Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

Information Processing Letters 88 (2003) 155–159

Information Processing Letters

www.elsevier.com/locate/ipl

On finding cycle bases and fundamental cycle bases with a shortest maximal cycle

Giulia Galbiati

Dipartimento di Informatica e Sistemistica, Università degli Studi di Pavia, 27100 Pavia, Italy

Received 20 April 2003; received in revised form 21 July 2003

Communicated by A.A. Bertossi

Abstract

An undirected biconnected graph G with nonnegative integer lengths on the edges is given. The problem we consider is that of finding a cycle basis B of G such that the length of the longest cycle included in B is the smallest among all cycle bases of G. We first observe that Horton's algorithm [SIAM J. Comput. 16 (2) (1987) 358–366] provides a fast solution of the problem that extends the one given by Chickering et al. [Inform. Process. Lett. 54 (1995) 55–58] for uniform graphs. On the other hand we show that, if the basis is required to be fundamental, then the problem is NP-hard and cannot be approximated within $2-\epsilon$, $\forall\epsilon > 0$, even with uniform lengths, unless P = NP. This problem remains NP-hard even restricted to the class of complete graphs; in this case it cannot be approximated within $13/11-\epsilon$, $\forall\epsilon > 0$, unless P = NP; it is instead approximable within 2 in general, and within $3/2$ if the triangle inequality holds.

Keywords: Algorithms; Approximation algorithms; Combinatorial problems; Computational complexity; Cycle basis

1. Introduction

Throughout this paper bet $G = (V, E)$ be an undirected biconnected graph without loops or multiple edges, with n vertices, m edges and a nonnegative integer length $w(e)$ assigned to each edge $e \in E$. (In the sequel word *length* and word *weight* will be used as synonyms.) Associated with G there is a vector space over $GF(2)$, called the *cycle space*, consisting of the incidence vectors of all cycles (including the null cycle) and of all unions of edge-disjoint cycles of G. The dimension of this space is $m-n+1$, called the *nullity* or *cyclomatic number* $\nu(G)$ of G. Let the length of a cycle be the sum of the lengths of its edges.

The Shortest Maximal Cycle Basis (SMCB) problem is that of finding a cycle basis B of G with the property that the length of the longest cycle included in B is the smallest among all bases of G. The interest in this problem has been motivated in [2] by a possible application as a preprocessing step in a Bayesian inference algorithm [5]. Here we observe that Horton's algorithm [3] provides a fast solution of the problem thus extending the one given in [2] for uniform graphs.

When G is connected there are special cycle bases that can be derived from the spanning trees of G, which we call *fundamental cycle bases*. If T is an arbitrary spanning tree of G, by adding anyone of the

E-mail address: giulia.galbiati@unipv.it (G. Galbiati).
URL: http://mate.unipv.it/~galbiati (G. Galbiati).

0020-0190/$ – see front matter
doi:10.1016/j.ipl.2003.07.003

$m-n+1$ edges of G which do not belong to T, the so-called chords, one creates a cycle and the set of these $m-n+1$ cycles forms a cycle basis, which is referred to as a fundamental cycle basis.

The Shortest Maximal Fundamental Cycle Basis (SMFCB) problem is that of finding a fundamental cycle basis B of G with the property that the length of the longest cycle included in B is the smallest among all fundamental bases of G. As far as we know this problem, addressed in [4], is new, and interesting both in a theoretical contest and in the practical context of electrical and communication networks. Matrix analysis of electrical networks provides examples of dynamic circuits (with capacities and inductances) whose state equations can be solved more or less rapidly depending on the choice of a fundamental cycle basis.

We show that this problem is NP-hard and cannot be approximated within $2-\epsilon$, $\forall\epsilon > 0$, even with uniform weights, unless P = NP. When restricted to the class of complete weighted graphs, the problem remains NP-hard and cannot be approximated within $\frac{13}{11}-\epsilon$, $\forall\epsilon > 0$, unless P = NP; it is instead approximable within 2 in general and within $3/2$ if the triangle inequality holds.

