COUNTRY: United Kingdom
REGISTRATION NUMBER: DLG5549

GROUP REF.: 220060028
PREF. SUBJECT: PS3
QUESTION N°: 3.1.1

Learning hidden patterns in data to isolate scenarios

The changing nature of the system means that System Operators are experiencing stability issues in parts of the network where previously there were no issues. A gradual decline in large transmission connected synchronous generators means the impact on system stability after a large disturbance is more pronounced than before (such as higher oscillations and longer settling time for rotor swings). The changes in system response after a disturbance are happening faster than ever and are more complex. It will be increasingly difficult to say definitively that the network behaviour will be guided by only a certain set of conditions or N-k contingencies.

Therefore, the current approach of studying a limited few scenarios (usually the winter peak condition in the context of the Great Britain system) will be inadequate and a probabilistic approach is more suited to account for all the uncertain factors in the system that can directly influence the system state.

However, due to the high dimensions of the problem (> 100) in a real system (i.e., the number of uncertain variables affecting the system response), a Monte-Carlo analysis will require studying lots of scenarios (> 70k) which is impractical for a national electricity transmission system (NETS) due to the associated computational burden. It is also difficult to represent the uncertainties by a joint distribution as the accuracy of the model decreases significantly as the number of dimensions increases beyond 20.

One way to approach this problem could be to borrow techniques from data science to capture the relationship between different uncertain variables (which will often have no closed form equation due to the complex nature of future systems) and identify the most onerous scenarios. This will help to reduce the size of the problem, both in terms of the number of scenarios to consider and the size of the network to simulate.

There are several techniques in data science which allow us to study the hidden patterns in data which is otherwise difficult to identify through engineering judgement alone. The underlying pattern in data (such as system state variables, generation and demand dispatch, line flows etc) provides means to identify the most important variables (or features in machine learning parlance). The features with high importance (a statistical measure based on a cost function) will provide an idea about the non-linear relationship of the system dynamics. This will also indicate the parts of the network which can be replaced by equivalents without compromising the accuracy of the solution. The feature selection method will therefore allow us to isolate the scenarios which are most important and reduce the total number of sample points to study.

A combination of a smaller pool of scenarios and a reduced network model along with automation and parallelisation of studies will significantly reduce the computational effort and will allow for stability studies to be undertaken in a probabilistic framework.