Fern Force 5
Dr. Pulimood
CSC 415 - Software Engineering

1 October 2013

Finding Potential Browfields through the Nearest Neighbor Algorithm

Problem Statement:

Habitat for Humanity requires information about the pollution history of plots of land in the Trenton area. If the organization receives a land donation, they would like an assessment of the quality of land. This pre analysis is to serve as a means to save on costs for soil contamination testing and time for planning, by directing Habitat for Humanity away from properties that will more than likely be too contaminated to pass the soil contamination test. If the soil is found to be contaminated, they are required by law to remediate the site before any further construction can take place. Testing and remediation has cost Habitat for Humanity a great deal of time and money in the past. In order to aid the foundation in future construction endeavors, this system will provide an estimation of the possibility of a specific site being polluted.

Objective:

We intend to implement the existing code from last semester that dealt with the nearest neighbor algorithm using the C++ language and the interface of the web application will be programmed using HTML. To achieve our goal, we will incorporate information from the database class. We will use data that pertains to local contaminated areas in order to determine the likelihood of whether or not the property in question is contaminated. We also will be working with Habitat for Humanity and the journalism class to complete the project. Habitat for Humanity will be involved to refine the product and ensure that it works as requested and the journalism class will be assisting with the gathering of data for the system.

Description of end product:

The end product should be able to determine, using the nearest neighbor algorithm, the potential pollution of the site in question based on negative influences in the area. The final product will be integrated within the SOAP website and also potentially within a mobile application from another group. Our specific component will work with other modules created by other groups within the class to create a system for Habitat for Humanity to assist with all future construction projects in the area.

Statement about the importance of the module:

This module is especially significant because it will be designed to determine the likelihood of whether or not a specific plot of land is a brownfield. As a result, this module has the potential to save Habitat for Humanity countless hours and a great deal of funding. Remediation of soil is very time consuming, labor intensive, and expensive. If Habitat for Humanity knows ahead of time whether or not soil is contaminated, they will be able to determine if it is advantageous to begin construction on that piece of land. The potential to discover brownfields before construction begins is incredibly beneficial to the foundation.

Other similar systems / approaches:

An alternate system which uses a similar nearest neighbor algorithm is estimated land value. Real estate companies will select a property and determine its value based on factors in the vicinity. These factors can be either positive, neutral or negative. The algorithm will be given data about these factors and will calculate accordingly and give an estimated value of the influences in the area which is then linked to the value of the land.

Other approaches in which the nearest neighbor algorithm has been used are in land-satellite imagery to determine characteristics of forests and determine their change over the years. This algorithm has even been used is to determine emotions using nodes on the face.

Explanation of why and what aspect of the module is innovative:

The module is innovative for Habitat for Humanity because they currently do not possess any systems to assist in their discovery of contaminated soil. The use of the nearest neighbor algorithm will allow us to take other contaminated sites and implement that information to help determine the likelihood of contamination in other nearby sites. This specific component of the overall system will provide insight into possible contaminated soil and has the ability to save the foundation a great deal of money.

Technologies and computer science concepts the team will need to learn/implement:

Our team will need to learn the basic concepts of the nearest neighbor algorithm to determine what data will need to be queried and how it is implemented in the algorithm. Our team will have to continue to learn and improve our existing C++ skills in order to understand the existing code written by the students last semester so we can make changes and integrate this specific component into the entire

system. We will also have to learn the syntax for HTML to incorporate the module into the web based application.

A diagrammatic representation: