

# Using Critical Sets to Solve the Maximum Independent Set Problem

Sergiy Butenko\*, Svyatoslav Trukhanov

*Department of Industrial and Systems Engineering, Texas A&M University,  
College Station, Texas, USA*

---

## Abstract

A method that utilizes the polynomially solvable critical independent set problem for solving the maximum independent set problem on graphs with a nonempty critical independent set is developed. The effectiveness of the proposed approach on large graphs with large independence number is demonstrated through extensive numerical experiments.

*Key words:* Maximum independent set, critical set, large-scale discrete optimization

---

## 1 Introduction

Let  $G = (V, E)$  be a simple undirected graph with the vertex set  $V = \{1, 2, \dots, n\}$ , the edge set  $E$  and a nonnegative weight  $w_i$  associated with each vertex  $i \in V$ . The weight of  $S \subseteq V$  is defined as  $w(S) = \sum_{i \in S} w_i$ . A subset  $I$  of  $V$  is called an independent (stable) set if the subgraph  $G(I)$  induced by  $I$  of  $G$  has no edges. An independent set is called *maximal* if every vertex from  $V \setminus I$  is adjacent to at least one vertex from  $I$ . A *maximum* weight independent set is an independent set of the largest weight in  $G$ . If all weights are equal to 1, then a maximum weight independent set is called maximum independent set and the cardinality of a maximum independent set is called the independence (stability) number, and is denoted by  $\alpha(G)$ .

---

\* Corresponding author. Address: 236E Zachry Engineering Center, TAMU-3131, College Station, TX 77843-3131, USA; tel.: +1-979-458-1089; fax: +1-979-847-9005.  
*Email addresses:* butenko@tamu.edu (Sergiy Butenko), slavik@tamu.edu (Svyatoslav Trukhanov).

For a set  $I \subseteq V$ , its neighborhood  $N(I)$  is the set of all vertices of  $G$  that are adjacent to at least one vertex of  $I$ . A subset  $U_c$  of vertices is called a *critical weighted set* if

$$w(U_c) - w(N(U_c)) = \max\{w(U) - w(N(U)) : U \subseteq V\}.$$

An independent set  $I_c$  (possibly empty) is a *critical weighted independent set* if

$$w(I_c) - w(N(I_c)) = \max\{w(I) - w(N(I)) : I \text{ is an independent set of } G\}.$$

Zhang [14] has shown that for a critical weighted set  $U_c$  and a critical weighted independent set  $I_c$ , we have

$$w(U_c) - w(N(U_c)) = w(I_c) - w(N(I_c)).$$

Moreover, the set of isolated vertices of the subgraph of  $G$  induced by  $U_c$  is a critical weighted independent set. If all vertices of  $G$  have the unit weight, we use the terms *critical set* and *critical independent set* for  $U_c$  and  $I_c$ , respectively. For a critical independent set  $I_c$ , we will denote by  $\alpha_c = |I_c| - |N(I_c)|$ .

The maximum weight independent set problem is a classical NP-hard problem [6] that has numerous important applications [4]. On the other hand, the problem of finding a critical weighted independent set is polynomially solvable [2, 14]. Therefore, an interesting question to investigate is, how useful the critical weighted independent sets of a graph can be for solving the maximum weight independent set problem in this graph.

The maximum weight independent set problem has a number of equivalent formulations as an integer programming problem, or as a continuous nonconvex optimization problem. We mention the simplest of the IP formulations, which is known as *edge formulation*:

$$\begin{aligned} \max \quad & \sum_{i=1}^n w_i x_i, \\ \text{s.t.} \quad & x_i + x_j \leq 1, \forall (i, j) \in E, \\ & x_i \in \{0, 1\}, i = 1, \dots, n. \end{aligned} \tag{1}$$

Nemhauser and Trotter [11] found that if a variable  $x_i$  had value 1 in an optimal solution to the linear relaxation of (1), then  $x_i = 1$  in at least one optimal solution to (1). In this paper, we will prove an analogous result that utilizes critical weighted independent sets instead of the solution of a linear programming relaxation. Moreover, we will show that our approach is very effective when applied to large-scale graphs with large independence number.

