Small diameter neighbourhood graphs for the traveling salesman problem: at most four moves from tour to tour

Gregory Gutin*
Department of Mathematics and Statistics
Brunel University of West London
Uxbridge, Middlesex, UB8 3PH, U.K.
z.g.gutin@brunel.ac.uk

Anders Yeo†
Department of Mathematics and Computer Science
Odense University, Odense DK-5230, Denmark
gyeo@imada.ou.dk

Abstract

Scope and Purpose – One of the main characteristics of a neighbourhood structure imposed on the solution set of a combinatorial optimization problem is the diameter of the corresponding neighbourhood graph. The diameter reflects the 'closeness' of one solution to another. We study the diameter of the neighbourhood graph of some exponential size neighbourhood structure for the TSP and show that the diameter of the graph is surprisingly small: at most four. This demonstrates a high potential of some exponential size neighbourhoods.

Abstract – A neighbourhood $N(T)$ of a tour T (in the TSP with n cities) is polynomially searchable if the best among tours in $N(T)$ can be found in time polynomial in n. Using Punnen's neighbourhoods introduced in 1996, we construct polynomially searchable neighbourhoods of exponential size satisfying the following property: for any pair of tours T_1 and T_5, there exist tours T_2, T_3 and T_4 such that T_i is in the neighbourhood of T_{i-1} for all $i = 2, 3, 4, 5$. In contrast, for pyramidal neighbourhoods considered by J. Carlier and P. Villon (1990), one needs up to $\Theta(\log n)$ intermediate tours to 'move' from a tour to another one.

*G. Gutin completed his Ph.D. under supervision of Professor N. Alon in School of Mathematics, Tel Aviv University. He has been a visiting member of Department of Mathematics and Computer Science, Odense University, Denmark. Now he is a lecturer at Department of Mathematics and Statistics, Brunel University, U.K. His research interests include combinatorial optimization and graph theory.

†A. Yeo is a PhD student in Department of Mathematics and Computer Science, Odense University, Denmark. His supervisors are J. Bang-Jensen and G. Gutin. His research interests include graph theory and combinatorial optimization.

1 Introduction

Let DK_n be a weighted complete directed graph on n vertices (the weights are assigned to the edges). In the asymmetric traveling salesman problem (TSP) one seeks a hamiltonian directed cycle (called a *tour*) of DK_n of minimum total weight.

Local search based heuristics and metaheuristics have been the most effective and efficient tool for finding near optimal solutions for large scale instances of the TSP (see e.g. [10, 12, 13]). Significant progress in local search methods for TSP has been made due to the use of more powerful neighbourhood structures (see e.g. [12, 13]). A *neighbourhood* of a tour T in DK_n is a set $N(T)$ of tours containing T. There are different types of neighbourhoods for the TSP; many of them can be found in [3, 6, 10, 11, 12, 13, 14]. The main property of a neighbourhood $N(T)$, which is usually imposed on it, is that the best among tours of $N(T)$ can be calculated in time polynomial in n. This is necessary to have local search efficient. Neighbourhoods satisfying this property are called *polynomially searchable.*

Given neighbourhood $N(T)$ for every tour T in DK_n (i.e., *neighbourhood structure*), the corresponding *neighbourhood graph* (of order n) is a directed graph with vertex set consisting of all tours in DK_n and arc set containing pairs (T', T'') such that $T'' \in N(T')$. When all neighbourhoods $N(T)$ are polynomially searchable, the corresponding graph is *polynomially searchable.* The diameter of the neighbourhood graph is one of the most important characteristics of the neighbourhood structure and the corresponding local search scheme [1, 4, 5, 6]. Clearly, a neighbourhood structure with neighbourhood graph of smaller diameter seem to be more powerful than one with neighbourhood graph of larger diameter, let alone that whose graph has infinite diameter (in the last case, some tours are not 'reachable' from the initial tour during local search procedure).

