Nearest State/County Finder

Group 4

Dian Jin, Yachen Wang, Qilong Wang, Tianze Li, Yu Guo

Project Overview

- Given the locations of cities and counties in the US as reference points
- Users could enter a latitude and longitude to find the nearest counties

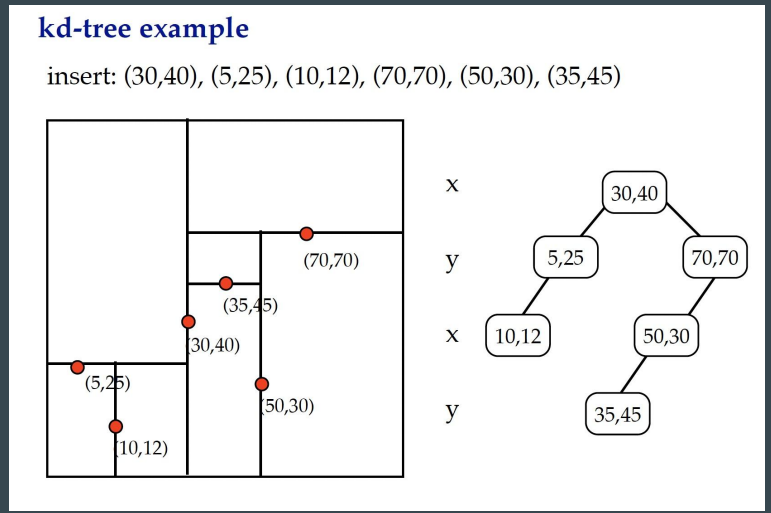
- Return the nearest K (1<=K<=10) counties and their states

- Dataset: The website of US Board on Geographic Names

Approach - KD Tree

- Data structure for organizing points in a K-dimensional space
- Binary search tree where data in each node is a K-Dimensional point in space

Build:

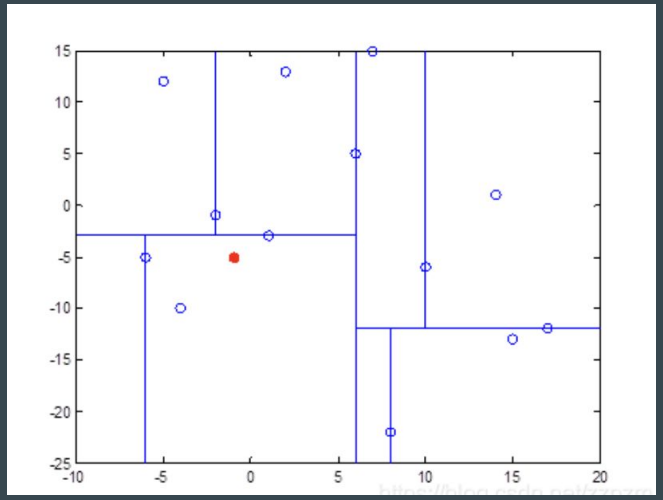


Step 1: choose the middle node on the x-axis and draw a vertical line.

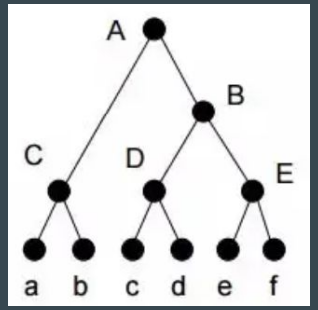
Step 2: choose the middle node on the y-axis and draw a horizontal line.

Step 3: repeat steps above until all the nodes have drawn lines.

Approach - KD Tree Search



Example
search nearest 3 points for (-1, -5)



1. DownSearch: Compare from the root to bottom: X-Y-X-Y-X...

2. Calculate the distance. Update or Discard.

3. Upsearch: Go to the upper level. Calculate the distance.

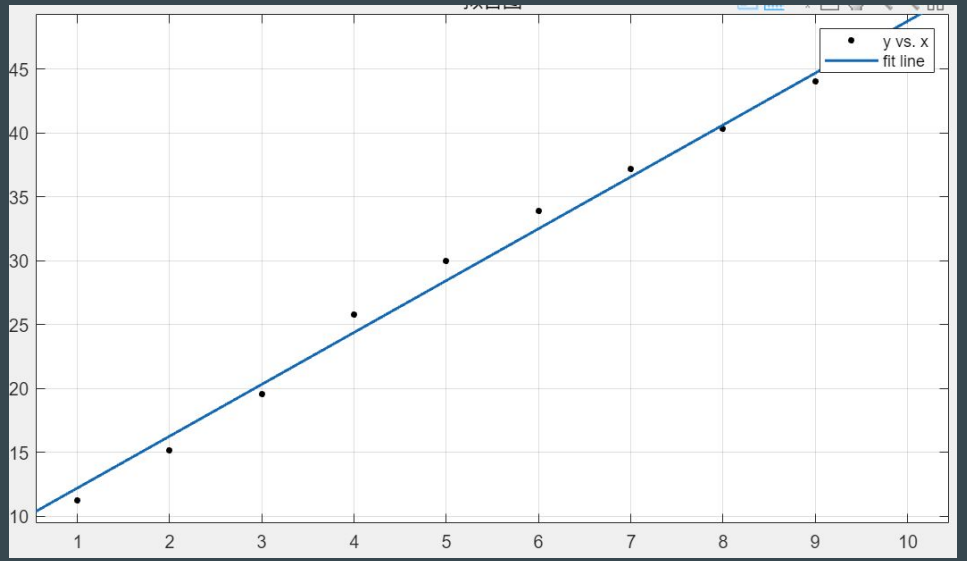
4. Decide whether to search other subtrees

