

Supplement: A Uniform Sampling Framework for Center-Based Clustering with Outliers

No Author Given

No Institute Given

1 Empirical Experiments

In this section, we provide several supplementary experimental results. In Figure 1, we show the results for k -median clustering with outliers. To illustrate the scalabilities of our proposed algorithms, we vary the data sizes of the synthetic datasets, and show the results in Figure 2 (for Algorithm 1 and 3, we keep the ratio $\tau = \frac{k+k'}{k} = 2$). In addition, we vary the outliers ratio from 2% to 10% and show the results in Figure 3 (the algorithm MALKOMES cannot run when the ratio is larger than 6% due to its high complexity). In general, our uniform sampling based approaches are much faster than the baseline algorithms, and can achieve comparable clustering qualities.

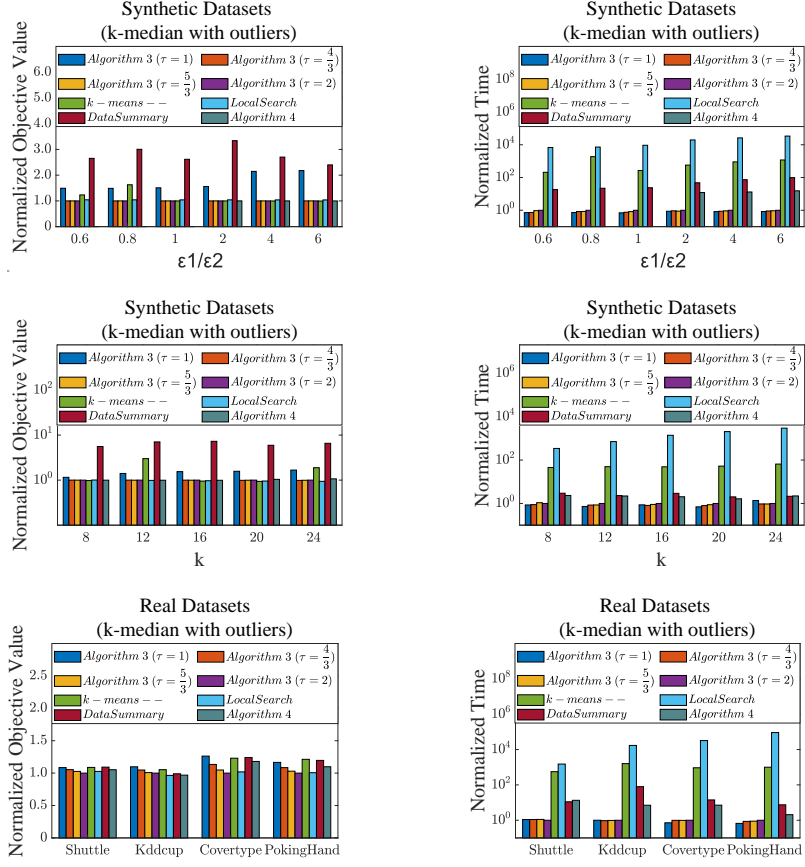


Fig. 1. The normalized objective values and running times on the synthetic and real datasets.

