Discussiones Mathematicae
Graph Theory 25 (2005) 427–433

ON A SPHERE OF INFLUENCE GRAPH IN A ONE-DIMENSIONAL SPACE

Zbigniew Palka and Monika Sperling

Department of Algorithmics and Programming
Adam Mickiewicz University
Umultowska 87, 61–614 Poznań, Poland

e-mail: palka@amu.edu.pl
e-mail: dwight@amu.edu.pl

Abstract

A sphere of influence graph generated by a finite population of generated points on the real line by a Poisson process is considered. We determine the expected number and variance of societies formed by population of n points in a one-dimensional space.

Keywords: cluster, sphere of influence graph.

2000 Mathematics Subject Classification: Primary 60D05; Secondary 60C05, 05C80.

1. Introduction

Let $X = \{X_1, X_2, \ldots, X_n\}$ be the set of n points of R^d chosen randomly and independently with the same probability. Let

$$r(X_i) = \min_{X_j \in X \setminus \{X_i\}} d(X_i, X_j)$$

denote the minimum distance between X_i and any other point in X. The open ball

$$B_i = \left\{ X \in R^d : d(X_i, X) < r_i \right\}$$

with center X_i and radius r_i is *the sphere of influence graph* at X_i $(i = 1, \ldots, n)$. The random sphere of influence graph $SIG(X)$ has vertex set X with edges corresponding to pairs of intersecting spheres of influence.

In other words two vertices, say X_i and X_j, are connected by an edge if and only if

$$r(X_i) + r(X_j) > d(X_i, X_j).$$

The definition of the sphere of influence graph was introduced in [10] by Touissant. These graphs have been widely investigated recently. It is known that on the Euclidean plane the sphere of influence graph always has a vertex of degree at most 18 (see [5], for related results see [1, 7]). Füredi [4] showed that the expected number of edges $E(n, \mathcal{N})$ of the random sphere of influence graph on n vertices in normed space $\mathcal{N}$ is equal to

$$E(n, \mathcal{N}) = C(d)n + o(n),$$

where $C(d)$ is a constant depending only on the dimension of the space and

$$\frac{\pi}{8} 2^d < C(d) < \Big(1 + \frac{1}{2d}\Big) \frac{\pi}{8} 2^d.$$

This result was also proved independently by Chalker et al in [2]. In [6] Hitczenko, Janson and Yukich proved analogue result for variance. They showed

$$c(d)n \leq Var(n, \mathcal{N}) \leq C(d)n,$$

where constants $c(d)$ and $C(d)$ depend only of the space dimension.

Consider a population of n points generated by some random process in R^d and its resulting sphere of influence graph. We thereby generate clusters of points that are connected by edges. We call these clusters *societies.* The following questions arise:

- Let M denote the number of societies formed. What is the distribution of M?
- Let N denote the size of society, i.e., the number of individuals (points) in a society. What is the distribution of N?
- Form the convex hull of each society. What is
  - the content (area, volume) covered by a society?
  - the fraction of R^d that is contained in some society, as $n \to \infty$?

In this paper our main concern is with the random variable M.

2. One-Dimensional Societies

Let the population consist of n points, $X_i, 1 \leq i \leq n$, generated on the real line by a Poisson process. Let $X_{(i)}$ denote the corresponding order statistics and let

$$A_i = X_{(i+1)} - X_{(i)}, \quad 1 \leq i \leq n-1,$$

denote the lengths of the spacings between adjacent points. Societies are determined by the relative magnitudes of the spacings. The A_i are identically distributed. Moreover, the distribution of the vector of ranks of the A_i is discrete uniform.

Consider now the number of societies M formed by a population of n points. Clearly, M satisfies $1 \leq M \leq \lfloor \frac{n}{2} \rfloor$. For fixed n, let

$$P_n(M = m) = P_n(m)$$

denote the distribution of M. Obviously, $P_2(1) = P_3(1) = 1$.

The following technical lemma will be helpful in the proof of the main theorem.

