Clustering and supervisory voltage control in power systems☆

Mayank Baranwal[a],*, Srinivasa Salapaka[b]

[a] University of Michigan, Ann Arbor, United States
[b] University of Illinois, Urbana-Champaign, United States

ABSTRACT

In this paper, the problem of decomposing a large interconnected power network into smaller loosely-coupled zones in order to facilitate easy and flexible management of power transmission systems is addressed. This decomposition enables secondary level voltage control at regional levels and controlled islanding, which can be used to prevent the spreading of large-area blackouts. An electrical power transmission system is viewed as a fully-connected, weighted directed graph, where nodes and edge-weights of the graph represent buses and quantifications of electrical similarity between any two buses, respectively. Unlike impedance or admittance based similarity measures which are largely restrictive and do not account for topology of power networks, the electrical similarity between any two buses in this work is considered in terms of their *influence* over the remainder of the network. In particular, the electrical similarity between two buses is quantified in terms of the respective voltage fluctuations over all the buses in the network as a result of reactive power perturbations at these buses. Moreover, quantification of electrical influence does not have significant bearing on the computational complexity since it is computed using jacobians obtained as byproducts of solving power flow equations. The resulting directed graph is then clustered into prespecified number of zones that are weakly coupled *electrically* using a graph-theoretic clustering algorithm. A rule-based decentralized control strategy is proposed for effective management of bus voltages in the weakly coupled zones that are obtained as a result of the clustering process. The proposed approach is then tested on IEEE test systems for applications such as supervisory voltage control and islanding, and results in excellent identification of mutually decoupled sub-networks within a large power network.

1. Introduction

It was only about fifteen years ago that the National Academy of Engineering (NAE) regarded the North American electrical grid as the most significant engineering achievement of the 20th century [1], and yet the modern power transmission system faces major challenges [2] due to ever increasing complex interconnections among multiple elements in the grid. Some of these challenges include avoiding cascading failures [3], ensuring network robustness [4], and reduction of large network into smaller systems for better analyzability [5]. Better strategies are required in order to manage and mitigate risks related to network failures.

An electrical power transmission system can be viewed as a weighted directed graph, where nodes and edge-weights of the graph represent buses and quantifications of electrical similarity between any two buses, respectively. This abstraction enables tools from graph theory to be employed for effective management of power systems through identification of cliques (or clusters). Such decomposition of a large interconnected power network into smaller loosely-coupled groups facilitates easy and flexible management of the power transmission systems by enabling secondary voltage control at regional levels [6] and controlled islanding that aims to prevent the spreading of large-area blackouts, thereby making the network robust to power and load fluctuations [7].

In this work, an interpretable classification of power networks is provided by identifying mutually decoupled (or loosely coupled) clusters (or zones). Such classification helps to control the spread of power outage and simultaneously identify the nodes that are most affected during any unforeseen event of blackout or catastrophic failure. A novel notion of *electrical similarity* between any two buses and an efficient graph clustering algorithm are proposed such that a bus is tightly coupled to other buses within its cluster, while bearing loose coupling with nodes in other clusters. This decomposition reveals the underlying topological structure in the network and enables synthesis of a
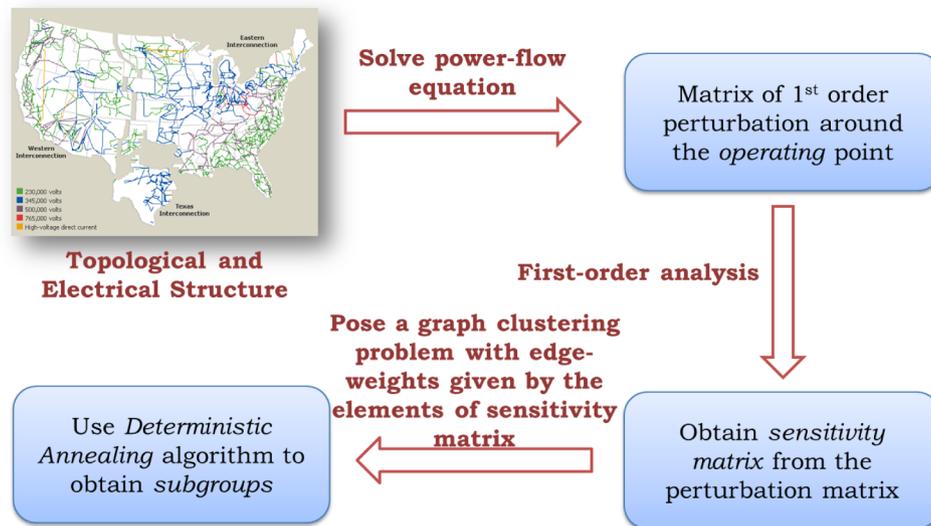
Fig. 1. Proposed approach for clustering of an electrical network.

supervisory voltage control scheme where it suffices to perform localized corrective actions (at cluster level) in case of any unforeseen events.

Several recent works have explored the graph-theoretic abstraction of power networks to come up with similarity measures and clustering algorithms for partitioning a large network. These similarity measures can be broadly classified as - (a) *Structural similarity*, which is purely based on graph-theoretic quantities such as degree distribution of nodes and degree assortativity [8], graph diameter [9] and characteristic path length [10], (b) *Topological similarity*, which incorporates underlying physics (such as Kirchhoff's laws) to derive *electrical distance* either from offline (non realtime) quantities such as nodal conductance matrix [11,12] or power flow matrix [6] and online quantities, such as derived time-series phase angle data from phase measurement units (PMUs). Graph theoretic measures of structural similarity are easy to obtain and are often useful in comparing a given power network with other graph structures. However, such measures often fail to capture electrical coupling among the buses of a network primarily due to not accounting for underlying dynamics. On the other hand, topological similarity measures derived from circuit laws and network theorems alleviate this problem. It is further preferred to employ offline measures of similarity since the online methods rely on the observed data after the disturbance has occurred. While majority of the above work focuses on hard partitioning of electrical networks, the authors in [13,14] employ a fuzzy clustering approach for PMU placements such that a 'dissimilarity' between PMU bus and non-PMU bus contingency response signals is minimized. The algorithm for network partitioning described in our work enjoys the best of both worlds. The proposed algorithm initializes with a fuzzy, unbiased estimate of network partitions, which are suitably *hardened* as the algorithm evolves.

The notion of electrical similarity proposed in this work is based on computation of first-order perturbation matrix obtained as a result of solving power flow equations [6]. In particular, two buses are considered *close*, if they have similar *influence* over the remainder of the network. Influence of a bus on another is characterized in terms of sensitivity of voltage fluctuations at one bus due to reactive power perturbations at another bus. The choice of quantification of influence can be more general to include perturbations in active power injections and phase angles. The characterization of *electrical* influence does not have any bearing on the clustering algorithm proposed in this manuscript. However since most networks exhibit active-reactive decoupling, it suffices to restrict the notion of influence in terms of fluctuations in voltage magnitudes due to reactive power perturbations. Consequently,

it is further shown that grouping of buses under the proposed notion of influence is such that the voltage fluctuations at a bus due to perturbations at buses within the same cluster are more than voltage fluctuations due to perturbations at buses from other clusters. That is, not only that perturbations at two buses in the same cluster have similar effects on the entire network, the resulting voltage fluctuations at buses from other clusters are much smaller than the voltage fluctuations at the buses from the same cluster. Thus the proposed notion of electrical similarity favors partitioning a network into loosely coupled zones.

Quantification of electrical distance necessitates development of an efficient graph clustering algorithm which is scalable, independent of initialization and has ability to avoid solutions comprising of non-robust clusters. In the proposed work, the problem of classification of power networks is presented as a combinatorial resource allocation problem. Similar combinatorial optimization problems have been studied in different areas such as minimum distortion problem in data compression [15], facility location problems [16], pattern recognition [17], neural networks [18], graph aggregation [19], motion coordination algorithms, coverage control [20] and mobile sensing network problems [21]. These problems are computationally complex, nonconvex and suffer from poor local optima that riddle the cost surface [22]. A variety of heuristics ranging from repeated optimization with different initialization, heuristics for good initialization, to heuristics for cluster splits and merges have been proposed in literature to address above difficulties. The graph partitioning algorithm proposed in this work is based on deterministic annealing (DA) algorithm [23], which avoids many poor local optima while maintaining a faster convergence rate when compared to approaches such as simulated annealing [24] or Lloyd's algorithm [25]. The DA algorithm shares connections with the computation of rate-distortion functions in information theory [26,15], where an effective rate-distortion function parameterized by an annealing variable is formulated and this function is deterministically optimized at successively increased values of the annealing parameter. Fig. 1 summarizes the proposed approach to clustering of electrical networks.

