SECONDARY SUBSTATION LOAD PROFILING – IDENTIFICATION AND VISUALISATION OF CHANGES

Cliff Walton
PPA Energy – UK
cliff.walton@ppaenergy.co.uk

Sarah Carter
PPA Energy – UK
sarah.carter@ppaenery.co.uk

ABSTRACT

This paper outlines the analytical and graphical techniques being used to classify secondary substation load profiles and identify when changes occur. The visualisation techniques being developed to assist Planning Engineers with network assessment of load growth and changing load type are illustrated. The paper also details the projects findings in respect of DG connected at low voltage (LV) and the masking of load.

INTRODUCTION

UK Power Networks (UKPN) has Remote Terminal Units (RTUs) in all its grid and primary substations. In London Power Networks, one of the license areas operated by UKPN, around 60% of its secondary substations (11/0.43 kV) are equipped with RTUs (approximately 9000 sites). These units not only provide telecontrol capability, but also have the functionality to measure analogues including current, voltage, power factor, and harmonic distortion on each LV phase as well as disturbances.

The Low Carbon Network Fund (LCNF) Tier 1 Distribution Network Visibility (DNV) project aims to demonstrate the business benefits of the smart collection, utilisation and visualisation of existing data (i.e. analogues available from RTUs). Areas under investigation are the identification of localised load growth, changes in load profiles and determination of where and when Distributed Generation (DG) is masking load, which can have an impact on planning, outage calculations and restoration actions after an outage. Understanding the effect of LV generation connections might potentially allow more generation to be connected without the need for reinforcement.

This paper outlines the analytical and graphical techniques being used to classify secondary substation load profiles and identify when changes occur. The visualisation techniques being developed to assist Planning Engineers with network assessment of load growth and changing load type are illustrated. The paper also details the projects findings in respect of DG connected at low voltage (LV) and the masking of load.

The LCNF DNV project is being undertaken in a consortium comprising of UKPN, PPA Energy and Capula.

Additional investigation of the technical feasibility of intelligent dynamic load profiling, geo-datamining and visualisation of network data was undertaken as part of a TSB funded project by PPA Energy and the University of Portsmouth, and is also discussed in this paper.

CONTEXT

As the expected changes to LV load and generation occur (PV, EV, heat pumps) the power flow on the distribution networks is expected to change significantly and the reinforcement cost to the network is expected to increase.

Visualisation of secondary substation load profiles types (commercial, retail, data centre, residential etc) provide the distribution planner with visibility of opportunities for power transfers between adjacent secondary substations. This will result in easier planning processes to determine load-driven reinforcement options e.g. transfer or upgrade circuits and to confirm the level of diversity between new load to be connected and the existing load on a given feeder or substation. Making better use of load diversity between adjacent items of plant may allow additional connections of more new load and generation to be made and the deferral of some HV feeder and transformer reinforcement schemes.

LOAD PROFILING TECHNIQUES

The aim is to determine daily /weekly load profiles for each secondary transformer which can categorised by comparison with a set of automated profiles to enable visualisation of clusters of similar types.

The DNV project presently uses ½ hour average kVA data which is normalised to allow for focus on the shape of the usage pattern. This enables the clustering algorithm to compare the consumption habits from two households or commercial areas of different sizes.

Industrial loads

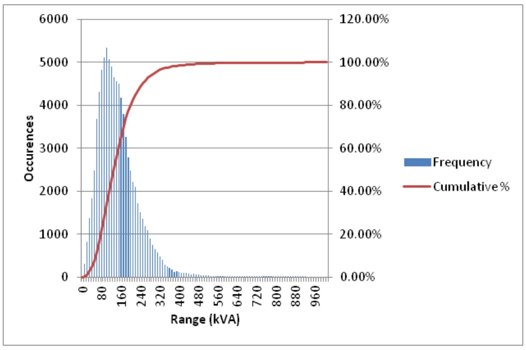
The load profiles of flat loads, which are typically industrial, are excluded from the clustering, but included at a later stage in the visualisation. A histogram of the daily range of data for the secondary substations associated with 5 primary substations is given in Figure 1. The limit range below which the load is considered as industrial load is defined as: $limit = \bar{x} - 2\sigma$

where: $\bar{x}$ is the average of the ranges

σ is the standard deviation

The limit calculated this way is 25 hence all data with a range less than 25 kVA were excluded from the clustering algorithm.

