Available online at www.sciencedirect.com
ScienceDirect
Information Processing Letters 104 (2007) 216–219
Information Processing Letters
www.elsevier.com/locate/ipl

Simulated Annealing versus Metropolis for a TSP instance

Klaus Meer [1]

Syddansk Universitet Odense, Department of Mathematics and Computer Science, Campusvej 55, DK-5230 Odense M, Denmark

Received 16 January 2007; received in revised form 8 June 2007; accepted 21 June 2007

Available online 6 July 2007

Communicated by P.M.B. Vitányi

Abstract

In a recent paper [I. Wegener, Simulated Annealing beats Metropolis in combinatorial optimization, in: L. Caires, G.F. Italiano, L. Monteiro, C. Palamidessi, M. Yung (Eds.), Proc. ICALP 2005, in: LNCS, vol. 3580, 2005, pp. 589–601] Wegener gave a first natural example of a combinatorial optimization problem where for certain instances a Simulated Annealing algorithm provably performs better than the Metropolis algorithm for any fixed temperature. Wegener's example deals with a special instance of the Minimum Spanning Tree problem. In this short note we show that Wegener's technique as well can be used to prove a similar result for another important problem in combinatorial optimization, namely the Traveling Salesman Problem. The main task is to construct a suitable TSP instance for which SA outperforms MA when using the well known 2-Opt local search heuristic.

Keywords: Analysis of algorithms; Simulated Annealing; Metropolis algorithm; 2-Opt heuristic for TSP

1. Introduction

In [1] the question is discussed whether there are natural examples of combinatorial optimization problems for which a Simulated Annealing (SA) algorithm outperforms any Metropolis algorithm (MA), i.e., SA with a fixed temperature. Though artificial examples have been known the first natural example giving a positive answer to the above question only recently was provided by Wegener [4]. There it is shown that SA beats MA for all fixed temperatures on special instances of the Minimum Spanning Tree problem.

The goal of this note is to give a similar positive example for another well known combinatorial optimization problem, namely the TSP problem.The heuristic we analyze is 2-Opt (see [2] and below). The paper is organized as follows: In Section 2 we recall basic notions about the 2-Opt heuristic and efficiency measures necessary to make the paper self-contained. The main Section 3 provides the construction of the TSP instance for which the result holds. Section 4 briefly outlines the proofs which technically closely follow those in [4].

2. The setting

We suppose the reader to be familiar with SA and MA. Let us briefly describe these algorithms for the situation we are interested in, i.e., the 2-Opt local search heuristic for TSP. For a weighted graph $G = (V, E)$ with n vertices the search points are Hamiltonian cycles.

E-mail address: meer@imada.sdu.dk.

[1] Partially supported by the IST Programme of the European Community, under the PASCAL Network of Excellence, IST-2002-506778 and by the Danish Agency for Science, Technology and Innovation FNU. This publication only reflects the author's views.

0020-0190/$ – see front matter
doi:10.1016/j.ipl.2007.06.016

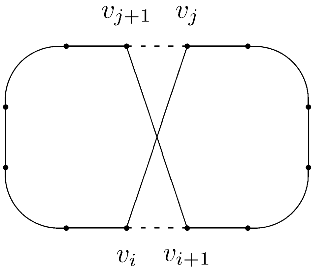

Fig. 1. 2-Opt heuristic.

Given as current search point such a cycle p which consists of an edge sequence $(v_1, v_2), (v_2, v_3), \ldots, (v_n, v_1)$ a local change is done as follows:

(1) Choose randomly two edges in the graph (for simplicity we allow also to choose twice the same); the probability to choose a particular pair of different edges thus is $\frac{2}{|E|^2}$.

(2.i) if the chosen edges are not two different edges of p do not change p; else

(2.ii) let $(v_i, v_{i+1}), (v_j, v_{j+1}), i < j$, be the chosen edges with weight sum s_1. If G does not contain the edges (v_i, v_j) and (v_{j+1}, v_{i+1}) do not change p. If G contains these two edges and their summed weight is s_2, then change p to $\tilde{p}$ according to Fig. 1.
This is done with probability 1 if $s_2 < s_1$ and with probability $\exp(-(s_2 - s_1)/T)$ else, where T is the current temperature used by SA or the fixed temperature used by MA, respectively.

Efficiency of SA or MA is measured as follows. For a longer discussion why to do it that way see [4]. Consider a randomized search heuristic $\mathcal{A}$ running a polynomial number $p(n)$ of rounds on a problem instance of size n. Denote by $S(n)$ the probability that $\mathcal{A}$ finds an optimal search point. $\mathcal{A}$ is called *successful* if $S(n) \geqslant \frac{1}{q(n)}$ for some polynomial q. It is *unsuccessful* if for all polynomials q the probability of success within $q(n)$ rounds is $o(n^{-k})$ for each $k \in \mathbb{N}$. Multistart versions of a successful $\mathcal{A}$ result in a success probability of order $1 - \exp(-\Omega(n^{\varepsilon}))$ for some $\varepsilon > 0$ in expected polynomial time.

