Simplified representation of a large transmission network for use in long-term expansion planning

Claude-Alexandra Fezeu[1], Keith Bell[2], Jiansong Ding[2], Patrick Panciatici[1] and Marie-Sophie Debry[1]
1. Réseau du Transport d'Electricité, France
2. University of Strathclyde, UK

***Abstract* — This paper concerns the development of a new approach to the simplification of representation of the spatial dimension of a large transmission network in order that the influences on bulk power transfers can be assessed in a practical way and the main routes that should be reinforced readily identified. The main challenge is to achieve a satisfactory clustering to deliver a number of zones that is small enough to make subsequent analysis of the expansion panning problem manageable but not so small as to neglect key regions of the original system. Two particular methods that have previously been proposed are described: a K-means algorithm and Dodu's mixed integer linear programming based approach. Each of them has some disadvantages, in particular that a direct interface between two zones might be derived that has no equivalent on the real network; or that it is difficult to control the number of zones. Hence, this paper describes a new hybrid method that ensures that resulting zonal delineations make engineering sense from the point of view of physical connections and allow some control over the number of zones. Results are presented in respect of the transmission network in Great Britain. Applications of the simplified network are discussed, not only in long-term planning but also in respect of the potential for use in transmission charging.**

Keywords— transmission planning, electrical distance, mixed integer linear programming, K-means clustering

I. INTRODUCTION

A transmission planner should design the network in order to facilitate its operation and avoid undue additional costs associated with insufficiency of network capacity. However, the operating conditions for which the network is being designed vary significantly, influenced by the variation of generation and demand of electricity. Those variations are three dimensional: temporal (hour, season, year...), spatial (regional, national…) and stochastic (best scenario, worst scenario…). In order to make sense of the design challenge and reach some decisions, approximations and simplifications are required. A collaboration between the French transmission system operator – Réseau de transport d'électricité (RTE) – and the University of Strathclyde aims to find ways to simplify a network around clusters of electrically close nodes.

Various simplified networks are used by Transmission System Operators (TSOs). For example, three different sets of zones are defined in the GB Ten Year Statement [1]: the first one is used for generation use of system charging; the second is used for demand use of system charging; and the final one is used for power flow studies in which boundaries between zones are intended to represent the main power flow bottlenecks. In studies concerned with secondary voltage control, control areas were built up and pilot nodes chosen for each. To define the areas, nodes that had similar voltage sensitivities were identified. This pilot node would be used to control voltages within the area [2, 3]. Thus, various methods were tried in order to simplify the network [4-6]. However, long-term planning requires some specific criteria for a simplified network: long term planning must take into account any planned changes in the future network; the simplified network should be both representative of the real network and flexible enough to include any significant changes to generation and demand patterns; and the method should be formal, in order to be used on any large transmission network. As electrical distance, unlike generation and demand, is constant over time for a given physical network, it is used in this paper to simplify the representation of a network.

The paper is organised as follows. Section II first introduces two existing algorithms: Dodu's linear integer mathematical program and the K-means clustering algorithm. Then, a new algorithm that combines the previous two, with the aim to overcome their individual downfalls, is proposed. Results of the new algorithm are analyzed in section III and Conclusions are presented in Section IV.

II. TWO EXISTING CLUSTERING METHODS

Clustering is the process of grouping data so that objects within a cluster have high similarity in comparison to one another, but are dissimilar to objects in other clusters [7]. In transmission networks, similarity can be associated with the electrical properties of nodes. In this work, two nodes are regarded as electrically similar if they are electrically close to each other. A simplified network in long-term planning can therefore be a set of clusters (each represented by a single node), where nodes are electrically close to each other within a cluster and significantly far from nodes in other clusters. To achieve such a clustering, two approaches based on electrical distance are used. The first one consists of interpreting the problem as a linear mathematical program. This point of view is inspired by some studies in France in the 1960s [6]. The second approach is the K-means clustering algorithm which is a popular top-down clustering algorithm [8]. The details of these two algorithms are introduced in this section. The main sample network used is the GB transmission network shown in Fig. 1. The voltages of this network are mainly 275kV (red lines) and 400kV (blue lines). However, while there is a significant number of 132kV (black lines) lines throughout Britain, in England and Wales they are classified as distribution while in

Scotland they are classified as transmission. Thus, black lines appear in Fig. 1 only in the north.

