Clustering point sets using quadtrees and applications

1. Let $S = \{p_1, p_2 \ldots p_n\}$ be a set of $n$ points in the plane and let $k$ be a positive integer. We would like to cover all the points using $k$ disks of diameter $D$, such that $D$ is minimum - let it be denoted by $D_o$.

   (a) Let $q_1, q_2 \ldots q_{k+1}$ be a subset of $S$ such that for all $j \geq 1$, $d_j \geq d_{j+1}$ where $d_j = \max_i ||q_{j+1} - q_i||$ $i \in \{1, 2 \ldots j\}$.
   Argue that the $D_o \geq ||q_{k+1} - q_k||$.

   Solution: Consider the alternate definition that the points in $q_1 \ldots q_{k+1}$ are separated at least by distance $d$. Then any optimal solution will consist of $k$ disks such that some disk must contain at least 2 points (from pigeon hole argument). Therefore $D_o \geq d$.

   (b) Let the points $q_i$ be defined in the following way. Start from an arbitrary point $q_1 \in S$. Let $q_2$ be the furthest point from $q_1$ and for $1 < i \leq k + 1$, $q_i = \max_{p \in S} d(p, Q_{i-1})$ where $Q_{i-1} = \{q_1, q_2 \ldots q_{i-1}\}$ and $d(p, Q)$ denotes the distance from $p$ to its closest neighbour in $Q$.

   Prove that you can cover the points of $S$ using $k$ disks of radius $D_o$, thereby obtaining a factor 2 approximation algorithm.

   Solution: Clearly all points are within a distance of $d_k$ from $Q_{k-1}$ (as $d_k$ is the furthest distance neighbour of $Q_{k-1}$). So the disks of radius $d_k$ (diameter $2d_k$) centered at $Q_{k-1}$ cover all points.

   To prove the approximation bound, first you must prove that the points $q_i$ defined this way satisfy the conditions of the part (a). Let $d_{i-1} = d(q_i, Q_{i-1})$ $i > 1$. By induction show that $d_{j+1} \leq d_j$. So, $D_o \geq d_k$ and the covering disks have diameter $2d_k$.

2. The weight of an $\varepsilon$WSPD is defined as $\sum_k (|A_i| + |B_i|)$ where $\{(A_1, B_1) \ldots (A_k, B_k)\}$ are the pairs of the WSPD. For a fixed $\varepsilon$, construct a set of points $P$ such that the weight of a valid WSPD of $P$ has weight $\Omega(n^2)$.

   Solution: Given $\varepsilon$, consider $n$ points on a line at coordinates $\alpha, 2\alpha \ldots 2^{n-1}\alpha$ where $\alpha = \lceil 1/\varepsilon \rceil$. We have to show that for this configuration of points, the WSPD will always have weight $\Omega(n^2)$ whatever be the algorithm used. Let $p_1, p_2$ be the points in increasing order.

   For this point set, all the WSP $(A_i, B_i)$ must be such that they are contained in non-overlapping intervals (otherwise distance is 0). Moreover, if the points in $B$ are larger than $A$ then $|B| = 1$. Otherwise the diameter is larger than the separation. The weight of such a pair is $|A_i|$ and it covers exactly $|A_i|$ pairs, therefore the total weight must be at least $\Omega(n^2)$ for covering all pairs.

   Note that it is not sufficient to show that some WSPD has weight $\Omega(n^2)$ as that is always true for the trivial WSPD.