Figure 1 daily range of data for the secondary substations



Clustering Method

The K-means method was used to determine a set of daily load profiles. The K-means method uses Euclidean distance calculated for centre c = (c_1; c_2; ...; c_n) and point p = (p_1; p_2; ...; p_n) as

$$distance = \sqrt{\sum_i (c_i - p_i)^2}$$

Each data point is assigned to one of the cluster centre locations by selecting the centre that is nearest to that data point. Once all the data points are assigned, each collection of points is considered, the new centre of the allocated points is calculated and the centre for that cluster is reassigned. The points are then reallocated to their new nearest centre and the algorithm continues as before until no changes are made to the allocations of points for an iteration. The K-means algorithm returns the n centres of the clusters (called centroids), and the classification of each point in the cluster it belongs to.

The method is highly dependent on the initial random allocation of cluster centres. The k-means algorithm can be run multiple times to reduce this effect.

Optimisation

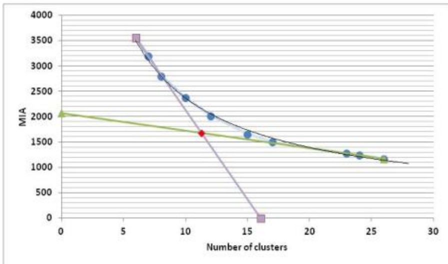
The optimal number of clusters depends on the shapes and the number of the points on which the algorithm is applied. A good clustering scheme will create clusters where the members of a particular cluster are closely grouped, but where the different clusters are well separated. The method used to determine the optimal number of clusters uses a measure assessing the quality of the clusters generated known as the Mean Index Adequacy (MIA):

$$MIA = \sqrt{\frac{1}{K}\sum_{k=1}^{K} d^2(r^{(k)}, C^{(k)})}$$

Where K is the number of clusters defined, $r^{(k)}$ is one of the load profiles assigned to cluster number k, and $C^{(k)}$ is the centroid of the cluster k.

The MIA assesses how the different points of the clusters are regrouped around their centroids. A low value of MIA indicates more compact clusters. Using this measure, we compared, for each group of data, stratified by season and day of the week, the MIA for different choices of n; as illustrated in figure 2. To determine the optimum number of clusters, two tangent lines were drawn from the points 5th to 6th clusters and 25th to 26th. The point of intersection gives an estimation of the elbow of the curve

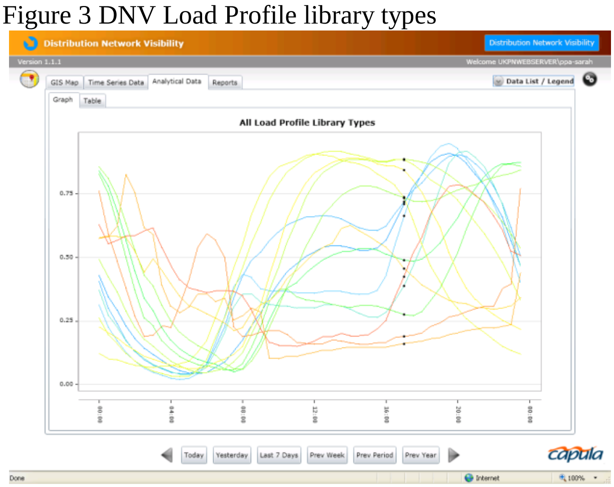
Figure 2 MIA by number of clusters



The optimum number of clusters determined in this example is between 10 and 15. The resulting 15 load profiles are given in figure 3. A range of automatically generated profiles from residential, commercial and loads with high night demand can be seen.

These load profiles are used in the DNV tool as the library types against which each daily load profile can be classified. This allows visualisation of adjacent substations and their profiles as well as highlighting any changes in profile over time.

Figure 3 DNV Load Profile library types



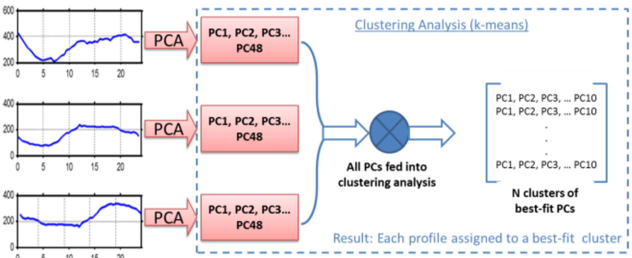
Principal Component Analysis

For the TSB Sprint project the Principal Component Analysis (PCA) technique was used to reduce data dimensionality so as to enable a 48-point daily profile (336 for weekly profiles) to be effectively explained using only a very small set of parameters. PCA analyses the data and extracts the most dominant components with the first few components usually containing most of the information contained within the data. When performed on daily 48-dimensional data, PCA extracts 48 components and lists them in order of dominance. The most dominant 10 Principal Components (PCs) are used to describe the profile.

