Doctoral School on Engineering Sciences
Università Politecnica delle Marche

Extended summary

Finding special structures in Integer Linear Programming problems

Curriculum: Ingegneria informatica, gestionale e dell'Automazione

Author

Angelo Parente

Tutor

Fabrizio Marinelli

Date: 15-02-2013

Abstract.

In general, Integer Linear Programming problems are computationally hard to solve. The design of efficient algorithms for them often takes advantage from the analysis of the problem underlying mathematical structure. Starting from the problem of finding the maximum embedded *reflected network* submatrix of a matrix with entries in $\{0, \pm 1\}$, this work deals with the equivalent problem of finding the maximum balanced induced subgraph of a *signed graph* (MBS, Max Balanced Subgraph). The contribution is twofold. In the first part of the work, a new heuristic for the MBS problem, the *Cycle-Contraction Heuristic* (CCH), has been proposed. The algorithm is based on a graph transformation rule that progressively reduces the lengths of cycles, preserving at the same time the feasibility of solutions for the *MBS* problem. CCH turns out to be more effective of the state-of-the-art heuristics. The efficiency and the effectiveness of CCH can be further improved by means of new rules of *data reduction*, i.e., by a procedure that simplifies instances and/or decrease their size while preserving the optimal solution of the problem.

In the second part of the work, a new exact approach for the MBS problem has been proposed. Such method is based on a polynomial transformation rule that turns a signed graphs into a simple *2-layer* graph. The transformation establishes a strong connection between MBS and the

well-known Maximum Independent Set problem (MIS) and allows to resort to a broad spectrum of (exact or heuristic) solution methods proposed in the literature for MIS.

Finally, the generalized counterpart on signed graphs of some well-known combinatorial problems have been investigated. In particular it has been proven that the *k-coloring* problem on a signed graph - the generalization on signed graph of the balancing problem and the generalization on a simple graph of the bipartization problem - is equivalent to MIS problem on a *k-layer* graph which is a simple graph obtained by generalizing the *2-layer* transformation.

Keywords. balanced subgraph, independent set, k-coloring, network submatrix, signed graph.

1 Problem statement and objectives

Nowadays, the most successful methods for solving an integer linear program works by iteratively tightening (cutting planes) or recursively decomposing (branch-and-bound) a polyhedron which represents a continuous formulation of the problem. Such algorithms stop as soon as an integer extreme point is reached and its optimality is proven. When one of such methods is applied, the cases where the polyhedron is integral (or at least has an integer extreme point that is optimal for a given objective function) is of particular interest because in that cases the integer linear problem boils down to a continuous linear problem. Indeed, the recognition of special structures in the coefficient matrix of (integer) linear programs on one hand helps in the solution of large-scale continuous models [1][2] and, on the other hand, can be exploited in the strategic choice of constraints to be relaxed (convexified) in a Lagrangian relaxation (Dantzig-Wolfe decomposition). A well-known families of such special structures are totally unimodular (TU) matrices: one of the famous theorems by Hoffman-Kruskal [3] states that the polyhedron $P = \{x \in \mathbb{R}^n \mid Ax \geq b, x \geq 0\}$ is integral for every integral vector $b \in \mathbb{R}^m$ if and only if $A(m \times n)$ is a TU matrix.

Total unimodularity can be checked in polynomial time [4] but the coefficient matrix of most of the combinatorial optimization problems are not naturally TU, hence we are interested in finding a maximum embedded TU submatrix [5][6]. An important subclass of TU matrices are that of the *pure network* matrices, i.e., the $\{0, \pm 1\}$-matrices where every column contains at most one element -1 and at most one element $+1$ [7].

DMERN problem (*Detecting a Maximum Embedded Reflected Network Matrix*): given a $\{0, \pm 1\}$-matrix A, find a maximum subset D of its rows such that D is a reflected network matrix. DMERN problem is known to be NP-hard [8].