Recently, a number of polynomially searchable neighbourhoods of exponential size for the TSP have been suggested (see e.g. [2, 3, 4, 5, 6, 7, 8, 9, 11, 15]). Computational experiments in [2, 4, 12] have already shown a very high potential of heuristics using such neighbourhoods. The most powerful among neighbourhoods of exponential size (at least from theoretical point of view) appear to be those of size $2^{\Theta(n \log n)}$ (note that $(n-1)! = 2^{\Theta(n \log n)}$ as well). Such neighbourhoods were introduced independently in [8, 11, 15] (some ideas on those neighbourhoods can be also found in [6]). Punnen's neighbourhoods [11] are the most general among them. Some features of Punnen's neighbourhoods were investigated in [9, 11].

To the best of our knowledge, the neighbourhood graph for polynomially searchable 'pyramidal' neighbourhoods introduced by Carlier and Villon in [4] has smallest known diameter $d_n = \Theta(\log n)$ (in [4], it was proved that $d_n \le \log n$, the lower bound $d_n = \Omega(\log n)$ follows from the facts that the cardinality of pyramidal neighbourhoods is $2^{\Theta(n)}$ and the total number of tours is $2^{\Theta(n \log n)}$.)

In this note, using Punnen's neighbourhoods, we construct certain polynomially searchable 'compound' neighbourhoods whose neighbourhood graphs have diameter bounded by a small constant. This result which we believe is of primal interest for the study of polynomially searchable neighbourhoods provides another theoretical evidence that the approach of [8, 11, 15] and its modifications have a high computational potential. Results of computational experiments with algorithms using Punnen's neighbourhoods and their modifications are to be described in future papers.

2 Neighbourhood graphs

We start by recalling some graph theory definitions. A digraph is *strongly connected* if, for any pair of vertices x and y there is a path from x to y and a path from y to x. The *distance* from a vertex v and to a vertex w in a strongly connected digraph G is the length of a shortest path from v to w. The *diameter* of G, $diam(G)$, is the least positive integer d such that the distance from a vertices to another vertex in G does not exceed d.

Let $C = x_1x_2...x_kx_1$ be a cycle in DK_n. The operation of *removal* of a vertex x_i $(1 \leq i \leq k)$ results in the cycle $x_1x_2...x_{i-1}x_{i+1}...x_kx_1$ (thus, removal of x_i is not deletion of x_i from C). Let y be a vertex of DK_n not in C. The operation of *insertion* of a vertex y into an arc x_ix_{i+1} results in the cycle $x_1x_2...x_iyx_{i+1}...x_kx_1$. An *insertion* of y in C is an insertion of y into x_ix_{i+1} for some $1 \leq i \leq k$. For a set $Z = \{z_1, ..., z_s\}$ $(s \leq k)$ of vertices not in C, an *insertion* of Z into C means an insertion of z_1 in C followed by an insertion of z_2 into the obtained cycle, etc. Note that we require that, in the cycle obtained after insertion of all vertices of Z into C, no pair of vertices from Z is adjacent.

Let $T = x_1x_2...x_nx_1$ be a tour in DK_n and let Z be a set of non-adjacent vertices of T. Punnen's neighbourhood of T with respect to Z, $N(T,Z)$, consists of the tours that can be obtained from T by removal of the vertices from Z one by one followed by an insertion of Z into the cycle derived after the removal. It was shown in [11] (see also [9]) that the best tour in $N(T,Z)$ can be found in time $O(n^3)$.

For a positive integer $k \leq n/2$, the *neighbourhood graph* $\Gamma(n,k)$ has vertex set formed by all tours in DK_n. A tour T is adjacent in $\Gamma(n,k)$ to a tour R if there exists a set Z of k non-adjacent vertices of T such that $R \in N(T,Z)$. Clearly, T is adjacent to R if and only if R is adjacent to T. Thus, $\Gamma(n,k)$ is a symmetric directed graph. We denote by $dist_k(T,R)$ the distance from T to R in $\Gamma(n,k)$. Clearly, if T is adjacent to R in $\Gamma(n,k)$, then $dist_k(T,R) = 1$.

