

Nearest-Neighbor Graphs

STEVEN FINCH

April 25, 2008

Consider a set P of n points that are independently and uniformly distributed in the d -dimensional unit cube. Let $p \in P$. There exists almost-surely $q \in P$ such that $q \neq p$ and $|p - q| < |p - r|$ for all $r \in P$, $r \neq p$, $r \neq q$. The point q is called the **nearest neighbor** of p and we write $p \prec q$. Note that $p \prec q$ does not imply $q \prec p$. Draw an edge connecting p and q if and only if $p \prec q$; the resulting graph of n vertices and $\leq n$ edges is called the **nearest-neighbor graph** G on P .

What is the probability, $\alpha(d)$, given $p \in P$, that $p \prec q$ implies $q \prec p$? Such a pair is **isolated** from the rest of G , in the sense that the only edge touching p or q is the edge that connects p and q . We have [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13]

$$\alpha(1) = \frac{2}{3}, \quad \alpha(2) = \frac{6\pi}{8\pi + 3\sqrt{3}} = 0.6215048968\dots, \quad \alpha(3) = \frac{16}{27}$$

and, more generally [9],

$$\alpha(d) = \begin{cases} \left[\frac{3}{2} + \frac{1}{2} \sum_{k=1}^{\ell} \frac{1 \cdot 3 \cdot \dots \cdot (2k-1)}{2 \cdot 4 \cdot \dots \cdot (2k)} \left(\frac{3}{4}\right)^k \right]^{-1} & \text{if } d = 2\ell + 1, \\ \left[\frac{4}{3} + \frac{\sqrt{3}}{2\pi} \left(1 + \sum_{k=1}^{\ell-1} \frac{2 \cdot 4 \cdot \dots \cdot (2k)}{3 \cdot 5 \cdot \dots \cdot (2k+1)} \left(\frac{3}{4}\right)^k \right) \right]^{-1} & \text{if } d = 2\ell. \end{cases}$$

Here is a variation of the preceding. Draw an edge connecting p and q if and only if $q \prec p$; the resulting graph of n vertices and $\leq n$ edges is called the **nearest-neighbor anti-graph** H on P . What is the probability, $\beta(d)$, that $p \in P$ is isolated from the rest of H ? That is, what proportion of points in P are not nearest neighbors of any other points? We have [14, 15, 16, 17, 18]

$$\beta(1) = \frac{1}{4}, \quad \beta(2) \approx 0.28, \quad \beta(3) \approx 0.30$$

but the latter two estimates are only simulation-based. To further understand $\beta(2)$ will occupy us for the remainder of this essay.

⁰Copyright © 2008 by Steven R. Finch. All rights reserved.

Define constants $C(0, d) = 1$ and

$$C(k, d) = \int_{\Omega(k, d)} \exp \left[-\text{Vol} \left(\bigcup_{j=1}^k S(x_j) \right) \right] dx_1 dx_2 \dots dx_k$$

for $k \geq 1$, where $S(x_j)$ is the ball in \mathbb{R}^d of radius $|x_j|$, centered at x_j , and

$$\Omega(k, d) = \left\{ (x_1, x_2, \dots, x_k) \in \mathbb{R}^{dk} : |x_i| \leq |x_i - x_j| \text{ for all } 1 \leq i \neq j \leq k \right\}.$$

It is known that [17, 19, 20, 21, 22]

$$\beta(2) = \sum_{k=0}^6 \frac{(-1)^k}{k!} C(k, 2), \quad \beta(3) = \sum_{k=0}^{12} \frac{(-1)^k}{k!} C(k, 3)$$

and clearly $C(1, d) = 1$, $C(2, 1) = 1/2$. The upper limits of summation are the *kissing numbers* in \mathbb{R}^2 and \mathbb{R}^3 , respectively. A proof that 24 is the kissing number in \mathbb{R}^4 was given only recently [23, 24]. Also, $C(6, 2) = 0$ since $\Omega(6, 2)$ is of measure zero.

Henze [21, 22] showed that

$$C(2, d) = \frac{2^{d+1} \pi^{d-1}}{\Gamma(d-1)} \int_0^\infty \int_0^\xi \int_{\theta_0}^\pi \xi^{d-1} \eta^{d-1} \sin(\theta)^{d-2} F_d(\xi, \eta) d\theta d\eta d\xi$$

where

$$\begin{aligned} \theta_0 &= \arccos \left(\frac{\eta}{2\xi} \right), \\ F_d(\xi, \eta) &= \exp [-f_d(\xi, \gamma) - f_d(\eta, \delta)], \\ \gamma &= \frac{\xi(\xi - \eta \cos(\theta))}{\sqrt{\xi^2 + \eta^2 - 2\xi\eta \cos(\theta)}}, \quad \delta = \frac{\eta(\eta - \xi \cos(\theta))}{\sqrt{\xi^2 + \eta^2 - 2\xi\eta \cos(\theta)}}, \\ f_d(x, y) &= \frac{\pi^{d/2} x^d}{2\Gamma(d/2 + 1)} \left[1 + I \left(\frac{y^2}{x^2}, \frac{1}{2}, \frac{d+1}{2} \right) \right] \end{aligned}$$

and I is the regularized beta function

$$I(z, a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^z w^{a-1} (1-w)^{b-1} dw.$$