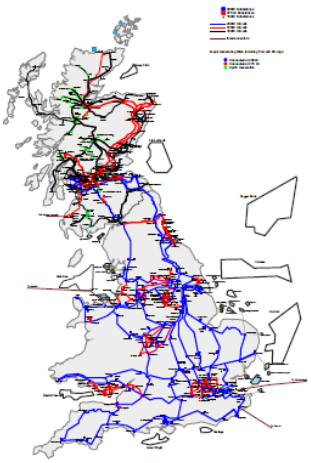

Fig. 1 GB existing transmission network [1]

A. Electrical Distance Matrix

The term 'electrical distance' is most commonly connected with short-circuit computing issues or, more precisely, determination of the short-circuit contribution of individual generators. In fact, electrical distance is a fictitious quantity that can be used to represent certain properties of a real network [8]. This fictional quantity is calculated using a sensitivity matrix, e.g. $[\partial v/\partial Q]$ or the impedance matrix $[Z_{bus}]$. (See, for example, [2, 3, 8, 9]). This study uses the impedance matrix only but, the method is equivalent if using $[\partial v/\partial Q]$ instead.

Given a small change in the current injected at a node k, the ratio of changes in voltages at nodes i and j can be found as $\frac{Z_{ki}}{Z_{kj}}$, $k \in [1,..,n]$ where n is the number of nodes in the network [2]. This ratio is argued in [2] to be reflective of the electrical coupling between nodes i and j, i.e. the electrical distance although the exact value found depends on the choice of k. In [2] and [10], it is further argued that the maximum attenuation between nodes i and j can be obtained by dividing the i, jth element of the impedance matrix by a corresponding diagonal element from the impedance matrix, e.g. :

$$\frac{\Delta V_i}{\Delta V_j} = \frac{Z_{ij}}{Z_{jj}} \tag{1}$$

where Z_{ij} is the i, jth element of impedance matrix $[Z_{bus}]$. However, although $[Z_{bus}]$ is symmetrical, it cannot be assumed that Z_{ii} is equal to Z_{jj}

If d_{ij} is the electrical distance between nodes i and j, d_{ij} is expected to show the following mathematical properties:

$$\begin{cases} d_{ij} = d_{ji} \rightarrow symmetric \\ d_{ij} = 0 \rightarrow i = j \\ d_{ij} = \infty \rightarrow i\ far\ from\ j \end{cases} \tag{2}$$

According to [10], to fulfill properties from (2), the electrical distance $[d_{ij}]$ may be defined as :

$$d_{ij} = -\ln\left(\frac{Z_{ij}}{Z_{jj}}\frac{Z_{ji}}{Z_{ii}}\right) \tag{3}$$

With this mathematical expression, the electrical distance is a quantity without units and is used in [2,3]. This expression has been chosen for this work because it is one of the more accurate ways to build electrical distance. Others ways are explained in [5]. However, more complex formulations make the computation slow and are judged not to be necessary for the work reported here.

B. Linear Integer Mathematical program

In 1960, when trying to plan generation and the high voltage transmission network, French engineers faced the same problems as described above of having several spatial and temporal variables to take into account. They looked for a method to reduce the spatial variables by grouping electrically close nodes into clusters. They used a mathematical program approach.

The aim of this approach is to cluster together nodes that are electrically close to each other. A threshold of electrical distance ε is introduced such that for two nodes i and j, with d_{ij} the electrical distance between i and j:

$$\begin{cases} d_{ij} \leq \varepsilon \rightarrow i\ and\ j\ are\ compatible \\ d_{ij} > \varepsilon \rightarrow i\ and\ j\ are\ incompatible \end{cases} \tag{4}$$

To say that nodes i and j are compatible means they are similar (close) enough to be in a same cluster; to say they are incompatible means they are dissimilar (far enough apart) and cannot be in the same cluster. An integer number t_i is introduced to represent the index of the cluster to which the node i belongs, i.e. if node i is in the third cluster, then $t_i = 3$.