.....

Time Complexity

Node numbers: $O(\sqrt{n})$

$$T(n) = 2^{k-1}T(n/2^k) + O(1)$$

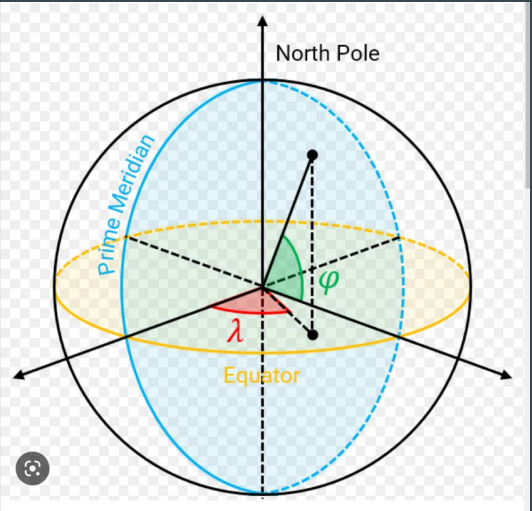


In our project, k = 2 $T(n) = \theta(k \cdot \sqrt{n})$

Difficulty encountered

- Building Tree - Sorting
- Search Tree - Upward Search
- K-nearest neighbors

Distance Calculation



Formula to calculate the distance between two points on Earth

$$x = (\lambda 2 - \lambda 1) * \cos((\phi 1 + \phi 2)/2);$$
$$y = (\phi 2 - \phi 1);$$
$$\text{Distance} = \sqrt{x*x + y*y} * R;$$

where ϕ is latitude, λ is longitude, R is earth's radius

Results

After entering a decimal latitude and a decimal longitude and a number K in the terminal, it will output the K nearest counties/states for this specific location.

```
PS C:\Users\dell\Desktop\ezyZip> g++ near.cpp -o near
PS C:\Users\dell\Desktop\ezyZip> ./near 40 -70 10
name: Siasconset, state id: MA, latitude: 41.263596, longitude: -69.971800, distance: 143.130816 km
name: West Chatham, state id: MA, latitude: 41.680423, longitude: -69.991800, distance: 187.007193 km
name: East Harwich, state id: MA, latitude: 41.708097, longitude: -70.033900, distance: 192.468616 km
name: Harwich Port, state id: MA, latitude: 41.672402, longitude: -70.064100, distance: 195.142107 km
name: Harwich Center, state id: MA, latitude: 41.692283, longitude: -70.069400, distance: 198.722361 km
name: North Eastham, state id: MA, latitude: 41.853915, longitude: -69.996800, distance: 206.165980 km
name: Northwest Harwich, state id: MA, latitude: 41.691710, longitude: -70.102600, distance: 210.492016 km
name: Dennis Port, state id: MA, latitude: 41.667703, longitude: -70.135800, distance: 223.847314 km
name: Madaket, state id: MA, latitude: 41.282618, longitude: -70.185500, distance: 228.963315 km
name: South Dennis, state id: MA, latitude: 41.705117, longitude: -70.153700, distance: 236.444263 km
```

Reference

- *Instruction of KNN Algorithm and KD Tree*. [Online]. Available: https://blog.csdn.net/zzpzm/article/details/88565645. [Accessed: 02-May-2023].
- *An Explanation of KD Tree*. [Online]. Available: https://zhuanlan.zhihu.com/p/53826008. [Accessed: 02-May-2023].
- B. Himite, “Calculating the distance between two points on Earth,” *Medium*, 29-May-2020. [Online]. Available: https://medium.com/swlh/calculating-the-distance-between-two-points-on-earth-bac5cd50c840. [Accessed: 02-May-2023].
- www.movable-type.co.uk Chris Veness, “Movable type scripts,” *Calculate distance and bearing between two Latitude/Longitude points using haversine formula in JavaScript*. [Online]. Available: http://www.movable-type.co.uk/scripts/latlong.html. [Accessed: 02-May-2023].