For a tour T in DK_n, let $\mathcal{I}_{nk}$ denote the family of all sets of k non-adjacent vertices in T. Clearly, the neighbourhood $N_k(T)$ of a tour T in $\Gamma(n,k)$ equals $\cup\{N(T,Z):\ Z \in \mathcal{I}_{nk}\}$. Thus if, for some k, $i(n,k) = |\mathcal{I}_{nk}|$ is polynomial in n, then $\Gamma(n,k)$ is polynomially searchable. Otherwise, $\Gamma(n,k)$ may be non-polinomially searchable. As polynomially searchable $\Gamma(n,k)$ are of our main interest, we start with evaluating $i(n,k)$ in Theorem

2.1. It follows from Theorem 2.1 that, for fixed k, $i(n,k)$ and $i(n,n-k)$ are polynomial.

Theorem 2.1 $i(n,k) = \binom{n-k}{k} + \binom{n-k-1}{k-1}$.

Proof: Let $T = x_1x_2...x_nx_1$ be a tour and let $Z \in \mathcal{I}_{nk}$. A set $R = \{(x_j, x_{j+1}) : j \in J\}$ of pairs of consecutive vertices in T is called *independent* if vertices in the pairs of R are all distinct. Observe that $R = \{(x_j, x_{j+1}) : j \in J\}$ is independent if and only if $\{x_j : j \in J\}$ is a set of non-adjacent vertices of T. Therefore, $i(n,k)$ is equal to the number of independent sets of k pairs of consecutive vertices in T.

Let $p(n,k)$ denote the analogous number for the path $x_1x_2...x_n$. All independent sets of k pairs of consecutive vertices in T can be partitioned into two families: the first one contains all the independent sets with pair (x_n, x_1) and the second one consists of all independent sets without pair (x_n, x_1). Observe that the number of sets in the second family is $p(n,k)$ and the number of sets in the first family is $p(n-2,k-1)$. Thus, $i(n,k) = p(n,k) + p(n-2,k-1)$.

Now, to prove this theorem, it is sufficient to show that $p(n,k) = \binom{n-k}{k}$. However, this follows from the fact that $p(n,k)$ is equal to the number of binary strings of length $n-k$ with k zeros. Indeed, a binary string $S = s_1s_2...s_{n-k}$ with k zeros generates an independent set $Z(S)$ of k pairs of consecutive vertices in T using the following algorithm:

put $j := 1$, $Z(S) := \emptyset$

for $i = 1$ **to** $n-k$ **do**

if $s_i = 0$ **then** (x_j, x_{j+1}) is added to $Z(S)$, $j := j+2$

otherwise $j := j+1$ **endif**

enddo

It is easy to verify that the 'inversion' of the above algorithm allows one to generate a binary string of length $n-k$ from an independent set of k pairs of consecutive vertices in T. □

Corollary 2.2 *If p is a non-negative fixed integer, then $\Gamma(n, p+1)$ and $\Gamma(n, \lfloor (n-p)/2 \rfloor)$ are polynomially searchable ($p < \lfloor n/2 \rfloor$).*

Proof: This follows from Theorem 2.1 taking into consideration that $\binom{m}{k} = \binom{m}{m-k}$. □

3 Diameter

It is easy to see that $\Gamma(n, n/2)$ (n is even) consists of an exponential number of connected components and, thus, its diameter is infinite (for example, $x_1x_2...x_nx_1$ and $x_1...x_{n-2}x_nx_{n-1}x_1$

belong to different components of this graph). Therefore, below we consider $\Gamma(n,k)$ for $k < n/2$ only.

Theorem 3.1 *$diam(\Gamma(n, \lfloor (n-1)/2 \rfloor) \leq 4$.*

Proof: We assume that $n \geq 5$ as for $2 \leq n \leq 4$ this theorem can be directly verified. Let $C = x_1x_2\ldots x_nx_1$ and $T = y_1y_2\ldots y_ny_1$ be a pair of distinct tours in DK_n ($(y_1, y_2, \ldots, y_n)$ is a permutation of $(x_1, x_2, \ldots, x_n)$). In the sequel, all indices are taken modulo n. Put $k = \lfloor (n-1)/2 \rfloor$. We will prove that $dist_k(T, C) \leq 4$ thus showing that $diam(\Gamma(n,k)) \leq 4$.

