ELSEVIER

Information Processing Letters 63 (1997) 63–67

Information Processing Letters

An approximation algorithm for maximum packing of 3-edge paths

Refael Hassin *, Shlomi Rubinstein [1]

Department of Statistics and Operations Research, School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel

Received 16 December 1996; revised 15 May 1997
Communicated by S. Zaks

Keywords: Analysis of algorithms; Maximum set packing

1. Introduction

Let $G = (V,E)$ be a complete graph with node set V and edge set E. For $(u,v) \in E$ let $w(u,v) \geqslant 0$ be its weight. Assume that $|V| = n = 4k$ for some integer k. A packing of 3-edge paths is a set of k node-disjoint paths of three edges (and thus four nodes) each. The subject of this note is the problem of computing a packing of 3-edge paths with maximum total edge weight. The problem is NP-hard [5].

The problem is a special case of the general set packing problem considered in [1,2] and the general results there imply a $\frac{1}{3}$ bound on the performance ratio. In this note we prove that a simple algorithm guarantees a bound of $\frac{3}{4}$. We also present related observations on the maximum symmetric traveling salesman problem (Max_TSP).

2. Max packing of 3-edge paths

We start by considering a more general problem. Suppose we want to partition V into k node-disjoint paths with $c_1, \ldots, c_k$, edges respectively, of maximum weight (where $n = k + \sum c_i$). The following algorithm guarantees a factor of $\alpha(1 - k/n)$, where α is the performance guarantee available for solving Max_TSP. The algorithm of Fisher, Nemhauser and Wolsey [3] gives $\alpha = \frac{2}{3}$ and an improved bound has recently been obtained by Kosaraju, Park and Stein [7].

- Approximate Max_TSP with factor α. Let the edges in this solution be $e_1, \ldots, e_n$ in this cyclic order.
- For every $i = 1, \ldots, n$: Construct a solution in which the jth path ($j = 1, \ldots, k$) consists of the edges $e_{l(i,j)}, \ldots, e_{r(i,j)}$, where indices are mod n, $l(i,j) = i + c_1 + \cdots + c_{j-1} + j$, and $r(i,j) = l(i,j) + c_j - 1$.
- Output the solution with maximum total edge weight among the n solutions computed above.

The stated bound results from the following observations. The n solutions constructed by the procedure use each edge of the tour exactly $n - k$ times, so that the average solution has weight $(n - k)/n = 1 - k/n$ of the tour's weight. The weight of the maximal of these solutions is at least as that of the average one.

* Corresponding author. Email: hassin@math.tau.ac.il.
[1] Email: shlom@math.tau.ac.il.

PII S0020-0190(97)00097-5

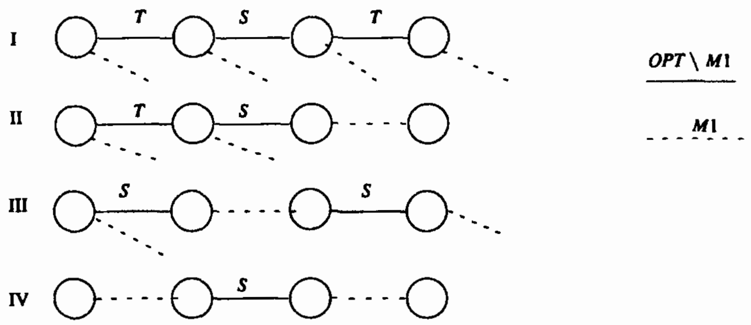

Fig. 1. Classification of *OPT* and partition of *OPT* \ *M*1.

Finally, the solution to Max_TSP is an upper bound on the solution of the 3-edge paths packing problem.

In the 3-edge paths packing problem, we have $k = \frac{1}{4}n$ so that the bound resulting from the above method is $\frac{3}{4}\alpha < \frac{3}{4}$. We will now suggest a different approach with a $\frac{3}{4}$ bound. We will denote by *OPT* an optimal solution and by *opt* its weight. Similarly, *APX* is an approximate solution and *apx* its weight.