The remainder of this paper is organized as follows. The relation between the the critical weighted independent set and the maximum weight independent set problems is studied in Section 2. The results of numerical experiments with an algorithm that utilizes a nonempty critical independent set in graphs with large independence number for solving the maximum independent set problem are reported in Section 3.

## 2 Results

This section establishes the results relating the critical weighted independent set problem to the maximum weight independent set problem. In particular, the main theorem states that any critical weighted independent set is a subset of a maximum weight independent set. This fact is then used to develop a scale-reduction procedure for the maximum independent set problem in graphs with a nonempty critical independent set.

**Lemma 1** *If  $I_c$  is a critical weighted independent set and a maximal independent set, then  $I_c$  is a maximum weight independent set.*

PROOF. Since  $I_c$  is a maximal independent set, we have  $N(I_c) = V \setminus I_c$ . Assume that there exists an independent set  $I$  with weight  $w(I) > w(I_c)$ . Then  $w(V \setminus I) < w(V \setminus I_c)$  and

$$w(I) - w(N(I)) \geq w(I) - w(V \setminus I) > w(I_c) - w(V \setminus I_c),$$

which contradicts to the fact that  $I_c$  is a critical weighted independent set.  $\square$

**Corollary 2** *If  $I_c$  is a critical weighted independent set and a maximal independent set, then  $w(I_c) \geq w(V)/2$ .*

PROOF. Follows from the nonnegativity of  $w(I_c) - w(N(I_c))$ .  $\square$

**Theorem 3** *If  $I_c$  is a critical weighted independent set, then there exists a maximum weight independent set  $I$ , such that  $I_c \subseteq I$ .*

PROOF. Let  $J$  be a maximum weight independent set in  $G$ , and  $I_c$  be a critical weighted independent set. Put

$$I = (J \cup I_c) \setminus N(I_c).$$

Then  $I$  is an independent set, and  $I_c \subseteq I$ . To prove that  $I$  is a maximum weight independent set, it suffices to show that  $w(I) \geq w(J)$ . Assume that  $w(I) < w(J)$ , then  $w(J \setminus I) > w(I \setminus J)$ . Since  $I \setminus J = I_c \setminus J$  and  $J \setminus I = N(I_c) \cap J$ , we obtain

$$w(N(I_c) \cap J) > w(I_c \setminus J).$$

Using the last inequality and the inequality

$$w(N(I_c)) \geq w(N(I_c) \cap J) + w(N(I_c \cap J)),$$

we have

$$\begin{aligned} w(I_c) - w(N(I_c)) &= w(I_c \setminus J) + w(I_c \cap J) - w(N(I_c)) \\ &< w(N(I_c) \cap J) + w(I_c \cap J) - w(N(I_c) \cap J) - w(N(I_c \cap J)) \\ &= w(I_c \cap J) - w(N(I_c \cap J)). \end{aligned}$$

We obtain a contradiction with the fact that  $I_c$  is a critical weighted independent set.  $\square$

Theorem 3 states that a nonempty critical weighted set is always a part of a maximum weight independent set, therefore it can be used in order to reduce the number of vertices to be analyzed when solving the maximum weight independent set problem. The following two lemmas provide some elementary properties of critical weighted sets and critical weighted independent sets in a graph.