Lemma 1. *Let $2 \leq m \leq \lfloor \frac{n}{2} \rfloor$. If for $n \geq 4$*

$$P_n(m) = \sum_{i=2}^{n-2} \frac{1}{4} P_i(1) P_{n-i}(m-1)$$

and for $n \geq 2$

$$P_n(1) = (n-1)2^{2-n}$$

then

$$P_n(m) = 2^{2-n} \binom{n-1}{2m-1}$$

for $n \geq 2$.

Proof. Let $m = 2$. Then

$$\begin{aligned} P_n(2) &= \sum_{i=2}^{n-2} \frac{1}{4} P_i(1) P_{n-i}(1) = \sum_{i=2}^{n-2} \frac{1}{4} 2^{2-i}(i-1)2^{2-n+i}(n-i-1) \\ &= 2^{2-n} \sum_{i=1}^{n-3} i(n-i-2) = 2^{2-n} \binom{n-1}{3}. \end{aligned}$$

Assume that lemma is true for $m \leq j$ and let $m = j+1$. Then by induction

$$\begin{aligned} P_n(j+1) &= \sum_{i=2}^{n-2} \frac{1}{4} P_i(1) P_{n-i}(j) = \sum_{i=2}^{n-2} \frac{1}{4} 2^{2-i}(i-1) 2^{2-n+i} \binom{n-i-1}{2j-1} \\ &= 2^{2-n} \sum_{i=1}^{n-3} i \binom{n-i-2}{2j-1} = 2^{2-n} \binom{n-1}{2j+1} \end{aligned}$$

which completes the proof. ∎

Theorem 2. *Let $E_n(M)$ and $Var_n(M)$ denote the mean and the variance of the number of societies formed in a population of n individuals. Then*

$$E_n(M) = \begin{cases} 2 & \text{for } n = 2, \\ \dfrac{n+1}{4} & \text{for } n \geq 3, \end{cases}$$

and

$$Var_n(M) = \begin{cases} 0 & \text{for } n = 2, 3, \\ \dfrac{n-1}{16} & \text{for } n \geq 4\,. \end{cases}$$

Proof. Let us assume that $A_{n-1} \geq A_{n-2}$. Then independently from the value of A_{n-3}, vertices $X_{(n-1)}$ and $X_{(n-2)}$ are connected by an edge. So by the above assumption the number of societies formed by population of n points is equal to one with probability

$$\frac{1}{2} P_{n-1}(1).$$

Now, let $A_{n-1} < A_{n-2}$. In this case the existence of only one society formed by n points, under condition that first $n-2$ points formed one society, depends on lengths $A_{n-3}, A_{n-2}, A_{n-1}$. Notice that two vertices $X_{(n-1)}$ and $X_{(n-2)}$ are not connected by an edge if the following inequality holds

$$A_{n-1} + A_{n-3} < A_{n-2}.$$

Assume that $A_{n-3} + A_{n-2} + A_{n-1} = l$. Then the probability of the event

$$A_{n-2} > \frac{1}{2} l,$$

i.e., probability that vertices $X_{(n-1)}$ and $X_{(n-2)}$ are not connected by an edge, is equal to

$$P\Big(A_{n-2} > \frac{1}{2}l\Big) = \frac{\frac{1}{8}l^2}{\frac{1}{2}l^2} = \frac{1}{4}\,.$$

Thus we obtain that if $A_{n-1} < A_{n-2}$, the number of societies formed by population of n points is equal to one with probability

$$\frac{1}{2}P_{n-1}(1) - \frac{1}{4}P_{n-2}(1).$$

Consequently population of n individuals forms one society with the probability

$$P_n(1) = P_{n-1}(1) - \frac{1}{4}P_{n-2}(1).$$

Solving this recurrence equation and considering boundary conditions we obtain

$$(*) \qquad P_n(1) = (n-1)2^{2-n}, \qquad n \geq 2.$$

Let B_i denote the event that two vertices, say $X_{(i)}$ and $X_{(i+1)}$, are the first ones that are not connected by an edge in the sphere of influence graph. It means that the number of societies formed by population of first i vertices is equal to one, while population of first $i+2$ vertices form two societies and the number of societies formed by population of last $n-i$ points is equal to $m-1$, assuming that $M=m$. Then

$$P_n(M=m|B_i) = P_{n-i}(M=m-1).$$

Therefore for $n \geq 4$

$$\begin{aligned} P_n(M=m) &= \sum_{i=2}^{n-2} P_n(m|B_i)P(B_i) \\ &= \sum_{i=2}^{n-2} \frac{1}{4}P_i(1)P_{n-i}(m-1)\,. \end{aligned}$$

This and $(*)$ imply (see Lemma 1) that

$$P_n(m) = 2^{2-n}\binom{n-1}{2m-1}$$

for $1 \leq m \leq \lfloor \frac{n}{2} \rfloor$.