With the electrical distance matrix $[d_{ij}]$ and a given threshold ε, the list of all pairs of incompatible nodes is built. That list is called C, thus:

$$(i,j) \in C \rightarrow t_i \neq t_j \tag{5}$$

If k is the total number of clusters, the aim of the study becomes:

Min (the number of clusters k)

Subject to:

$$\begin{gathered} |t_i - t_j| \geq 1 \\ 1 \leq t \leq k \end{gathered} \tag{6}$$

The problem can be solved by a linear integer program [11]. The objective function aims to obtain the minimum number of clusters such that two incompatible nodes are not in the same cluster. The approach is detailed in [6]. AMPL, a comprehensive and powerful algebraic modeling language for linear and nonlinear optimization problems, in discrete or continuous variables [12], has been combined with a solver of integer linear programs, developed and used at RTE, PNE, to implement and test the clustering algorithm. .

This algorithm has been tested first on a small network (22 nodes) and then on the whole French network (7000 nodes). Results for the small network were found after a couple of minutes, while after 72 hours of computation, there was still no result for the entire French network. The speed of the algorithm needs to be improved. Especially for large networks, the number of possible combinations of nodes into clusters should be reduced in order to improve the computation time.

Some operations may be done upstream from the program to reduce the number of nodes to be processed. A process dubbed 'pre-colouring' is introduced to find sets of nodes that are incompatible with each other the nodes that must be in different clusters. Nodes in different clusters have different 'colours'. The 'pre-colouring' proceeds as below and as shown in Fig. 2.

Before starting the operation, the electrical distance matrix L needs to be calculated and the threshold of the electrical distance ε needs to be defined. The details of this operation are: compare the element l_{ij} of matrix L with electrical distance ε. If $l_{ij} > \varepsilon$, nodes i and j are incompatible. Otherwise, they are compatible. Since the electrical distance matrix L is a symmetrical matrix, only the upper diagonal (or the lower diagonal) elements need to be checked. Once each element in either the upper or lower half has been checked, the number of other nodes that are incompatible with each node can be known. This number is defined as the "degree" of each node. 'Pre-colouring' now proceeds as follows:

1. All the nodes are sorted from the highest degree to the lowest degree.
2. Each node has two statuses: checked status (1 for checked and 0 for unchecked) and colour status (0 for uncoloured and an integer value larger than 0 representing the 'colour' of the node). Initially, all nodes are marked as unchecked and uncoloured.
3. Starting with the node with the highest degree, a value of 1 is assigned to its colour status and it is marked as checked.
4. Move to the next unchecked node and compare it with all colored nodes. If the current node is not compatible with any of the colored nodes, assign a new color to this node. Mark it as checked. If a node is compatible with more than one coloured node, its colour is set to that of the first compatible coloured node that was found.
5. Continue the process until all nodes have been checked.

With the above pre-processing operation, the number of nodes to be clustered by using the linear programming method described in Section II.B has been decreased. The computation is faster. It now took less than 10 minutes to run the French network and a couple of minutes for the GB network. For the latter, superimposed on a map of GB network taken from [1], the resulting clusters are shown in Fig. 3.

Despite the algorithm being fast and logically correct, the result is not physically correct. Cluster boundaries in Fig. 3 are drawn in red. Green is used to represent some particular areas inside clusters. It can be seen from Fig. 3 that Area 1 and Area 2 have been allocated by the algorithm to be in the same cluster. However, although, for the given threshold value ε, they are electrically close (in this case by virtue of the 400kV system), they are not physically directly connected to each other but have a connection via Area 3 which is in another cluster. This result is not acceptable because it does not reflect the real network in which there is no physical connection between Areas 1 and 2. So, even after speed improvement, this method has drawbacks.

