Joanna McDonald
Algorithms Homework 2

Problem 1

Let us consider a long, straight country road with n houses scattered sparsely along it. (We can picture the road as a long line segment, with a western endpoint and an eastern endpoint, where the western point is at position 0, the eastern point is at position $L > 0$, and the n houses are at positions $x_1, x_2, \cdots, x_n$, respectively, where $0 \leq x_1 < x_2 < \cdots < x_n \leq L$.)

You want to place cell phone base stations at certain points along the road, so that every house is within (that is, less than or equal to) k miles of at least one of the base stations. Design an efficient algorithm that achieves this goal, using as few base stations as possible. Prove the correctness of your algorithm, and analyze its time complexity.

Main idea:

The main idea of this algorithm is to create a set for any optimal base station positions, B, for any set of n house locations X = (X_1, X_2, …, X_n). For the first house at X_1, the greedy algorithm will add the position of X_1 to k and add that position to B. This is the most greedy option, because it creates the greatest possible distance away from the current house position to the base station that will cover it. This maximizes the coverage of each base station. For each subsequent house visited, the position of the house should be compared to the position of the most recently added base station. If the current house's position is less than or equal to the most recent position in B plus k, then the current house is covered by the most recent base station. Otherwise, if the current house's position is greater than the most recent position in B plus k, then the current house is not covered by the most recent base station. So, the current house's position should be added to k, and this value of the new base station position should be inserted in B. This cycle will continue through all of the houses from X_1 to X_n, and then the set B with the optimal base station positions will be returned.

Pseudocode:

The input is the set of house locations X = (X_1, X_2, …, X_n) for n houses.
Initialize an empty set B to contain the base station locations
For each h in the set of all houses X:
    If index of h in X = 0:
        Insert the sum of the first house position and k (h+k) to the set B
    Else if current value of h in X > sum of last location in set B and k:
        Insert the sum of the current house position h and k (h+k) to the set of base station locations B.
Return set of base station locations B.

Prove its correctness:

Proof 1:

Proposition: Assume the algorithm is correct for any set of at most n houses X = (X_1, X_2, …, X_n) and there exists an optimal solution set of base station locations B = (B_1, B_2,…,B_m) for n houses.

Consider any set of n+1 houses H = (X_1, X_2, …, X_{n+1}).
$0 \leq X_1 < X_2 < \ldots < X_n < X_{n+1} \leq L$ where 0 is the westernmost point and L is the easternmost point on the road.
Given any optimal solution O = {B_1, B_2, …, B_p} for any set of n+1 houses H, the house at X_{n+1} is covered by the final base station at location B_p, which is equivalent to either B_m or $B_{m+1.}$
I will prove that B is an optimal solution using proof by induction on n+1.
Proof.
Given the proposition, we know that X_n is covered by the final base station at location B_m. For X_{n+1}, whether B_m covers it depends on the value of X_n + 2k. If $X_n + 2k \geq X_{n+1}$, then B_m covers X_{n+1} as well. This means that $B_m = B_p$. If $X_n + 2k < X_{n+1}$, then a new base station, B_p, will be placed at X_{n+1} + k. In this case, $B_m \neq B_p$. Rather, $B_m = B_{p-1}$, which can also be stated as $B_p = B_{m+1}$. Therefore, the house at X_{n+1} is covered by either the same final base station as X_n, B_m, or the base station after, B_{m+1}. Consequently, the algorithm will create an optimal solution for both X_n and X_{n+1}.■

Proof 2:
Proposition. For any set of at most n houses X = (X_1, X_2, …, X_n), there exists an optimal solution for base station locations B = (B_1, B_2,…,B_m) where base station B_1 is located at position X_1 + k.
$0 \leq X_1 < X_2 < \ldots < X_n \leq L$ where 0 is the westernmost point and L is the easternmost point on the road. I will prove this using proof by contradiction.
Proof. For the sake of contradiction, suppose that base station B_1 is not located at position X_1 + k.
If B_1 is located before X_1 + k, then the range of distance that base station B_1 covers is shifted to the left in comparison to if B_1 is located at X_1 + k. By the coverage of base station B_1 shifting to the left, there would be less coverage towards the right. Consequently, the locations of all following base stations (B_2, …, B_m) would also be shifted to the left, leading to an overall lower amount of base coverage to the right side. Therefore, the houses would not be covered optimally, and the set of base station locations would not be an optimal solution.
In the case that B_1 is located after X_1 + k, then X_1 would not be covered by the base station B_1. This is because X_1 will not be within k miles of a base station.
Therefore, B_1 cannot be located before or after X_1 + k in order to have an optimal set of base station locations. Thus, it must be the case that our assumption that B_1 is not located at position X_1 + k is false, so B_1 must be located at position X_1 + k.■

Time complexity:
The time complexity will be O(n). The algorithm performs a parse through a set X as many times as the number of houses, n, which takes O(n) time. In each parse, comparison statements are used, which will take O(1) time. Therefore the total time complexity for the algorithm is O(n).