3. A particular TSP instance for which SA beats MA

The main idea is to construct an instance for which it is easy to control the possible exchange steps when 2-Opt is applied. The given instance is academic in that the optimal tour easily is determined. For sake of simplicity we consider a non-complete graph. However, our results hold as well if lacking edges are included with a large enough weight such that the probability of ever including them in a tour will be sufficiently small.

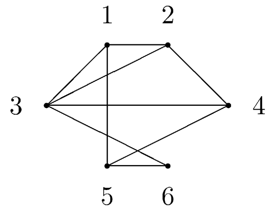

Fig. 2. Skeleton graph S.

To construct the particular instance graph we start from a skeleton graph S that contains 6 vertices and 9 edges. Copies of this graph (with different weights for the edges) will then be joined to obtain the input graph. The skeleton graph S is given in Fig. 2.

The following is easily verified.

Proposition 1. *In S there are precisely three paths starting in vertex* 1 *and leading to vertex* 6 *which traverse all the vertices in* $\{2, 3, 4, 5\}$ *exactly once. Those paths are*[2]

$$p_{\text{mid}} := 123456, \quad p_{\text{wst}} := 132456, \quad p_{\text{opt}} := 154236.$$

Applying the 2-Opt *heuristic in S it is possible to move from p_{wst} to either p_{mid} or p_{opt}. It is not possible to move in one local search step from p_{mid} to p_{opt}* (*or vice versa*). *Finally, the edge* 12 *only occurs in p_{mid}, the edge* 13 *only in p_{wst} and the edge* 15 *only in p_{opt}.*

The input graph G is constructed from several copies of S (see Fig. 3).

Here, both the H_i's and the L_i's, $1 \leqslant i \leqslant n$, are copies of S. For each i the nodes $v_{i1}, \ldots, v_{i6}$ are identified with nodes 1 to 6 of S, respectively; similarly for the $w_{i1}, \ldots, w_{i6}$. The resulting graph G contains $m := 20n$ edges and $12n$ vertices. Finally we attach weights to all edges of G:

- in the graphs H_i the edge (v_{i1}, v_{i2}) gets weight $m^3 - m^2 + 1$, and the edge (v_{i1}, v_{i3}) gets weight $m^3 + 1$;
- in the graphs L_i the edge (w_{i1}, w_{i2}) gets weight m and edge (w_{i1}, w_{i3}) gets weight $m + 1$;
- all other edges in G get weight 1.

The important observation for the analysis below is the following: In each *heavy* subgraph H_i the increases

[2] The meaning of the subscripts wst, mid and opt will become clear below when weights are assigned to the edges of S.

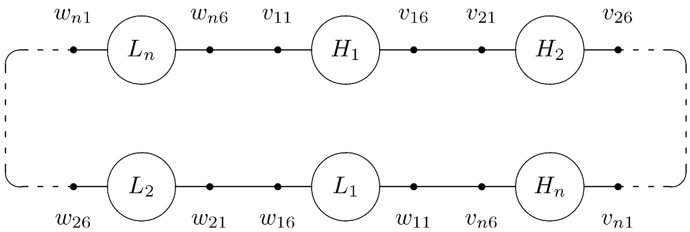

Fig. 3. Input graph G.

in weight from path p_{opt} to p_{mid} to p_{wst} are $m^3 - m^2$ and m^2, respectively. In each *light* subgraph L_i the corresponding increases are $m-1$ and 1, respectively.

Note also that our weights are chosen in such a way that the proofs from [4] can be applied with as few changes as possible.

4. Results and proofs

Here are the main results comparing SA and MA in the above setting:

Theorem 1. *Let G be as above and let T be a fixed temperature. Apply the Metropolis algorithm to G using temperature T and the* 2-Opt *heuristic. As starting point we choose the Hamiltonian tour that in all H_i and L_i consists of the corresponding paths p_{wst}. Then the probability to compute the minimal cost Hamiltonian tour in G within* $\exp(cm)$ *steps (where $c > 0$ is a small enough constant, $m = 20n$, and n the number of vertices in G) is at most* $\exp(-\Omega(m))$. *Thus MA is unsuccessful.*

Theorem 2. *Under the same assumptions as in Theorem* 1 *apply Simulated Annealing using the cooling schedule $T(1) := m^3, T(t+1) = \alpha \cdot T(t)$, where $\alpha := (1 - \frac{1}{cm^2})$ for $c > 0$ small enough. Then the probability to compute the minimal cost Hamiltonian tour in* $O(m^2 \cdot \log m)$ *rounds is at least $1 - \frac{1}{p(m)}$ for some polynomial $p, m := 20n$ as above. Thus SA is successful.*

In the remaining part we briefly outline the proofs. Since they closely follow those in [4] we just collect the necessary intermediate results together with a short hint how to establish them. The interested reader will be able to fill in details.