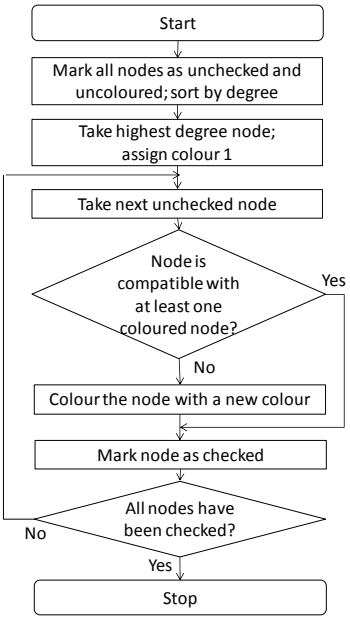

Fig. 2 The flowchart of "pre-coloring"

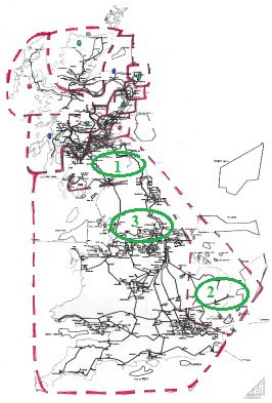

Fig. 3 The clustering result of the GB network

C. *K-means Clustering*

The K-means algorithm is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem [8]. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume *k* clusters) fixed a priori. The main idea is to define *k* nodes called centroids, one for each cluster. Then, each point belonging to a given data set is associated with the nearest centroid to form a cluster. After the initial centroids have been selected randomly, an initial grouping can be determined by calculating the distance between each node and each centroid. Each node is associated with the centroid to which it is closest. All the nodes associated with a particular centroid form a particular cluster.

In order to improve the clustering, the centroids are adjusted. For each cluster *i*, this involves finding the node l_i within the cluster for which the sum of distances to all the other nodes in the cluster is least, i.e.:

$$\sum_{l_i, m \in G_i, l_i \neq m} d_{l_i m} = \min_{(n \in G_i)} \sum_{(m \in G_i, m \neq n)} d_{nm} \qquad (7)$$

$$i \in (1, \ldots, k)$$

Each new central node replaces the old centroid of the cluster it belonged to. In this way, a new list of k-centroids is built. Each node is then tested again to see which centroid it is nearest to. If it turns out that a node is nearer to the centroid for cluster *j* than for its initial cluster *i*, it is removed from *i* and put into *j*. When each node has been tested, there is a new set of clusters. If any of the clusters have changed, equation (7) should be applied again to find the new centroid for the cluster. The process continues until no centroids move.

This method has its inconveniences. In particular, the optimization depends on the initial number of centroids, *k*, and where they are initially placed. The algorithm adjusts nodes and clusters according to the first clusters obtained with the first centroids. To highlight the problem, the algorithm has been tested on a small network of ten nodes shown in Fig. 4 and Fig. 5.

Consider the diamond, triangle and square nodes in Fig. 4 In Fig. 4, the square and diamond nodes were chosen as initial centroids (case (a)). As the triangle node is closer to the diamond node than the square node, it is associated with the diamond node. In Fig. 5, the square and the diamond nodes were chosen as initial centroids (case (b)). This time the diamond nodes is closer to the square node. But, in reality, those three nodes are so close together that they should have been grouped in the same cluster. This result would have been obtained if only one of them had been an initial centroid.

To ensure that nodes in different clusters will be dissimilar and node in the same cluster will be similar, initial clusters have to be as far away from each other as possible [4]. For the previous example in Fig. 4 and Fig. 5, an optimal clustering is represented in Fig. 6.

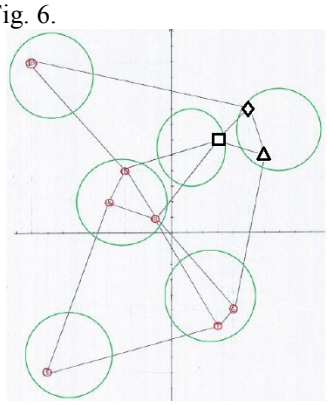

Fig. 4 An example network of 10 nodes, simplified using the K-means clustering algorithm, with the first centroids chosen randomly – case (a)