**Lemma 4** *Let  $U$  be a critical weighted set of the simple undirected graph  $G = (V, E)$ . Then  $U' = U \cup W$ , where  $W = V \setminus (U \cup N(U))$  is also a critical weighted set of  $G$  and  $U' \cup N(U') = V$ .*

PROOF. By the definition of  $W$ ,  $U \cap W = \emptyset$  and  $N(U) \cap W = \emptyset$ , hence

$$N(W) \subseteq V \setminus U \subseteq W \cup N(U).$$

Since  $U' = U \cup W$ , we have

$$N(U') = N(U) \cup N(W) \subseteq N(U) \cup W \cup N(U) = W \cup N(U).$$

Thus,

$$\begin{aligned} w(U') - w(N(U')) &= w(U) + w(W) - w(N(U')) \\ &\geq w(U) + w(W) - w(W \cup N(U)) \\ &= w(U) - w(N(U)), \end{aligned}$$

so  $U'$  is also a critical weighted set of  $G$ .

Note that  $U' \cup N(U') = U \cup W \cup N(U) \cup N(W) = V$ .  $\square$

**Lemma 5** *Let  $U$  be a critical weighted set of  $G = (V, E)$ , such that  $U \cup N(U) = V$ ,  $I \subseteq U$  be a critical weighted independent set obtained from  $U$  by taking isolated vertices of  $G(U)$ . Then*

$$I = U \setminus N(U), \quad N(I) = N(U) \setminus U,$$

and

$$V \setminus (I \cup N(I)) = U \cap N(U) = U \setminus I$$

PROOF. We first show that  $I = U \setminus N(U)$ . Consider  $v \in I$ . By the definition of  $I$ ,  $v \in U$ . Assume that  $v \in N(U)$ , then there exists  $w \in U, w \neq v : (v, w) \in E$ , so  $v$  is not isolated in  $G(U)$ , therefore  $v \notin I$ . So,  $I \subseteq U \setminus N(U)$ . On the other hand, if  $v \in U \setminus N(U)$  then  $v \in U$  and is isolated in  $G(U)$ , so  $v \in I$ . Thus,  $U \setminus N(U) \subseteq I$ , so  $I = U \setminus N(U)$ .

Next we show that  $N(I) = N(U) \setminus U$ . Note that since  $U \cup N(U) = V$  and  $N(I) \cap U = \emptyset$ , we have  $N(I) \subseteq N(U) \setminus U$ . Recall that from [2] we know that  $w(I) - w(N(I)) = w(U) - w(N(U))$ , but

$$\begin{aligned} w(U) - w(N(U)) &= w(U \setminus N(U)) + w(U \cap N(U)) - w(N(U) \setminus U) - w(U \cap N(U)) \\ &= w(I) - w(N(U) \setminus U). \end{aligned}$$

So,  $w(N(I)) = w(N(U) \setminus U)$  and, since  $N(I) \subseteq N(U) \setminus U$ , we have

$$N(I) = N(U) \setminus U.$$

Finally,

$$\begin{aligned} V \setminus (I \cup N(I)) &= (U \cup N(U)) \setminus ((U \setminus N(U)) \cup (N(U) \setminus U)) \\ &= (U \cup N(U)) \setminus (U \Delta N(U)) \\ &= U \cap N(U) \\ &= U \setminus I, \end{aligned}$$

where  $\Delta$  denotes the symmetric difference. □

### 3 Numerical Experiments

We tested the critical weighted set approach to the maximum weight independent set problem on graphs  $G = (V, E)$  with unit vertex weight and  $\alpha(G) > |V|/2$ , since in this case a critical independent set  $I_c$  is guaranteed to be nonempty. Note that it is easy to prove that the maximum independent set problem remains NP-hard even if restricted to graphs with  $\alpha(G) > |V|/2$ .

