Labelling the power system states

There are 4 possible states for the system under analysis

- High load rate during peak hours -> Cluster 4
- Shut down of generator for maintenance -> Cluster 2
- Low load rate during night -> Cluster 1
- Disconnection of a line for maintenance -> Cluster 3

Once the K-Means Clustering algorithm is applied, the points are assigned in four clusters. A point represents a state of the system for a specific time, hence, it has 18 attributes that define it (9 voltage values and 9 angle values for each bus). The table below shows the centroids for each one of the clusters. The numeration of the cluster is done according the results of the code.

Bus	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1 Clark	$1\angle 0$	$1\angle 0$	$1\angle 0$	$1\angle 0$
2 Amherst	$1\angle -19.3$	$1\angle -21.1$	$1\angle 14.76$	$1\angle -23.56$
3 Winlock	$1\angle -23.7$	$1\angle -25.4$	$1\angle 9.48$	$1\angle -27.84$
4 Bowman	$0.90\angle -14.7$	$0.89\angle -15.8$	$0.98\angle -1.27$	$0.88\angle -16.31$
5 Troy	$0.86\angle -26.47$	$0.85\angle -27.70$	$0.97\angle -3.37$	$0.83\angle -29.73$
6 Maple	$0.96\angle -26.70$	$0.95\angle -28.41$	$1.01\angle 7.64$	$0.95\angle -30.85$
7 Grand	$0.91\angle -31.30$	$0.90\angle -33.30$	$0.99\angle 6.51$	$0.89\angle -35.91$
8 Wautaga	$0.94\angle -25.50$	$0.94\angle -27.40$	$0.99\angle 8.90$	$0.93\angle -29.83$
9 Cross	$0.82\angle -28.50$	$0.81\angle -30.30$	$0.96\angle -1.07$	$0.79\angle -32.40$

It can be seen that cluster 3 has positive angle values for most of the buses where the other clusters have not. This is because the reactive power flow for the states (points) gathered in cluster 3 is different from the other clusters. Since the generators do not produce reactive power, it has to be produced by the lines. Therefore, disconnection of a line for maintenance is the state associated to cluster 3.

Low load rate has as consequence the increase of the bus voltages, therefore, cluster 1 is the one associated to that state since the voltage of all the buses is always higher in that cluster.

Concerning the two remaining states, they have similar consequences regarding the voltage levels (both states tend to drop the voltage down), hence, voltage will not be a good indicator to label the two remaining clusters. The good indicator here will be the power flow. High load rate implies higher power flows than the case in where a generator is shutted down. The power flow over one line can approximately be calculated as follows.

$$P = \frac{E \cdot V}{X} \cdot sin\delta$$

Where E and V are the absolute voltage values of each terminal of the line, X is the line impedance and δ is the angular difference between the two so-mentioned voltages. Assuming the lines to have similar impedance, an estimation of the power flow for each line can be done considering the voltage values of the centroids. The calculations show that for almost every line the power flow through the lines in cluster 4 is higher than in cluster 2, therefore, high load rate corresponds to cluster 4 and shut down of generator corresponds to cluster 2.