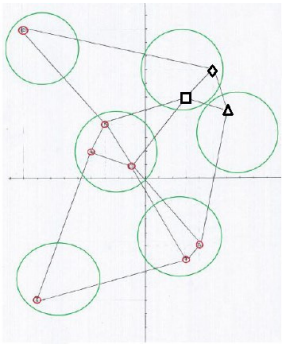

Fig. 5 An example network of 10 nodes, simplified using the *K*-means clustering algorithm, with the first centroids chosen randomly – case (b)

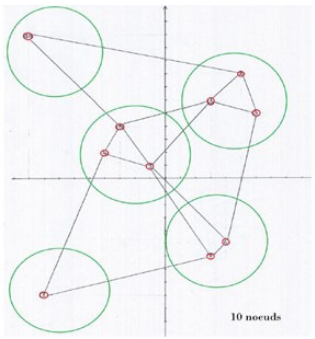

Fig. 6 Example of 10 nodes network, optimally clustered by the *K*-means algorithm, k=5

III. NEW HYBRID METHOD

A. Outline of the method

The mathematical program described in section II.B gives results that are logically exact but electrically incoherent. The K-means results depend on the initial clusters. A new algorithm is therefore proposed that combines the previous two, with the aim to overcome their individual downfalls.

A more optimal choice of *k* centroids is to place these initial nodes as far away from each other as possible [13]. At the very least, in the context of the power network simplification problem described above, they should all be incompatible with each other. If the pre-colouring algorithm described in section II.B is used here, according to a given threshold value, it gives a list of pre-coloured nodes in which nodes of a certain colour are incompatible with nodes of a different colour. Moreover, for the given electrical distance threshold ε, the list gives the minimal number of clusters needed, which can help answer the question: what is an appropriate value of *k* in the K-means algorithm. The first set of *k* centroids can be chosen as one node of each colour where each colour represents a separate cluster. In this way, the electrical distance between each of the initial centroids will be at least ε.

Only those nodes near the boundary between two clusters and at a comparable distance between two centroids might be in different cluster. Clusters are iteratively refined using the K-mean algorithm as described above.

B. Threshold of electrical distance

A limitation of the above approach is that the pre-colouring needs the threshold of electrical distance to be given. This value is an input chosen by the user. The effect of this is illustrated in the following example.

The hybrid method is applied to the GB network shown in Fig. 1 with two values of threshold (ε) which are 30 *p.u.* and 20 *p.u.*. The results are shown in Fig. 7 and Fig. 8 respectively in which the clusters are represented by dashed purple lines. From these two figures, it can be observed that the smaller the threshold is, the more likely nodes are to be incompatible to each other which leads to the network having more clusters. As a consequence, more centroids are required to group the network. A criterion is required to choose the threshold suitable for the specific study

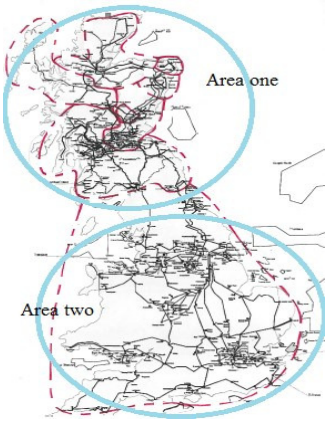

Fig. 7 The clusters of GB network by using the hybrid method with $\varepsilon = 30\ p.u.$

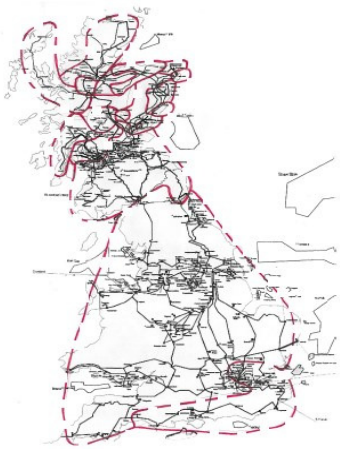

Fig. 8 The clusters of GB network by using the hybrid method with $\varepsilon = 20\ p.u.$

In order to find a criterion to determine an electrical distance threshold, the effect of different values of the threshold on the number of clusters obtained is observed below and two new indicators α and β are introduced. α is the normalized value of the number of clusters, i.e. number of clusters relative to the number of nodes in the original network, and β is the normalized value of the threshold, i.e. the value of the threshold relative to maximum electrical distance between any two nodes in the network. α and β are found by:

$$\begin{cases} \alpha = \dfrac{k}{n} \\ \beta = \dfrac{\varepsilon}{d_{max}} \end{cases} \tag{8}$$

where:

n: the number of network nodes

d_{max}: the maximum electrical distance between two nodes

ε: the threshold of electrical distance between two nodes

k: the number of clusters obtained after computation.