To find a critical independent set  $I_c$  in  $G$ , we first computed a critical set  $U_c$  of  $G$  and then put  $I_c = U_c \setminus N(U_c)$  [2]. To find a critical set  $U_c$  in the unweighted case, we first used a reduction of the critical set problem to the maximum matching problem in a bipartite graph proposed by Ageev [2] and then applied a standard algorithm for computing the maximum matching in bipartite graph in  $O(|E||V|)$  time (see, e.g., the proof of König's theorem

in [5]). In the general case, critical weighted set problem may be reduced to the selection problem [2], which is equivalent to finding a minimum cut in a bipartite graph [3,12]. In summary, the proposed algorithm for reduction of a maximum independent set instance  $G = (V, E)$  consists of the following steps:

- (1) Compute a critical set  $U_c$  of  $G$  using the reduction of the critical set problem to the maximum matching problem in a bipartite graph.
- (2) Compute a critical independent set of  $G$  by putting  $I_c = U_c \setminus N(U_c)$ .
- (3) If  $I_c \neq \emptyset$ , put  $V = V \setminus (I_c \cup N(I_c))$ ,  $G = G(V)$  and go to step (1).
- (4) Denote the remaining graph  $G$  by  $G_r = (V_r, E_r)$  and output  $G_r$ .

To test the proposed approach, we generated a number of graphs with  $\alpha(G) > |V|/2$  using Sanchis generator of maximum clique instances available from the DIMACS ftp server at <ftp://dimacs.rutgers.edu/pub/challenge/> (see also [10]). Sanchis graph generator details may be found in [8,13]. We took complements of graphs obtained using the generator, and tested them for connectivity. All graphs considered in our experiments were connected. Since the Sanchis generator produces graphs with a predetermined maximum clique size  $\omega(G)$ , the size  $\alpha(\bar{G})$  of the maximum independent set of its complement  $\bar{G}$  is known,  $\alpha(\bar{G}) = \omega(G)$ . The number of vertices in the graphs used in our computations ranged from 1000 to 18000. Due to their large size, the maximum independent set problem in these graphs cannot be solved using standard exact algorithms. However, the critical set approach presented in this paper was extremely effective with all of the considered cases. The results of computations are summarized in Table 1. Here each row corresponds to one graph  $G = (V, E)$  and other notations used in the table are defined as follows.  $U_c$  denotes the computed critical set,  $I_c$  is the critical independent set obtained from  $U_c$  by taking the isolated vertices of  $G(U_c)$ , and, as before,  $\alpha_c(G) = |I_c| - |N(I_c)|$ . By  $G_r = (V_r, E_r)$  we denote the output of our algorithm, i.e. the graph obtained from  $G$  after recursively removing a critical independent set and its neighborhood. Finally, the last column reports the CPU time (in seconds) of execution of a C implementation of the proposed algorithm. The programs were compiled with the GCC 3.3.6 compiler and run on a Dell Inspiron 8600 computer running Linux 2.6 and configured with Pentium-4M 1400 MHz processor and 512 MB of RAM.

In another set of numerical tests we experimented with Erdős collaboration networks available from Batagelj's Networks/Pajek Graph Files website at <http://vlado.fmf.uni-lj.si/pub/networks/pajek/data/gphs.htm>, as well as complements of the graph coloring problem instances from the Trick's graph coloring page at <http://mat.gsia.cmu.edu/COLOR04/>. The results of these experiments are presented in Tables 2 and 3, where all the notations used are the same as in Table 1. In Erdős collaboration networks considered the vertices represent authors who are "connected" to Paul Erdős through a short path of co-authors. More specifically, Table 2 considers instances with names in the

Table 1  
Results of experiments with Sanchis graphs.