We suggest the following algorithm.

- Compute in G a maximum weight perfect matching $M1$.
- Form a complete graph $G' = (V', E')$. The nodes of V' correspond to the edges of $M1$. The weight of $(u,v) \in E'$, where u corresponds to $(a_u, b_u) \in M1$ and v corresponds to $(a_v, b_v) \in M1$, is defined as $w'(u,v) = \max\{w(a_u,a_v), w(a_u,b_v), w(b_u,a_v), w(b_u,b_v)\}$.
- Compute a maximum weight perfect matching in G'. Let $M2$ be the edges corresponding to this matching (through the definition of w') in G.
- Output $APX = M1 \cup M2$.

Theorem 1. $apx \geqslant \frac{3}{4}opt$.

Proof. Partition *OPT* into four classes according to its intersection with $M1$, as described in Fig. 1.

We will now describe a process that constructs three matchings S, T and M with the following properties:

(1) S, T partition $OPT \setminus M1$.
(2) Each edge of T is adjacent to two edges of $M1 \setminus OPT$.
(3) $M \subset S$ corresponds to a matching in G'.
(4) $w(M) \geqslant \frac{1}{2}w(S)$.

We start the construction process with the initial sets S, T as follows (see Fig. 1):

I- and II-paths: Assign the middle edge to S and the other edges to T.

III-paths: Assign both edges to S.

IV-paths: Assign the edge to S.

Consider now the subgraph $H = (V, S \cup M1)$ of G. Since both S and $M1$ are matchings in G, H consists of a collection of node-disjoint paths and simple cycles.

Define the *S-length* of a path in H as the number of S-edges it contains. Call a path (or a cycle) *odd* if its S-length is odd. Otherwise, call it *even*. We observe that edges from II- and IV-paths are not contained in any cycle of H while each III-path contributes 2 to the S-length of each cycle or path component in H that intersects it.

We will describe now how odd cycles can be eliminated from H by changing the way S and T edges are defined for some I-paths. Suppose that H contains an odd cycle C. Since a III-path contributes 2 S-edges to at most one cycle that intersects it, C must contain an S-edge, say s, from a I-path. We now modify the sets S and T by moving the middle edge of this I-path to T and its end edges to S. A new odd cycle may be formed from the union of $C \setminus \{s\}$ and an odd path. Fig. 2 illustrates such a case. However, as we observed, the odd path contains an S-edge from a I-path. The process is repeated with that edge. After a finite number of steps, the odd cycle is eliminated. We repeat this process for each odd cycle in H.

We now form the edge set $M \subset S$ satisfying property (4).

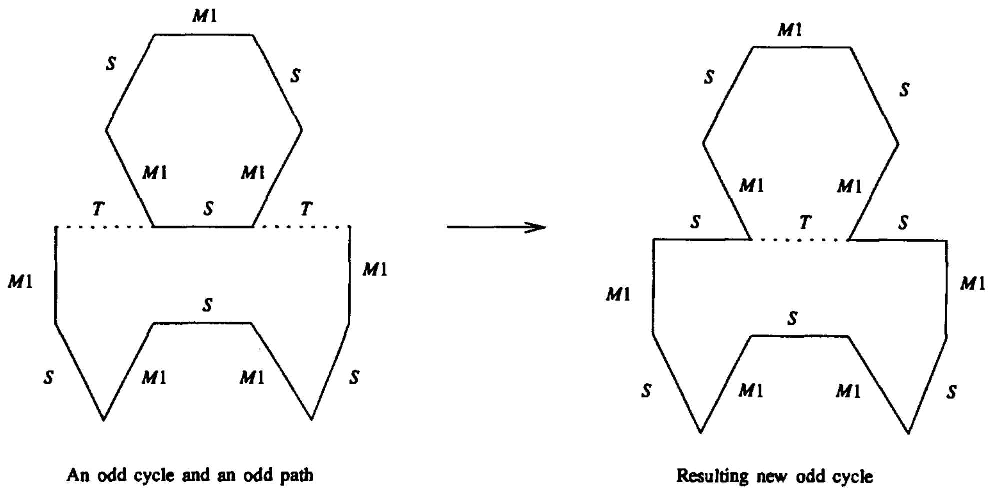

Fig. 2. Elimination of odd cycles.