Two network models are used to evaluate the relationship between α and β: the GB network which has multiple voltage levels, and a part of the French transmission network which has 1126 nodes most of which are at 400kV. The results for these two networks are shown in Fig. 9 and Fig. 10 respectively which show the value of α for given values of β. The red dotted line is the trend of the black curve. For GB network, d_{max} is equal to approximately 200 *p.u.* while the French network's d_{max} is 7 *p.u.*. The maximum electrical distance is much longer in the GB model due to the inclusion of 132kV circuits in the north plus 275kV. For comparison, the trend of α versus β for the northern part of GB is shown in Fig. 11. (The maximum electrical distance in this network model is still around 200 *p.u.*).

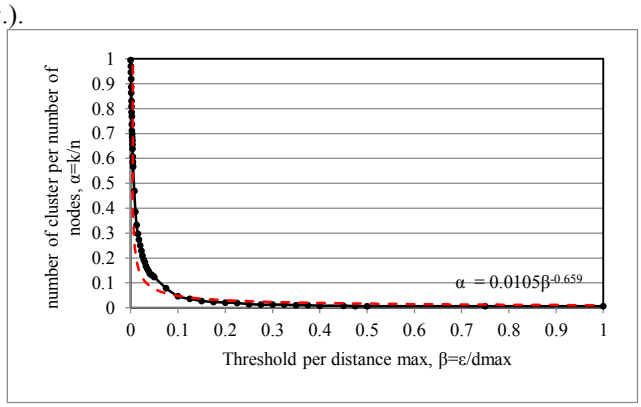

Fig. 9 The relationship between α and β for the GB network whose maximum electrical distance is approximately 200 *p.u.*

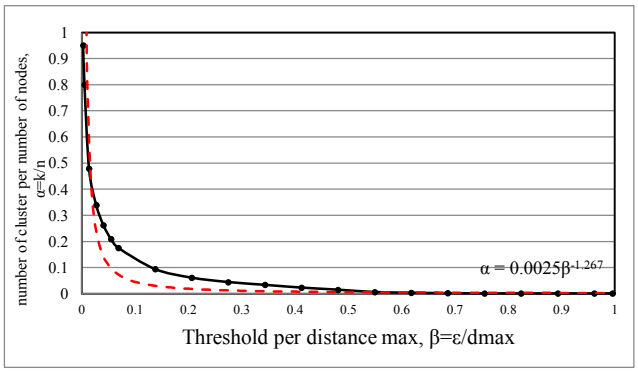

Fig. 10 The relationship between α and β for the French network whose maximum electrical distance is 7 *p.u.*

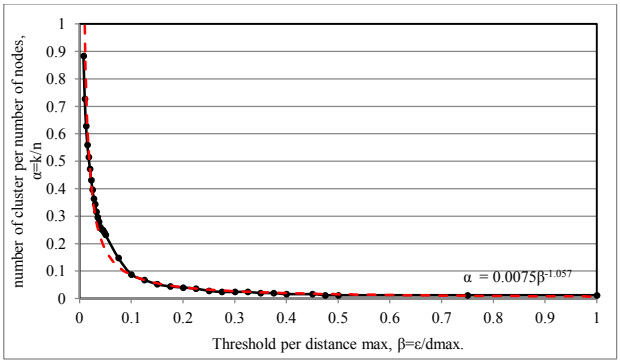

Fig. 11 The relationship between α and β for the northern part of the GB network, mainly 132kV, whose maximum electrical distance is approximately 200 *p.u*