Gulpinar et al. [9] showed that the *DMERN* problem is closely related to that of balancing subgraphs in *signed graphs*.

A signed graph $S(G, \sigma)$ is an undirected (multi)graph $G(V, E)$ whose edges are labelled either with $+1$ (positive edges) or with -1 (negative edges) by the signature function $\sigma(E)$. The graph G is called the underlying graph of S, see [10] for a comprehensive survey on signed graphs. A subset of edges of a signed graph, e.g., a path or a cycle, is positive (resp. negative) if it contains an even (resp. odd) number of negative edges. From a theoretical perspective, signed graphs and their incident matrices are deeply connected with matroid theory [11].

A signed graph S_A can be defined for any given $\{0, \pm 1\}$-matrix A. S_A has a vertex for each row of A and an edge (u, v) if A has a column with non-zero elements a_u and a_v on rows u and v. The edge (u, v) is:

- negative if a_u and a_v have the same sign (rows *u* and *v* are in *conflict*);
- positive if a_u and a_v have opposite signs (rows *u* and *v* are in *potential conflict*).

Following the definition by Harary [12], a signed graph is said to be *balanced* if it has no negative cycles (cycles with an odd number of negative edges). Harary also proved that a signed graph is balanced if and only if the set of negative edges is empty or is a cut. Given a signed graph, the Maximum Balanced induced Subgraph problem (MBS) asks for finding a balanced induced subgraph of maximum order.

MBS is polynomial on series parallel graphs but in general is NP-hard as it admits maximum independent set and vertex bipartization problems as special cases. In particular, the maximum independent set problem on a graph G corresponds to MBS on the signed graph obtained from G by doubling its edges and assigning opposite signs to each pair of multiple edges (notice that a pair of multiple edges of opposite signs always forms a negative cycle). Moreover, negative cycles correspond to odd cycles when S has negative edges only. Then, a minimum set of vertices to be removed to obtain a balanced subgraph of S corresponds to a minimum odd cycle transversal of G.

The notion of balancing arises in many different contexts. On one hand, it generalizes well-known combinatorial structures such as the bipartite induced subgraph and the independent set, the former with applications, e.g., in VLSI design, DNA sequencing and computational biology [13][14], the latter arising as a subproblem or relaxation of many 0-1 integer problems, e.g., general preprocessing and probing techniques perform a vertex packing on the so-called conflict graph.

On the other hand, problems originating in different domains, even far away from each other, can be modelled as MBS. In the social networking, for example, relationships between individuals can be represented by a signed graph where nodes are persons and positive (negative) edges express friendships (hostility). Based on the structural balance theory [15], positive cycles are supposed to indicate stable social situations, whereas negative cycles are supposed to be unstable. The maximum balanced subgraph gives a measure of the cohesion of the social group as well as, when friendship or hostility lie to siding with the same or opposite fronts, indicate who has to change his/her orientation in order to have the largest group with the same opinion.

Another different application lies in the context of portfolio analysis; in that case, the nodes of a signed graph denote the stocks and the positive (negative) edges represent the direct (inverse) correlation between the stocks. It is shown that the largest is the maximum balanced induced subgraph, the most predictable is the behaviour of the portfolio.

2 Research planning and activities

After the study of the heuristic (SGA), which is the state-of-the-art approach for MBS problem, the main goal of the research was the study of a new heuristic that is based on a transformation of the graph driven by a local analysis of *negative-cycles*, rather than on node-switch operations (e.g. rows reflection on A) as SGA does.

The new idea is the transformation of the graph S in a new graph S^C according to rules that assure that the solution on S^C is also a solution for S. The aim of the transformation is to improve the solution of the independent set problem on S^C both in terms of efficiency and accuracy

The graph transformation is obtained by applying the *cycle contraction rule* (CCR) on some edges of the graph.