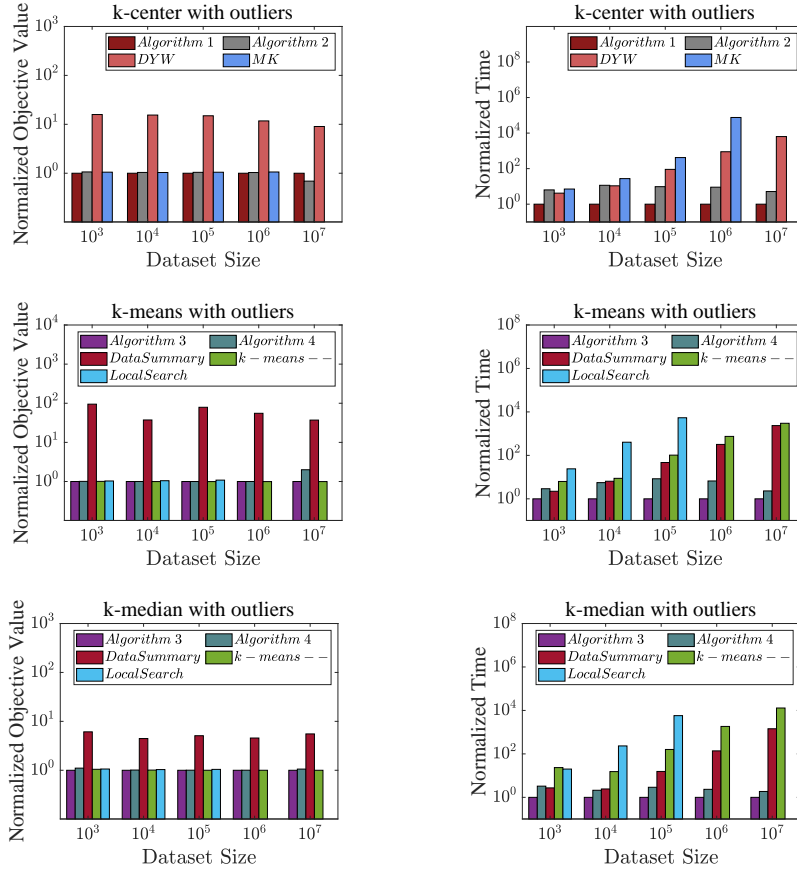


Fig. 2. The normalized objective values and running times on the synthetic datasets with varying n .

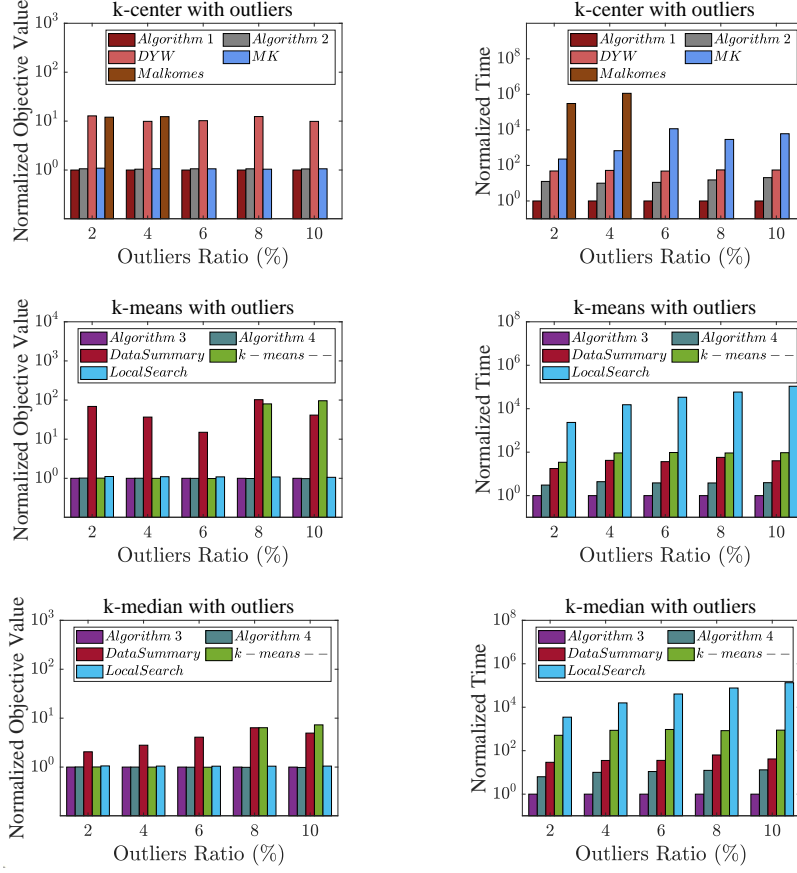


Fig. 3. The normalized objective values and running times on the synthetic datasets with varying z .