The PCA analysis means that two years' worth of half hourly kVA data can effectively be described with only a few most significant components, e.g. 2 components hold approximately 94% of the information contained in each daily load from the data set, with just 5 components holding 98% of the information etc.

The PC counterparts are fed into a k- means clustering algorithm which analyses all of the input PCs and outputs N sets (defined by user) of 'best-fit' PCs. For an input set of 5000 daily profiles, 5000 sets of PCs are processed and the clustering algorithm creates N clusters. Each input profile is then assigned to a cluster which best describes the input. For the purpose of this study, the number of clusters was varied (N = 15, 25 and 50). Figure 4 shows an example of this process.

Figure 4 PCA and k-means clustering explanation



A 'cluster database' can then be designed so that each cluster describes a type of load, with the input profile being mathematically classified using its' principal components, and indicating the type of consumer(s) supplied.

Daily and weekly profiling

It was determined that the use of weekly profiles is likely to provide greater benefits using PCA as the accuracy of such techniques improve with increasing input values.

Analyses with and without normalisation were examined. It was concluded that as the amplitude of a profile represents size/scale, it would be disadvantageous to ignore this information in clustering hence the analysis proceeded without the use of normalisation.

Classification of daily and weekly load profiles was investigated and clustered (using K-means) into 15, 25 and 50 numbers of groups. The 'best fit' clusters were further classified into the following 5 categories; no data, residential, commercial, night load and industrial.

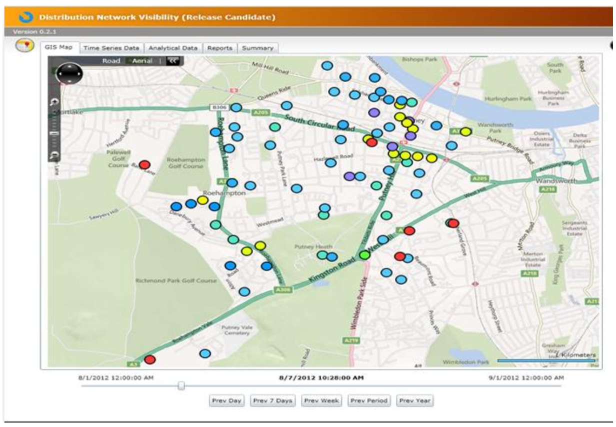
An interesting observation was that that the RTU clocks remain on GMT, therefore, all data between the start and the end of BST need to be time shifted by one hour to ensure correct time of day profiling.

CASE STUDIES

Load Profiling

An example of the DNV profile results is shown in figure 5. The coloured dots on the map represent a secondary substation, with the colours matching the load profile types in figure 3. The red dots illustrate a substation with significant night load such as a hospital (lighting). Blue dots represent residential profiles with yellow dots being associated with commercial centres.

Figure 5 Load Profiles for a group of secondary substations in London

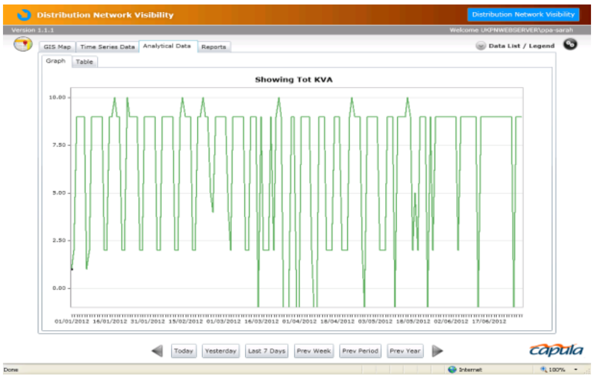


In order to meet a new load connection, such as a residential development, a Planning Engineer would be able to review historical load data from substations surrounding the new development, to visualize the load profiles and associated trends to optimise the use of existing assets and manage the system peak.