Consider an even cycle in H. Its edges alternate between $M1$ and S. The edges of S in this cycle can be decomposed (alternately according to their order in the cycle) into two disjoint subsets, such that each subset forms with the edges of $M1$ in this cycle a set of 3-edge paths. We assign to M the subset of higher total weight.

Similarly, the edges of a path in H alternate between $M1$ and S, with its end edges belonging to $M1$. Its S-edges can be partitioned into two disjoint subsets each forming with $M1$ edges sets of 3-edge paths. We assign to M the subset of higher total weight. We end up with a subset $M \subset S$ satisfying property (4) and since $M2$ is an edge set of maximum weight of this type, it follows that also

$$w(M2) \geqslant \tfrac{1}{2}w(S). \tag{1}$$

We consider now the final partition S, T as constructed above. Each T-edge is adjacent to two edges of $M1 \setminus OPT$. It follows that the weight of $M1 \setminus OPT$ is at least as that of T, since otherwise replacing $M1 \setminus OPT$ by T will give a matching of weight greater than $w(M1)$. Thus,

$$w(M1 \setminus OPT) \geqslant w(T). \tag{2}$$

Since the S-edges are node-disjoint they form a (not necessarily perfect) matching. By the optimality of $M1$ (and non-negativity of the weights),

$$w(M1) \geqslant w(S). \tag{3}$$

From (1), (2), (3), and a trivial identity, we obtain

$$4w(M2) \geqslant 2w(S),$$
$$3w(M1 \setminus OPT) \geqslant 3w(T),$$
$$w(M1) \geqslant w(S),$$
$$3w(M1 \cap OPT) = 3w(M1 \cap OPT).$$

Summation gives

$$\begin{aligned} 4apx &= 4(w(M1) + w(M2)) \\ &\geqslant 3(w(S) + w(T) + w(M1 \cap OPT)) \\ &= 3opt, \end{aligned}$$

as claimed. □

Example 2. Consider 8 nodes on a cycle with edge weights 1,2,1,0,1,2,1,0 in this cyclic order, and all the edges not on the cycle are of zero weight. Clearly, $opt = 8$. A possible choice for $M1$ is the two edges of weight 2 and the two edges of weight 0 from the cycle. In this case $M2$ consists of two edges of unit weight. Thus,

$$apx = w(M1) + w(M2) = 4 + 2 = 6.$$

This example demonstrates that the bound of Theorem 1 is tight.

3. Relation to Max_TSP

We now determine the performance guarantee of the algorithm we used for packing 3-edge paths when applied to Max_TSP. Thus, we denote by *OPT* an optimal solution to this problem, and similarly for the other notation.

We observe that a tour can be covered three times by four 3-edge paths packings. Thus, an optimal solution to the 3-edge paths packing problem can be completed to a $\frac{3}{4}$ approximation for Max_TSP. Consequently we obtain as a corollary to Theorem 1 that $w(M1) + w(M2) \geqslant \frac{9}{16} opt$. We now strengthen this result:

Theorem 3. $w(M1) + w(M2) \geqslant \frac{5}{8} opt$.

Proof. Consider an optimal tour *OPT*. The edges of $OPT \setminus M1$ form node-disjoint paths, or *OPT* itself. Each such path contains at least one edge. It is possible to partition $OPT \setminus M1$ into disjoint subsets *S* and *T* so that the following properties hold:

(1) The edges of $T \cup (OPT \cap M1)$ are node disjoint.

(2) *S* consists of node disjoint 1-edge and 2-edge paths.

By the first property and maximality of $M1$ it follows that

$$w(M1 \setminus OPT) \geqslant w(T), \tag{4}$$

since otherwise, by replacing $M1 \setminus OPT$ by *T*, we get a matching with a greater weight than $M1$.