As described earlier, the primary goal for partitioning a large power network is to facilitate design of local controllers and allow for controlled islanding so as to prevent the spreading of large-area blackouts. A rule-based approach for decentralized voltage control is proposed, which exploits mutual decoupling of the partitioned network and requires only local control actions only at the local level. Similar rule-based expert system for voltage control is proposed in [27], albeit the control actions were not confined at local level. The proposed approach

combining clustering and rule-based control is tested for scenarios such as overloading and controlled islanding, and demonstrates faster voltage recovery as compared to [27]. The main contributions of this work can be summarized as:

a. Quantification of electrical distance: A novel notion of electrical distance between two buses is proposed. In this notion, two buses are considered close if they have similar influence over the remainder of the network. The similarity is quantified in terms of the respective voltage fluctuations over all the buses in the network as a result of reactive power perturbations at these buses (primarily due to active-reactive decoupling in inductive networks). In the case of systems where the hypothesis of active-reactive decoupling does not hold valid, the measure of electrical distance proposed in this manuscript can be easily modified to include more general measure also comprising of active-power perturbations and fluctuations in phase angles.

b. Graph-Theoretic clustering algorithm: The algorithm proposed for partitioning an electrical network is based on aggregating buses (nodes) of the underlying graph into nodes of a representative *super-graph*. Each node of a supergraph represents a cluster (zone) and comprises of mutually strongly coupled electrical buses, whereas two zones are practically decoupled with each other. Such an aggregation defines a partition of the underlying electrical network.

c. Rule-based voltage control: A decentralized rule-based voltage control scheme is proposed for effective management of power networks, where violations in bus voltage magnitudes can be corrected for using only 'local' measurements and control actions.

d. Validation on IEEE test systems: The proposed clustering algorithm along with the rule-based control strategy is employed on IEEE test systems. Several critical scenarios ranging from undervoltage situations to potential cascading failures due to severe underloading are considered and successfully alleviated by the proposed scheme.

2. Some preliminaries and quantification of electrical similarity

In this work, quantification of electrical similarity between two buses is based on computation of the Jacobian matrix obtained by solving power flow equations [28]. We first describe some preliminaries on fundamental electrical quantities and matrix equations through a toy-network containing four buses in Fig. 2. These buses can be of different types - Slack bus, Generator bus (or *PV* bus) and Load bus (or
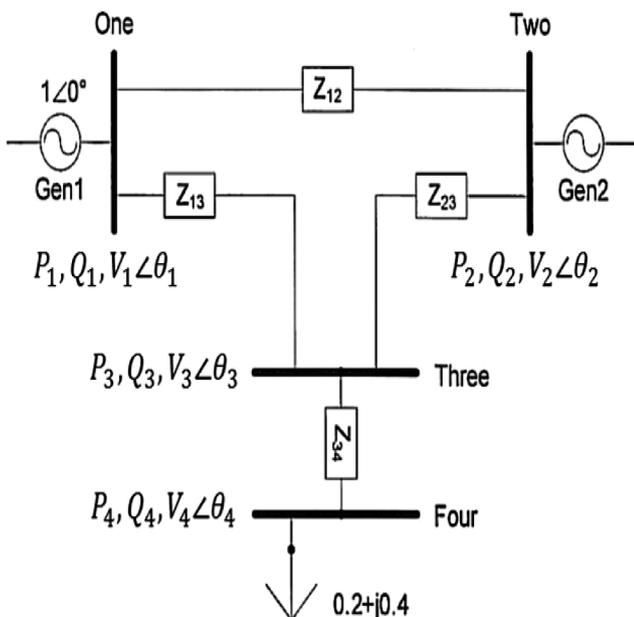

Fig. 2. An example of a four-bus network, where buses 1,2,3 and 4 denote *slack* bus, *generator* (*PV*) bus, *connecting* bus and *load* (*PQ*) bus, respectively.

PQ bus). Each bus i is completely specified by four physical quantities - voltage magnitude V_i, phase angle θ_i, real power flow P_i, and reactive power flow Q_i. Link between two buses i and j are specified in terms of line impedance Z_{ij}. The impedance matrix of the network is denoted by $\mathbf{Z}_{bus}$. The admittance matrix $\mathbf{Y}_{bus} \triangleq [Y_{ij}]$ is the inverse of $\mathbf{Z}_{bus}$. The current flow through bus i is given by I_i. $\mathbf{V}$ and $\mathbf{I}$ denote the column vectors of voltage and current magnitudes, respectively. Similarly, $\mathbf{P}$, $\mathbf{Q}$, Θ are the column vectors depicting real power flows, reactive power flows and the voltage phase angles at the buses of an electrical network.

The admittance matrix $\mathbf{Y}_{bus}$ is generally complex with real part (conductance) G_{ij} and imaginary part (susceptance) B_{ij}, i.e. $Y_{ij} = G_{ij} + jB_{ij}$. These fundamental electrical quantities are related to each other using Kirchoff's laws and power balance equations, and can be described as:

$$\mathbf{I} = \mathbf{Y_{bus}}V, \qquad \mathbf{V} = \mathbf{Z_{bus}}I,$$

$$P_j = \sum_{k=1}^{N} V_k V_j (G_{kj}\cos(\theta_k - \theta_j) + B_{kj}\sin(\theta_k - \theta_j)),$$

$$Q_j = \sum_{k=1}^{N} V_k V_j (G_{kj}\sin(\theta_k - \theta_j) - B_{kj}\cos(\theta_k - \theta_j)), \tag{1}$$

where N is the number of nodes (buses) in the network and $j \in \{1, ..., N\}$. The last two equations are called the power flow equations, and are typically solved to address the *power flow problem* [28]. The goal of a power-flow problem is to obtain complete voltages magnitude and phase information for each bus in an electrical network with specified load and generator real power and voltage conditions. *Electrical distance* between two buses is quantified by considering first-order perturbations around the operating point (power-flow solution at nominal loading conditions). Considering first-order perturbations in (1) results in following linearized equations:

$$\mathbf{\Delta I} = \mathbf{Y_{bus}}\mathbf{\Delta V}, \qquad \mathbf{\Delta Q} = [\partial\mathbf{Q}/\partial\mathbf{V}]\mathbf{\Delta V},$$

$$\mathbf{\Delta V} = \mathbf{Z_{bus}}\mathbf{\Delta I}, \qquad \mathbf{\Delta V} = [\partial\mathbf{V}/\partial\mathbf{Q}]\mathbf{\Delta Q}. \tag{2}$$

The matrix $[\partial\mathbf{Q}/\partial\mathbf{V}] \in \mathbb{R}^{N \times N}$ appears in load-flow computation using Newton-Raphson method. Its inverse matrix $[\partial\mathbf{V}/\partial\mathbf{Q}] \in \mathbb{R}^{N \times N}$ (also known as *sensitivity matrix*) reflects the propagation of voltage variations through the power network due to reactive power injection at a bus. Note that it is possible to consider a more general form of sensitivity matrix where voltage variations depend on both reactive as well as active power perturbations. However, most high power networks exhibit active-reactive decoupling where fluctuations in voltage magnitudes are tied predominantly to reactive power perturbations and therefore it suffices to consider the above notion of sensitivity matrix. For ease of illustration, this work investigates the sensitivity of voltage fluctuations caused at a bus with respect to reactive power fluctuations at another bus.