We call a vertex v *even* (*odd*) if $v = x_k$, where $1 \leq k \leq n$ and k is even (odd). For a set of vertices X of DK_n, X_{odd} (X_{even}) is the set of odd (even) vertices in X.

First we consider the case of even n, i.e. $k = n/2 - 1$. The proof of this case consists of two steps. In the first step, we show that there exists T'' whose vertices alternate in parity and such that $dist_k(T, T'') \leq 2$. Moreover, T'' has a pair of consecutive vertices which are also consecutive in C. In the second step, we easily see that $dist_k(T'', C) \leq 2$ as the odd and even vertices of T'' (except for the vertices of the above pair) can be separately reordered to form C. Thus, we will conclude that $dist_k(T, C) \leq 4$. Now, we consider the proof in more detail.

Clearly, T has a pair y_j, y_{j+1} such that y_{j+1} is odd and y_j is even. Let $Z = \{y_{j+2}, y_{j+4}, \ldots, y_{j+2k}\}$ and let $|Z_{odd}| = s$. Remove the vertices of Z from T and then insert s odd vertices of Z into the arcs $y_{j+1}y_{j+3}, \ldots, y_{j+2s-1}y_{j+2s+1}$ and $k - s$ even vertices of Z into the arcs $y_{j+2s+1}y_{j+2s+3}, y_{j+2s+3}y_{j+2s+5}, \ldots, y_{j+2k-1}y_{j+2k+1}$. We have obtained a tour

$$T' = y_jy_{j+1}v_{j+2}y_{j+3}v_{j+4}y_{j+5}\ldots y_{j+2k-1}v_{j+2k}y_{j+2k+1}y_j,$$

where $\{v_{j+2}, \ldots, v_{j+2k}\} = Z$.

Let $Z' = \{y_{j+3}, y_{j+5}, \ldots, y_{j+2k+1}\}$ and let $|Z'_{even}| = t$. As the number of odd vertices in $V(DK_n) - \{y_j, y_{j+1}\}$ is equal to $k = |Z_{odd}| + |Z'_{odd}| = s + k - t$, we obtain that $s = t$. Remove Z' from T' and insert t even vertices of Z' into the arcs $y_{j+1}v_{j+2}, v_{j+2}v_{j+4}, v_{j+6}v_{j+8}, \ldots, v_{j+2s-2}v_{j+2s}$ and $k-s$ odd vertices of Z' into the arcs $v_{j+2s+2}v_{j+2s+4}, \ldots, v_{j+2k-2}v_{j+2k}, v_{j+2k}y_j$. We have derived a tour $T'' = u_1u_2\ldots u_nu_1$. Clearly, the vertices of T'' alternate in parity, i.e., for every m, if u_m is odd, then u_{m+1} is even.

Now we prove that the processes of insertion of Z and Z' can be performed in such a way that T'' contains a pair of consecutive vertices which are also consecutive in C (i.e. there exist indices p and q such that $u_p = x_q$ and $u_{p+1} = x_{q+1}$). As $1 < |Z'| < n$, there exists a pair of distinct indices i, m such that $x_i, x_m \in Z'$ and $x_{i+1}, x_{m-1} \notin Z'$. Without loss of generality, we assume that i is odd. We consider two cases.

Case 1: $|Z'_{odd}| \geq 2$. We prove that we may choose index $q = i$. Since $x_{i+1} \notin Z'$ and $i+1$ is even, either $y_j = x_{i+1}$ or $x_{i+1} \in Z_{even}$. If $x_{i+1} \in Z_{even}$, in the process of insertion of Z, we insert x_{i+1} into $y_{j+2k-1}y_{j+2k+1}$, i.e. $x_{i+1} = v_{j+2k}$. In the process of insertion of Z', we insert x_i into $v_{j+2k}y_j$ if $x_{i+1} = y_j$ or into $v_{j+2k-2}v_{j+2k}$, otherwise (i.e. $x_{i+1} = v_{j+2k}$).

Case 2: $|Z'_{odd}| = 1$. Thus, m is even. As $n \geq 6$, we obtain that $|Z'_{even}| \geq 2$. Analogously to Case 1, we can show that q may be chosen as $q = m - 1$.

