



Average-case approximation ratio of the 2-opt algorithm for the TSP

Christian Engels, Bodo Manthey*

Saarland University, Department of Computer Science, Postfach 151150, 66041 Saarbrücken, Germany

ARTICLE INFO

Article history:

Received 25 November 2008

Accepted 19 December 2008

Available online 30 January 2009

Keywords:

Traveling salesman problem

2-opt

Average-case analysis

Approximation ratio

ABSTRACT

We show that the 2-opt heuristic for the traveling salesman problem achieves an expected approximation ratio of roughly $O(\sqrt{n})$ for instances with n nodes, where the edge weights are drawn uniformly and independently at random.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The *traveling salesman problem* (TSP) is one of the most important problems in combinatorial optimization: Given a complete graph with edge weights, the goal is to find a Hamiltonian cycle (also called a *tour*) of minimum weight. 2-opt is probably the most widely used local search heuristic for the TSP. It incrementally improves an initial tour by exchanging two edges of the tour with two other edges, until a local optimum is reached. More formally: Let w be the edge weights. If $\{a, b\}$ and $\{c, d\}$ are two edges of the current cycle such that a, b, c, d appear in that order in the cycle, then we can improve the tour by replacing $\{a, b\}$ and $\{c, d\}$ by $\{a, c\}$ and $\{b, d\}$, provided that $w(\{a, c\}) + w(\{b, d\}) < w(\{a, b\}) + w(\{c, d\})$. On randomly generated instances, 2-opt comes within a small percentage of the global optimum [1]. Chandra et al. [2] analyzed 2-opt's worst-case approximation ratio: On instances that fulfil the triangle inequality it is $O(\sqrt{n})$, where n is the number of nodes. This means that the worst local optimum is within a factor of $O(\sqrt{n})$ of the global optimum. For Euclidean instances, 2-opt's worst-case approximation ratio is $O(\log n)$. Englert et al. [3] showed that the expected approximation ratio is $O(\sqrt[4]{\phi})$ for d -dimensional Euclidean instances that are drawn according to density functions bounded by ϕ .

To explain the good performance of *subtour patching* for TSP, Karp [4] analyzed its approximation performance in a simple probabilistic setting: All edge weights are drawn uniformly and independently at random from the interval $[0, 1]$. In this setting, the triangle inequality is usually not fulfilled. In the worst case, TSP

cannot be approximated at all without triangle inequality, and also 2-opt cannot provide any approximation guarantee.

We use Karp's probabilistic model [4] to analyze the approximation performance of 2-opt. Let WLO_n be the weight of the worst, i.e., heaviest, locally optimal tour of a graph of n nodes with random edge weights, and let OPT_n be the weight of an optimum tour. We prove an upper bound for WLO_n that holds with high probability (Theorem 2), and we bound the expected approximation ratio (Theorem 4).

2. Approximation performance of 2-opt

First, we bound the probability that a specific tour is locally optimal, provided that it contains enough "heavy" edges. This lemma is the crucial ingredient for Theorem 2.

Lemma 1. *Let H be any fixed Hamiltonian cycle, and let $\eta \in (0, 1]$. Assume that H contains at least $m \geq 4$ edges of weight at least η . Then*

$$\mathbb{P}(H \text{ is locally optimal}) \leq \exp(-\eta^2 m^2 / 16).$$

Proof. The tour H contains m heavy edges. For simplicity, we assume that m is even. (Odd m can be handled similarly.) Thus, we can find at least $m/2$ pairwise non-adjacent edges among them. Consider any two edges e, e' of them. Let f, f' be the two replacement edges for e and e' . If both $w(f) < \eta$ and $w(f') < \eta$, then surely replacing e, e' by f, f' improves the tour, and H would not be locally optimal. By independence, $\mathbb{P}(w(f), w(f') < \eta) = \eta^2$.

There are $\binom{m/2}{2} = \frac{m^2 - 2m}{8} \geq \frac{m^2}{16}$ possible choices for e and e' , and all of them result in different replacement candidates f and f' .

* Corresponding author.

E-mail address: manthey@cs.uni-sb.de (B. Manthey).

