

Knn algorithm using k-d tree

Estudiante: Andrés Sánchez López

Carné: B26214

1. Introducción

Cada día se requieren hacer búsquedas cada vez más rápidas y eficientes, esto requiere de la búsqueda de algoritmos que utilicen estructuras de datos organizadas (como el k-d tree) que organicen la información de tal manera que se logre hacer más rápidas las búsquedas. Y las búsquedas muchas veces necesitan encontrar cuáles son los k puntos más cercanos a otro ya que muchos programas y algoritmos requieren este tipo de búsqueda en su código.

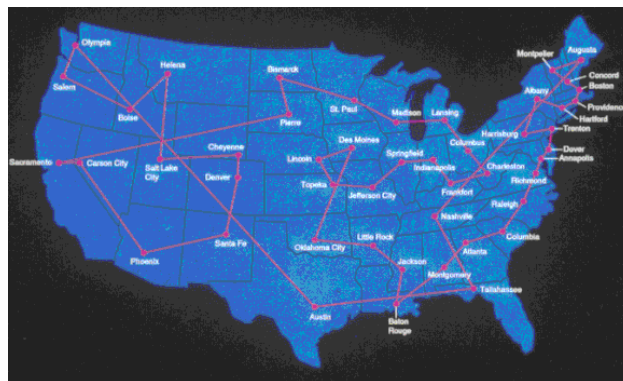


Figura 1: Mapa USA

Por ejemplo se puede ver en la figura 1 el mapa de USA con distintas locaciones marcadas y se busca la manera de recorrer todas las ciudades de la manera más rápida posible empezando por una ciudad cualquiera(A esto se le llama problema del viajante y tiene muchas aplicaciones), se puede buscar algún algoritmo que busque los vecinos más cercanos y que este algoritmo vaya iterando en cada nodo, buscando el vecino más cercano de cada uno, de tal manera que encuentre cuál es la ruta más rápida posible. EL algoritmo Knn es precisamente un algoritmo capaz de encontrar los k-vecinos más cercanos a un punto(K nearest neighbors, de ahí el nombre "knn"), por lo tanto se le puede usar para este tipo de problema. Este algoritmo, originalmente, no funciona utilizando un k-d tree pero se puede hacer un algoritmo que haga casi lo mismo que el algoritmo knn pero mediante la utilización de un k-d tree.



UNIVERSIDAD DE COSTA RICA
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA ELÉCTRICA
**ESTRUCTURAS ABSTRACTAS DE DATOS Y
ALGORITMOS PARA INGENIERÍA**

IE-0217
I CICLO 2014

PROPUESTA DE INVESTIGACIÓN BIBLIOGRÁFICA 2



2. Objetivos

2.1. Objetivo General

El objetivo general consiste en analizar el funcionamiento y los alcances del algoritmo knn cuando se le programa mediante un k-d tree.

2.2. Objetivos Específicos

Los objetivos específicos son:

1. Identificar las principales características del algoritmo knn así como algunos de sus usos.
2. Describir el funcionamiento del algoritmo knn utilizando un k-d tree.
3. Entender el algoritmo knn utilizando un k-d tree así como su importancia en la actualidad.
4. Aplicar el conocimiento que se obtenga sobre el algoritmo knn en alguna aplicación.

3. Metodología

Lo que se realizará es una investigación bibliográfica sobre una variación del algoritmo knn(utilizando k-d trees) utilizando fuentes primarias, para esto se leerá toda la bibliografía necesaria para poder llegar a entender de la mejor manera como es que funciona esta variación del algoritmo knn y por qué este es necesario, posteriormente se comparará con otros algoritmos utilizados para la misma función, para esto se buscará bibliografía sobre algoritmos que tengan el mismo fin que el knn para poder llegar a entender bien estos algoritmos y poder compararlos como debe de ser con el knn y ver cuáles son ventajas y desventajas de estos otros algoritmos con respecto al knn, y por lo tanto así determinar de una manera más precisa cuando es que la utilización de esta variación del algoritmo del knn es la mejor opción.

También, se buscará alguna implementación del algoritmo del knn utilizando un k-d tree para demostrar en clase, o se llevará a cabo la implementación de un prototipo demostrativo, para esto se deberá buscar bibliografía que sea de tipo práctica para poder entender ejemplos de problemas o aplicaciones en los que se utilice el algoritmo del knn para resolver ese problema o para lograr la aplicación requerida según sea el caso. Todo lo descrito anteriormente se realizará mediante el siguiente cronograma:

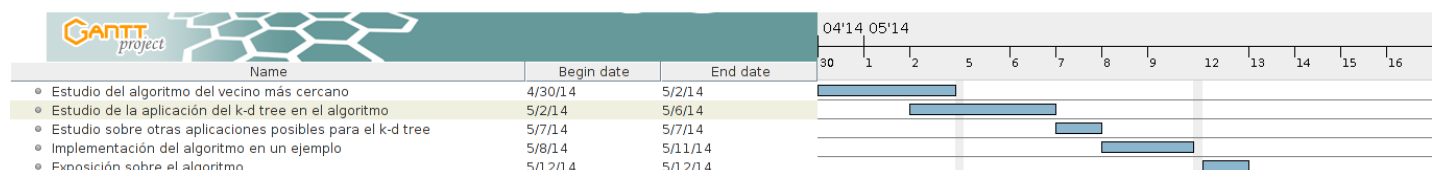


Figura 2: Cronograma



UNIVERSIDAD DE COSTA RICA
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA ELÉCTRICA
**ESTRUCTURAS ABSTRACTAS DE DATOS Y
ALGORITMOS PARA INGENIERÍA**

IE-0217

I CICLO 2014

PROPUESTA DE INVESTIGACIÓN BIBLIOGRÁFICA 2



4. Referencias

1. Hanan Samet. *The design and analysis of special data structures*. Addison Wesley, 1990.
2. Algoritmos y Estructuras de Datos para Búsqueda de Objetos Similares. (n.f). Recuperada 27, Abril, 2014, de:
<http://www.dcc.uchile.cl/gnavarro/ps/cita98.pdf>.
3. How to use a KdTree to search. (n.f). Recuperada 27, Abril, 2014, de:
http://pointclouds.org/documentation/tutorials/kdtree_search.php
4. k-Nearest Neighbors. (n.f). Recuperada 27, Abril, 2014, de:
<http://www.statsoft.com/textbook/k-nearest-neighbors>.