Thus, without loss of generality, we assume that $u_{n-1} = x_i,\ u_n = x_{i+1}$.

As $\{u_2, u_4, ..., u_{2k}, x_{i+1}\} = C_{even}$, we can delete $\{u_2, u_4, ..., u_{2k}\}$ from T'' and insert it into the obtained cycle to get the tour $C' = u_1x_{i+3}u_3x_{i+5}u_5...u_{2k-1}x_{i-1}u_{n-1}x_{i+1}u_1$. Analogously, we can delete $\{u_1, u_3, ..., u_{2k-1}\}$ from C' and insert it into the obtained cycle to get C.

Clearly, $dist_k(T, C) \leq 4$.

Now let n be odd. Then $k = (n - 1)/2$. Notice that, without loss of generality, we assume that $x_n = y_n$ (to 'fix' the initial labelings of T and C). Consider tours $X = x_1x_2...x_nx_{n+1}x_1$ and $Y = y_1y_2...y_{n-1}y_ny_{n+1}y_1$ in DK_{n+1}, where $y_n = x_n,\ y_{n+1} = x_{n+1}$. If we assume that $j = n,\ j + 1 = n + 1$, we can obtain, analogously to the case of even n, a tour Y'' such that the vertices of Y'' alternate in parity (with respect to their indices in X), x_{n+1} follows x_n in Y'' and $dist_k(Y, Y'') \leq 2$. Now if we assume that $i = n$ and $i+1 = n+1$, we can show, similarly to the case of even n, that $dist_k(Y'', X) \leq 2$. Clearly, $dist_k(Y, X) \leq 4$. Notice that in the whole process of constructing X from Y, we have never removed x_n and x_{n+1} or inserted any vertex into the arc x_nx_{n+1}. Thus, we could contract the arc x_nx_{n+1} to x_n and obtain C from T in four 'steps'. This shows that $dist_k(T, C) \leq 4$. $\square$

Theorem 3.2 *Let $dist_k(T, C) = 1$ for tours T and C and let m be an integer smaller than k. Then, $dist_m(T, C) \leq \lceil k/m \rceil$.*

Proof: Let Z be a set of k non-adjacent vertices of T such that $C \in N(T, Z)$. Let $Z = Z_1 \cup Z_2 \cup ... \cup Z_{\lceil k/m \rceil}$ be a partition of Z (i.e. $Z_i \cap Z_j = \emptyset$ for any $i \neq j$) such that $|Z_i| = m$ for every $i = 1, 2, ..., \lfloor k/m \rfloor$. For a vertex $z \in Z$, denote by $p(z)$ and $q(z)$, respectively, the vertex that is followed by z in C and the vertex that follows z in C, respectively.

Clearly, if $p(z)$ and $q(z)$ are not adjacent in T, then there is a vertex $s(z)$ between them in T (only one such vertex) and $s(z) \in Z$. Put $T^0 = T$. For $i = 1, 2, ..., \lceil k/m \rceil$, let T^i be a tour derived from the previous tour T^{i-1} by removal of Z_i followed by insertion of Z_i into the obtained cycle such that a vertex $z \in Z_i$ is inserted into the arc $p(z)p^+(z)$, where $p^+(z)$ is the vertex that follows $p(z)$ in T^{i-1} (clearly, $p^+(z) \in \{s(z), q(z)\}$.) As every vertex $s(z)$ $(z \in Z)$ is removed while forming one of the tours T^i, every vertex z is between $p(z)$ and $q(z)$ in T^m. Therefore, $C = T^m$.

Clearly, $T^i \in N(T^{i-1}, Z_i)$ for all $i = 1, 2, ..., \lceil k/m \rceil$. In particular, T^{i-1} and T^i are adjacent in $\Gamma(n, k)$ for all $i = 1, 2, ..., \lfloor k/m \rfloor - 1$. If m does not divide k, $|T_{\lceil k/m \rceil}| < m$. In this case, it does not follows directly that $T^{\lceil k/m\rceil -1}$ and $T^{\lceil k/m\rceil}$ are adjacent in $\Gamma(n,k)$. However, if we enlarge $Z_{\lceil k/m\rceil}$ by adding some vertices of Z_1 to it (those vertices are to be returned to the same places they were removed from), we see that indeed $T^{\lceil k/m\rceil -1}$ and $T^{\lceil k/m\rceil}$ are adjacent. □

Corollary 3.3 *For every positive m,*

$$diam(\Gamma(n,m)) \leq 4\lceil \lfloor (n-1)/2 \rfloor /m \rceil.$$

In particular, if p is a positive integral constant, then $diam(\Gamma(n,\lfloor (n-p)/2\rfloor)) \leq 8$ for every $n \geq 2p+1$.

