A Simple Clique Camouflaging Against Greedy Maximum Clique Heuristics *

Stanislav Busygin
busygin@gmail.com
http://www.stasbusygin.org

Abstract

Taking a small graph, on which the randomized New-Best-In maximum clique heuristic fails to find the maximum clique, we construct on its basis a class of graphs exemplifying the inefficiency of SM greedy heuristics considered in [2]. We show that a $7(k+1)$-vertex graph from this class is enough to provide a counterexample for the SM^k heuristic. On the other hand, two recent continuous heuristics – Max-AO [3] and QUALEX-MS [4] – successfully find exact solutions on the considered graphs.

Keywords: maximum clique, greedy heuristics, forbidden subgraphs, continuous approach, algorithms, NP-hard.

1 Introduction

Let $G(V,E)$ be a simple undirected graph, $V = \{1, 2, \ldots, n\}$. The *adjacency matrix* of G is a matrix $A_G = (a_{ij})_{n\times n}$, where $a_{ij} = 1$ if $(i,j) \in E$, and $a_{ij} = 0$ if $(i,j) \notin E$. The set of vertices adjacent to a vertex $i \in V$ will be denoted by $N(i) = \{j \in V : (i,j) \in E\}$ and called the *neighborhood* of the vertex i. A *clique* Q is a subset of V such that any two vertices of Q are adjacent. The maximum clique problem asks for a clique of the maximum cardinality. This cardinality is called the *clique number* of the graph and denoted by $\omega(G)$.

The maximum clique problem is NP-hard [1], so it is considered unlikely that an exact polynomial time algorithm for its solution exists. Approximation of large cliques is also hard. It was shown in [5] that unless $NP = ZPP$ no polynomial time algorithm can approximate the clique number within a factor of $n^{1-\epsilon}$ for any $\epsilon > 0$. Recently this margin was tightened in [6] to $n/2^{(\log n)^{1-\epsilon}}$.

To find practically a possibly good substitution for the maximum clique, many heuristic methods were developed. One basic heuristical principle used is to prefer vertices of higher degrees (i.e. number of neighbors) to vertices of lower degrees. Heuristics based on it are called *greedy*. The *New-Best-In* greedy heuristic successively chooses a vertex of the maximum degree in the subgraph composed of the vertices joined with all vertices already included into the formed clique.

Algorithm 1 (New-Best-In)

Input: *a graph* $G(V,E)$.

Output: *a maximal clique* Q.

1. *Construct the vector of vertex degrees* $d \in \mathbb{R}^n$ *such that* $d_i = |N(i)|$.

2. *Set* $V_1 := V; \quad k := 1; \quad Q := \oslash$.

3. *Choose a vertex* $v_k \in V_k$ *such that* d_{v_k} *is greatest.*

4. *Set* $Q := Q \cup \{v_k\}$.

5. *Set* $V_{k+1} := V_k \cap N(v_k)$.

6. *For each* $j \in V_{k+1}$, $d_j := d_j - |(V_k \setminus V_{k+1}) \cap N(j)|$.

7. *If* $V_{k+1} \neq \oslash$, *then* $k := k+1$ *and go to 3.*

8. *STOP.*

Obviously, it runs in $O(n^2)$ time. When considered with respect to maximum independent set finding, this algorithm is called *MIN*. It always finds a maximum independent set (correspondingly, maximum clique on the complementary graph) unless certain (*forbidden*) subgraphs exist in the graph [7]. We may see that there is an uncertainty on Step 3 of the algorithm if more than one vertex of the maximum degree exist. We will suppose that v_k is chosen among them in this case randomly.

A natural attempt to improve the New-Best-In routine is to run it many times starting from all possible cliques of a size k. These heuristics are called SM^k [2].

Algorithm 2 (SM^k**)**

Input: *a graph* $G(V, E)$.

Output: *a maximal clique* Q.

1. $Q := \oslash$.

2. *For all cliques* $Q_0 \subseteq V$ *such that* $|Q_0| = k$:

2.1. *construct the subgraph* $\mathcal{N}_G^{Q_0}$ *induced by* $\cap_{i \in Q_0} N(i)$;

2.2. *run New-Best-In for* $\mathcal{N}_G^{Q_0}$, *assign the result to* Q_1;

2.3. $Q_1 := Q_0 \cup Q_1$;

2.4. *if* $|Q_1| > |Q|$, *assign* $Q := Q_1$.