We start by collecting basic properties concerning certain transition probabilities for the 2-Opt heuristic in our setting. Since G has $m = 20n$ edges the probability to select one pair of different edges is $\frac{2}{m^2}$ (since we allow to choose twice the same edge). If the selected edges allow a 2-Opt exchange step the transition probabilities are as follows: In each H_i moving

- from p_{opt} to p_{wst} has probability $\frac{2}{m^2}\exp(-\frac{m^3}{T})$;
- from p_{wst} to p_{mid} has probability $\frac{2}{m^2}$; same for moving from p_{wst} to p_{opt};
- from p_{mid} to p_{wst} has probability $\frac{2}{m^2}\exp(-\frac{m^2}{T})$.

In each L_i moving

- from p_{opt} to p_{wst} has probability $\frac{2}{m^2}\exp(-\frac{m}{T})$;
- from p_{wst} to p_{mid} has probability $\frac{2}{m^2}$; same for moving from p_{wst} to p_{opt};
- from p_{mid} to p_{wst} has probability $\frac{2}{m^2}\exp(-\frac{1}{T})$.

The main technical tool in most of the proofs is Chernoff's bound, see [3].

Theorem 3 *(Chernoff). Let $X_1, \dots, X_n$ be independent Poisson trials such that* $\Pr(X_i = 1) =: p_i, 1 \leqslant i \leqslant n$, *where $p_i \in (0,1)$. Define $X := \sum X_i$ and $\mu := E[X]$.*

(a) *For* $0 < \delta \leqslant 1$ *it is* $\Pr(X < (1-\delta)\mu) < \exp(-\mu\frac{\delta^2}{2})$ (*lower tail*).
(b) *For* $0 < \delta < 2e - 1$ *it is* $\Pr(X < (1+\delta)\mu) < \exp(-\mu\frac{\delta^2}{4})$ (*upper tail*).
(c) *For* $\delta \geqslant 2e - 1$ *it is* $\Pr(X < (1+\delta)\mu) < 2^{-\mu\delta}$ (*upper tail*).

Proof of Theorem 1. (*Sketch*) Recall that the starting point is the tour which in all L_i and H_i takes path p_{wst}. For low temperatures $T \leqslant m$ and $1 \leqslant i \leqslant n$ define stochastic variables X_i as taking value 0 iff the first time where p_{wst} in H_i is changed it is changed into p_{opt}, and $X_i = 1$ if it is changed into p_{mid}. For $X = \sum X_i$ we have $\mu^* = E[X] = \frac{1}{2}n$. Chernoff's bound (lower tail with $\delta = \frac{2}{3}$) implies $\Pr(X < \frac{n}{6}) \leqslant \exp(-\frac{n}{9})$. Thus, with probability $1 - \exp(-\Omega(m))$ in the run of MA for at least $\frac{n}{6}$ of the H_i's the path p_{mid} is part of a used search point. The probability of repairing p_{mid} is $\leqslant \frac{2}{m^2}\exp(-m)$ since first an increasing step to p_{wst} has to be chosen. Repairing at least $\frac{n}{6}$ of these paths in an exponential number $\exp(\frac{m}{2})$ of rounds has a probability $\leqslant \exp(-\Omega(m))$; this follows from Chernoff by taking

X as the number of repaired heavy graphs in $\exp(\frac{m}{2})$ rounds and noticing that $E[X] \leqslant \exp(-\Omega(m))$.

Finally, for large temperatures $T \geqslant m$ consider the light graphs L_i; define X_t as number of L_i's after t steps in which p_{opt} is chosen. With $\Pr(X_{t+1} = a+1 \mid X_t = a) \leqslant \frac{2(n-a)}{m^2}$ and $\Pr(X_{t+1} = a-1 \mid X_t = a) \geqslant \frac{2a}{m^2} \cdot \exp(-\frac{m}{T}) \geqslant \frac{2a}{3m^2}$ it follows (using the gambler's ruin problem as in [4]) that the probability of increasing a from $\frac{10}{11}n$ to n (i.e. to optimize the tour through all L_i's) is at most $\exp(-\Omega(m))$.

