natural classes for which MIS remains NP-complete were also found [7]. In the nineties, approximateness have received much attention. With the advance of study of relationship between nonapproximation and interactive proofs, the results on hardness of MIS to approximate have been improved. It is known that for MIS there is a polynominal time approximation algorithm with a ratio of $O\left(\frac{n}{(\log n)^2}\right)$ (which is of the form $n^{1-o(1)}$) [4]. Recently Håstad have shown that MIS is hard to approximate within $n^{1-\epsilon}$ for any ϵ [13]. For graphs with degree bounded by Δ, MIS is known to be MAX SNP-complete [15] and can be approximable within $\frac{\Delta+3}{5+\epsilon}$ for every $\epsilon > 0$ [2]. For bounded degree weighted graphs, a polynominal time approximation algorithm with a ratio of $\frac{\Delta+2}{3}$ is known. [11].

Greedy strategy is one of the most common heuristic method for optimization problems. For MIS, two simple greedy algorithms were investigated. One is called *MIN*, which selects a vertex of minimum degree, removes it and its neighbors from the graph, and iterates this process on the remaining graph until no vertex remains. The other is called *MAX*, which deletes a vertex of maximum degree until no edge remains. Halldórsson and Radhakrishnan showed that *MIN* achieves approximation ratio $\frac{\Delta+2}{3}$ and this bound is tight [12]. Griggs [10] and Chvátal and C. McDiarmid [5] proved independently that *MAX* outputs an independent set of size at least $\sum_{v\in V}\frac{1}{d(v)+1}$ for any graph G. This implies that the approximation ratio is at most $\Delta + 1$. Halldórsson and Radhakrishnan showed that the ratio is at least $\frac{\Delta+1}{2}$ [12].

In this paper, we consider three simple greedy algorithms and two parallel algorithms for MWIS. In section 2, we review terminology and concepts used throughout the paper. In section 3, first we give two simple algorithms *GWMIN* and *GWMAX* which are generalization of *MIN* and *MAX* respectively. Then we show that both greedy algorithms output an independent set of weight $\geq \sum_{v\in V(G)}\frac{W(v)}{d(v)+1}$. This can be considered as a natural extension of Turán's theorem. We also give another simple greedy algorithm, which outputs an independent set of weight $\geq \sum_{v\in V(G)}\frac{W(v)^2}{\sum_{u\in N_G^+(v)}W(u)}$. This can be also thought as an extension of Turán's theorem. In section 4, we present two parallel algorithms *PWMIN* and *PWMAX*, which are parallelization of *MIN* and *MAX* respectively.

---

[*] The last author was supported by the Scientific Grant-in-Aid for Encouragement of Young Scientists from Ministry of Education, Science, Sports and Culture of Japan.

2 Definitions

Let $G = (V, E, W)$ be a weighted undirected graph without loops and multiple edges, where V is the set of vertices, E is the set of edges, and W is the vertex weighting function such that $W : 2^V \rightarrow N^+$ (i.e. each vertex has a positive integral weight), $W(u) \geq 0$ for all $u \in V$ and $W(S) = \sum_{u \in S} W(u)$ for $S \subseteq V$. We also use the notation $V(G)$ and $E(G)$ to denote the set of vertices and edges in G. For a subset $S \subseteq V$, $W(S)$ and $|S|$ are referred as the *weight* and *size* of S respectively. G is *unweighted* if $W(u) = 1$ for all $u \in V$. A subset $I \subseteq V$ is an *independent set* of G if for any two vertices $u, v \in I$, $\{u, v\} \notin E$. An independent set I of G is *maximum* if there is no independent set I' of G such that $W(I) < W(I')$. We denote the weight of maximum independent set of G by $\alpha(G)$ (i.e. $W(I) = \alpha(G)$ for a maximum independent set I of G). Let $G[V']$ denote the subgraph of G induced by V', $d_G(u)$ the degree of vertex u, Δ_G the maximum degree of vertex in G, $\bar{d}_G$ the average degree of G, $N_G(v)$ the neighborhoods of v, and $N_G^+(v)$ $\{v\} \cup N_G(v)$. If G is understood, then we often omit the inscription G in $d_G(u)$, Δ_G, $\bar{d}_G$, $N_G(v)$, and $N_G^+(v)$. For an independent set algorithm A, $A(G)$ is the weight of the solution obtained by A on graph G. The *performance ratio* ρ_A of A is defined by $\rho_A = \max_G \frac{\alpha(G)}{A(G)}$.