(In [21], the definitions of γ and δ were mistakenly reversed; also, the expression within square brackets for $f_d(x, y)$ was unclear.) We obtain

$$C(2, 2) = 0.63317\dots = 2(0.316585\dots), \quad C(2, 3) = 0.70888\dots$$

Tao & Wu [17] independently showed that

$$\begin{aligned}
 C(2, 2) &= \pi \int_{\pi/2}^{3\pi/2} \int_0^\infty \frac{\tau}{(g(\tau, \theta) + \tau^2 h(\tau, \theta))^2} d\tau d\theta \\
 &\quad + 2\pi \int_{\pi/3}^{\pi/2} \int_{2 \cos(\theta)}^{1/(2 \cos(\theta))} \frac{\tau}{(g(\tau, \theta) + \tau^2 h(\tau, \theta))^2} d\tau d\theta
 \end{aligned}$$

where

$$\begin{aligned}
 g(\tau, \theta) &= \pi - \varphi + \frac{1}{2} \sin(2\varphi), & h(\tau, \theta) &= \pi - \psi + \frac{1}{2} \sin(2\psi), \\
 \varphi &= \arcsin \left| \frac{\tau \sin(\theta)}{\sqrt{1 + \tau^2 - 2\tau \cos(\theta)}} \right|, & \psi &= \arcsin \left| \frac{\sin(\theta)}{\sqrt{1 + \tau^2 - 2\tau \cos(\theta)}} \right|.
 \end{aligned}$$

(In [17], the absolute value signs in the definitions of φ and ψ were mistakenly omitted.) Even more elaborate integral formulas apply for $C(3, 2)$, $C(4, 2)$, $C(5, 2)$. Given the discrepancy between our estimate of $C(2, 2)$ and their estimate (see the Table), it seems doubtful that their approximation $\beta(2) = 0.284051\dots$ is entirely correct.

Table 1 *Old and New Calculations of Constants*

k	Tao & Wu estimate of $C(k, 2)/k!$	Current estimate of $C(k, 2)/k!$
2	0.3163335...	0.316585...
3	0.0329390...	?
4	0.0006575...	?
5	0.0000010...	?

A discrete version of the latter problem appears in [25, 26, 27, 28]. Let all the vertices of the lattice \mathbb{Z}^d be initially occupied by particles which can annihilate one-by-one their $2d$ nearest neighbors. More precisely, for each unit-length edge $\{u, v\}$ of the lattice, there is a Uniform $[0, 1]$ random variable $T_{\{u, v\}}$ representing the time of an attack along the edge. If vertices u, v are both occupied immediately prior to time $T_{\{u, v\}}$, then at time $T_{\{u, v\}}$ either vertex u or vertex v (each with probability $1/2$) becomes vacant (that is, one particle annihilates the other). If u, v are not both occupied at time $T_{\{u, v\}}$, then there is no change. Once a vertex becomes vacant, it remains vacant permanently. The variables $T_{\{u, v\}}$, considered over all unit-length edges $\{u, v\}$, are independent. By time 1, no two surviving particles can be adjacent. When $d = 1$, the probability that a given vertex remains occupied is $1/e = 0.3678794411\dots$. When $d = 2$, this probability is known to be in the interval $(0.227, 0.306)$ and is approximately 0.25 via simulation. Greater accuracy is desired.

0.1. Appeal for Help. Any assistance in completing Table 1, using the formulation in [17], would be deeply appreciated! We note similarity between this problem and others in [29, 30].