3. *STOP.*

The complexity of SM^k is $O(n^{k+2})$. Brockington and Culberson showed in [2] how to construct certain graphs where maximum cliques are effectively hidden from SM^k routines. The purpose of this paper is similar, however, we will show that such graphs can be very small. In fact, a $7(k+1)$-vertex graph is enough to “cheat” SM^k.

2 The SM^k Counterexamples

First of all, we provide a 7-vertex graph, where New-Best-In never can find the maximum clique. This graph is complementary to F_{10} graph from [7], so we denote it by $\overline{F_{10}}$. Its adjacency matrix is

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

Obviously, New-Best-In algorithm always selects the first vertex on the first step, but the maximum clique is $\{5, 6, 7\}$.

We construct a sequence $\mathcal{G}$ of SM^k counterexamples as follows. $\mathcal{G}_k$, $k = 0, 1, 2, \ldots$ consists of $k+1$ copies of $\overline{F_{10}}$ totally connected to each other: any two vertices from different copies of $\overline{F_{10}}$ are always joined by an edge. That is, $\mathcal{G}_k$ is a $7(k+1)$-vertex graph. To describe its adjacency matrix formally, we introduce the designations:

$$c(i) = \lfloor (i-1)/7 \rfloor,$$
$$\iota(i) = ((i-1)\mathrm{mod}7) + 1.$$

So,

$$A_{\mathcal{G}_k} = \begin{cases} 0, & \text{if } c(i) = c(j) \text{ and } A_{\iota(i),\iota(j)} = 0 \\ 1, & \text{otherwise.} \end{cases}$$

The maximum clique of $\mathcal{G}_k$ consists of $3(k+1)$ vertices – three last vertices of each $\overline{F_{10}}$ copy should be selected.

Theorem 1 *If $k' \geq k \geq 0$, then SM^k finds only a $(2(k'+1)+k)$-vertex clique of the graph $\mathcal{G}_{k'}$.*

Proof. Obviously, New-Best-In algorithm cannot select a correct vertex of $\mathcal{G}_{k'}$ in any copy of $\overline{F_{10}}$ unless one of the correct vertices is already chosen. That is, a branch of SM^k giving the best result necessarily starts from preselection of k correct vertices from k different $\overline{F_{10}}$ copies. Then, the other two correct vertices in each of these copies will be chosen by the subsequent New-Best-In run immediately – they are connected to all vertices of the residual subgraph. However, $k' - k + 1$ copies of $\overline{F_{10}}$ remain unaltered, so in each of them New-Best-In will lose one vertex for the maximum clique and thus, only a $(2(k'+1)+k)$-vertex clique will be find as the result of SM^k. QED.

We remark that the swapping modification of SM algorithms introduced in [8] does not help on $\mathcal{G}_k$ graphs – there is no room for the useful swaps throughout a New-Best-In run. However, two recent continuous algorithms – Max-AO [3] and QUALEX-MS [4] – successfully find exact solutions on the considered graphs.

References

[1] M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman & Co., 1979).

[2] M. Brockington and J.C. Culberson, Camouflaging independent sets in quasi-random graphs, in: D. Johnson and M.A. Trick, eds., *Cliques, Coloring and Satisfiability, DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* **26** (AMS, Providence, RI, 1996) 75–88.

[3] S. Burer, R.D.C. Monteiro, and Y. Zhang, Maximum Stable Set Formulations and Heuristics based on Continuous Optimization, to appear in *Mathematical Programming*.

[4] S. Busygin, A New Trust Region Technique for the Maximum Weight Clique Problem, a manuscript submitted to *Special Issue of Applied Discrete Mathematics: Combinatorial Optimization 2002*, available at `http://www.busygin.dp.ua/npc.html`.

[5] J. Håstad, Clique is hard to approximate within $n^{1-\epsilon}$, in: *Proc. 37th Annual IEEE Symposium on the Foundations of Computer Science (FOCS)* (1996) 627–636.

[6] S. Khot, Improved inapproximability results for maxclique, chromatic number and approximate graph coloring, in: *Proc. 42nd Annual IEEE Symposium on the Foundations of Computer Science (FOCS)* (2001) 600–609.

[7] J. Harant, Z. Ryjáček, and I. Schiermeyer, Forbidden subgraphs and MIN-algorithm for independence number, a preprint accepted in *Discrete Mathematics*, available at `http://imath.mathematik.tu-ilmenau.de/~harant/x.ps`.

[8] M. Locatelli, I.M. Bomze, and M. Pelillo, Swaps, diversification, and the combinatorics of pivoting for the maximum weight clique, a submitted manuscript available at `http://www.di.unito.it/~locatell/combpivot6.ps`.