3 Greedy algorithms and extension of Turán's theorem

3.1 Known results

The following theorem is known as Turán's theorem [3].

Theorem 3.1 *For any unweighted graph G,*

$$\alpha(G) \geq \frac{n}{\bar{d}_G + 1}.$$

Erdös showed that for unweighted graphs *MIN* attains the above bound [6]. The following extension of Theorem 3.1 was proved first by Wei and later by Alon and Spencer in a different way from Wei.

Theorem 3.2 *For any unweighted graph G,*

$$\alpha(G) \geq \sum_{v \in V} \frac{1}{d(v) + 1}.$$

Wei demonstrated that *MIN* outputs an independent set of at least $\sum_{v \in V} \frac{1}{d(v)+1}$ vertices [18]. Alon and Spencer gave an elegant probabilistic proof of Theorem 3.2 [1], and Selkow improved the probabilistic proof [17]. The probabilistic proof is nonconstructive, however. We can apply the probabilistic proof to the case where graphs are weighted. Thus we have the next theorem.

Theorem 3.3 *For any weighted graph G,*

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{W(v)}{d_G(v) + 1}.$$

3.2 An extension of *MIN*

Let us consider the following framework of *MIN* type algorithm.

```
Algorithm WMIN
  INPUT   :  A weighted graph $G$
  OUTPUT  :  A maximal independent set in $G$.

    begin
        $I := \emptyset$; $i := 0$; $G_i := G$;
        while $V(G_i) \neq \emptyset$ do
            Choose a vertex, say $v_i$, in $G_i$;
            $I := I \cup \{v_i\}$;
            $G_{i+1} := G_i[V(G_i) - N^+_{G_i}(v_i)]$;
            $i := i + 1$;
        od
        Output $I$;
    end.
```

Theorem 3.4 *In WMIN, if each* v_i $(0 \leq i \leq |I|)$ *satisfies* $\sum_{u \in N^+_{G_i}(v_i)} \frac{W(u)}{d_{G_i}(u)+1} \leq W(v_i)$ *(there exists such a node for any graph), then WMIN outputs an independent set of weight at least* $\sum_{v \in V} \frac{W(v)}{d_G(v)+1}$.

□

We refer to a simple greedy algorithm (based on *WMIN*) in which a vertex v maximizing $\frac{W(u)}{d_{G_i}(u)+1}$ over all $u \in V(G_i)$ is selected in each iteration as *GWMIN*.

Corollary 3.5 *GWMIN outputs an independent set of weight at least* $\sum_{v \in V} \frac{W(v)}{d_G(v)+1}$.

Theorem 3.6 $\Delta - 1 \leq \rho_{GWMIN} \leq \Delta + 1$.

□

It seems to be worth noting that we cannot guarantee the performance if we pick up a vertex maximizing $\frac{W(v)}{d_{G_i}(v)}$. The graph depicted in Fig.1 is a counterexample. In the graph, $\sum_{v \in V} \frac{W(v)}{d(v)+1} = 14$. If we choose the vertex v_2 (which maximizes $\frac{W(v)}{d(v)+1}$), we get the independent set $\{v_2\}$, and the total weight is 30. On the other hand, if we choose the vertex v_1 (which maximizes $\frac{W(v)}{d(v)}$), we get the independent set $\{v_1, v_3, v_4\}$, and the total weight is 13.

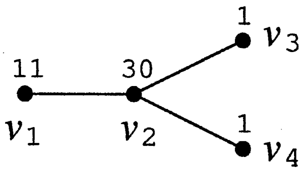

Figure 1: A counterexample.

3.3 An extension of *MAX*

Let us consider the following framework of *MAX* type algorithm.

```
Algorithm WMAX
 INPUT    :  A weighted graph $G$
 OUTPUT   :  A maximal independent set in $G$.

   begin
      $I := \emptyset$; $i := 0$; $G_i := G$;
      while $E(G_i) \neq \emptyset$ do
          Choose a vertex, say $v_i$, in $G_i$;
          $G_{i+1} := G_i[V(G_i) - \{v_i\}]$;
          $i := i + 1$;
      od
      $I := V(G)$;
      Output $I$;
   end.
```

The next theorem is a generalization of the result of Griggs and Chvátal and C. McDiarmid. The essence of the proof is the same as Griggs's proof.

