Worst Case and Probabilistic Analysis of the 2-Opt Algorithm for the TSP

Matthias Englert    Heiko Röglin    Berthold Vöcking

Department of Computer Science
RWTH Aachen

Oberwolfach 2007
Traveling Salesperson Problem

Traveling Salesperson Problem (TSP)

- Input: weighted (complete) graph
  \[ G = (V, E, d) \text{ with } d : E \to \mathbb{R}. \]
Traveling Salesperson Problem

Traveling Salesperson Problem (TSP)

- Input: weighted (complete) graph $G = (V, E, d)$ with $d : E \rightarrow \mathbb{R}$.
- Goal: Find Hamiltonian cycle of minimum length.
Theoretical Results

- General TSP
  - Strongly NP-hard.
  - Not approximable \textit{within any polynomial factor}.

- Metric TSP
  - Strongly NP-hard.
  - $3/2$-approximation \cite{Christofides1976}
  - APX-hard: lower bound of $220/219$ \cite{PapadimitriouVempala2000}

- Euclidean TSP
  - Cities $\subset \mathbb{R}^d$
  - Strongly NP-hard ($\Rightarrow$ no FPTAS) \cite{Papadimitriou1977}
  - PTAS exists \cite{Arora1996, Mitchell1996}.
Theoretical Results

- **General TSP**
  - Strongly NP-hard.
  - Not approximable within any polynomial factor.

- **Metric TSP**
  - Strongly NP-hard.
  - 3/2-approximation [Christofides, (1976)]
Theoretical Results

- **General TSP**
  - Strongly NP-hard.
  - Not approximable within any polynomial factor.

- **Metric TSP**
  - Strongly NP-hard.
  - $3/2$-approximation [Christofides, (1976)]

- **Euclidean TSP**
  - Cities $\subset \mathbb{R}^d$
  - Strongly NP-hard ($\Rightarrow$ no FPTAS) [Papadimitriou (1977)]
  - PTAS exists [Arora (1996), Mitchell (1996)].
Numerous experimental studies.

- TSPLIB contains “real-world” and random (Euclidean) instances.
- DIMACS Implementation Challenge [Johnson and McGeoch (2002)].
Experimental Results

- Numerous experimental studies.
  - TSPLIB contains “real-world” and random (Euclidean) instances.
  - DIMACS Implementation Challenge [Johnson and McGeoch (2002)].

Some conclusions:

- Worst-case results are often too pessimistic.
- The PTAS is too slow on large scale instances.
- The most successful algorithms (w.r.t. quality and running time) in practice rely on local search.
Start with an arbitrary tour.
2-Opt Heuristic

1. Start with an arbitrary tour.
2. Remove two edges from the tour.
2-Opt Heuristic

1. Start with an arbitrary tour.
2. Remove two edges from the tour.
3. Complete the tour by two other edges.
2-Opt Heuristic

1. Start with an arbitrary tour.
2. Remove two edges from the tour.
3. Complete the tour by two other edges.
4. Repeat steps 2 and 3 until no local improvement is possible anymore.
Why 2-Opt?

Experiments on Random Euclidean Instances

[Johnson and McGeoch (2002)]

Approximation Ratio

- Christofides (for $n \leq 10^5$): $\approx 1.1$
- 2-Opt (for $n \leq 10^6$): $\approx 1.05$
Why 2-Opt?

Experiments on Random Euclidean Instances

[Johnson and McGeoch (2002)]

Approximation Ratio

- Christofides (for $n \leq 10^5$): $\approx 1.1$
- 2-Opt (for $n \leq 10^6$): $\approx 1.05$

Number of Local Improvements of 2-Opt

- Greedy Starts: Probably $O(n)$
- Random Starts: Probably $O(n \log n)$
Running time of 2-Opt: Known and New Results

<table>
<thead>
<tr>
<th></th>
<th>General TSP</th>
<th>Euclidean metric</th>
<th>Manhattan metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>smoothed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>worst-case</td>
<td>$2^{\Omega(n)}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Average-case results: [Chandra, Karloff, Tovey (1999)].
Worst-case results: [Lueker (1975)].
## Running time of 2-Opt: Known and New Results