The *influence* of a bus on another bus is captured by the magnitude of voltage coupling between the two buses, and is quantified in terms of *matrix of attenuation* $\boldsymbol{\alpha} \triangleq [\alpha_{ij}] \in \mathbb{R}^{N \times N}$, where

$$\Delta V_i = \alpha_{ij}\Delta V_j, \qquad \text{where} \quad \alpha_{ij} := \left[\frac{\partial V_i}{\partial Q_j}\right] \bigg/ \left[\frac{\partial V_j}{\partial Q_j}\right], \tag{3}$$

which quantifies the voltage fluctuation at bus i per unit voltage fluctuation at j^{th} bus, when reactive perturbations are applied at bus j. Here the normalization in the definition of α_{ij} has two distinct advantages - (i) making the quantities dimensionless, (ii) assigning equal importance to all the buses (i.e. $\alpha_{ii} = 1$ for all i). If $\boldsymbol{\alpha}_i$, $\boldsymbol{\alpha}_j$ denote the ith and j^{th} columns of the matrix of attenuation, respectively, then the *electrical distance* between nodes i and j is defined as:

$$d\left(i, j\right) = \left\|\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j\right\|_2^2 = \sum_{k=1}^{N} (\alpha_{ki} - \alpha_{kj})^2. \tag{4}$$

Qualitatively, two buses i and j are *close*, when the influences of these

buses on the entire network (including the buses i and j themselves) are similar. Note that from the definition (3), the diagonal terms of the attenuation matrix satisfy $\alpha_{kk} = 1$, for all $1 \leqslant k \leqslant N$, and therefore for any $\epsilon > 0$, if

$$d(i, j) < \epsilon \quad \Rightarrow |\alpha_{ii} - \alpha_{ij}| = |1 - \alpha_{ij}| < \epsilon$$

Similarly, we have $|1 - \alpha_{ji}| < \epsilon$. Therefore $|\alpha_{ij} - \alpha_{ji}| < 2\epsilon$. Therefore, if two buses i and j are *close*, then as a consequence the influence of perturbations at buses i and j on each other are similar. Thus, if the buses of a network are partitioned in terms of how similar they are in influencing the network, then the influence of buses on each other from the same cell in a partition will be large, that is close to 1.

Thus, an electrical network can be viewed as a weighted, directed graph, where electrical buses are presented by nodes of the graph and the corresponding edge-weights denote elements α_{ij} of the sensitivity matrix. This makes it amenable to a graph aggregation method developed in [19], where a large weighted directed graph $\mathcal{G}_x$ with N nodes is approximated by a smaller weighted directed graph $\mathcal{G}_y$ with $K \ll N$ nodes, such that the smaller graph maintains least amount of representation error quantified in terms of a dissimilarity measure. Each node (supernode) of the smaller representative graph $\mathcal{G}_y$ can be viewed as a set of nodes on the larger graph $\mathcal{G}_x$; in fact, the algorithm explicitly gives the set of nodes in $\mathcal{G}_x$ that each node of $\mathcal{G}_y$ represents. Thus this graph aggregation can be used to cluster nodes in $\mathcal{G}_x$ into K clusters, for a given notion of distance between nodes. Accordingly the graph aggregation method is used to group the buses in the electrical network into clusters for the above notion of electrical distance. This graph aggregation algorithm and its important features are briefly presented in the next section. A more rigorous and exhaustive treatment can be found in [19]. An important aspect of this article is the reinterpretation this algorithm in terms of a specific information theoretic view point.

3. Graph clustering for determining loosely-coupled zones

An electrical network can be viewed as a directed, weighted graph, where nodes of the graph represent buses and edge-weights between two nodes denote electrical similarity between the associated buses. A weighted directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathbf{W})$ is described in terms of $\mathcal{V}, \mathcal{E} \in \mathcal{V} \times \mathcal{V}$ and $\mathbf{W} \in \mathbb{R}_+^{|\mathcal{V}| \times |\mathcal{V}|}$ which represent the set of nodes, edges and the edge-weight matrix, respectively. Furthermore, $|\mathcal{V}| = N \in \mathbb{N}$ and the relative node weights are denoted by $\{p_i\}$, $i \in \{1, ..., N\}$, which satisfy $p_i \geqslant 0$ with $\sum_i p_i = 1$. The *incoming vector* of the ith node is described by the weights of its incoming edges and is denoted by $\mathbf{W}_i \triangleq [W_{1i}, ..., W_{Ni}]^T$, the ith column of the matrix $\mathbf{W}$. Distance between two nodes i and j is considered based on *edge connectivity* given by $d(\mathbf{W}_i, \mathbf{W}_j)$. Note that this distance measures *similarity* between nodes; for example, small value of $d(\mathbf{W}_i, \mathbf{W}_j)$ implies that nodes i and j have similar connectivity in the graph. In the context of the problem of clustering power networks, buses represent the nodes of the graph, while the edge-weight matrix of the underlying graph is represented by the matrix of electrical similarities between buses.

In a graph aggregation problem, a small representative graph $\mathcal{G}_y$ with $|\mathcal{V}_y| = K$ of a large graph $\mathcal{G}_x$ with $|\mathcal{V}_x| = N \gg K$ is sought, where similar nodes in $\mathcal{V}_x$ are aggregated into K *supernodes* and the resulting edge-weights among these supernodes are to be determined (see Fig. 3). These supernodes in turn represent the aggregated zones in an electrical network and the edge-weights between the supernodes describe their electrical coupling. Such an aggregation defines a map (also known as partition function) $\phi\colon \mathcal{V}_x \to \mathcal{V}_y$ such that for any $1 \leqslant j \neq l \leqslant K$, following properties hold true: (i) $\phi^{-1}(j) \subset \mathcal{V}_x$ is non-empty, (ii) $\phi^{-1}(j) \cap \phi^{-1}(l) = \varnothing$, and (iii) $\cup_{j=1}^{K} \phi^{-1}(j) = \mathcal{V}_x$. Each partition function ϕ defines an *aggregation matrix* $\Phi \in \{0, 1\}^{N \times K}$ as

$$\Phi_{ij} := [\Phi]_{i,j} = \begin{cases} 1 & \text{if } \phi(i) = j, \\ 0 & \text{otherwise.} \end{cases} \qquad (5)$$

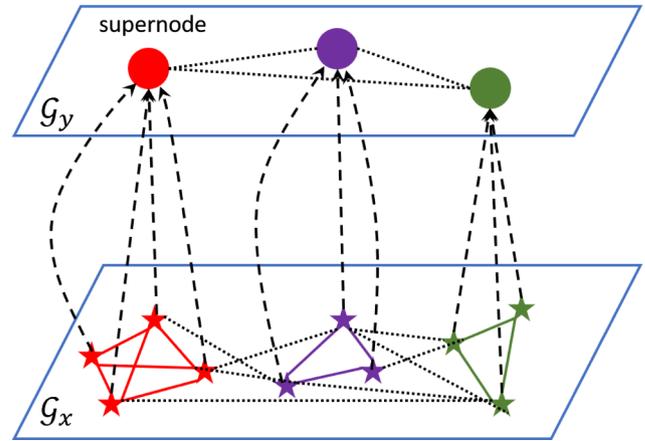Let us consider a simple example for ease of exposition before we

Fig. 3. Schematic of aggregating a large graph $\mathcal{G}_x$ into smaller representative graph $\mathcal{G}_y$ with three supernodes. These supernodes represent loosely coupled zones in an electrical network.
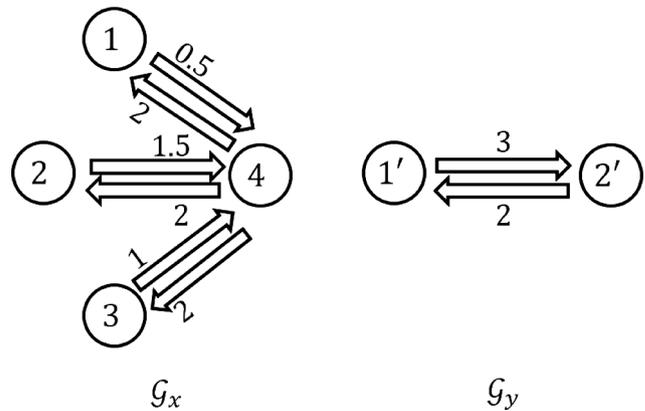
Fig. 4. Case when graph $\mathcal{G}_y$ is a contraction of $\mathcal{G}_x$.

dive deeper into concepts of graph aggregation. Consider the graph $\mathcal{G}_x$ shown in Fig. 4 with $\mathcal{V}_x = \{1, 2, 3, 4\}$ with $|\mathcal{V}_x| = N = 4$ nodes. The corresponding edge-weight matrix is given by

$$\mathbf{X} = \begin{bmatrix} 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1.5 \\ 0 & 0 & 0 & 1 \\ 2 & 2 & 2 & 0 \end{bmatrix}.$$

Suppose we want to determine a graph $\mathcal{G}_y$ with two supernodes ($|\mathcal{V}_y| = K = 2$), that is $\mathcal{V}_y = \{1', 2'\}$, which aggregates the graph $\mathcal{G}_x$. Note that in this example, $\mathbf{X}$ contains duplicated columns, which indicates $\{1, 2, 3\}$ are similar; in fact have identical edge-weight connections. Thus it is easy to verify that one of the supernodes (say $1'$) must correspond to the nodes 1, 2 and 3 in the original graph, and supernode $2'$ must be associated with node 4 in the original graph, i.e., the partition function is given by $\phi\colon \{1, 2, 3, 4\} \to \{1', 2'\}$. The representative smaller graph $\mathcal{G}_y$ with edge-weight matrix $\mathbf{Y}$ is obtained by aggregating nodes of $\mathcal{G}_x$. The associated aggregation matrix Φ and edge-weight matrix $\mathbf{Y}$ are:

$$\Phi = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{Y} = \begin{bmatrix} 0 & 3 \\ 2 & 0 \end{bmatrix}.$$