Let S be a signed graph and E^P the set of parallel edges, the CCR consists in the following operations:

1. choose an edge $(u, v) \notin E^P$ (according to a given strategy);
2. choose an endpoint of (u, v) as a pivot vertex (for example u);
3. consider the neighborhood $N(u)$;
4. for each vertex $w \notin (N(u)\setminus \{v\})$, if $(u, w) \notin E^P$ add the edge (v, w) with sign $s(u, v) \cdot s(u, w)$;
5. remove the edge (u, w) from S.

It is easy to see that all the cycles containing a given edge (v, w) are contracted and preserve the signs when CCR is applied on (v, w)

Theorem: *if a subset of vertices* $W \subseteq N$ *induces a balanced subgraph on* S^C *then* W *also induces a balanced subgraph on* S.

Therefore a DMERN solution on S^C is also a DMERN solution on S

DMERN solution on S^C is feasible for S.

Corollary: *If* S *is balanced,* S^C *can be transformed in an empty graph*, i.e. repeated cycle-contraction on a balanced graph will deletes all edges (an heuristic based on CCR is able to find the optimal solution whenever S is balanced.)

The new heuristic, called Cycle-Contraction Heuristic (CCH), exploits the transformation rule in order to *isolate* a chosen node by removing all its incident edges. The node processing sequence is based on the number of negative *chordless* cycles traversing the nodes [16]. Due the complexity of such computation, the evaluation is performed only on a local bases (a given neighbourhood).

Table 1 shows the comparison between the new heuristics CCH and the heuristic SGA. The benchmark consists in 26 instances from NetLib library whose respective graphs are not balanced. The last two columns show the gap from optimal solutions. On these instances the average gap goes from 4.58 by SGA to 0.2 by CCH, an improvement of more than 95%. In addition, CCH finds the optimal solution in 19 cases.

3 Analysis and discussion of main results

New results on MBS problem are been achieved both toward the design of heuristics algorithms, and for the research of exact solutions. Work outcomes can be recap in four principal points.

The first is the ideation of a new transformation rule (Cycle Contraction Rule) for signed graphs that can be exploited in the development of new effective heuristic algorithms. One instance of such kind of algorithms is the Cycle Contraction Heuristic (CCH) whose computational results in general considerably improves the best results known.

A second result consists in a new exact formulation model that transform DMERN problem in a maximum independent set problem (MIS). Being MIS one of the basic combinatorial problems, such transformation allows to solve DMERN by resorting to the tools provided by an extensive bibliography. Indeed, the polynomial transformation of a signed graph in a simple graph (2-layer graph) allows to apply some of the numerous heuristic algorithms for the maximum independent set existing in the literature. It was also demonstrated the equivalence between the signed graph balancing property and the bipartition of the corresponding 2-layer graph.

The third important point is the definition of a new *data reduction* algorithm that reduces the number of edges of a signed graph by exploiting the properties of chordless cycles, and without compromise the possibility of finding the optimal solution to the problem DMERN. The reduction rate has been very high on instances related to real problems, whose graphs are not very dense.

Table 1: CCH vs SGA heuristics

name	|V|	Exact	SGA	CCH	gap SGA	gap CCH
25fv47	224	212	199	212	6,13	**0**
80bau3b	1374	1346	1340	1345	0,45	0,07
agg2	141	62	58	62	6,45	**0**
agg3	141	62	58	62	6,45	**0**
bnl1	275	261	260	261	0,38	**0**
bnl2	1418	1367	1329	1367	2,78	**0**
cycle	505	504	504	504	0,00	**0**
czprob	719	718	717	718	0,14	**0**
d2q06c	844	804	778	804	3,23	**0**
degen3	1503	796	711	777	10,68	2,39
dfl001	6019	4525*	3148	4419	30,43	2,34
fffff800	157	125	118	125	5,60	**0**
ganges	631	534	518	534	3,00	**0**
greenbea	915	890	872	889	2,02	0,11
greenbeb	915	890	873	889	1,91	0,11
maros	305	300	297	300	1,00	**0**
modszk1	148	130	124	130	4,62	**0**
perold	150	143	141	143	1,40	**0**
pilot	275	252	248	252	1,59	**0**
pilot87	339	306	303	306	0,98	**0**
pilotnov	209	203	202	203	0,49	**0**
scfxm2	282	252	234	252	7,14	**0**
seba	402	140	133	140	5,00	**0**
shell	483	482	432	482	10,37	**0**
stocfor2	1262	1106	1072	1104	3,07	0,18
stocfor3	9632	8516	8186	8512	3,88	0,05
*upper bound				**Avg:**	**4,58**	**0,20**