The curves in Figs. 9-11 are exponential. The equations of the Fig. 10 and Fig. 11 curves are approximately:

$$\alpha = 0.0025 \times \beta^{-1.3}$$

$$\alpha = 0.0075 \times \beta^{-1.1}$$

They are very close to each other. Although it has not been rigorously tested for many different transmission networks, it may be expected that all curves have approximately the same trend, at least when the networks comprise only one or two voltage levels (discussed in the next section); then, in order to achieve a network simplification to give approximately a certain number of clusters, users should calculate the maximum electrical distance in the network for which the number of nodes is already known, and use the curves to identify an appropriate threshold value ε. For example, if the desired number of clusters is one twentieth of the number of nodes, then $\alpha = 0.05$, which leads to the corresponding value of β, and hence the value of ε to use in the simplification process.

C. *Voltage level*

In the results shown in Figs. 7 and 8, it can be seen that, unlike when Dodu's linear integer programming method was applied in its original form, the nodes are physically linked to each other within same cluster. In addition, the results are determined by voltage level.

In Fig. 7, the northern and southern areas of the system are shown (by mean of solid, blue circles). The northern network operates at a lower voltage level (132kV and 275kV with some 400kV) than the southern network (all 275kV and 400kV). It can be seen that there are many more clusters in area one than in area two, although area two is physically bigger than area one.

The main purpose of network simplification for long-term planning is to identify the main regions that tend to import or export power under different future operational conditions and market scenarios. The areas found by a clustering algorithm such as those described above should contain the main generation and demand centres as well as showing the strength of connection between. The electrical distance method succeeds in doing the latter but fails in the former if multiple voltage levels are included in the original network model, and these voltage levels are not well spread across the network.

A system planner using a network simplification method may choose to exclude lower voltages, e.g. sub-transmission and distribution voltages, when calculating the electrical distance of the transmission network. For many systems, this represents little contradiction between the function of the simplification process and the aims of long-term planning since the distribution or sub-transmission systems generally do not interconnect main generation or demand centres. However, for both geographical and historical reasons, in the north of Britain especially in the north of Scotland, large areas are covered by 132kV lines that do form part of the main interconnected network. This will not always be the case, though, as the connection of wind generation in the north is driving investment in additional transmission capacity, including uprating to 275kV and 400kV [1]. A 'naïve' application of the electrical distance would actually have helped to reveal this need – a threshold value of over 20 *p.u.* ($\beta \geq 0.1$) in Fig. 9 leads to some clusters in lower voltage level while nodes at higher voltage levels are hardly partitioned at all. The threshold needs to be less than 10 *p.u.* ($\beta \leq 0.05$) for the high voltage network in Fig. 9 to be partitioned. However, one of the things the electrical distance based clustering can do is to reveal the likely main bottlenecks restricting power flow; the density of clusters in the north of Britain would have confirmed what is true, that the network there is relatively weak.

IV. CONCLUSIONS

The study reported here has sought to develop a methodology to reduce the size of a large network in order that long-term network planning studies can be more easily carried out. In addition, in order that planning studies can be carried out by a number of different transmission owners responsible for different parts of a large network and to enable a new round of studies after any changes to the network, the methodology developed should be easy to apply in a consistent way.

In order to avoid dependency on a particular scenario of generation capacity or its dispatch, the methodology developed is based on electrical distance with the objective being to identify groups of nodes that are electrically close to each other and can each be represented in a simplified planning model by a single node.

It is concluded that two clustering approaches described in the literature, one based on a linear integer programming approach to the minimization of the number of clusters in which nodes that are closer to each other than a certain threshold are associated with each other, the other on K-means clustering, have significant limitations. In the case of the former, it is that nodes can be placed within the same cluster even when they are physically connected only via nodes in a different cluster. In the case of the latter, it is that the results depend to a large extent on the choice of the initial set of centroids.

A hybrid method has been proposed that uses the idea of a threshold of electrical distance from the first method to determine the initial set of centroids for the K-means clustering.