(The inequality holds since $m \geq 4$.) This yields

$$\mathbb{P}(H \text{ is locally optimal}) \leq (1 - \eta^2)^{m^2/16} \leq \exp(-\eta^2 m^2/16). \quad \square$$

Theorem 2. For any $c > 0$, we have

$$\mathbb{P}(\text{WLO}_n \geq (17 + c) \cdot \sqrt{n} \cdot (\log n)^{3/2}) \leq \exp(-cn \log n).$$

Proof. Let $\eta = (17 + c) \cdot \sqrt{\log n/n}$. Let $m_i = 2^{-i}n$, and let $\eta_i = 2^i \eta$. If $i \geq \log n$, then $m_i < 4$ and $\eta_i > 1$. Thus, it suffices to consider $i \in \{0, \dots, \log n - 1\}$ in the following. If for all i , a tour H does not contain more than m_i edges of weight at least η_i , then

$$w(H) \leq \sum_{i=0}^{\log n - 1} m_i \eta_{i+1} = (17 + c) \cdot (\log n)^{3/2} \cdot \sqrt{n}.$$

Fix any tour H . The probability that H is locally optimal, provided that H contains at least m_i edges of weight at least η_i for some fixed i , is $\exp(-\eta^2 n^2/16)$ by Lemma 1. By Boole's inequality, the probability that H is locally optimal, provided that there exists an $i \in \{0, \dots, \log n - 1\}$ for which H contains at least m_i edges of weight at least η_i , is at most $\log n \cdot \exp(-\eta^2 n^2/16)$. Again by Boole's inequality, the probability that one of the $n!$ possible tours is locally optimal, provided that it contains at least m_i edges of weight η_i for some i , is at most

$$n! \cdot \log n \cdot \exp(-\eta^2 n^2/16) \leq \exp(-cn \log n),$$

which is the desired bound. \square

Since OPT_n and WLO_n are not independent, we do not have $\mathbb{E}\left(\frac{\text{WLO}_n}{\text{OPT}_n}\right) = \frac{\mathbb{E}(\text{WLO}_n)}{\mathbb{E}(\text{OPT}_n)}$. In order to bound the expected approximation ratio, we need the following lower bound for OPT_n . Combining this lower bound with Theorem 2 yields our second result (Theorem 4).

Lemma 3. For any $n \geq 2$ and $c \in [0, 1]$, we have $\mathbb{P}(\text{OPT}_n \leq c) \leq c^n$.

Proof. Fix any tour H . By independence, $\mathbb{P}(w(H) \leq c) = \frac{c^n}{n!}$. (This can be proved by induction on n .) Using Boole's inequality, the probability that there exists a tour H with $w(H) \leq c$ is bounded as claimed. \square

Theorem 4. We have

$$\mathbb{E}\left(\frac{\text{WLO}_n}{\text{OPT}_n}\right) \in O(\sqrt{n} \cdot (\log n)^{3/2}).$$

Proof. Assume that $\text{WLO}_n / \text{OPT}_n$ exceeds $2c^2 \cdot \sqrt{n} \cdot (\log n)^{3/2}$ for $c \geq 17$. Then $\text{WLO}_n \geq (17 + c) \cdot \sqrt{n} \cdot (\log n)^{3/2}$ or $\text{OPT}_n \leq \frac{1}{c}$. The probability that any of these events happens is at most $c^{-n} + \exp(-cn \log n) = P_c$. By substituting $x = 2c^2$, we obtain

$$\begin{aligned} \mathbb{E}\left(\frac{\text{WLO}_n}{\text{OPT}_n}\right) &\leq \sqrt{n} \cdot (\log n)^{3/2} \cdot \int_{578}^{\infty} P_{\sqrt{x/2}} dx \\ &\quad + O(\sqrt{n} \cdot (\log n)^{3/2}) \in O(\sqrt{n} \cdot (\log n)^{3/2}) \end{aligned}$$

since the integral evaluates to $O(1)$ for sufficiently large n . \square

References

- [1] David S. Johnson, Lyle A. McGeoch, The traveling salesman problem: A case study, in: Emile Aarts, Jan Karel Lenstra (Eds.), Local Search in Combinatorial Optimization, John Wiley & Sons, 1997 (Chapter 8).
- [2] Barun Chandra, Howard Karloff, Craig Tovey, New results on the old k -opt algorithm for the traveling salesman problem, SIAM Journal on Computing 28 (6) (1999) 1998–2029.
- [3] Matthias Englert, Heiko Röglin, Berthold Vöcking, Worst case and probabilistic analysis of the 2-opt algorithm for the TSP, in: Proc. of the 18th Ann. ACM-SIAM Symp. on Discrete Algorithms, SIAM, 2007, pp. 1295–1304.
- [4] Richard M. Karp, A patching algorithm for the nonsymmetric traveling-salesman problem, SIAM Journal on Computing 8 (4) (1979) 561–573.