Theorem 3.7 *In WMAX, if each* v_i $(0 \le i \le |V(G) - I|)$ *satisfies* $\sum_{u \in N_{G_i}(v)} \frac{W(u)}{d_{G_i}(u)(d_{G_i}(u)+1)} \ge \frac{W(v)}{d_{G_i}(v)+1}$ *and* $d_{G_i}(v_i) \neq 0$ *(there exists such a node for any graph G such that* $E(G) \neq \emptyset$*), then WMAX outputs an independent set of weight at least* $\sum_{v \in V} \frac{W(v)}{d_G(v)+1}$.

□

We refer to a simple greedy algorithm (based on *WMAX*) in which a vertex v minimizing $\frac{W(u)}{d_{G_i}(u)(d_{G_i}(u)+1)}$ for all $u \in V(G_i)$ is selected in each iteration as *GWMAX*.

Corollary 3.8 *GWMAX outputs an independent set of weight at least* $\sum_{v \in V} \frac{W(v)}{d_G(v)+1}$.

Corollary 3.9 *A simple greedy algorithm (based on WMAX) in which outputs an independent set of weight at least* $\sum_{v \in V} \frac{W(v)}{d_G(v)+1}$.

Theorem 3.10 $\Delta \le \rho_{GWMAX} \le \Delta + 1$.

□

3.4 New greedy algorithm

Theorem 3.11 *A simple greedy algorithm (based on WMIN) in which a vertex v maximizing* $\frac{W(u)}{\sum_{w \in N^+_{G_i}(u)} W(w)}$ *over all* $u \in V(G_i)$*, is selected in each iteration, outputs an independent set of weight at least* $\sum_{v \in V(G)} \frac{W(v)^2}{\sum_{u \in N^+_G(v)} W(u)}$.

□

If $W(v) = 1$ for all $v \in V(G)$ (i.e. unweighted case), then $\sum_{v \in V(G)} \frac{W(v)^2}{\sum_{u \in N^+_G(v)} W(u)}$ is equal to $\sum_{v \in V(G)} \frac{1}{d(v)+1}$ which is the bound of Turán's theorem.

4 Parallel algorithms

Goldberg and Spencer gave a parallel algorithm for MIS which finds an independent set of size at least $\frac{n}{d_G+1}$ [9]. In [12], Halldórsson and Radhakrishnan parallelized *MIN*. In this section, we show that *WMIN* and *WMAX* can be parallelized using the same technique in [12]. Due to limitations of space, we give only *WMIN* type parallel algorithm. It is easy to see that we can construct *WMAX* type parallel algorithm in the same way. We assume the PRAM model.

Let $D_G(v) = W(v) - \sum_{u \in N_G^+(v)} \frac{W(u)}{d_G(u)+1}$ and $\overline{D}_G(v) = \sum_{u \in N_G(v)} \frac{W(u)}{d_G(u)(d_G(u)+1)} - \frac{W(v)}{d_G(v)+1}$. In order to estimate roughly the number of vertices satisfying $D_G(v) \geq 0$ in Lemma 4.3 ($\overline{D}_G(v) \geq 0$ in Lemma 4.4), we need the following two propositions.

Proposition 4.1 $\mathbf{E}[D_G(v)] = 0$.

$\square$

The next proposition can be shown in the same way as proposition 4.1.

Proposition 4.2 $\mathbf{E}[\overline{D}_G(v)] = 0$.

Lemma 4.3 *There are at least* $\frac{|V(G)|}{1+(\Delta+1)!\, W_{max}}$ *vertices* v *which satisfies* $D_G(v) \geq 0$*, where* $\Delta = \max_{v \in V(G)} d_G(v)$*,* $W_{max} = \max_{v \in V(G)} W(v)$*.*

$\square$

Then the next lemma can be shown in the same way as lemma 4.3.

Lemma 4.4 *There are at least* $\frac{2|V(G)|}{2+\Delta(\Delta+1)!\, W_{max}}$ *vertices* v *which satisfies* $\overline{D}_G(v) \geq 0$*, where* $\Delta = \max_{v \in V(G)} d_G(v)$*,* $W_{max} = \max_{v \in V(G)} W(v)$*.*

If there are vertices u and v such that $D_G(v) \geq 0$, $D_G(u) \geq 0$, and the distance between u and v in G is greater than three, then we can select u and v in parallel. Because the selection of u does not affect the selection of v. This suggests the following natural *WMIN* type parallel algorithm, where G^3 denote the graph obtained by taking the adjacency matrix of G to the third power.

```
Algorithm PWMIN

 INPUT    :  A weighted graph $G$
 OUTPUT   :  A maximal independent set in $G$.

   begin
       $I := \emptyset$;
       while ($V(G) \neq \emptyset$) do
           $SAT := \{v \in V(G) : D_G(v) \geq 0\}$;
           $H := G^3[SAT]$;
           $NEW :=$ a maximal independent set of $H$;
           $I := I \cup NEW$;
           $G := G[V(G) - \bigcup_{v \in NEW} N_G^+(v)]$;
       od
       Output $I$;
   end.
```