2. Preliminaries

Let G be as above and let $B = \{b_1, \ldots, b_{m-n+1}\}$ be a cycle basis for G. The *length* $W(b_i)$ of cycle b_i is defined as the sum of the lengths of all the edges in the cycle, whereas the *global length* $W(B)$ of base B is defined as the maximum among the lengths of its cycles. If $C \subseteq B$ we denote by G_C the subgraph of G consisting of the cycles in C. In [6] two useful characterizations of fundamental cycle bases are given.

Theorem 1. *A cycle basis B of G is fundamental if and only if B contains no cycle which consists entirely of edges belonging to other cycles of B.*

Theorem 2. *A cycle basis B of G is fundamental if and only if $\nu(G_C) = |C|$ for every $C \subseteq B$.*

3. The Shortest Maximal Cycle Basis problem

The algorithm that we present in this section for solving the SMCB problem on graph G is essentially the one given in [3], it extends the one given in [2] and works if there is a unique shortest path between any two vertices of G. If this is not the case it is enough, before applying the algorithm, to use standard perturbation techniques in order to guarantee uniqueness (see [1]). For instance, let $\sigma : E \to \{1, \ldots, |E|\}$ be an arbitrary permutation of the edges of G. Define the perturbation on an edge $e \in E$ to be $\varepsilon(e) = 2^{\sigma(e)-|E|-1}$. It is easy to see that $\sum_{e\in E}\varepsilon(e) < 1$ and that, for all subsets $E_1, E_2 \subseteq E$ we have $\sum_{e\in E_1}\varepsilon(e) \neq \sum_{e\in E_2}\varepsilon(e)$ if and only if $E_1 \neq E_2$. Thus if one increase the integral length $w(e)$ of any edge $e \in E$ by $\varepsilon(e)$ different paths get different lengths and every shortest path is unique.

Algorithm 1.

(1) Find the shortest path $P(x, y)$ between each pair of vertices x, y of G.
(2) For each vertex v and edge $\{x, y\}$ in G, create the cycle $C(v, x, y) = P(v, x) + P(v, y) + \{x, y\}$, and calculate its length. Degenerate cases in which $P(v, x)$ and $P(v, y)$ have vertices other than v in common can be omitted.
(3) Order the cycles by increasing lengths.
(4) Use the greedy algorithm to find from this reduced set of cycles a cycle basis of G, i.e., add to an initial empty basis the next shortest cycle if it is independent from the already selected ones.

Theorem 3. *Algorithm* 1 *solves the problem of finding in a graph G a cycle basis B with shortest global length if any two vertices of G are joined by a unique shortest path.*

Proof. It is well known that the cycle space of graph G is a matroid; the reduced set of cycles used by Algorithm 3, being a finite subset of a vector space, is also a matroid. It is known that in a matroid the greedy algorithm finds a basis that simultaneously minimizes the sum of the lengths of its elements and the maximum among the lengths of its elements. Theorem 4 in [3] proves that, if all shortest paths in G are unique, the reduced set of cycles used by the algorithm contains all cycles appearing in any cycle basis of G that minimizes the sum of the lengths of its elements. Hence it follows easily that the greedy algorithm in step (4) finds a basis for the cycle space of G that minimizes the maximum among the lengths of its elements. □

4. The Shortest Maximal Fundamental Cycle Basis problem

In this section we investigate the complexity of the SMFCB problem. We prove that it is NP-hard even when restricted to uniform graphs, i.e., having $w(e)$ equal to 1 for each edge $e \in E$.

Theorem 4. *The problem of finding in a uniform graph G a fundamental cycle basis B with shortest global length is* NP-*hard.*

Proof. We prove the theorem by exhibiting a reduction of the Satisfiability problem to the recognition form of our problem. Given an instance I of Satisfiability, i.e., a CNF formula F on a set U of boolean variables, we define an instance I' for the recognition form of the SMFCB problem, i.e., a graph G and an integer k, such that I is satisfiable iff there exists in G a fundamental cycle basis of global length at most k.

