Module 5: Disk-based Index Structures

Lecture 25: Analysis of Index Structures

Prev topic Next topic Prev page
Next page

The Lecture Contains:

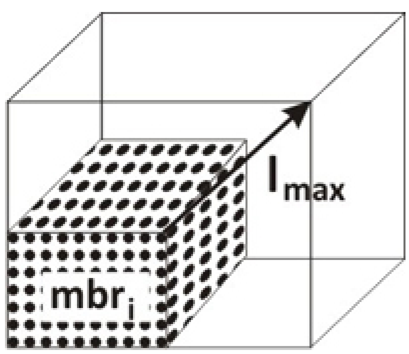
- Space-partitioning structures
- Data-partitioning structures: Hyper-cubes
- Data-partitioning structures: Hyper-spheres
- General structures

Space-partitioning structures

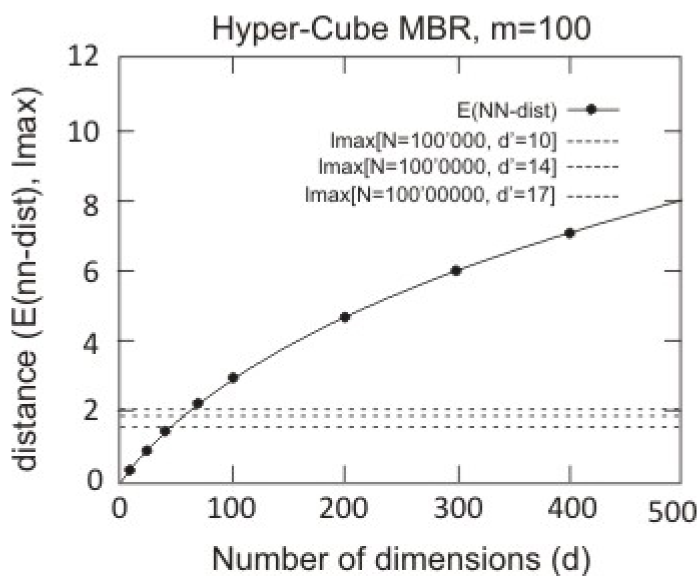
- Assume partition at only d' dimensions such that $2^{d'} \geq N/m$
- An MBR then contains d' sides of length 0.5 and sides of length 1
- Maximum distance of any point from an MBR, therefore, is

$$l_{max} = (1/2)\sqrt{d'} = (1/2)\sqrt{\lceil \lg N/m \rceil}$$

- l_{max} does not depend on d
- Hence, at some dimensionality, $E[nn^{dist}] > l_{max}$
- Volume of MS of MBR extended by $E[nn^{dist}]$ is then > 1



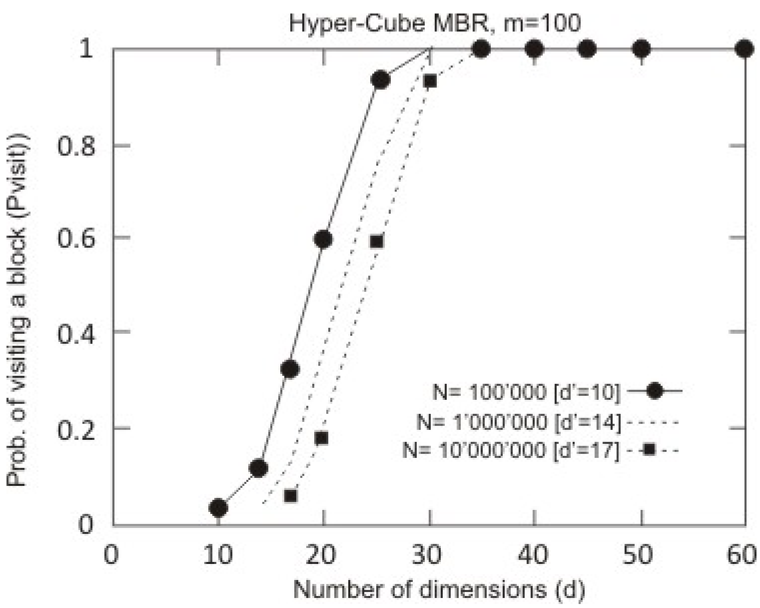
Space-partitioning structures



Prev topic Next topic Prev page
Next page

Data-partitioning structures: Hyper-cubes

- Analysis similar to that of space-partitioning structures



Data-partitioning structures: Hyper-spheres

- MBR consist of center C and its $m-1$ nearest neighbors

$$P_{visit} = vol(MS(sp^d(C, nn^{dist,m-1})))$$

$$= vol(sp^d(C, nn^{dist,m-1} + E[nn^{dist}]))$$

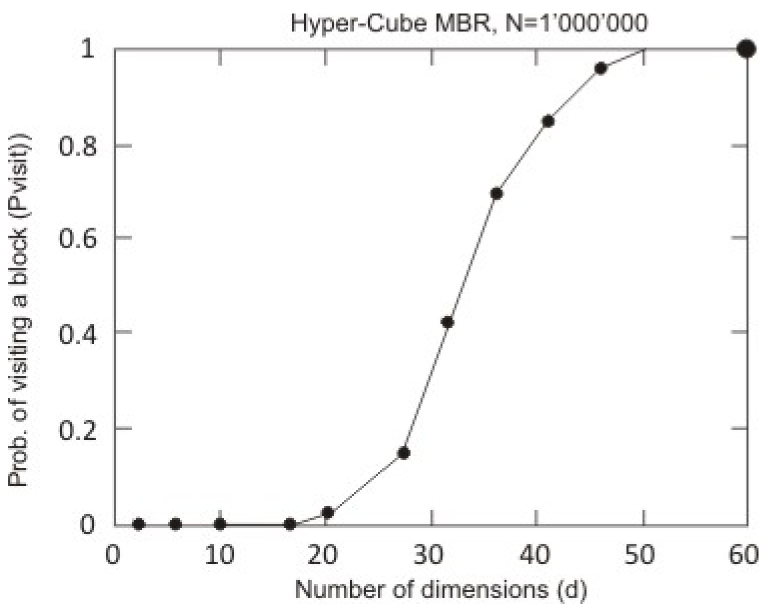
$$\geq vol(sp^d(C, nn^{dist,1} + E[nn^{dist}]))$$

$$P_{visit}^{avg} = vol(sp^d(2.E[nn^{dist}]))$$

- Volume increases with dimensionality
- At some dimensionality, volume > 1

Prev topic Next topic Prev page
Next page

Data-partitioning structures: Hyper-spheres



Module 5: Disk-based Index Structures

Lecture 25: Analysis of Index Structures

General structures

- Each MBR covers all its objects
- An MBR contains at least 2 objects
- MBR is convex
- For each MBR, pick two points and join them by a line
- Volume is lower bounded by minimum distance pairs