<table>
<thead>
<tr>
<th></th>
<th>General TSP</th>
<th>Euclidean metric</th>
<th>Manhattan metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>smoothed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>worst-case</td>
<td>$2^{\Omega(n)}$</td>
<td>$2^{\Omega(n)}$</td>
<td>$2^{\Omega(n)}$</td>
</tr>
</tbody>
</table>

Average-case results: [Chandra, Karloff, Tovey (1999)].
Worst-case results: [Lueker (1975)].
Our results.
### Running time of 2-Opt: Known and New Results

<table>
<thead>
<tr>
<th></th>
<th>General TSP</th>
<th>Euclidean metric</th>
<th>Manhattan metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
<td></td>
<td>$\tilde{O}(n^{10})$</td>
<td>$\tilde{O}(n^6)$</td>
</tr>
<tr>
<td>smoothed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>worst-case</td>
<td>$2^{\Omega(n)}$</td>
<td>$2^{\Omega(n)}$</td>
<td>$2^{\Omega(n)}$</td>
</tr>
</tbody>
</table>

Average-case results: [Chandra, Karloff, Tovey (1999)].  
Worst-case results: [Lueker (1975)].  
Our results.
Running time of 2-Opt: Known and New Results

<table>
<thead>
<tr>
<th></th>
<th>General TSP</th>
<th>Euclidean metric</th>
<th>Manhattan metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
<td>$n^{3+o(1)}$</td>
<td>$\tilde{O}(n^{10})$</td>
<td>$\tilde{O}(n^6)$</td>
</tr>
<tr>
<td>smoothed</td>
<td></td>
<td>$\tilde{O}(n^{4.33})$ [(\tilde{O}(n^{3.83}))]</td>
<td>$\tilde{O}(n^4)$ [(\tilde{O}(n^{3.5}))]</td>
</tr>
<tr>
<td>worst-case</td>
<td>$2\Omega(n)$</td>
<td>$2\Omega(n)$</td>
<td>$2\Omega(n)$</td>
</tr>
</tbody>
</table>

Average-case results: [Chandra, Karloff, Tovey (1999)].
Worst-case results: [Lueker (1975)].
Our results.
Introduction

Lower Bound

Upper Bound

Extensions and Open Problems
Theorem

For every $n \in \mathbb{N}$, there is a graph in the Euclidean plane with $8n$ vertices on which 2-Opt can make $2^{n+3} - 14$ steps.

Possible States of a Gadget:

(Long,Long), (Long,Short), (Short,Long), (Short,Short)
\((\text{Long,Long}) = 0\)

\((\text{Long,Short}) = 1\)

\((\text{Short,Short}) = 2\)
Lower Bound

\[(\text{Long}, \text{Long}) = 0\]

\[(\text{Long}, \text{Short}) = 1\]

\[(\text{Short}, \text{Short}) = 2\]

\[
\begin{array}{cc}
0 & 2 \\
\end{array}
\]
Lower Bound

\begin{align*}
(Long, Long) &= 0 \\
(Long, Short) &= 1 \\
(Short, Short) &= 2
\end{align*}
Lower Bound

\[(\text{Long,Long}) = 0\]

\[(\text{Long,Short}) = 1\]

\[(\text{Short,Short}) = 2\]
Lower Bound

\((\text{Long},\text{Long}) = 0\)

\((\text{Long},\text{Short}) = 1\)

\((\text{Short},\text{Short}) = 2\)
Lower Bound

\[(\text{Long},\text{Long}) = 0\]

\[(\text{Long},\text{Short}) = 1\]

\[(\text{Short},\text{Short}) = 2\]

0 \[\rightarrow\] 2
1 \[\downarrow\] 0
2 \[\downarrow\] 1
2 \[\downarrow\]