Let I be a collection $C = C_1, \ldots, C_h$ of h disjunctive clauses of literals, where a literal is a variable or a negated variable in $U = \{u_1, \ldots, u_n\}$.

First we define a graph G having arcs with lengths equal to 1 or to a large integer M to be defined later and we prove the result for this graph; then we observe that the result is not affected if we replace each arc having length M with a chain of M arcs of unitary length.

We start the construction of G from the graph G' given in Fig. 1 where the only lengths indicated are those equal to M.

Then, in order to obtain G from G', for each clause C_i we add to G' two vertices c_i and c^i and the edge $\{c_i, c^i\}$ with length equal to M; moreover if C_i contains the variable u_j or its negation we add the edge $\{c^i, v_j\}$; finally if C_i contains the variable u_j (respectively $\bar{u}_j$) we add the edge $\{c_i, u_j\}$ (respectively $\{c_i, \bar{u}_j\}$) (see Fig. 2). We complete the reduction by setting k equal to $M + 3$.

Now if I is satisfiable there exists a truth assignment for U that satisfies each clause; we show that we can find a spanning tree T of G having a fundamental set of cycles of global length at most $M + 3$.

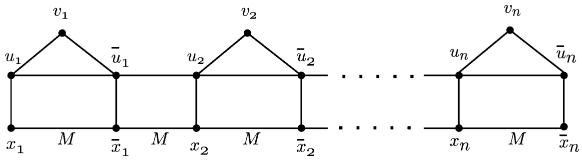

Fig. 1. Graph G'.

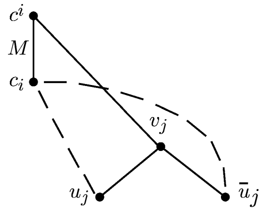

Fig. 2. Additional edges and vertices for clause C_i.

We start the construction of tree T from the tree T' consisting of the edges $\{u_j, \bar{u}_j\}$, $\{u_j, x_j\}$, $\{\bar{u}_j, \bar{x}_j\}$, for all $j = 1, \ldots, n$ and of the edges $\{\bar{u}_j, u_{j+1}\}$, for all $j = 1, \ldots, n - 1$. Then in order to obtain T we add to T', for each variable u_j set to true (respectively false) in the assignment, the edge $\{v_j, u_j\}$ (respectively $\{v_j, \bar{u}_j\}$); moreover for each clause C_i we choose a literal that satisfies the clause and if the chosen literal is variable u_j (respectively negated variable $\bar{u}_j$) we add the edge $\{c^i, v_j\}$ and the edge $\{c_i, u_j\}$ (respectively $\{c_i, \bar{u}_j\}$). It is easy to verify that, if M is chosen to satisfy the inequality $2n + 3 \leqslant M + 3$, the set of fundamental cycles with respect to T has cycles of length at most $M + 3$.

Conversely, suppose that there exists in G a fundamental cycle basis of global length at most $M + 3$, with $M = 2n$. Observe that all cycles that are fundamental cycles with respect to the chords of T' (these chords have length equal to M) must belong to the basis; moreover for each clause C_i the edge $\{c_i, c^i\}$ must belong to a cycle in the basis that goes through a vertex v_j, for some $j = 1, \ldots, n$; call this cycle A_j (respectively $\bar{A}_j$) if it goes also through vertex u_j (respectively $\bar{u}_j$). It is crucial to notice that all cycles of the basis containing the edges $\{c_i, c^i\}$, for all $i = 1, \ldots, h$, cannot contain both A_j and $\bar{A}_j$, for some index j, otherwise Theorem 2 would be violated: the A_j and $\bar{A}_j$ plus the cycle that goes through the vertices $\{u_j, \bar{u}_j, x_j, \bar{x}_j\}$ would represent a set S of cycles such that $\nu(G_S) = |S| + 1$, since the additional cycle through the vertices $\{u_j, v_j, \bar{u}_j\}$ would be generated.