Thus, for each fixed temperature T the probability that MA finds the optimal tour in $\exp(cm)$ rounds ($c > 0$ small enough) is bounded by $\exp(-\Omega(m))$. □

Proof of Theorem 2. (*Sketch*) Using the cooling schedule $T(k) := (1 - \frac{1}{cm^2})^k \cdot T(1), T(1) := m^3$ and c suitably chosen it is shown that SA optimizes with a high enough probability in $O(m^2 \ln(m))$ many steps the path through the H_i's, when $T \in [m^2, m^{5/2}]$. Thereafter, with high probability SA will not change any more the (optimal) path through the H_i's. Finally, for $T \in [1, \sqrt{m}]$ SA with high probability optimizes the path through the light subgraphs L_i.

More detailed: The cooling schedule guarantees SA to need $cm^2 \ln(m)$ rounds for reducing T by a factor of $\sqrt{m}$.

(1.1) For $T \in [m^2, m^{5/2}]$ the probability of destroying a p_{opt} in an H_i is at most $\frac{2}{m^2}\exp(-\sqrt{m})$. Consequently, the probability of not destroying a p_{opt} within k rounds is at least $1 - \frac{2k}{m^2}\exp(-\frac{1}{2}\sqrt{m})$ and thus for any fixed polynomial $k(m)$ bounded from below by $1 - \exp(-\Omega(\sqrt{m}))$ for large enough m.

(1.2) For each H_i the probability of changing p_{mid} to p_{wst} is $\frac{2}{m^2}\exp(-\frac{m^2}{T})$ and the probability of changing p_{wst} (either to p_{mid} or p_{opt}) is $\frac{4}{m^2}$. Thus the expected number of changes of a non-optimal path in an H_i within $t := cm^2 \ln m$ steps is $\geqslant \frac{2c}{e}\ln m$. Chernoff's bound implies that the probability of changing in an H_i within t rounds at least $c' \ln m$ times a non-optimal path is bounded from below by $1 - \exp(-\Omega(\ln m)) = 1 - m^{-\text{const}}$ for a small enough $c' > 0$. By choosing c and c' suitably the above constant const > 0 in m's exponent can be made arbitrarily large.

(1.3) The probability of changing in H_i either p_{wst} or p_{mid} within two consecutive steps to p_{opt} is $\geqslant (\frac{2}{m^2})^2 \cdot \frac{1}{e}$. Now (1.2) implies that the probability of not at least once changing within $\frac{t}{2}$ sequences of two steps a non-optimal path into p_{opt} is $\leqslant (\frac{1}{m^2 e})^{c' \ln m}$ and thus smaller than m^{-k} for arbitrary k, m large enough.

Altogether, the probability of not optimizing the paths through the H_i's in a polynomial number of steps is bounded by $\frac{1}{p_1(m)}$ for an arbitrary polynomial p_1, m large enough.

Now towards the light subgraphs L_i and low temperatures:

(2.0) The probability to destroy a p_{opt} in an H_i for a temperature $T \leqslant m^{2.5}$ is $\leqslant \frac{2}{m^2}\exp(-\sqrt{m})$; in t steps it follows that the probability of destroying a p_{opt} in an H_i is at most $\frac{1}{p_2(m)}$ for a suitable polynomial p_2.

Next consider $T \in [1, \sqrt{m}]$:

(2.1) Similarly, for $T \leqslant \sqrt{m}$ the probability of destroying a p_{opt} in an L_i is $\leqslant \frac{2}{m^2}\exp(-\sqrt{m})$.

(2.2) The probability to change either a p_{wst} or a p_{mid} in an L_i is $\geqslant \frac{2}{m^2 e}$.

(2.3) Using (2.2) the probability to change a non-optimal path into an optimal one within two consecutive steps is $\geqslant \frac{4}{m^4 e}$.

Applying Chernoff's bound as under (1.2) and (1.3) finally implies that SA optimizes the light subgraphs for low temperatures with a probability $\geqslant 1 - \frac{1}{p_3(m)}$ for a suitable polynomial p_3. □

References

[1] M. Jerrum, A. Sinclair, The Markov chain Monte Carlo method. An approach to approximate counting and integration, in: D. Hochbaum (Ed.), Approximation Algorithms for NP-Hard Problems, PWS Publishing, 1996, pp. 482–522.

[2] D.S. Johnson, L.A. McGeoch, The traveling salesman problem: A case study, in: E. Aarts, J.K. Lenstra (Eds.), Local Search in Combinatorial Optimization, Princeton University Press, 2003, pp. 215–310.

[3] R. Motwani, P. Raghavan, Randomized Algorithms, Cambridge Univ. Press, 1995.

[4] I. Wegener, Simulated Annealing beats Metropolis in combinatorial optimization, in: L. Caires, G.F. Italiano, L. Monteiro, C. Palamidessi, M. Yung (Eds.), Proc. ICALP 2005, in: LNCS, vol. 3580, 2005, pp. 589–601.