LSH is based on the simple idea that, if two points are close together, then after a “projection”

operation these two points will remain close together.

在高维的空间中相似的（相近），经过投影后依然相似。相似的点会落在同一个bucket中。

Well, they always say with the advent of the internet. Records insides databases become billions. And objects are becoming more and more complicated, they are in high dimensional space, described by a vector. Sequential compare will give the exactly correct answer. But we do not have enough compute power. To reduce the linear time required by the sequential search, special data structure is invented to solve the problem. Traditional way are trees and hashing. The LSH is a kind of hash that has two important properties:

1. If two objects are close, they will have high probability to collision
2. If two objects are far apart, they will have low probability to collision

So if a query q is coming. We use several(N) LSH functions to Hash the q, get the has code and find them in the Hash table, object appears in the table buckets imply that they are close to each other.