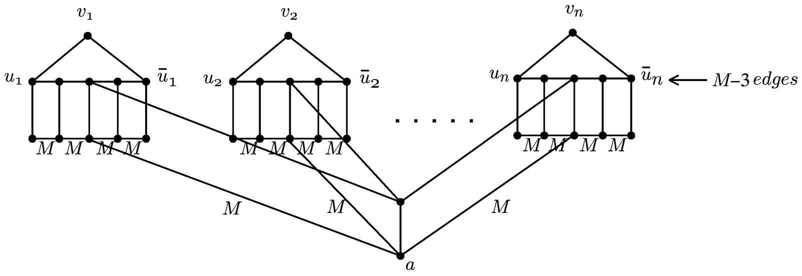

Fig. 3. Graph G''.

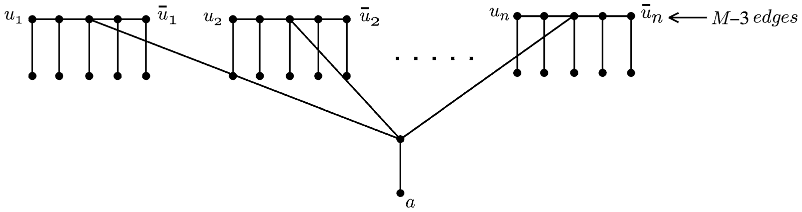

Fig. 4. Tree T''.

Now it is easy to conclude that all the A_j or $\bar{A}_j$ containing the edges $\{c_i, c^i\}$, for each $i = 1, \ldots, h$, allow to identify a truth assignment for U that satisfies all clauses in I. □

The next theorem proves a non-approximability result for the SMFCB problem, again for uniform graphs.

Theorem 5. *The problem of finding in a uniform graph G a fundamental cycle basis B with shortest global length cannot be approximated within $2 - \epsilon$, $\forall \epsilon > 0$, unless* $\mathrm{P} = \mathrm{NP}$.

Proof. We prove the theorem by giving a more sophisticated reduction, from Satisfiability to the optimization form of our problem, which exhibits a gap. More precisely, we show that yes-instances of Satisfiability are mapped into instances that have an optimum solution of global length at most $M + 3$, whereas no-instances are mapped into instances that have an optimum solution of global length at least $2M$. From this we will be able to conclude that the problem cannot be approximated within $2 - \epsilon$, $\forall \epsilon > 0$, unless $\mathrm{P} = \mathrm{NP}$.

The reduction constructs a graph G from a starting graph in a way identical to the one used in Theorem 4; here the starting graph is not G' but the graph G'' given in Fig. 3; the addition of edges and vertices to G'' is identical to the one for G'.

It is easy to see that yes-instances of Satisfiability are mapped into instances that exhibit a fundamental cycle basis of global length at most $M + 3$ and hence into instances whose shortest global length is at most $M + 3$. Such a basis is the set of fundamental cycles with respect to the tree T built as in Theorem 4, but starting here from the tree T'' illustrated in Fig. 4. The only necessary requirement for M is to satisfy $M - 3 \geqslant 2$.

We now show that if G had a fundamental cycle basis where all cycles have length less then $2M$, then it would be possible to satisfy instance I. In fact, in such a case, all cycles that are fundamental with respect to the chords of T'' with length equal to M and no vertex a as an endpoint would belong to the basis; we group these cycles naturally in n groups, called $B_1, \ldots, B_n$. We observe that the cycles of the basis that include the edges $\{c_i, c^i\}$, for all $i = 1, \ldots, h$, cannot include both edges $\{v_j, u_j\}$ and $\{v_j, \bar{u}_j\}$ for some j, otherwise Theorem 2 would be violated, because of the cycles in B_j; and this would be sufficient to identify a true assignment satisfying instance I. Hence no-instances are mapped into instances whose shortest global length is at least $2M$. At this point we may conclude that the problem cannot be approximated within $2M/(M + 3) - \epsilon'$, $\forall \epsilon' > 0$, unless $\mathrm{P} = \mathrm{NP}$.