The above three results become even more important when applied to the instances of either the equivalent problem of finding the maximum induced balanced subgraph (MBS), or the special case of determining the maximum bipartite induced subgraph of a simple graph.

Finally the solution methods based on the construction of the 2-layer graph have been generalized (k-layer construction of the graph) and applied to the problem of k-coloring of a signed graph and in particular to the problem of determining, given a parameter k, the maximum k-colorable induced subgraph (a problem recently presented with the name k-MBS).

4 Conclusions

The first research issue to pursue in the future work is the improvement of the MIS heuristics on 2-layer graphs. In particular, it is of great interest to assess the benefits that could be achieved by replacing the simple minimum-degree greedy algorithm currently used with one of the most sophisticated existing heuristic for the MIS problem. On the other hand, the same techniques used in these heuristics may be analyzed and exploited to improve the processing strategy of new heuristic proposed. An attempt would be to apply the CCR transformation rule, in order to avoiding the formation of cliques of parallel edges and facilitate the formation of stars of parallel edges instead.

Another point that deserve further analysis concerns the effects of data reduction rules, both in terms of quality of the solutions and impact on the performance of exact algorithms and/or heuristics to problems of bipartition and balancing.

Finally, it would be interesting to transfer the proposed methods to balancing problems based on edge deletion.

References

[1] R.E. Bixby and W.H. Cunningham. Converting linear programs to network problems. Mathematics of Operations Research, 1980.
[2] R.E. Bixby and R. Fourer. Finding embedded network rows in linear programs I. Extraction heuristics. Management Science, 1988.
[3] A. Hoffman and J. Kruskal, Integral boundary points of convex polyhedra, in Linear Inequalities and Related Systems. Princeton University Press, 1956.
[4] K. Truemper. A decomposition theory for matroids v. testing of matrix total unimodularity. Journal of Combinatorial Theory, 1990.
[5] G.G. Brown and W.G. Wright. Automatic identification of embedded network rows in large-scale optimization models, Mathematical Programming, 1984.
[6] R. Figueiredo, M. Labbè and C. de Souza. An exact approach to the problem of extracting an embedded network matrix, Computers & Operations Research, 2011.
[7] P.D. Seymour. Decomposition of regular matroids. Journal of combinatorial theory, 1980.
[8] J.J. Bartholdi et al. A good submatrix is hard to find. *Operations Research Letters*, 1982.
[9] N. Gülpinar, G. Gutin, G. Mitra and A. Zverovitch. Extracting pure network sub matrices in linear programs using signed graphs. Discrete Applied Mathematics, 2004.
[10] T. Zaslavsky. Signed graphs. Discrete Applied Mathematics, 1982.
[11] A. Schrijver, Theory of Linear and Integer Programming. John Wiley, 1998.
[12] F. Harary. On the notion of Balance of a Signed Graph. Michigan Mathematical Journal, 1953.
[13] H.A. Choi, K. Nakajima, and C.S. Rim. Graph bipartization and via minimization. 1987.
[14] F. Huffner, Algorithm engineering for optimal graph bipartization, Experimental and Efficient Algorithms, Springer, 2005.
[15] D. Cartwright and F. Harary. Structural balance: a generalization of heider's theory. Psychological review, 1956.
[16] F. Barahona and A.R. Mahjoub. Facets of the balanced (acyclic) induced subgraph polytope. Mathematical Programming, 1989.