Now we can calculate the expected value of number of societies formed by n points. For $n \geq 3$ we have

$$\begin{aligned} E_n(M) &= \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} i P_n(i) = \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} 2^{2-n} \binom{n-1}{2i-1} i \\ &= 2^{1-n} \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n-1}{2i-1} 2i = 2^{1-n} \left(2^{n-2} + (n-1)2^{n-3}\right) \\ &= \frac{n+1}{4} \,. \end{aligned}$$

Consequently

$$E_n(M) = \begin{cases} 2 & \text{for } n = 2, \\ \dfrac{n+1}{4} & \text{for } n \geq 3\,. \end{cases}$$

The second moment (for $n \geq 4$) is equal to

$$\begin{aligned} E_n(M^2) &= \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} i^2 P_n(i) = \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} 2^{2-n} \binom{n-1}{2i-1} i^2 \\ &= 2^{-n} \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n-1}{2i-1} (2i)^2 \\ &= 2^{-n} \left(n2^{n-2} + (n-1)2^{n-3} + (n-1)(n-2)2^{n-4}\right) \\ &= \frac{n(n+3)}{16} \,. \end{aligned}$$

And thus we obtain

$$Var_n(M) = \begin{cases} 0 & \text{for } n = 2, 3, \\ \dfrac{n-1}{16} & \text{for } n \geq 4\,. \end{cases} \qquad \blacksquare$$

Although we formulated the problem for R^d, we provided results only for the one-dimensional case. Even for simpler model of nearest neighbour graph (see [11] and [3]), higher-dimensional situations become complex enough to require simulation.

References

[1] P. Avis and J. Horton, *Remarks on the sphere of influence graph*, in: ed. J.E. Goodman, *et al.* Discrete Geometry and Convexity (New York Academy of Science, New York) 323–327.

[2] T. Chalker, A. Godbole, P. Hitczenko, J. Radcliff and O. Ruehr, *On the size of a random sphere of influence graph*, Adv. in Appl. Probab. **31** (1999) 596–609.

[3] E.G. Enns, P.F. Ehlers and T. Misi, *A cluster problem as defined by nearest neighbours*, The Canadian Journal of Statistics **27** (1999) 843–851.

[4] Z. Fűredi, *The expected size of a random sphere of influence graph*, Intuitive Geometry, Bolyai Math. Soc. **6** (1995) 319–326.

[5] Z. Fűredi and P.A. Loeb, *On the best constant on the Besicovitch covering theorem*, in: Proc. Coll. Math. Soc. J. Bolyai **63** (1994) 1063–1073.

[6] P. Hitczenko, S. Janson and J.E. Yukich, *On the variance of the random sphere of influence graph*, Random Struct. Alg. **14** (1999) 139–152.

[7] L. Guibas, J. Pach and M. Sharir, *Sphere of influence graphs in higher dimensions*, in: Proc. Coll. Math. Soc. J. Bolyai **63** (1994) 131–137.

[8] T.S. Michael and T. Quint, *Sphere of influence graphs: a survey*, Congr. Numer. **105** (1994) 153–160.

[9] T.S. Michael and T. Quint, *Sphere of influence graphs and the L_∞-metric*, Discrete Appl. Math. **127** (2003) 447–460.

[10] Toussaint, *Pattern recognition of geometric complexity*, in: Proceedings of the 5th Int. Conference on Pattern Recognition, (1980) 1324–1347.

[11] D. Warren and E. Seneta, *Peaks and eulerian numbers in a random sequence*, J. Appl. Prob. **33** (1996) 101–114.

Received 9 September 2004
Revised 4 May 2005