DISTRIBUTION OF DISTANCES IN THE TRAVELLING SALESMAN PROBLEM ON A SQUARE LATTICE

***Zbigniew Domański*[1, 2], *Jarosław Kęsy*[1]**

[1] *Institute of Mathematics and Computer Science, Czestochowa University of Technology*
[2] *Polonia University, Częstochowa*

Abstract. The Manhattan distance between two points is defined as the sum of the horizontal distance along streets and the vertical distance along avenues. We derive an exact formula for the number of zigzag distances travelled by a salesman on a finite $N \times N$, square lattice. In the limit $N \to \infty$, we obtain the density function of distances on a unit square.

1. Introduction

Consider random geometrical points, i.e. points with uncorrelated positions, occupied vertices of a square lattice. We address the following question: what is the mean distance $\langle r_N \rangle$ between a given reference point and other randomly chosen point, where points are distributed uniformly.

This is essentially the problem of geometrical importance but the distribution of the quantity $\langle r_N \rangle$ is also important in certain physical and computational problems. For example in physics and in optimisation theory the quantity $\langle r_N \rangle$ is important in determining the statistical properties of systems composed of objects whose interactions are proportional to the Manhattan distance between objects. In the field of computer science the pairwise distance between processors is the number of vertical communication hops plus the number of horizontal communication hops. That is, the allocation of processors to parallel programs in a supercomputer grid consisting of a large number of processors also relays on the Manhattan distances between processors [1].

The sum of the pairwise distances between points successively visited by the salesman is strongly correlated with the total length of a path required to complete his task. Thus, the knowledge of the Manhattan distance distribution can yield a valuable information needed for estimating the optimal path in the travelling salesman problem (TSP) [2, 3], a typical, well-known optimization problem which consists of finding the shortest closed tour connecting all cities in a map.

2. Results

We begin with finite value of N. In Figure 1 we present such a case for $N = 11$.

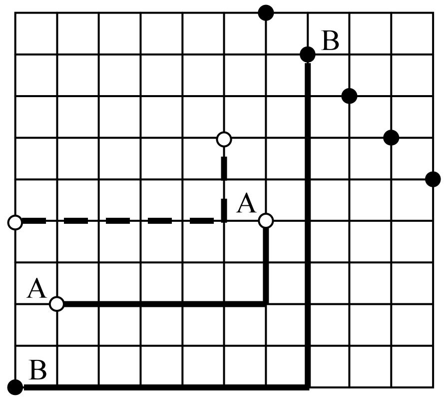

Fig. 1. A-A (open circles) and B-B (filled circles) are pairs of points on a square lattice of size $N = 11$. The Manhattan distances: $R_N(A,A) < N$, and $N \le r_N(B,B) << 2N - 2$

Consider a pair of points A-A, whose distance $q < N$. It is easy to see, that the total number of such pairs is equals to a number of arrangements of A-A segment on the lattice, i.e.

$$2 \times \sum_{j=0}^{q-1} (N - q + j)(N - j) \tag{1}$$

Multiplication by 2 in Equation (1) comes from the segments obtained by counter-clockwise rotation of the A-A segments. Similar consideration for B-B segments yields

$$2 \times \sum_{j=1}^{p+1} j(p - j + 2), \quad \text{with} \quad q = 2(N-1) - p \tag{2}$$

An auxiliary quantity $p = 0,1,..., N - 2$, in Equation (2) measures the distance between right end of the segment B-B and the upper right corner of the square. Collecting the terms in Equations (1) and (2) we obtain the following formula for the number of pair of points with the shortest path's length equals to q

$$\Delta(q) = \begin{cases} 2N(N-q)q + \dfrac{(q-1)q(q+1)}{3} & \text{for} \quad q = 1,2,..., N-1 \\ \dfrac{1}{3}(2N - q - 1)(2N - q)(2N - q + 1) & \text{for} \quad q = N,...,2N-2 \end{cases} \tag{3}$$

After some elementary algebra the above equation can be rewritten with the help of variable $x_q = q/N$

$$\frac{\Delta(x_q)}{N^3} = \begin{cases} 2x_q(1-x_q) + \frac{1}{3}x_q(x_q^2-1); & x_q = 1/N, 2/N, ..., 1-1/N \\ \frac{1}{3}(2-x_q-1/N)(2-x_q)(2-x_q+1/N); & x_q = 1, ..., 2-2/N \end{cases} \quad (4)$$

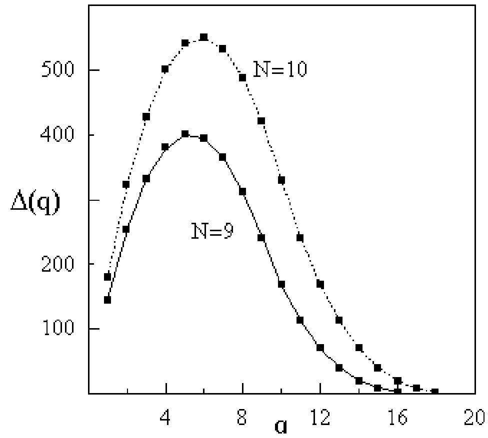

Fig. 2. The Manhattan-distance multiplicities given by Equation (3)

Now we consider the limiting case of a dense line packing and we assume that the number of lines, separated by δ goes to infinity in a way that $N \cdot \delta = 1$. Within this limit, from Equation (4), we get final expression for the density of zigzag distances on a unit square:

$$D(x) = \begin{cases} 4(1-x)x + \frac{2}{3}x^3 & \text{for} \quad 0 < x \leq 1 \\ \frac{2}{3}(2-x)^3 & \text{for} \quad 1 < x \leq 2 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

3. Conclusions

For the both, discrete (Equation (3)) and continuous (Equation (5)), distributions of Manhattan distances on squares we can compute the moments of an arbitrary order. Especially, the mean distance $\langle r_N \rangle = 2N/3$, for the discrete case, and its value is 2/3 for the continuous distribution. In a different way the same average pairwise distance value was obtained and reported in [4]. The second moment equals to $(5N^2 - 2)/9$ and it is related to the minimization of the average of the squares of the pairwise distances in clustering applications.

Similar distributions of distances can be derived for lattices with different symmetries, such as triangular or honeycomb.

References

[1] Chang C., Mohapatra P., J. Parallel Distrib. Comput. 1998, 52, 40-68.

[2] Michalewicz Z., Genetics Algorithms + Data Structures = Evolution Programs, Springer Verlag, Berlin, Heidelberg 1996.

[3] Whitley D., Starkweather T., Fuquay D.A., Scheduling Problems and Traveling Salesman: The Genetic Edge Recombination Operator, Proc. of the 3th Int. Conf. on Genetic Algorithms, ed. J. Schaffer, Morgan Kaufmann Publishers 1988, 133-140.

[4] Bender C.M., Bender M.A., Demaine E.D., Fekete S.P., Journal of Phys. A: Math. Gen. 2004, 37, 147-159.