It follows that $\forall\epsilon > 0$, if we choose ϵ' to be less than or equal to $\epsilon/2$ and M is chosen in such a way that $6/(M+3) \leqslant \epsilon/2$, then the inequality $2M/(M+3) - \epsilon' \geqslant 2 - \epsilon$ becomes true and the conclusion follows. □

4.1. *A special case*

In this subsection we consider complete graphs. Of course the only significant case is the nonuniform case. The next theorem shows that this case is just as interesting as the general case of uniform graphs.

Theorem 6. *The problem of finding, in a complete weighted graph G, a fundamental cycle basis B with shortest global weight is* NP-*hard and cannot be approximated within* $\frac{13}{11} - \epsilon$, $\forall\epsilon > 0$, *unless* P = NP.

Proof (*Sketch*). It is enough to follow the lines of the preceding proofs, using a third reduction from Satisfiability. This time the starting graph G''', from which graph G is built, is illustrated in Fig. 5 where only weights different from 1 are indicated. Graph G is built from G''' by first adding edges and vertices as specified in the proofs of the preceding theorems and as illustrated in Fig. 2, then by completing the resulting graph with edges of weights that we now specify. Precisely, the weights assigned are set equal to $y = 4$ for all edges $\{u_j, z_j\}$, $\{x_j, y_j\}$, $\{y_j, \bar{x}_j\}$, $\{z_j, \bar{u}_j\}$, $\{b, z_j\}$, $\{a, y_j\}$, $j = 1, \ldots, n$, and for all edges $\{c_j, v_j\}$, if clause C_i contains variable u_j or its negation, and for all edges $\{c^i, u_j\}$ (respectively $\{c^i, \bar{u}_j\}$), if C_i contains the variable u_j (respectively $\bar{u}_j$); the remaining edges receive weights set equal to $z = 3$. It can be seen that if one sets $M = 8$ the conclusion follows. □

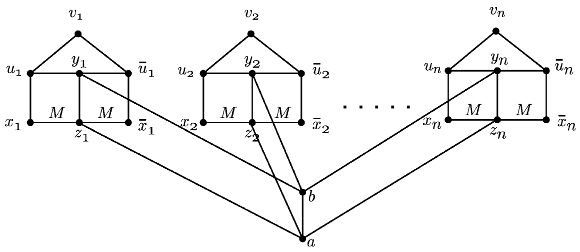

Fig. 5. Graph G'''.

We conclude this paper by observing that it is easy to prove that any spanning tree of minimum diameter in a complete weighted graph is a solution to the SMFCB problem within 2 in general and within 3/2 if the weights satisfy the triangle inequality.

Acknowledgements

The author wishes to thank Francesco Maffioli and Edoardo Amaldi for drawing her attention to this problem and for generously devoting time to discussing its various aspects.

References

[1] A. Bley, M. Grotschel, R. Wessaly, Design of broadband virtual private networks: Model and heuristics for the B-win, Konrad-Zuse-Zentrum für Informationstechnik, Berlin, 1998, Preprint.
[2] D.M. Chickering, D. Geiger, D. Heckerman, On finding a cycle basis with a shortest maximal cycle, Inform. Process. Lett. 54 (1995) 55–58.
[3] J.D. Horton, A polynomial-time algorithm to find the shortest cycle basis of a graph, SIAM J. Comput. 16 (2) (1987) 358–366.
[4] G. Galbiati, On optimum cycle bases, in: J. Nesetril, M. Noy, O. Serra (Eds.), Electronic Notes in Discrete Mathematics, Vol. 10, Elsevier Science Publishers, Amsterdam, 2001.
[5] S.L. Lauritzen, D.J. Spiegelhalter, Local computations with probabilities on graphical structures and their applications to expert systems, J. Roy. Statist. Soc. Ser. B 50 (1988) 157–224.
[6] M.M. Syslo, On cycle bases of a graph, Networks 9 (1979) 123–132.