From Lemma 4.3, $|SAT| \geq \frac{|V(G)|}{1+(\Delta+1)!\, W_{max}}$ in each iteration. Since the maximum degree of H is at most $\Delta + \Delta(\Delta-1) + \Delta(\Delta-1)^2 \leq \Delta^3$, and H is a maximal independent set, the size of NEW is at least $\frac{|V(G)|}{(1+(\Delta+1)!\, W_{max})(\Delta^3+1)}$. Thus, the number of iterations is $O((1+(\Delta+1)!\, W_{max})(\Delta^3+1)\log n)$. It is known that there is an algorithm which finds a maximal independent set of a graph G in time $O((\log \Delta_G)((\Delta_G)^2 + \log^* n))$ using a linear number of processors [8]. As Δ and W_{max} do not depend on n, we have the next theorem.

Theorem 4.5 *Let* Δ *and* W_{max} *be fixed integers. For a graph for which the maximum degree is at most* Δ *and each vertex has a positive integral weight bounded by* W_{max}*, PWMIN outputs an independent set of weight at least* $\sum_{v \in V} \frac{W(v)}{d_G(v)+1}$ *in time* $O((\Delta^6 + \log^* n)(\log \Delta)(1+(\Delta+1)!\, W_{max})(\Delta^3+1)\log n)$ *using linear number of processors in the EREW model.*

References

[1] N.Alon, J.H. Spencer, The probabilistic method, *Wiley, New York*, (1992).

[2] P. Berman and M. Fürer, Approximating maximum independent set in bounded degree graphs, it Proc. Fifth Ann. ACM-SIAM Symp. on Discrete Algorithms, (1994), pp.365-371.

[3] C. Berge, "Graphes et Hypergraphes", Dunod, Paris, (1970).

[4] R. Boppana and M. Halldórsson, Approximating maximum independent sets by excluding subgraphs, *BIT*, Vol. 32, (1992), pp.180-196.

[5] V. Chvátal and C. McDiarmid, Small transversals in hypergraphs, *Combinatorica*, 12 (1), pp.19-26, (1992).

[6] P. Erdös, On the graph theorem of Turán (in Hungarian), Mat. Lapok, 21, pp.249-251, (1970).

[7] Michael R. Garey, David S.Johnson, Computers and Intractability, A Guide to the Theory of NP-Completeness, *"H.FREEMAN San Francisco"*, (1979).

[8] A. V. Goldberg, S. A. Plotkin and G. E. Shannon, Parallel symmetry-breaking in sparse graphs, it SIAM J. Disc. Math., Vol.1, (1988), pp.434-446.

[9] M. Goldberg and T. Spencer, An efficient parallel algorithm that finds independent sets of guaranteed size, *SIAM J. Disc. Math.*, Vol. 6, (1993), pp.443-459.

[10] J. R. Griggs, Lower bounds on the independence number in terms of the degrees, *J. Combin. Theory Ser. B*, 34, pp.22-39, (1983).

[11] Magnús M. Halldórsson, Approximation via partitioning, *Technical Report IS-RR 95-0003F, School of Information Science, Japan Advanced Institute of Science and Technology, Hokuriku*, (1995).

[12] Magnús M. Halldórsson, Jaikumar Radhakrishnan, Greed is Good: Approximating Independent Sets in Sparse and Bounded-degree Graphs, *STOC*, (1994), pp.439-448.

[13] Johan Håstad, Clique is hard to approximate within $n^{1-\epsilon}$, *FOCS*, (1996), pp.627-636.

[14] L. Lovász, Stable set and polynomials, *Discrete Mathematics*, 124, (1994), pp.137-153.

[15] C. H. Papadimitriou and M. Yannakakis, Optimization, approximation, and complexity classes, *J. Comput. System Sci.* 43, (1991), pp.425-440.

[16] P. M. Pardalos and Jue Xue, The Maximum Clique Problem, *Journal of Global Optimization* 4, (1994), pp.301-328.

[17] S.M. Selkow, A probabilistic lower bound on the independence number of graphs, *Discrete Mathematics*, (1994), pp.363-365.

[18] V. K. Wei, A lower bound on the stability number of a simple graph, Technical Memorandum No. 81-11217-9, Bell Laboratories, (1981).