However, it must be noted that $\mathbf{X}$ contains two distinct incoming vectors that enable to aggregate $\mathcal{G}_x$ into $\mathcal{G}_y$. These distinct incoming vectors can be viewed as *resources* that need to be allocated to the nodes of original graph. This viewpoint allows us to define a weight matrix $\mathbf{Z} \in \mathbb{R}^{N \times K}$ of distinct resources given by

$$\mathbf{Z} = \begin{bmatrix} 0 & 0.5 \\ 0 & 1.5 \\ 0 & 1 \\ 2 & 0 \end{bmatrix}.$$

Here the column $\mathbf{Z}_{\phi(i)}$ approximates the ith column of $\mathbf{X}$; in fact, they are exactly the same in this example. Moreover, element Z_{kl} in the matrix $\mathbf{Z}$ can be interpreted as the directed weight from the kth supernode to the lth node of $\mathcal{G}_x$. The edge-weight matrix $\mathbf{Y}$ of the aggregated graph is related to $\mathbf{Z}$ through aggregation matrix Φ by $\mathbf{Y} = \Phi^T \mathbf{Z}$. Note that in this example, the problem of aggregating $\mathcal{G}_x$ into $\mathcal{G}_y$ is equivalent to finding the double $(\Phi, \mathbf{Z})$ such that the cost function $\sum_i d(\mathbf{X}_i, \mathbf{Z}_{\phi(i)})$ is minimized. Accordingly a general problem of aggregating a large graph $\mathcal{G}_x(\mathcal{V}_x, \mathcal{E}_x, \mathbf{X})$ with $|\mathcal{V}_x| = N$ into a graph $\mathcal{G}_y(\mathcal{V}_y, \mathcal{E}_y, \mathbf{Y})$ with $|\mathcal{V}_y| = K < N$ is given by:

$$\min_{\Phi \in \chi, \mathbf{Z} \in \mathbb{R}^{N \times K}} \sum_{i=1}^{N} p_i d\left(\mathbf{X}_i, \mathbf{Z}_{\phi(i)}\right),$$

(6)

where χ represents the set of all $\{0, 1\}^{N \times K}$ aggregation matrices; the edge-weight matrix $\mathbf{Y}$ is then given by $\mathbf{Y} = \Phi^T \mathbf{Z}$. Here $\{p_i\}$, with $\sum_i p_i = 1$ have been added in the problem formulation to represent relative weights of the nodes of $\mathcal{G}_x$, which are known a priori; in the case where all nodes are equally important, one can choose $p_i = \frac{1}{N}$ for $1 \leqslant i \leqslant N$.

The discrete optimization problem in (6) is computationally hard (NP-hard) and heuristics, such as Lloyd's algorithm [25] that aim to solve (6) suffer from the curse of poor local minima and initialization. In this context, a deterministic annealing (DA) based algorithm for graph clustering is described. The DA algorithm is independent of initialization and has ability to avoid poor local minima. The algorithm is described in Section 4.

4. Determining loosely-coupled electrical zones using graph-theoretic clustering algorithm

In a graph aggregation problem, a given graph with edge-weight matrix $\mathbf{X} = \{\mathbf{X}_i\}_{i=1}^{N}$ is required to be represented by an equivalent graph with much smaller set of nodes $\mathbf{Z} = \{\mathbf{Z}_j\}_{j=1}^{K}$ with $K \ll N$. The objective is to find mapping $\phi: \{1, ..., N\} \rightarrow \{1, ..., K\}$ ascribing to each node i in the original graph a supernode $\phi(i)$ such that the cumulative representation error is minimized. The representation error is often referred as a distortion (primarily due to its similarity to distortion function in the information theory literature) and the associated optimization problem is given by:

$$\min_{\mathbf{Z}, \phi} \underbrace{\sum_{i=1}^{N} p_i d\left(\mathbf{X}_i, \mathbf{Z}_{\phi(i)}\right)}_{D(\mathbf{X}, \mathbf{Z})},$$

(7)

where $D(\mathbf{X}, \mathbf{Z})$ is the distortion between egde-wieght matrices of the original graph and its supergraph. Note that the optimization problem in (7) is equivalent to the optimization problem in (6). Thus finding an optimal aggregation is equivalent to finding the optimal 'hard' (0–1) association. There are combinatorially many ways to associate nodes to supernodes, and complexity of any exact algorithm to find optimal associations is exponential in the worst case. These associations can be 'softened' by ascribing a probability distribution over nodes of the original graph. More precisely, the probability $p_{j|i}$ of associating a node i with supernode j is defined. The probabilistic association results in following modified distortion:

$$\overline{D}\left(\mathbf{X}, \mathbf{Z}\right) := \sum_{i=1}^{N} \sum_{j=1}^{K} p_i p_{j|i} d\left(\mathbf{X}_i, \mathbf{Z}_j\right).$$

(8)

Note that the optimization problem

$$\min_{\{\mathbf{Z}_j\}, \{p_{j|i}\}} \overline{D}\left(\mathbf{X}, \mathbf{Z}\right)$$

(9)

is equivalent to the optimization problem in (7) if probability distributions $\{p_{j|i}\}$ are restricted to be hard, i.e., $p_{j|i} \in \{0, 1\}$ for all $1 \leqslant i \leqslant N, 1 \leqslant j \leqslant K$. However, the probability distributions are not known apriori. In this work, deterministic annealing (DA) algorithm is used to estimate them. The DA algorithm seeks to maximize entropy, H, between nodes and supernodes without exceeding a given distortion $\overline{D}^*$.

Note that the entropy is maximized if all the nodes are mapped to the same supernode, since knowing the supernode does not reduce any uncertainty about the knowledge of nodes of the original graph. On the other hand, representing all the nodes by distinct supernodes maintains maximum representation error (distortion). The trade-off between maximizing the entropy and minimizing the distortion is achieved by minimizing the Lagrangian given by:

$$F\left(\mathbf{Z}, \left\{p_{j_i}\right\}\right) \triangleq \frac{1}{\beta} H + \overline{D}\left(\mathbf{X}, \mathbf{Z}\right),$$

(10)

where β is the Lagrange multiplier, referred to as the *annealing* parameter. Minimizing F with respect to the association probability $p_{j|i}$ yields a Gibbs distribution:

$$p_{j|i} = \frac{\exp\{-\beta d(\mathbf{X}_i, \mathbf{Z}_j)\}}{\sum_{j=1}^{K} \exp\left\{-\beta d\left(\mathbf{X}_i, \mathbf{Z}_j\right)\right\}}.$$

(11)

By substituting the association probabilities (11) into the expression for Lagrangian, (10) simplifies to

$$F(\mathbf{Z}) = -\frac{1}{\beta} \sum_{i=1}^{N} p_i \log\left(\sum_{j=1}^{K} \exp\left\{-\beta d\left(\mathbf{X}_i, \mathbf{Z}_j\right)\right\}\right).$$

(12)

Let us revisit the Lagrangian in (10). At low values of β, minimizing the Lagrangian is equivalent to minimizing mutual information, which indeed is minimized when the association probabilities are uniform. Lagrange multiplier β defines a homotopy between the mutual information and distortion function. As β is increased, minimization of the underlying Lagrangian results in the minimization of the modified distortion $\overline{D}(\mathbf{X}, \mathbf{Z})$. However, it should be remarked that at large values of the annealing parameter β, the association probabilities $\{p_{j|i}\}$ in (11) are approximately 0 or 1, i.e., $\{p_{j|i}\}$ are hard. Therefore as β increases, from $p_{j|i} \xrightarrow{\beta \rightarrow \infty} \Phi_{ij}$, where Φ_{ij} is an element of the (hard) aggregation matrix Φ, and from the equivalence between (7) and (9) for hard partitions, minimization of (12) with respect to $\{\mathbf{Z}_j\}$ results in minimization of the original distortion function $D(\mathbf{X}, \mathbf{Z})$. In the DA algorithm, the Lagrangian in (12) is deterministically optimized at successively increased values of β over repeated iterations (For more details on the DA algorithm see [29,23]).

