Constructing a Small Compact Binary Model for the Travelling Salesman Problem

J. Fabian Meier

Abstract

A variety of formulations for the Travelling Salesman Problem as Mixed Integer Program have been proposed. They contain either non-binary variables or the number of constraints and variables is large. We want to give a new formulation that consists solely of binary variables; the number of variables is of order \(O(n^2 \ln(n)^2)\) and the number of constraints is of order \(O(n^2 \ln(n))\).

Keywords: Integer Program, Landau function, Traveling Salesman Problem

1. Introduction

The Travelling Salesman Problem is the classical discrete optimization problem. The main difficulty in formulating the Travelling Salesman Problem as (Mixed) Integer Program is the avoidance of subtours. The most prominent approaches to avoid subtours are the subtour elimination constraints of Dantzig \[DFJ54\] (which are exponentially many) and the Miller-Tucker-Zemlin formulation \[MTZ60\] which has few variables, but weak constraints with real-valued variables. The first model is numerically superior. In the following section we will define a small model formulation and use the subsequent sections to give an estimate of its size.

2. The Model Formulation

As usual, we consider as set \(V\) of \(n\) given points with distances \(c_{ij}, i, j \in V\). For the construction of the model, we need a virtual split of one of the vertices (we call it 1) into a source and a sink, which we call \(1\) and \(\overline{1}\). This is necessary because the following constraints for subtour elimination forbid any kind of circle. To avoid excessive writing, we define \(\overline{V}\) to be all vertices except \(\overline{1}\) and vice versa for \(\overline{V}\). Let \(E\) be the resulting set of arcs, i.e. all pairs \((i, j)\) with \(i \in \overline{V}, j \in V, i \neq j\) except for \((1, \overline{1})\).
Following the usual model definition, we define the binary variables $x_{ij}$ to be one if the arc $i \rightarrow j$ is used and zero if not. We use the standard objective function and the two well-known constraints:

$$\text{Min } \sum_{(i,j) \in E} c_{ij} x_{ij}$$

$$\sum_{i: (i,j) \in E} x_{ij} = 1 \quad \forall j \in V$$  \hspace{1cm} (1)

$$\sum_{j: (i,j) \in E} x_{ij} = 1 \quad \forall i \in V$$  \hspace{1cm} (2)

$$x_{ij} \in \{0,1\} \quad \forall (i,j) \in E$$  \hspace{1cm} (3)

To construct our model formulation, we use some simple number theory. Let $p_1, \ldots, p_\alpha$ be pairwise relative prime numbers with $\prod_k p_k \geq n$. Then we define, for every $(i,j) \in E$, the following variables:

$$x^1_{ij}, x^1_{ji}, \ldots, x^1_{ij}$$

$$x^2_{ij}, x^2_{ji}, \ldots, x^2_{ij}$$

$$\vdots$$

$$x^\alpha_{ij}, x^\alpha_{ji}, \ldots, x^\alpha_{ij}$$

Because we have $|E| = n(n-1)$, we get $n(n-1)(\sum_k p_k + 1)$ variables in total. The additional constraints are the following:

$$\sum_a x^{k,a}_{ij} = x_{ij} \quad \forall (i,j) \in E, k \leq \alpha$$  \hspace{1cm} (4)

$$\sum_{i: (i,j) \in E} x^{k,a}_{ij} = \sum_{i: (j,i) \in E} x_{ji}^{k,(a+1 \mod p_k)} \quad \forall j \in V \cap \bar{V}, k \leq \alpha, a \leq p_k - 1$$  \hspace{1cm} (5)

The number of constraints is even smaller than the number of variables, in total $n(n-1)\alpha + (n-1)\sum_k p_k$.

Why does this formulation work? Assume a solution forms a circle. Let us consider the resulting $x^{k,a}_{ij}$ variables for a fixed value $k$ and a given circle $i_0, i_1, \ldots, i_m, i_0$. Constraint (4) implies that for every $i_\beta, i_{\beta+1}$ and $k$, there is a unique $a$ with $x^{k,a}_{i_\beta i_{\beta+1}} = 1$. Assume that for $\beta = 0$ we have $a = 0$, so that $x^{k,0}_{i_0 i_1} = 1$. Constraint (5) now implies that $x^{k,1}_{i_1 i_2} = 1$. This implication can be furthered until we come back to the beginning. In total, this shows that a circle can only exist if its length is divisible by $p_k$.