REFERENCES

- [1] P. J. Clark and F. C. Evans, On some aspects of spatial pattern in biological populations, *Science* 121 (1955) 397–398.
- [2] P. J. Clark, Grouping in spatial distributions, *Science* 123 (1956) 373–374.
- [3] M. F. Dacey, Proportion of reflexive n^{th} order neighbors in spatial distribution, *Geographical Analysis* 1 (1969) 385–388.
- [4] G. F. Schwarz and A. Tversky, On the reciprocity of proximity relations, *J. Math. Psych.* 22 (1980) 157–175; MR0609119 (82f:92050).
- [5] T. F. Cox, Reflexive nearest neighbours, *Biometrics* 37 (1981) 367–369; MR0673043 (83k:62122).
- [6] D. P. Shine and J. Herbert, Birds on a wire, *J. Recreational Math.* 11 (1978-79) 227–228; 15 (1982-83) 232.
- [7] S. Morris, Competition winners, *Omni* v. 2 (1980) n. 9, p. 108.
- [8] C. Kluepfel, Birds on a wire, cows in the field, and stars in the heavens, *J. Recreational Math.* 13 (1980-81) 241–245.
- [9] D. K. Pickard, Isolated nearest neighbors, *J. Appl. Probab.* 19 (1982) 444–449; MR0649985 (83g:60063).
- [10] M. F. Schilling, Mutual and shared neighbor probabilities: finite- and infinite-dimensional results, *Adv. Appl. Probab.* 18 (1986) 388–405; MR0840100 (87k:60041).
- [11] N. Henze, On the probability that a random point is the j^{th} nearest neighbour to its own k^{th} nearest neighbour, *J. Appl. Probab.* 23 (1986) 221–226; MR0826925 (87k:60133).
- [12] D. Eppstein, M. S. Paterson and F. F. Yao, On nearest-neighbor graphs, *Discrete Comput. Geom.* 17 (1997) 263–282; available online at <http://www.ics.uci.edu/~eppstein/pubs/EppPatYao-DCG-97.pdf>; MR1432064 (98d:05121).

- [13] C. A. S. Tercariol, F. de Moura Kiipper and A. Souto Martinez, An analytical calculation of neighbourhood order probabilities for high dimensional Poissonian processes and mean field models, *J. Phys. A* 40 (2007) 1981–1989; cond-mat/0609210; MR2316309 (2008a:82035).
- [14] F. D. K. Roberts, Nearest neighbours in a Poisson ensemble, *Biometrika* 56 (1969) 401–406.
- [15] R. Abilock and M. Goldberg, N riflemen, *Amer. Math. Monthly* 74 (1967) 720; 75 (1968) 1009; 89 (1982) 274–275.
- [16] S. Morris, Rifle puzzle, *Omni* v. 8 (1986) n. 4, p. 113; v. 9 (1987) n. 7, p. 141.
- [17] R. Tao and F. Y. Wu, The vicious neighbour problem, *J. Phys. A* 20 (1987) L299–L306; MR0888078 (88d:82021).
- [18] E. G. Enns, P. F. Ehlers and T. Misi, A cluster problem as defined by nearest neighbours, *Canad. J. Statist.* 27 (1999) 843–851; MR1767151 (2001b:60017).
- [19] C. M. Newman, Y. Rinott and A. Tversky, Nearest neighbors and Voronoi regions in certain point processes, *Adv. Appl. Probab.* 15 (1983) 726–751; MR0721703 (85m:60023).
- [20] C. M. Newman and Y. Rinott, Nearest neighbors and Voronoi volumes in high-dimensional point processes with various distance functions, *Adv. Appl. Probab.* 17 (1985) 794–809; MR0809431 (87d:60048).
- [21] N. Henze, Über die Anzahl von Zufallspunkten mit typ-gleichem nächsten Nachbarn und einen multivariaten Zwei-Stichproben-Test, *Metrika* 31 (1984) 259–273; MR0773815 (86i:62075).
- [22] N. Henze, On the fraction of random points with specified nearest-neighbour interrelations and degree of attraction, *Adv. Appl. Probab.* 19 (1987) 873–895; MR0914597 (89c:60063).
- [23] O. R. Musin, The kissing number in four dimensions, math.MG/0309430.
- [24] F. Pfender and G. M. Ziegler, Kissing numbers, sphere packings, and some unexpected proofs, *Notices Amer. Math. Soc.* 51 (2004) 873–883; MR2145821 (2006a:52015).
- [25] M. O’Hely and A. Sudbury, The annihilating process, *J. Appl. Probab.* 38 (2001) 223–231; MR1816125 (2001m:60226).

- [26] A. Sudbury, Inclusion-exclusion methods for treating annihilating and deposition processes, *J. Appl. Probab.* 39 (2002) 466–478; MR1928883 (2003k:60266).
- [27] A. Sudbury, The annihilating process on random trees and the square lattice, *J. Appl. Probab.* 41 (2004) 816–831; MR2074826 (2005g:82090).
- [28] M. D. Penrose and A. Sudbury, Exact and approximate results for deposition and annihilation processes on graphs, *Annals Appl. Probab.* 15 (2005) 853–889; MR2114992 (2005k:60307).
- [29] S. R. Finch, Poisson-Voronoi tessellations, unpublished note (2005).
- [30] S. R. Finch, Hermite’s constants, *Mathematical Constants*, Cambridge Univ. Press, 2003, pp. 506–508.