Coming back to the problem of partitioning a power network into loosely coupled zones, the matrix of attenuation $\alpha = [\alpha_{ij}]$ in (3) is considered as edge-weight matrix $\mathbf{X}$ of the underlying graph $\mathcal{G}_x$, i.e., $\mathbf{X} = \alpha$. Application of the DA algorithm for graph clustering results in a set of codewords $\mathbf{Z}$, aggregation matrix Φ and a smaller representative graph $\mathcal{G}_y$ with edge-weight matrix $\mathbf{Y} = \Phi^T \mathbf{Z}$. Partitions of the network are uniquely determined by the columns of the aggregation matrix Φ through the associated partition function $\phi: \mathcal{V}_x \rightarrow \mathcal{V}_y$. The inverse map $\phi^{-1}(j) \subset \mathcal{V}_x$ for all $j \in \mathcal{V}_y$ defines a zone (or set of buses aggregated in a cluster) in the network.

Remark: Identifying a suitable number of supernodes (K) is a fundamental problem in clustering analysis. Many methods, such as gap statistics [30] and information-theoretic [31] are suggested to address this fundamental problem in clustering analysis. In fact in our recent

work [32], it is shown that the proposed clustering method has a natural way of determining natural clusters and is used here for determining suitable number of supernodes.

5. Rule-based supervisory voltage control

The proposed clustering algorithm provides a classification of a power network intended for easy and flexible management of bus voltages in the network. The "local" voltage control is achieved using a rule-based expert system, similar to the work proposed in [27]. However in our work, we rely only on local (comprising of buses belonging to the same zone in a network) measurements and control actions to achieve desired voltage control. Thus the task of a large network-wide voltage control is reduced to control of many sub-networks with very few number of buses.

Primary task of the proposed rule-based strategy is to ensure that the voltage magnitudes at buses stay within the tolerable limits. This is achieved by altering the generator voltages, adding/removing shunt capacitances, and adjusting tap changers in the priority order - Generator $\prec$ Shunt $\prec$ Tap Changer, where $a \prec b$ indicates that a is prioritized over b. The rationale behind the proposed priority order is understood as follows. Generators being the sources of power injection in the network, they must be adjusted first during instants of network failure (voltage fluctuations). Unlike transformers that interconnect two buses (possibly belonging to different zones), shunt capacitors exist between individual buses and ground. Thus, adjusting shunt capacitance results in large voltage variations primarily in the zone the shunt capacitor belongs to.

The rule-based actions rely only on "local" information in any event of failure (voltage fluctuations). Consider the overloading scenario in the IEEE-14 bus network shown in Table 4. The excessive overloading results in drop in voltage magnitudes below allowable lower limit (0.9 p.u.) at buses 9, 10 and 14. All the affected buses belong to green (G) zone, as suggested by the proposed clustering algorithm. As a consequence, corrective actions are limited only to buses belonging to green zone. Since a generator (PQ) bus has the largest priority, the voltage of the generator bus 6, which belongs to the green zone, is adjusted until voltage at bus 6 reaches maximum allowable limit (1.1 p.u.). This control action is followed by adjusting the shunt capacitor at bus 9, which again belongs to green zone. Thus clustering the electrical network makes it feasible to employ corrective actions only at local (zonal) level and facilitates easy management of the associated power transmission system. The overall rule-based strategy is implemented as a set of IF-THEN rules, as described in Fig. 5.

6. Results on IEEE test systems

The graph clustering algorithm described in Section 4 in combination with rule-based control 5 is tested on some standard network configurations - IEEE-14 bus system and IEEE-30 bus system. Figs. 6a and 6b show the network configuration of the IEEE-14 bus and IEEE-30 bus test cases, respectively. The IEEE-14 bus test case represents a portion of the American Electric Power System (in the Midwestern US) as of February, 1962. The test case includes all different kinds of buses - Slack, PV and PQ comprising of 5 generator buses, 3 tap changers and 1 shunt capacitor. The IEEE-30 bus test case represents a simple approximation of the American Electric Power system as it was in December 1961, and comprises of 8 generators, 4 tap changers and 2 shunt capacitors. The matrices of attenuation $[\alpha_{ij}]$ are first obtained by load-flow computations using Newton-Raphson method for the two test cases. The obtained matrices are then clustered into 3 partitions for the IEEE-14 bus test case and 2 partitions for the IEEE-30 bus test case. Tables 1 and 2 denote the clustering results for the two test systems. These partitions are marked by different colors in the 'Bus Type' columns and also indicated by corresponding initials. The results for various overloading and islanding scenarios are summarized below.

Moreover, quantities pertaining to clustering of IEEE-14 test system - matrix of attenuation $[\alpha_{ij}]$, matrix of associations Φ and locations of supernodes $\mathbf{Z}$ are listed in Appendix A for clarity of presentation.

6.1. Effect of perturbations on inter and intra-cluster elements

The power-flow solutions in per unit (p.u.) at nominal loading conditions for the two test systems are indicated in column 3 of Tables 1 and 2, whereas columns 4, 5 and 6 indicate the effects of perturbing generator voltages at different buses. It is observed that the influence of these perturbations is larger at the buses belonging to the same group (cluster) where the perturbations originate. For instance, in the IEEE-14 bus system, doubling the generator voltage at bus 2 results in change in voltage magnitudes at buses 4 and 5 by about 0.4 p.u. The effect of this perturbation is less severe at other buses, which do not belong to the group formed by the buses 2, 3, 4 and 5. Note that bus 3 is a generator bus (PV bus) where voltage is set a priori, and hence there is no change in its voltage magnitude. Similar effects are seen in columns 5 and 6 when perturbing the generator voltages at buses 6 and 3, respectively. Interestingly, buses {6, 9, 10, 11, 12, 13, 14} in the IEEE-14 bus system are labeled as low-voltage (LV) buses, whereas buses 7 and 8 are marked as tertiary-voltage (TV) buses. Remaining buses are indicated as high-voltage (HV) buses. This underlying electrical structure is naturally captured by the proposed clustering algorithm.

The proposed approach generalizes to larger bus systems too. Similar to the 14-bus test system, perturbing generator buses in the IEEE-30 bus test system result in large perturbations in buses belonging to the same cluster where the perturbations originate. While the algorithm was also tested for the IEEE-300 bus system, the details of it are excluded in this manuscript for the sake of brevity.

6.2. Effect of perturbations at buses within the same cluster over the remainder of the network

By construction, we have that two buses are considered close (electrically) when they have similar influence over the entire network. This is very well captured in the resulting partitions for the two test systems. Generator voltages at buses 2 and 3 are perturbed separately in the IEEE-14 bus system. Both these buses belong to the same cluster and result in similar perturbations over the entire network. For instance, both these buses have a very small influence on bus 12, affecting the voltage magnitudes by 0.013 p.u. and 0.008 p.u., respectively. However, the effect is large on buses such as bus 7, where the changes in voltage magnitudes are 0.176 p.u. and 0.100 p.u., respectively.

Similar conclusions can be drawn for the IEEE-30 bus test system. Doubling generator voltages at buses 2 and 11, both of which belong to the same cluster, have similar effect over the entire network. Their effects are large on buses such as bus 28, where the changes in voltage magnitudes are 0.148 p.u. and 0.052 p.u., respectively.

Comparison with existing methods: In [33,34], a spectral clustering algorithm is proposed for clustering of power networks. Their method employs absolute value of inverse of the admittance matrix as a measure of pairwise electrical distances. The underlying graph is then partitioned using spectral clustering and results in following partition of the IEEE-14 bus: {1, 2, 3, 10, 11, 13, 14}, {8} and {4, 5, 6, 7, 9, 12}. Clearly from Table 1, the admittance matrix based approach (coupled with spectral clustering) does not identify a suitable grouping of electrical buses. For instance, the existing approaches groups buses 4 and 6 into same zone, however they appear mutually decoupled during events of perturbations.