Lower Bound

\[
\begin{align*}
(Long, Long) &= 0 \\
(Long, Short) &= 1 \\
(Short, Short) &= 2
\end{align*}
\]
Lower Bound

\[(\text{Long,Long}) = 0\]

\[(\text{Long,Short}) = 1\]

\[(\text{Short,Short}) = 2\]

---

Gadget \(G_i\) is reset \(i - 1\) times to \((\text{Long,Long}) = 0\).
Gadget $G_i$ is reset $2^{i-1}$ times to $(\text{Long,Long}) = 0$. 
Euclidean Embedding of the Gadgets
Upper Bound

Introduction

Lower Bound

Upper Bound

Extensions and Open Problems
Theorem

Assume that \( n \) points are placed \textit{independently, uniformly} at random in the unit square \([0, 1]^2\). The expected number of 2-Opt steps is bounded by \( O(n^{4+1/3} \cdot \log n) \) (for every initial tour and every pivot rule).
Theorem

The expected number of 2-Opt steps is bounded by \( O(n^7 \log^2 n) \).
Theorem

The expected number of 2-Opt steps is bounded by $O(n^7 \log^2 n)$.

Proof.

- Consider a 2-Opt step $(e_1, e_2) \rightarrow (e_3, e_4)$.
- $\Delta(e_1, e_2, e_3, e_4) = l(e_1) + l(e_2) - l(e_3) - l(e_4)$.
Theorem

The expected number of 2-Opt steps is bounded by $O(n^7 \log^2 n)$.

Proof.

- Consider a 2-Opt step $(e_1, e_2) \rightarrow (e_3, e_4)$.
- $\Delta(e_1, e_2, e_3, e_4) = l(e_1) + l(e_2) - l(e_3) - l(e_4)$.
- $\Delta = \min_{e_1, e_2, e_3, e_4} |\Delta(e_1, e_2, e_3, e_4)|$. 
Theorem

The expected number of 2-Opt steps is bounded by $O(n^7 \log^2 n)$.

Proof.

- Consider a 2-Opt step $(e_1, e_2) \rightarrow (e_3, e_4)$.
- $\Delta(e_1, e_2, e_3, e_4) = l(e_1) + l(e_2) - l(e_3) - l(e_4)$.

\[
\Delta = \min_{e_1, e_2, e_3, e_4} |\Delta(e_1, e_2, e_3, e_4)|.
\]

\[
\# \text{ 2-Opt Steps} \leq \frac{\sqrt{2n}}{\Delta}.
\]
Theorem

*The expected number of 2-Opt steps is bounded by* $O(n^7 \log^2 n)$.

**Proof.**

- Consider a 2-Opt step $(e_1, e_2) \rightarrow (e_3, e_4)$.
- $\Delta(e_1, e_2, e_3, e_4) = l(e_1) + l(e_2) - l(e_3) - l(e_4)$.

$$\Delta = \min_{e_1, e_2, e_3, e_4} |\Delta(e_1, e_2, e_3, e_4)|.$$

- $\# \text{ 2-Opt Steps} \leq \frac{\sqrt{2n}}{\Delta}$.

- Bound $\Delta$ by a union bound: There are $O(n^4)$ different 2-Opt steps, analyze $\Delta(e_1, e_2, e_3, e_4)$ for one of them. $\Rightarrow \Delta \approx 1/(n^4 \log n)$. 

Heiko Röglin (RWTH Aachen)

Worst Case and Prob. Analysis of 2-Opt

Oberwolfach 2007
The bound is too pessimistic: Not every step yields the smallest possible improvement $\Delta \approx 1/(n^4 \log n)$. 
Idea for Improvement

- The bound is too pessimistic: Not every step yields the smallest possible improvement $\Delta \approx 1/(n^4 \log n)$.

- Consider two consecutive steps: They yield $\Delta + \Delta_2 > 2\Delta$. 
The bound is too pessimistic: Not every step yields the smallest possible improvement $\Delta \approx 1/(n^4 \log n)$.

