

Υπολογιστική Γεωμετρία

3ο Μέρος (γ): High-Dimensional Nearest neighbors

Γιάννης Εμίρης

Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστήμιο Αθηνών

Εαρινό 2014

Exact NN

Given set $P \subset \mathbb{R}^d$, and query point q , its **NN** is point $p_0 \in P$:

$$\text{dist}(p_0, q) \leq \text{dist}(p, q), \quad \forall p \in P.$$

Approximate NN

Given set $P \subset \mathbb{R}^d$, approximation factor $1 > \epsilon > 0$, and query point q , an **ϵ -NN**, or ANN, is any point $p_0 \in P$:

$$\text{dist}(p_0, q) \leq (1 + \epsilon) \text{dist}(p, q), \quad \forall p \in P.$$

Sort/store the points, use binary search for queries, then:

- Preprocessing in $O(n \log n)$ time
- Data structure requiring $O(n)$ space
- Answer the query in $O(\log n)$ time