6.3. Rule-based supervisory voltage control

Each test system consists of generators, tap changers and shunt capacitors. The ranges and steps in which these control actions are varied is shown in Table 3.
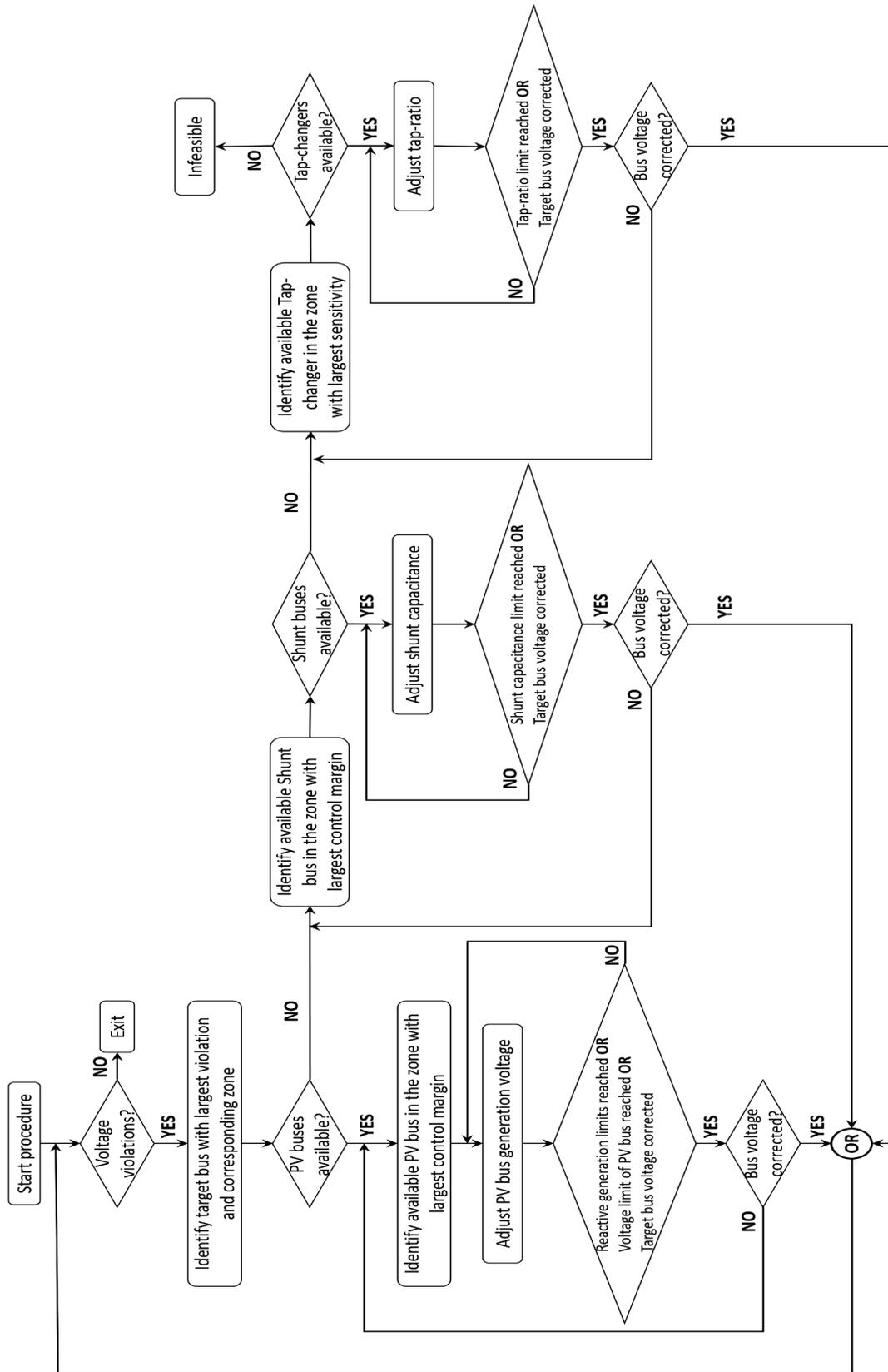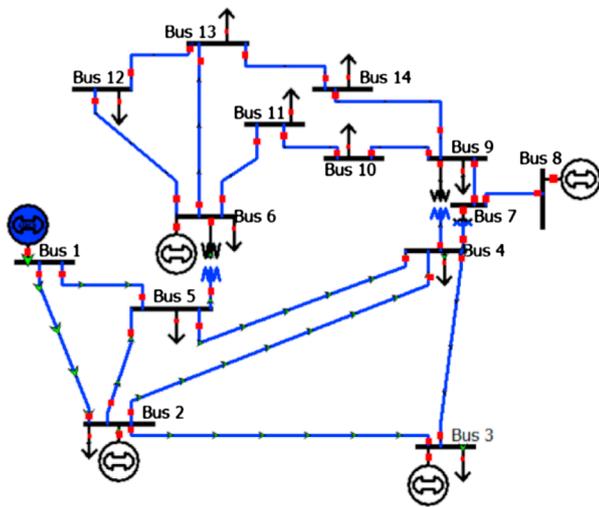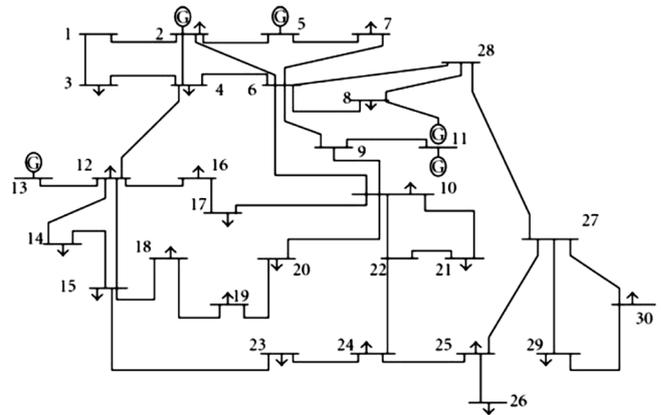
Fig. 5. Proposed decentralized rule-based scheme for voltage regulation in power systems. The scheme can be interpreted as a set of IF-ELSE rules.

(a) IEEE-14 bus test case

(b) IEEE-30 bus test case

Fig. 6. IEEE test systems approximating American Electric Power system.

Table 1
Clustering Results For IEEE-14 Bus Data

Bus #	Bus Type	Volt mag. Operating pt	Volt mag. $(2\times V_2)$	Volt mag. $(1.5\times V_6)$	Volt mag. $(2\times V_3)$
1	Slack (B)	1.060	1.060	1.060	1.060
2	PV (Y)	1.045	2.090	1.045	1.045
3	PV (Y)	1.010	1.010	1.010	2.020
4	PQ (Y)	1.018	1.408	1.107	1.241
5	PQ (Y)	1.020	1.411	1.130	1.152
6	PV (G)	1.070	1.070	1.605	1.070
7	PQ (B)	1.062	1.238	1.188	1.162
8	PV (B)	1.090	1.090	1.090	1.090
9	PQ (G)	1.056	1.226	1.280	1.151
10	PQ (G)	1.051	1.192	1.331	1.130
11	PQ (G)	1.057	1.129	1.462	1.097
12	PQ (G)	1.055	1.068	1.572	1.063
13	PQ (G)	1.050	1.075	1.546	1.064
14	PQ (G)	1.036	1.144	1.382	1.096

CASE 1: IEEE-14 bus system with 3.5 times the active load and 4.2 times the reactive load

Desired Voltage:	0.90 p.u. to 1.10 p.u.

Fault:	Voltages at buses 9, 10 and 14 are 0.886 p.u., 0.884 p.u. and 0.833 p.u., respectively and are below the lower limit of 0.90 p.u. (shown in column 3 of Table 4)
Action:	Control actions and their steps are shown in Table 5.
Results:	The bus voltages after correction are within the desired range and the corresponding magnitudes are shown in column 4 of Table 4.

As seen in Table 4, a sudden increase in active and reactive loads results in large violations in the bus voltages. In particular, voltage magnitudes at buses 9, 10 and 14 fall below the allowable limit of 0.9 p.u.. These violations are subsequently corrected though a set of control actions indicated in Table 5. Note that the *faulty* buses belong to green (G) zone. The operating voltage of generator at bus 6, which lies in the fault zone, is increased to compensate for low voltages at buses 9, 10 and 14. Once the generator bus voltage reaches the maximum allowable limit of 1.1 p.u., the shunt capacitor at bus 9 is increased in steps of 0.1

Table 2
Clustering results for IEEE-30 bus data.

