Joseph S. Dulny
Booz Allen Hamilton
Quantum Annealing Enabled Cluster Analysis

Unsupervised machine learning applications are of critical importance in many fields given the lack of labeled training data. Clustering is a powerful unsupervised learning technique that involves dividing data points into groups that share “similar” characteristics. Many clustering algorithms aim to minimize a cost function: The sum of “within-the-cluster” distances between points. A straightforward approach which guarantees a global minimum involves examining all the possible assignments of points to each of the clusters. The number of possible assignments scales quickly with the number of data points and becomes computationally intractable even for very small datasets. In order to circumvent this issue, cost function minima are found using popular heuristic approaches such as k-means and hierarchical clustering. Due to their greedy nature, such techniques do not guarantee a global minimum will be found and can lead to sub-optimal clustering assignments. In this work, we describe how quantum annealing can be used to carry out clustering in an optimal manner. We map the problem to a quadratic binary optimization problem and discuss two clustering algorithms which are then implemented on commercially-available quantum annealing hardware. The first algorithm assigns N data points to K clusters, and the second is used to perform binary clustering in a hierarchical manner. We present our results in the form of benchmarks against well-known classical clustering methods and discuss the issues faced while implementing these algorithms on actual hardware.