Consider two consecutive steps: They yield $\Delta + \Delta_2 > 2\Delta$.

Consider linked pair: $(e_1, e_2) \rightarrow (e_3, e_4)$ and $(e_3, e_5) \rightarrow (e_6, e_7)$. 
Idea for Improvement

- The bound is too pessimistic: Not every step yields the smallest possible improvement $\Delta \approx 1/(n^4 \log n)$.
- Consider two consecutive steps: They yield $\Delta + \Delta_2 > 2\Delta$.
- Consider linked pair: $(e_1, e_2) \rightarrow (e_3, e_4)$ and $(e_3, e_5) \rightarrow (e_6, e_7)$.
- Sequence of $t$ consecutive steps, contains $\Omega(t)$ linked pairs:
The bound is too pessimistic: Not every step yields the smallest possible improvement $\Delta \approx 1/(n^4 \log n)$.

Consider two consecutive steps: They yield $\Delta + \Delta_2 > 2\Delta$.

Consider linked pair: $(e_1, e_2) \rightarrow (e_3, e_4)$ and $(e_3, e_5) \rightarrow (e_6, e_7)$.

Sequence of $t$ consecutive steps, contains $\Omega(t)$ linked pairs:

$$\begin{align*}
S_1 & \rightarrow S_2 \rightarrow S_3 \rightarrow S_4 \rightarrow S_5 \rightarrow S_6 \rightarrow S_7 \rightarrow S_8 \rightarrow S_9 \\
(S_1, S_4) & \rightarrow (S_2, S_5) \rightarrow (S_6, S_9)
\end{align*}$$

$\Delta_{\text{Linked}} \approx 1/(n^{3+1/3} \log^{2/3} n)$. 
1 Introduction

2 Lower Bound

3 Upper Bound

4 Extensions and Open Problems
Smoothed Analysis

Each point $i \in \{1, \ldots, n\}$ is chosen independently according to a probability density $f_i : [0, 1]^2 \rightarrow [0, \phi]$. 
Smoothed Analysis

Each point $i \in \{1, \ldots, n\}$ is chosen independently according to a probability density $f_i : [0, 1]^2 \to [0, \phi]$. 

![Diagram showing the distribution of points with 1/√φ scale]
## Smoothed Analysis

<table>
<thead>
<tr>
<th></th>
<th>General TSP</th>
<th>Euclidean metric</th>
<th>Manhattan metric</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>average</strong></td>
<td>$n^{3+o(1)}$</td>
<td>$\tilde{O}(n^{4.33})$</td>
<td>$\tilde{O}(n^4)$</td>
</tr>
<tr>
<td><strong>smoothed</strong></td>
<td>$m \cdot n^{1+o(1)} \cdot \phi$</td>
<td>$\tilde{O}(n^{4.33} \cdot \phi^{2.67})$</td>
<td>$\tilde{O}(n^4 \cdot \phi)$</td>
</tr>
<tr>
<td><strong>worst-case</strong></td>
<td>$2^{\Omega(n)}$</td>
<td>$2^{\Omega(n)}$</td>
<td>$2^{\Omega(n)}$</td>
</tr>
</tbody>
</table>
Approximation Ratio

- Worst Case: $O(\log n)$
- Worst Case: $\Omega(\log n / \log \log n)$
Approximation Ratio

- Worst Case: $O(\log n)$
- Worst Case: $\Omega(\log n / \log \log n)$
- Average Case: $O(1)$
Approximation Ratio

- Worst Case: $O(\log n)$
- Worst Case: $\Omega(\log n / \log \log n)$
- Average Case: $O(1)$
- Smoothed: $O(\sqrt{\phi})$
Open Problems

Worst-Case Analysis

- Analyze the **diameter** of the 2-Opt state graph.
- Analyze **particular pivot rules** like “largest improvement”.

Probabilistic Analysis

- Show **exact bounds** on the running time of 2-Opt and $k$-Opt.
- Show **small constant approximation ratio** for 2-Opt on random Euclidean instances.
Thanks!
Questions?