Bus #	Bus Type	Volt mag. Operating pt	Volt mag. $(2\times V_2)$	Volt mag. $(2\times V_{11})$	Volt mag. $(2\times V_{13})$
1	Slack (G)	1.060	1.060	1.060	1.060
2	PV (Y)	1.045	2.090	1.045	1.045
3	PQ (Y)	1.021	1.234	1.057	1.094
4	PQ (Y)	1.012	1.286	1.057	1.103
5	PV (Y)	1.010	1.010	1.010	1.010
6	PQ (Y)	1.011	1.203	1.066	1.063
7	PQ (Y)	1.003	1.117	1.036	1.034
8	PV (Y)	1.010	1.010	1.010	1.010
9	PQ (G)	1.051	1.165	1.488	1.187
10	PQ (G)	1.045	1.174	1.343	1.296
11	PV (Y)	1.082	1.082	2.164	1.082
12	PQ (G)	1.057	1.070	1.150	1.654
13	PV (G)	1.071	1.071	1.071	2.142
14	PQ (G)	1.043	1.159	1.160	1.597
15	PQ (G)	1.038	1.157	1.177	1.550
16	PQ (G)	1.045	1.165	1.224	1.497
17	PQ (G)	1.040	1.167	1.303	1.354
18	PQ (G)	1.028	1.153	1.225	1.452
19	PQ (G)	1.026	1.152	1.256	1.396
20	PQ (G)	1.030	1.157	1.277	1.371
21	PQ (G)	1.033	1.164	1.318	1.295
22	PQ (G)	1.034	1.164	1.314	1.299
23	PQ (G)	1.027	1.154	1.202	1.451
24	PQ (G)	1.022	1.156	1.242	1.325
25	PQ (G)	1.018	1.161	1.177	1.230
26	PQ (G)	1.000	1.146	1.162	1.215
27	PQ (G)	1.024	1.172	1.144	1.177
28	PQ (Y)	1.007	1.155	1.059	1.061
29	PQ (G)	1.004	1.155	1.126	1.160
30	PQ (G)	0.992	1.145	1.116	1.150

Table 3
Ranges and steps of control measures.

Serial #	Control measure	Range of control (p.u.)	Control step (p.u.)
1	Generator voltage	0.9–1.1	0.01
2	Tap changer	0.9–1.1	0.02
3	Shunt capacitor	0.0–0.5	0.10

Table 4
Supervisory control for IEEE-14 bus network.

Bus #	Bus Type	Volt magnitude (Before correction)	Volt magnitude (After correction)
1	Slack (B)	1.060	1.060
2	PV (Y)	1.045	1.045
3	PV (Y)	1.010	1.010
4	PQ (Y)	0.945	0.956
5	PQ (Y)	0.954	0.965
6	PV (G)	1.070	1.100
7	PQ (B)	0.956	0.995
8	PV (B)	1.090	1.090
9	PQ (G)	0.886	0.960
10	PQ (G)	0.884	0.952
11	PQ (G)	0.960	1.011
12	PQ (G)	0.998	1.034
13	PQ (G)	0.966	1.007
14	PQ (G)	0.833	0.900

Table 5
Control actions for 14-bus system.

Serial #	Control action	Steps
1	Generator at 6	3
2	Shunt capacitor at 9	4
3	Transformer between 4 and 9	4

Table 6
Supervisory control for IEEE-30 bus network.

Bus #	Bus Type	Volt magnitude (Before correction)	Volt magnitude (After correction)
1	Slack (G)	1.060	1.060
2	PV (Y)	1.045	1.045
3	PQ (Y)	0.966	0.978
4	PQ (Y)	0.956	0.970
5	PV (Y)	1.010	1.010
6	PQ (Y)	0.969	0.978
7	PQ (Y)	0.961	0.967
8	PV (Y)	1.010	1.010
9	PQ (G)	0.973	1.057
10	PQ (G)	0.914	1.034
11	PV (Y)	1.082	1.082
12	PQ (G)	0.972	1.034
13	PV (G)	1.071	1.100
14	PQ (G)	0.925	1.003
15	PQ (G)	0.908	0.999
16	PQ (G)	0.924	1.013
17	PQ (G)	0.901	1.013
18	PQ (G)	0.873	0.979
19	PQ (G)	0.862	0.975
20	PQ (G)	0.872	0.988
21	PQ (G)	0.874	1.009
22	PQ (G)	0.875	1.013
23	PQ (G)	0.866	0.997
24	PQ (G)	0.837	1.019
25	PQ (G)	0.844	0.992
26	PQ (G)	0.785	0.943
27	PQ (G)	0.877	0.999
28	PQ (Y)	0.958	0.974
29	PQ (G)	0.815	0.947
30	PQ (G)	0.780	0.918

p.u. until the capacitance cannot be increased further (max 0.5 p.u.). Finally, the tap-changer between buses 4 and 9 is adjusted until all the bus voltages are within the allowable limits.

Thus, the proposed rule-based scheme relies only on "local" inputs and control actions for voltage correction and achieves the desired

Table 7
Control actions for 30-bus system.

Serial #	Control action	Steps
1	Generator at 13	3
2	Shunt capacitor at 10	4
3	Shunt capacitor at 24	5
4	Transformer between 6 and 9	4
5	Transformer between 28 and 27	2

performance in very few number (11 - see Table 5) of steps as compared to similar rule-based schemes proposed in [27,35] which requires 25 steps for voltage correction.

CASE 2: IEEE-30 bus system with 2.1 times the active load and 3.1 times the reactive load

Desired Voltage:	0.90 p.u. to 1.10 p.u.
Fault:	Voltages at buses 18–27 and 29–30 are below the lower limit of 0.90 p.u. (shown in column 3 of Table 6)
Action:	Control actions and their steps are shown in Table 7.
Results:	The bus voltages after correction are within the desired range and the corresponding magnitudes are shown in column 4 of Table 6.

Similar to the 14-bus system, increase in active and reactive loads result in large violations in voltage magnitudes at buses 18–27, 29 and 30, all of which belong to green (G) zone. As before, corrective action can be localized to green (G) zone for voltage control in the IEEE-30 bus system using fewer number (18) of control actions. These control actions are indicated in Table 7. On the other hand, prior methods in [27,35] require 21 steps for less severe perturbations.

6.4. Controlled islanding

Islanding is required whenever there is a fault and whenever the maintenance is required in a power network. A controlled islanding can not only prevent damage to customer equipment due to continued excessive generation, but also avoid widespread blackouts. Our classification approach is very well suited for operations such as controlled islanding. A simulated scenario is considered, where there is a fault at one of the load buses in green (G) zone in the IEEE-14 bus system and it is desired to avoid any cascading failure by appropriately isolating a major part of the network. The identification obtained using the proposed clustering algorithm provides a natural way to prevent such cascading failure. Since, the clustering algorithm partitions a network into mutually decoupled (loosely coupled) zones, it is natural to isolate the zones not containing the faulty bus. As a consequence, zones (B) and (Y) in the IEEE-14 bus are isolated system from zone (G). The effects of isolating the buses belonging to zones (B) and (Y) are shown in Table 8. Even though the 14-bus system is practically reduced to a 7-bus system, the effect of such an isolation is minimal in terms of changes in p.u. bus voltages. The largest increase in p.u. voltage is only 0.021 p.u. at bus 5.

Table 8
Controlled islanding for IEEE-14 bus network.

Bus #	Bus Type	Volt mag. before islanding	Volt mag. after islanding
1	Slack (B)	1.060	1.060
2	PV (Y)	1.045	1.045
3	PV (Y)	1.010	1.010
4	PQ (Y)	1.018	1.037
5	PQ (Y)	1.020	1.041
7	PQ (B)	1.062	1.076
8	PV (B)	1.090	1.090

7. Conclusion

In this article, the problem of partitioning an electrical network into loosely coupled groups is considered. Such a classification enables easy and flexible management of power systems. A novel notion of electrical similarity between two buses is proposed and quantified and the classification problem is transformed into an equivalent combinatorial graph clustering problem. A graph-theoretic DA algorithm is employed to efficiently partition the resulting graph. A rule-based expert control system is then suggested for synthesizing local control actions during events such as overloading, underloading and islanding. The combined approach (clustering + supervisory control) is then tested for various scenarios for the IEEE-14 and IEEE-30 bus test systems and the results corroborate the effectiveness of this approach. The algorithm is computationally scalable [36] and can easily handle more complex systems, such as IEEE-300 bus systems.