Let M_S be the subset of $M1 \setminus OPT$ of edges that have at least one end node incident to two *S* edges (that is, this end node is a "center" of a 2-edge path of *S*-edges). If we contract the edges of M_S (by identifying their end nodes), the edges of *S* define a simple graph with edge set *S* and maximum node degree of 4. Moreover, in this graph there are no edges connecting two nodes whose degree is 4 so that by a theorem of Fournier [4] (see also [6]), it is 4-edge colorable. Let *M* be the set of edges corresponding to the color class whose total edge length is maximal. Then, $w(M) \geqslant \frac{1}{4} w(S)$. Now, each edge in *M* connects two distinct edges of $M1$ and the 3-edge paths formed this way are node disjoint. Since $M2$ is a maximum weight subset of this type also $w(M2) \geqslant \frac{1}{4} w(S)$. With (4) we obtain

$$\begin{aligned} w(M2) &\geqslant \tfrac{1}{4} w(S) \\ &= \tfrac{1}{4}[w(OPT) - w(M1 \cap OPT) - w(T)] \\ &\geqslant \tfrac{1}{4}[opt - w(M1 \cap OPT) - w(M1 \setminus OPT)] \\ &= \tfrac{1}{4}[opt - w(M1)], \end{aligned}$$

or,

$$\tfrac{1}{4} w(M1) + w(M2) \geqslant \tfrac{1}{4} opt. \tag{5}$$

By assumption *n* is even, so that *OPT* can be partitioned into two edge-disjoint matchings. Thus, $w(M1) \geqslant \frac{1}{2} opt$, or $\frac{3}{4} w(M1) \geqslant \frac{3}{8} opt$. Adding this inequality to (5) we get

$$w(M1) + w(M2) \geqslant \tfrac{5}{8} opt. \qquad \square$$

Example 4. Consider for a positive integer *k*, a $3k$-node graph with a cycle of $3k$ edges whose weights are $1, 1, 2, 1, 1, 2, \ldots$ in cyclic order. All the other weights are 0. Then, $opt = 4k$. A maximum matching has a weight of $2k$ that can be achieved in several ways. Suppose that $M1$ selects the *k* edges of weight 2 together with $\frac{1}{2}k$ zero weight edges. Then, $M2$ cannot select more than one unit weight edge from each adjacent pair of such edges. Thus, $w(M2) = \frac{1}{2}k$ and $w(M1) + w(M2) = \frac{5}{2}k$. This shows that the bound proved above is the best possible.

One may consider a natural enhancement of the algorithm. After computing $M1$ and $M2$ continue the process by computing $M3$, a maximum weight matching of end nodes of the 3-edge paths obtained. Then compute $M4$ to match the end nodes of the resulting 7-edge paths. The tour is finally constructed from the union of $M1, \ldots, Ml$ for $l = 3, \ldots, \lceil \log n \rceil + 1$. Note that the $+1$ relates to a last edge needed to turn a Hamiltonian path into a cycle. However, the bound may at best improve to $\frac{2}{3}$ as can be verified by constructing examples with this ratio.

References

[1] E. Arkin and R. Hassin, On local search for weighted *k*-set packing, 1996.

[2] V. Bafna, B. Narayanan and R. Ravi, Nonoverlapping local alignments (weighted independent sets of axis parallel rectangles), in: *Proc. Workshop for Algorithms and Data Structures (WADS)*, 1995.

[3] M.L. Fisher, G.L. Nemhauser and L.A. Wolsey, An analysis of approximations for finding a maximum weight Hamiltonian circuit, *Oper. Res.* 27 (1979) 799–809.
[4] J.C. Fournier, Colorations des arêtes d'un graphe, *Cahiers Centre Études Rech. Opér.* 15 (1973) 311–314.
[5] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, 1978).
[6] T.R. Jensen and B. Toft, *Graph Coloring Problems* (Wiley, New York, 1995).
[7] C. Stein, Private communication, May 1997.