In testing on the GB transmission network, it is revealed that, in a network with multiple voltage levels, the highest voltage level might all be included within a single cluster

whereas the lowest level is split into many. When one of the purposes of long-term network planning is to reveal the effects on bulk system transfers of different scenarios regarding the development of generation and demand and different operating patterns, this can mean that some key generation and demand centres are contained within a single cluster. This has the result that power transfers between them are not revealed. On the other hand, it does show the weakest areas of the system that, depending in particular on how generation is developed, require significant transmission expansion.

A further issue concerns the choice of electrical distance threshold. Through a number of tests on the GB and France transmission networks, it has been found that the relationship between the degree of simplification of the network and the distance threshold relative to the maximum found on the network can be quite well generalized. Although tests should be done using other networks to confirm this generalization, it suggests that, if a planner has an idea of how many equivalent nodes they want in their simplified network, they can identify an appropriate value of electrical distance threshold for the simplification algorithm.

As well as being useful for long-term network planning in order that the planner can devote more time to the practical exploration of stochastic aspects of generation expansion and operational patterns that affect bulk power flows, the network simplification methodology can be used for identification of system zones for transmission charging in which (for generation) electrical distance from demand centres is key or (for demand centres), distance from generation is key.

.

REFERENCES

[1] NationalGrid. *Electricity Ten Year Statement 2012*. Available: http://www.nationalgrid.com/NR/rdonlyres/DCCBEA81-166A-48BF-928D-3B5807FF1F79/57729/ETYS_2012_Appendix_A1.pdf

[2] P. Lagonotte, J. C. Sabonnadiere, J. Y. Leost, and J. P. Paul, "Structural analysis of the electrical system: application to secondary voltage control in France," *Power Systems, IEEE Transactions on,* vol. 4, pp. 479-486, 1989.

[3] Zhong Jin, E. Nobile, A. Bose, and K. Bhattacharya, "Localized reactive power markets using the concept of voltage control areas," in *Power Engineering Society General Meeting, 2006. IEEE*, 2006, p. 1 pp.

[4] Zhang Yanfeng, Xu Xiaofei, and Yunming Ye, "NSS-AKmeans: An Agglomerative Fuzzy K-means clustering method with automatic selection of cluster number," in *Advanced Computer Control (ICACC), 2010 2nd International Conference on*, 2010, pp. 32-38.

[5] Liu Hong, Bao Hai, and Liu Lei, "A new method about calculating electrical distance," in *Power Engineering and Automation Conference (PEAM), 2011 IEEE*, 2011, pp. 382-385.

[6] J. C. Dodu, Ludot, J. P., Pouget, J., "Sur le regroupement optimal des sommets dans un réseau électrique," *RAIRO - Operations Research - Recherche Opérationnelle,* vol. 3, pp. 17-37, 1969.

[7] Wang Haixia, Liu Rao, Li Weidong, and Zhao Caihong, "Power Flow Tracing with Consideration of the Electrical Distance," in *Power and Energy Engineering Conference, 2009. APPEEC 2009. Asia-Pacific*, 2009, pp. 1-4.

[8] S. Blumsack, P. Hines, M. Patel, C. Barrows, and E. Cotilla Sanchez, "Defining power network zones from measures of electrical distance," in *Power & Energy Society General Meeting, 2009. PES '09. IEEE*, 2009, pp. 1-8.

[9] Wang Yansong, Zhao Jinli, Zhang Fei, and Lei Binghui, "Study on structural vulnerabilities of power grids based on the electrical distance," in *Innovative Smart Grid Technologies - Asia (ISGT Asia), 2012 IEEE*, 2012, pp. 1-5.

[10] P. Lagonotte, "Probabilistic approach of voltage control based on structural aspect of power systems," in *Probabilistic Methods Applied to Electric Power Systems, 1991., Third International Conference on*, 1991, pp. 208-213.

[11] Patrick Caron, Alain Juhel, and Franck Vandevelde, *Programmation linéaire méthodes et applications*. Paris: Dunod, 1988.

[12] G. Michel and M. Michel, *Graphes et algorithmes (4e ed.)*, 2009.

[13] Na Shi, Xumin Liu, and Yong Guan, "Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm," in *Intelligent Information Technology and Security Informatics (IITSI), 2010 Third International Symposium on*, 2010, pp. 63-67.