As customers begin to alter consumption behaviour by moving from gas to electric heating and heat pumps, EV charging points or small-scale distributed generation, is it expected that geo-spatial clusters of such loads will begin to form and may lead to extreme loading conditions. Visualisation of such clusters will warn of the changes assisting the Network Planner to manage the change effectively.

In order to view the load profile over time a slider can be used on the geographic profile view, or the user can view the change in load profile at each secondary substation graphically as illustrated in figure 6.

Figure 6 Graphical view of change in load profile: types 9 & 10 commercial (weekdays); type 2 residential (weekends)

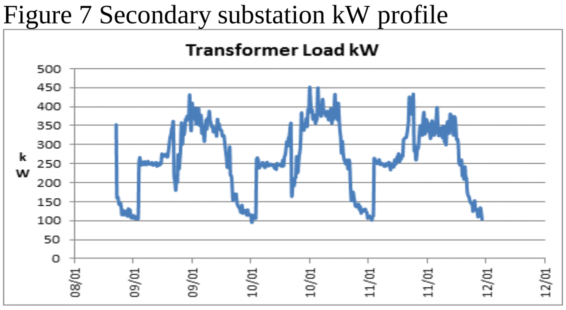


Embedded Generation

Profiling techniques can be used to automatically identify transformers with unconventional profiles. Further processing using step change quanta allows the load seen by the substation transformer to be decomposed to allow hidden DG and load masked by DG to be both identified, quantified and trended.

The substation transformer load profile in figure 7 can be seen to have some profile discontinuities.

Figure 7 Secondary substation kW profile



By processing the step changes in power between each time series measurement, +ve and –ve step changes of approximately 160 kW are identified in LV circuit P4 shown in figure 8.

Quanta can then used to automatically be used to decompose the load seen by the substation transformer into:

- a masked site load of 350kW, and
- a CHP generator of 160kW

that together produce the net demand of 200 kW seen by the P4 circuit (figure 9).

Figure 8 Step changes in load (~160 kW) on LV feeder P4

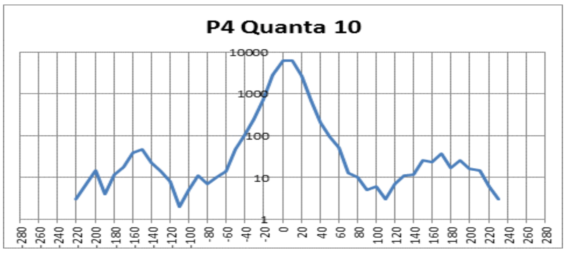
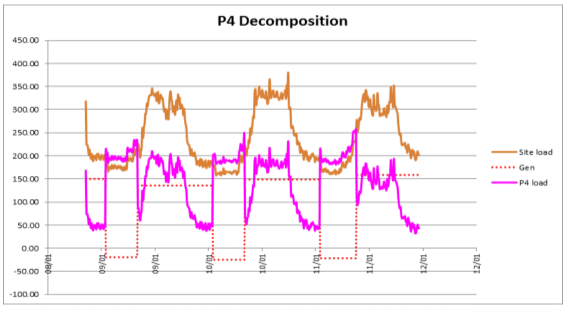


Figure 9 Decomposition of load (kW) of LV feeder P4



The same process can also be automatically applied to reactive power.

CONCLUSIONS

This paper has detailed the techniques presently used in the Distribution Network Visibility tool to normalise load profiles and assign them to a standard library type using a k-means clustering technique. It also outlines recent advances made in the Sprint system using PCA which reduces the data dimensionality so as to enable a load profile to be effectively explained using only a very small set of parameters. This finding is equally valid for other sets of profile data.

The load profiles can be visualised on a geographic basis and over time to assist Planning Engineers with network assessment of load growth, changing load type and optimising new connections so as to manage the system peak. In instances of reasonably sized LV connected generation it appears to be possible to identify the presence of generation and the associated masking of load which occurs when they are operational.

ACKNOWLEDGEMENTS

The authors would like to thank and acknowledge the following individuals and companies for their contribution to this work:

Matthieu Michel and Omer Khan, UKPN;
Alistair Frith, Andrew Varney, Chris Bradbury, Capula;
Borsu Shahnavaz, John Snow, Thomas Payet, PPA Energy;
Dr Branislav Vuksanovic, University of Portsmouth.