Proof: The first inequality follows directly from the above two theorems and the triangle inequality for distances in graphs. The first inequality implies the second one. Indeed, $n \geq 2p+1$ implies

$$\frac{(n-1)/2}{(n-p-1)/2} \leq 2, \;\; Q = \frac{\lfloor (n-1)/2 \rfloor}{\lfloor (n-p)/2 \rfloor} \leq 2.$$

Thus, $\lceil Q \rceil \leq 2$. □

Acknowledgements

This work was partially supported by Brunel University grant BRIEF. Part of this paper was written when GG was visiting A. Punnen at University of New Brunswick, St. John, Canada and F. Glover at University of Colorado, Boulder, USA. We thank J. Orlin, G. Woeginger and the referee for helpful suggestions.

References

[1] E. Angel, Talk at the Fifth Twente Workshop on Graphs and Combinatorial Optimization, May, 1997.

[2] E. Balas and N. Simonetti, Linear time dynamic programming algorithms for some new classes of restricted TSP's. *Proc. IPCO V*, LNCS **1084**, Springer Verlag, 1996, 316-329.

[3] R.E. Burkard, V.G. Deineko and G.J. Woeginger, The travelling salesman problem and the PQ-tree. *Proc. IPCO V*, LNCS **1084**, Springer Verlag, 1996, 490-504.

[4] J. Carlier and P. Villon, A new heuristic for the traveling salesman problem. *RAIRO* **24,** 245-253 (1990).

[5] V. Deineko and G.J. Woeginger, A study of exponential neighbourhoods for the travelling salesman problem and for the quadratic assignment problem. Technical Report Woe-05, July (1997).

[6] F. Glover, Ejection chains, reference structures, and alternating path algorithms for traveling salesman problem. Research Report, Univ. of Colorado-Boulder, Graduate School of Business (1992). A shortened version appeared in *Discrete Appl. Math.* **65,** 223-253 (1996).

[7] F. Glover and A. Punnen, The traveling salesman problem: New solvable cases and linkages with the development of approximation algorithms. *J. Oper. Res. Soc.* **48,** 502-510 (1997).

[8] G. Gutin, On an approach for solving the traveling salesman problem. In *Proc. The USSR Conference on System Research*, Moscow, 184–185 (1984) (in Russian).

[9] G. Gutin, Exponential neighbourhood local search for the traveling salesman problem. To appear in Special Issue of *Computers and OR* on the TSP.

[10] D.S. Johnson and L.A. McGeoch, The traveling salesman problem: a case study in local optimization. *Local Search in Combinatorial Optimization,* E.H.L. Aarts and J.K. Lenstra (eds.), Wiley, N.Y., 215-310 (1997).

[11] A.P. Punnen, The traveling salesman problem: new polynomial approximation algorithms and domination analysis. Manuscript, December (1996).

[12] A.P. Punnen and F. Glover, Implementing ejection chains with combinatorial leverage for the TSP, submitted.

[13] C. Rego, Relaxed tours and path ejections for the traveling salesman problem. Technical Report, Lab. PRiSM, Univ. de Versailles (1996).

[14] G. Reinelt, *The Traveling Salesman: Computational Solutions for TSP Applications.* Springer-Verlag, Berlin (1994).

[15] V.I. Sarvanov and N.N. Doroshko, Approximate solution of the traveling salesman problem by a local algorithm with scanning neighbourhoods of factorial cardinality in cubic time. *Software: Algorithms and Programs,* Math. Institute of Belorussian Acad. of Sci., Minsk, No. 31, 11-13 (1981) (in Russian).