Appendix A. Relevant quantities for clustering of IEEE-14 bus test system

For the IEEE-14 bus test system, the matrix of attenuation $\boldsymbol{\alpha}$, the matrix of association $\boldsymbol{\Phi}$ and the location of supernodes $\mathbf{Z}$ are obtained as:

$$\boldsymbol{\alpha} = \begin{bmatrix} 1 & 1.03 & 1.02 & 1.03 & 1.03 & .90 & .94 & .90 & .92 & .89 & .87 & .84 & .87 & .85 \\ .94 & 1 & .99 & .99 & .99 & .87 & .91 & .87 & .89 & .86 & .84 & .81 & .84 & .82 \\ .90 & .95 & 1 & .96 & .95 & .84 & .88 & .84 & .86 & .83 & .81 & .78 & .81 & .80 \\ .92 & .97 & .98 & 1 & .98 & .87 & .91 & .87 & .89 & .86 & .84 & .80 & .84 & .82 \\ .93 & .97 & .98 & .99 & 1 & .88 & .91 & .87 & .89 & .86 & .85 & .81 & .85 & .83 \\ .98 & 1.03 & 1.03 & 1.06 & 1.05 & 1 & .99 & .94 & .97 & .95 & .95 & .92 & .96 & .92 \\ .94 & .99 & .99 & 1.02 & 1.01 & .91 & 1 & .96 & .95 & .92 & .89 & .84 & .88 & .88 \\ .91 & .96 & .96 & .99 & .98 & .88 & .96 & 1 & .92 & .89 & .86 & .81 & .85 & .84 \\ .97 & 1.02 & 1.02 & 1.05 & 1.04 & .94 & 1.01 & .96 & 1 & .97 & .93 & .88 & .92 & .91 \\ .97 & 1.02 & 1.03 & 1.05 & 1.04 & .96 & 1.01 & .96 & .99 & 1 & .95 & .88 & .92 & .92 \\ .97 & 1.02 & 1.03 & 1.06 & 1.05 & .98 & .99 & .95 & .98 & .98 & 1 & .90 & .94 & .92 \\ .98 & 1.02 & 1.03 & 1.05 & 1.05 & .99 & .99 & .94 & .97 & .95 & .95 & 1 & .97 & .92 \\ .98 & 1.03 & 1.03 & 1.06 & 1.05 & .99 & .99 & .95 & .98 & .95 & .95 & .94 & 1 & .94 \\ .97 & 1.02 & 1.03 & 1.05 & 1.04 & .96 & 1 & .96 & .99 & .96 & .94 & .90 & .95 & 1 \end{bmatrix}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{Z} = \begin{bmatrix} 0.9218 & 0.9503 & 0.9728 \\ 0.9718 & 0.9919 & 1.0226 \\ 0.9853 & 0.9950 & 1.0295 \\ 0.9896 & 1.0148 & 1.0536 \\ 0.9815 & 1.0042 & 1.0461 \\ 0.8649 & 0.8974 & 0.9756 \\ 0.9060 & 0.9700 & 0.9964 \\ 0.8655 & 0.9527 & 0.9518 \\ 0.8798 & 0.9306 & 0.9834 \\ 0.8555 & 0.9019 & 0.9651 \\ 0.8368 & 0.8755 & 0.9510 \\ 0.8002 & 0.8313 & 0.9176 \\ 0.8340 & 0.8679 & 0.9521 \\ 0.8168 & 0.8579 & 0.9322 \end{bmatrix}$$

Appendix B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ijepes.2019.02.025.

References

[1] Wulf WA. Great achievements and grand challenges. Bridge 2000;30(3):4.
[2] Albert R, Albert I, Nakarado GL. Structural vulnerability of the north american power grid. Phys Rev E 2004;69(2):025103.
[3] Hines P, Balasubramaniam K, Sanchez EC. Cascading failures in power grids. IEEE Potentials 2009;28(5):0.
[4] Wang J-W, Rong L-L. Robustness of the western united states power grid under edge attack strategies due to cascading failures. Saf Sci 2011;49(6):807–12.
[5] Holmgren ÅJ. Using graph models to analyze the vulnerability of electric power networks. Risk Anal 2006;26(4):955–69.
[6] Lagonotte P, Sabonnadiere J, Leost J-Y, Paul J-P. Structural analysis of the electrical system: application to secondary voltage control in France. IEEE Trans Power Syst 1989;4(2):479–86.
[7] Van Cutsem T, Vournas C. Voltage stability of electric power systems vol. 41. Springer Science & Business Media; 1998.
[8] Barabási A-L, Albert R. Emergence of scaling in random networks. Science 1999;286(5439):509–12.
[9] Barabási A-L. Statistical mechanics of complex networks. Rev Mod Phys 2002;74(1):47.
[10] Watts DJ, Strogatz SH. Collective dynamics of small-worldnetworks. Nature 1998;393(6684):440–2.
[11] Hogan E, Cotilla-Sanchez E, Halappanavar M, Wang S, Mackey P, Hines P, et al. Towards effective clustering techniques for the analysis of electric power grids. Proceedings of the 3rd international workshop on high performance computing, networking and analytics for the power grid. ACM; 2013. p. 1.
[12] Cotilla-Sanchez E, Hines PD, Barrows C, Blumsack S, Patel M. Multi-attribute partitioning of power networks based on electrical distance. IEEE Trans Power Syst 2013;28(4):4979–87.
[13] Kamwa I, Pradhan AK, Joós G. Automatic segmentation of large power systems into fuzzy coherent areas for dynamic vulnerability assessment. IEEE Trans Power Syst 2007;22(4):1974–85.
[14] Kamwa I, Pradhan AK, Joos G, Samantaray S. Fuzzy partitioning of a real power system for dynamic vulnerability assessment. IEEE Trans Power Syst 2009;24(3):1356–65.
[15] Cover TM, Thomas JA. Elements of information theory. John Wiley & Sons; 2012.
[16] Drezner E. Facility location: a survey of applications and methods. J Oper Res Soc 1996;47(11):1421.
[17] Therrien CW. Decision estimation and classification: an introduction to pattern recognition and related topics. John Wiley & Sons, Inc.; 1989.
[18] Haykin S, Network N. A comprehensive foundation. Neural Networks 2004;2.
[19] Xu Y, Salapaka SM, Beck CL. Aggregation of graph models and markov chains by deterministic annealing. IEEE Trans Autom Control 2014;59(10):2807–12.
[20] Xu Y, Salapaka SM, Beck CL. Clustering and coverage control for systems with acceleration-driven dynamics. IEEE Trans Autom Control 2014;59(5):1342–7.
[21] Cortes J, Martinez S, Karatas T, Bullo F. Coverage control for mobile sensing networks. IEEE international conference on robotics and automation, 2002. Proceedings. ICRA'02. vol. 2. IEEE; 2002. p. 1327–32.
[22] Gray R, Karnin E. Multiple local optima in vector quantizers (corresp.). IEEE Trans Inf Theory 1982;28(2):256–61.
[23] Rose K. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. Proc IEEE 1998;86(11):2210–39.
[24] Kirkpatrick S, Gelatt CD, Vecchi MP, et al. Optimization by simulated annealing. Science 1983;220(4598):671–80.
[25] Lloyd SP. Least squares quantization in PCM. IEEE Trans Inf Theory

1982;28(2):129–37.

[26] Rose K. A mapping approach to rate-distortion computation and analysis. IEEE Trans Inf Theory 1994;40(6):1939–52.

[27] Wagner WR, Keyhani A, Hao S, Wong TC. A rule-based approach to decentralized voltage control. Power industry computer application conference, 1989. PICA'89, conference Papers. IEEE; 1989. p. 163–71.

[28] Monticelli A. Power flow equations. State estimation in electric power systems. Springer; 1999. p. 63–102.

[29] Rose K. Deterministic annealing, clustering, and optimization [Ph.D. thesis]. California Institute of Technology; 1990.

[30] Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. J Roy Stat Soc: Ser B (Stat Methodol.) 2001;63(2):411–23.

[31] Sugar CA, James GM. Finding the number of clusters in a dataset: an information-theoretic approach. J Am Stat Assoc 2003;98(463):750–63.

[32] Srivastava A, Baranwal M, Salapaka S. On the persistence of clustering solutions and true number of clusters in a dataset, arXiv preprint arXiv: <1811.00102v2>.

[33] Hines P, Blumsack S, Sanchez EC, Barrows C. The topological and electrical structure of power grids. 2010 43rd Hawaii International Conference on System Sciences (HICSS). IEEE; 2010. p. 1–10.

[34] Blumsack S, Hines P, Patel M, Barrows C, Sanchez EC. Defining power network zones from measures of electrical distance. Power & Energy Society general meeting, 2009. PES'09. IEEE. IEEE; 2009. p. 1–8.

[35] Singh S, Raju G, Gupta A. Sensitivity based expert system for voltage control in power system. Int J Electric Pow Energy Syst 1993;15(3):131–6.

[36] Parekh PM, Katselis D, Beck CL, Salapaka SM. Deterministic annealing for clustering: Tutorial and computational aspects. American Control Conference (ACC), 2015. IEEE; 2015. p. 2906–11.