$ V $	$ E $	$\alpha(G)$	$ U_c $	$ I_c $	$\alpha_c(G)$	$ V_r $	$ E_r $	Time
1000	181256	524	524	524	476	0	0	0.07
1000	186723	505	505	505	495	0	0	0.04
2000	686341	1103	1103	1103	897	0	0	0.97
2000	711955	1067	1067	1067	933	0	0	0.44
3000	944175	1535	1535	1535	1465	0	0	0.58
3000	954717	1563	1563	1563	1437	0	0	0.87
4000	1014603	2069	2069	2069	1931	0	0	7.08
4000	1090563	2309	2309	2309	1691	0	0	11.27
5000	1533472	2717	2717	2717	2283	0	0	11.43
5000	720845	3132	3132	3132	1868	0	0	47.74
6000	1775988	3302	3305	3259	2741	46	44	86.97
6000	1815973	3412	3412	3412	2588	0	0	33.64
7000	890777	4493	4493	4493	2507	0	0	154.07
8000	3193335	4394	4394	4394	3606	0	0	103.30
8000	481800	5249	5249	5249	2751	0	0	263.44
9000	4040615	4927	4930	4887	4113	43	41	273.70
9000	681131	5899	5899	5899	3101	0	0	381.69
10000	3775385	5811	5813	5799	4201	14	12	625.58
10000	4908379	5507	5508	5478	4522	30	29	249.76
11000	2546883	6862	6862	6862	4138	0	0	594.24
11000	6528244	5901	5902	5868	5132	34	33	220.11
12000	4862197	7098	7097	7075	4925	22	20	1041.04
12000	5549355	6973	6973	6973	5027	0	0	463.30
13000	1339999	8474	8474	8474	4526	0	0	1131.41
13000	5638263	7698	7705	7640	5358	67	58	1865.82
14000	10772525	7417	7423	7346	6654	77	72	1045.77
14000	3371180	8844	8844	8844	5156	0	0	1330.66
15000	4207335	9413	9417	9386	5614	31	27	2288.17
15000	6912205	8993	8994	8983	6017	11	10	1523.01
16000	14346706	8401	8407	8360	7640	47	41	2923.90
16000	4807361	10042	10042	10042	5958	0	0	2215.34
17000	10748092	9898	9901	9862	7138	39	36	2666.40
17000	913028	11239	11239	11239	5761	0	0	2596.95
18000	2038675	11782	11782	11782	6218	0	0	3249.14
18000	5106081	11412	11412	11402	6594	14	10	3830.01

Table 2

Results of experiments with Erdős networks.

Graph	$ V $	$ E $	$\alpha(G)$	$ U_c $	$ I_c $	$\alpha_c(G)$	$ V_r $	$ E_r $	Time
ERDOS.97.1	472	1314	254	317	179	58	158	271	11.96
ERDOS.97.2	5488	8972	5047	5034	5034	4606	26	13	183.52
ERDOS.98.1	485	1381	261	366	152	58	229	526	1.00
ERDOS.98.2	5822	9505	5368	5356	5356	4914	24	12	213.02
ERDOS.99.1	492	1417	263	375	144	56	246	614	0.42
ERDOS.99.2	6100	9939	5639	5629	5629	5178	20	10	239.61

Table 3

Results of experiments with coloring problem instances

Graph	$ V $	$ E $	$\alpha(G)$	$ U_c $	$ I_c $	$\alpha_c(G)$	$ V_r $	$ E_r $	Time
anna	138	493	80	75	61	29	45	45	0.01
david	87	406	36	81	13	9	70	305	0.01
fpsol2.i.1	496	11654	307	496	227	227	269	11654	0.31
fpsol2.i.2	451	8691	261	348	172	120	223	977	0.33
fpsol2.i.3	425	8688	238	322	146	94	223	974	0.27
huck	74	301	27	34	16	4	46	143	0.01
inithx.i.1	864	18707	566	741	430	363	367	11079	2.21
inithx.i.2	645	13979	365	383	257	144	273	642	0.97
inithx.i.3	621	13969	360	349	259	132	221	423	0.86
jean	80	254	38	39	26	15	35	108	0.01
zeroin.i.1	211	4100	120	211	85	85	126	3775	0.02
zeroin.i.2	211	3541	127	143	111	59	48	207	0.04
zeroin.i.3	206	3540	123	140	108	56	46	206	0.04