As we have shown this for every $k$, and the values $p_k$ are relative prime, we know that a circle has at least the length $p_1 p_2 \cdots p_k \geq n$ which is impossible.
3. Estimating the Size of the Model

In the last section we constructed a model for which the number of variables is bounded by \(n(n-1)\left(\sum k p_k + 1\right)\) for values \(p_k\) with \(\prod_k p_k \geq n\). Of course, one would pick the \(p_k\) so that the sum \(\sum k p_k\) is minimized. But how small is this sum, compared to \(n\)? For that, we make the following definition

\[P_n = \{P \subset \mathbb{N} | \text{lcm}_{p \in P}(p) \geq n\},\]

where lcm denotes the least common multiple. Then we define

\[f(n) = \min_{P \in P_n} \sum_{p \in P} p \quad (6)\]

Obviously, the optimal \(P \in P_n\) consists of pairwise relative prime numbers. Furthermore, these numbers are prime powers because \(k \cdot m \geq k + m\) for every \(k, m \geq 2\). For practical purposes, it is easy to give an estimate by explicitly giving values for \(p_k\).

\[
\begin{array}{c|c|c}
   n  & \text{UB } f(n) & \\
10  & 3 + 4 & \\
100 & 9 + 11 & \\
1000& 9 + 11 + 13 & \\
10000& 9 + 11 + 13 + 17 & \\
100000& 7 + 9 + 11 + 13 + 17 & \\
\end{array}
\]

We see that the size of \(f(n)\) is rather small and grows roughly on a logarithmic scale. Can we give a general estimate for high values of \(n\)?

For that we consider the Landau function \(g(m)\). It is defined for every natural number \(m\) to be the largest order of an element of the symmetric group \(S_m\). Equivalently, \(g(m)\) is the largest least common multiple of any partition of \(m\). Let us prove that

\[f(g(m)) \leq m.\]

For that let \(m = p_1 + \ldots + p_\alpha\) be a partition of \(m\) with largest least common multiple. Then \(\{p_1, \ldots, p_\alpha\}\) forms a valid set for \(f(n)\) with \(n = g(m)\). As \(f(x)\) and \(g(x)\) are monotone functions, the inequality holds.

Our aim is to use this relation to construct an upper bound for \(f(n)\) for large \(n\). For that we exploit some properties of \(g(m)\). From [Nic97] we know that

\[\lim_{n \to \infty} \frac{g(m+1)}{g(m)} = 1\]

which implies that for every \(\hat{m}\) there is a constant \(C_{\hat{m}}\) with

\[g(m+1) \leq C_{\hat{m}} g(m) \quad \forall m \geq \hat{m} \quad (7)\]
Furthermore, \cite{MNR89} shows that
\[
\ln g(m) \geq \sqrt{m \ln m} \quad \forall m \geq 906 \quad (8)
\]
The estimate \cite{MNR89} clearly implies
\[
(\ln g(m))^2 \geq m \quad \forall m \geq 906 \quad (9)
\]
Now, for a given \(n\) large enough, choose \(m\) with
\[
g(m) \leq n \leq g(m+1)
\]
Note that \(g(m)\) is clearly unbounded, see e.g. \cite{MNR89}. Then we have, using the monotonicity of \(f, g\) and \(\ln\):
\[
f(n) \leq f(g(m+1))
\]
\[
\leq m + 1
\]
\[
\leq (\ln g(m+1))^2
\]
\[
\leq (\ln C_m g(m))^2
\]
\[
\leq (\ln C_m n)^2
\]
\[
= (\ln C_m + \ln n)^2
\]
This shows that \(f(n)\) is of size \(O(\ln(n))^2\), so that the total number of variables can be estimated by \(O(n^2 \ln(n))^2\). Note furthermore that
\[
https://oeis.org/A000793/b000793.txt
\]
lists the first 10000 values of \(g(m)\). From this is would be possible to derive good upper bounds for \(f(n)\) for every reasonable value of \(n\) (considering that \(g(10000) = 83724831478115383622262752790863851979860821403995464606889515707038262291218453286312511300740953193197240510692638318426636506681182400\)).

For the number of constraints, we see that the term \(n(n-1)\alpha\) clearly dominates the other terms. Assuming that \(P \in P_n\) is chosen to minimize \(\sum_{p \in P} p\) as before, what can be said about the size of \(\alpha = |P|\) in relation to \(n\)? For that, we establish a small lemma:

**Lemma 1.** If \(P \in P_n\) is chosen to minimize \(\sum_{p \in P} p\), then \(\prod_{p \in P} p < 2n\).

**Proof.** As noted before, all \(p \in P\) are prime powers. Let \(q_0\) be the smallest involved prime number and \(p_1 = q_0^\alpha\). Because \(\hat{p}_1 = (q_0 - 1)q_0^\alpha - 1\) is also relative prime to \(p_2, \ldots, p_\alpha\) or equal to one, we know that \(\hat{p}_1 p_2 \ldots p_\alpha < n\) (otherwise we could choose a better \(P\), contradicting the optimality). Now we have
\[
p_1 p_2 \ldots p_\alpha = \frac{q_0}{q_0 - 1}\hat{p}_1 p_2 \ldots p_\alpha
\]
\[
< \frac{q_0}{q_0 - 1} n \leq 2n
\]
\[\square\]
Now we know:

\[ 2^\alpha \leq p_1 \ldots p_\alpha < 2n \]
implying

\[ \alpha < C \ln(n) \]

for some constant \( C \). In total, we see that the number of constraints is \( O(n^2 \ln(n)) \).

4. Conclusion

This paper explains a new modelling approach for the TSP. It consists of \( O(n^2 \ln(n)^2) \) binary variables and \( O(n^2 \ln(n)) \) constraints. In its pure form, it will probably not useful for solving the problem numerically, but it may form the starting point for a different view on subtour elimination constraints.

References