form ERDOS. $x.y$ , where  $x$  represents the last two digits of the year for which the network was constructed, and  $y$  represents the largest Erdős number of a scientist represented by a vertex in the graph. For example, the vertices in graph ERDOS.99.1 represent 472 co-authors of Paul Erdős, and the vertices in graph ERDOS.99.2 correspond to 6100 researchers who co-authored a paper either with Erdős or with at least one of his co-authors. We considered such networks for years 1997-1999 and  $y = 1$  and 2. Note that in all considered instances of Erdős collaboration networks their independence number exceeds half of the number of vertices, which is typical for social networks, as well as for large sparse networks arising in many other applications [1, 9]. Naturally, the approach proposed in this paper is a very effective step in solv-

ing the maximum independent set problem for such networks. On the other hand, complements of most of the standard DIMACS maximum clique instances have relatively small independence numbers [10] and empty critical independent sets, thus the critical independence set approach is useless for these instances. The same can be said about the collections of test instances for the maximum independent set problem available online at

<http://www.research.att.com/~njas/doc/graphs.html> and  
<http://www.nlsde.buaa.edu.cn/~kexu/benchmarks/graph-benchmarks.htm>

For all of these instances, with exception of several instances that had isolated vertices, the computed critical independent sets were empty. The results reported in Table 3 show that for a critical independent set to be nonempty, the independence number does not necessarily need to be very large, as for some of the considered graphs  $\alpha(G) < |V|/2$ . The problem of characterizing the graphs with nonempty critical independent sets in general is an interesting problem for future research. Since a critical weighted independent set is a part of a maximum weight independent set, another relevant question is, how to find the largest critical weighted independent set in a graph?

## Acknowledgements

We would like to thank an anonymous referee whose suggestions helped us to improve this paper.

## References

- [1] J. Abello, P. M. Pardalos, and M. G. C. Resende. On maximum clique problems in very large graphs. In J. Abello and J. S. Vitter, editors, *External Memory Algorithms*, volume 50 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 119–130. American Mathematical Society, Providence, RI, 1999.
- [2] A. A. Ageev. On finding critical independent and vertex sets. *SIAM J. Discrete Mathematics*, 7:293–295, 1994.
- [3] M. L. Balinski. On a selection problem. *Management Science*, 17:230–231, 1970.
- [4] I. M. Bomze, M. Budinich, P. M. Pardalos, and M. Pelillo. The maximum clique problem. In D.-Z. Du and P. M. Pardalos, editors, *Handbook of Combinatorial Optimization*, pages 1–74. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

- [5] W.J. Cook, W.H. Cunningham, W.R. Pulleyblank, and A. Schrijver. *Combinatorial Optimization*. John Wiley and Sons, New York, 1998.
- [6] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-completeness*. W.H. Freeman and Company, New York, 1979.
- [7] J. Grossman, P. Ion, and R.D. Castro. The Erdős Number Project. Available online at <http://www.oakland.edu/enp/>. Accessed April 2005.
- [8] J. Hasselberg, P. Pardalos, and G. Vairaktarakis. Test case generators and computational results for the maximum clique problem. *Journal of Global Optimization*, 3:463–482, 1993.
- [9] B. Hayes. Graph theory in practice. *American Scientist*, 88:9–13 (Part I), 104–109 (Part II), 2000.
- [10] D. S. Johnson and M. A. Trick, editors. *Cliques, Coloring, and Satisfiability: Second DIMACS Implementation Challenge*, volume 26 of *DIMACS Series*. American Mathematical Society, Providence, RI, 1996.
- [11] G.L. Nemhauser and L.E. Trotter. Vertex packings: Structural properties and algorithms. *Mathematical Programming*, 8:232–248, 1975.
- [12] J. M. W. Rhys. A selection problem of shared fixed costs and network flows. *Management Science*, 17:200–207, 1970.
- [13] L. Sanchis and A. Jagota. Some experimental and theoretical results on test case generators for the maximum clique problem. Technical Report 69, DIMACS, 1993.
- [14] C.-Q. Zhang. Finding critical independent sets and critical vertex subsets are polynomial problems. *SIAM J. Discrete Mathematics*, 3:431–438, 1990.