Advanced Engineering Informatics 38 (2018) 232–241

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Full length article

Community detection in national-scale high voltage transmission networks using genetic algorithms

Manuel Guerrero[a], Francisco G. Montoya[b], Raúl Baños[b,*], Alfredo Alcayde[b], Consolacíon Gil[a]

[a] CeiA3, Department of Informatics, University of Almería, Carretera de Sacramento s/n, 04120 Almería Spain
[b] CeiA3, Department of Engineering, University of Almería, Carretera de Sacramento s/n, 04120 Almería Spain

ARTICLE INFO

Keywords:
Electric power system
Power grid
High voltage transmission networks
Contingency analysis
Community detection
Genetic algorithms

ABSTRACT

The large-scale interconnection of electricity networks has been one of the most important investments made by electric companies, and this trend is expected to continue in the future. One of the research topics in this field is the application of graph-based analysis to identify the characteristics of power grids. In particular, the application of community detection techniques allows for the identification of network elements that share valuable properties by partitioning a network into some loosely coupled sub-networks (communities) of similar scale, such that nodes within a community are densely linked, while connections between different communities are sparser. This paper proposes the use of competitive genetic algorithms to rapidly detect any number of community structures in complex grid networks. Results obtained in several national- scale high voltage transmission networks, including Italy, Germany, France, the Iberian peninsula (Spain and Portugal), Texas (US), and the IEEE 118 bus test case that represents a portion of the American Electric Power System (in the Midwestern US), show the good performance of genetic algorithms to detect communities in power grids. In addition to the topological analysis of power grids, the implications of these results from an engineering point of view are discussed, as well as how they could be used to analyze the vulnerability risk of power grids to avoid large-scale cascade failures.

1. Introduction

The growing demand for electricity has involved that high voltage transmission networks have become one of the most important infrastructures in our society societies [1]. Furthermore, the complexity of power grids has increased with economic development, necessitating the application of robust control and optimisation strategies to manage large-scale systems [22].

Different studies have shown that graph-based network analysis is a powerful tool for describing many real systems in a variety of fields [4], including engineering tasks [11,38,43]. Community structure is an important feature of graphs representing real systems, since many real networks have clusters, with many edges connecting nodes within the same cluster, and comparatively few edges connecting to nodes in different clusters. Finding the optimal partition of the vertices of a graph into clusters such that the corresponding modularity [27] is maximized is an NP-hard problem [10,25]. As community detection is a difficult problem, complex computational and mathematical techniques are needed.

Some authors have applied community detection techniques to manage small and medium-sized power grids [5,29,30,37], but a little attention has been paid to solve this problem in national-scale electrical networks. To cover this gap in the literature, this paper analyses the performance of evolutionary approaches for solving the community detection problem with applications to several national-scale high voltage power grids. These algorithms, which are guided by the modularity index [27] and consider different *degrees of abstraction* (i.e. the methods are able to detect any number -*k*- of communities), allow for a flexible and adaptive analysis of the grid by considering different levels of detail.

The remainder of the paper is organised as follows: Section 2 introduces the problem of community detection in graphs, and revises some previous studies that have been applied to electrical grids. Section 3 presents the main characteristics of two genetic algorithms used to solve the community detection problem in graphs. Section 4 presents an empirical study that compares the performance of these methods in detecting communities in five national-scale grids, and also in a well-known benchmark. Section 5 discusses the results and implications of this work into the engineering field, emphasizing the possibility of using the proposed methods for line outage contingency analysis,

* Corresponding author.
E-mail addresses: mguerrero@fundacionual.es (M. Guerrero), pagilm@ual.es (F.G. Montoya), rbanos@ual.es (R. Baños), aalcayde@ual.es (A. Alcayde), cgilm@ual.es (C. Gil).

https://doi.org/10.1016/j.aei.2018.07.001
Received 8 December 2017; Received in revised form 11 May 2018; Accepted 6 July 2018

including the system islanding scenario. Finally, the conclusions of this work are presented in Section 6.

2. Related work

Electricity networks have been built since the end of the XIX century [24]. Electrification still continues today, leading to a high degree of interconnection spanning states and now reaching a continental scale. In the past, the distribution system was unidirectional, distributing electricity from a small number of large power plants down to end users, whose demand was generally regarded as rigid and exogenous. However, grid operation has changed significantly in recent decades for several reasons, including the integration of variable output renewable energy sources [8]. The popularity of renewable energy has led electricity generated in power plants to be complemented by renewable power sources, some of which are located in industrial installations and residential buildings, with the result that the distribution system has become bidirectional. Moreover, the use of renewable sources has presented an alternative to grid extension for remote village electrification [21]. Unfortunately, the uncertainty and variability of wind and solar generation affects the grid operations, although some recent studies have shown that these inconveniences can be mitigated by balancing the variability of renewable sources using the transmission grid and balancing with storage [35].

The growing worldwide demand for electricity, together with the inclusion of new power plants, requires increasing grid connectivity and applying complex control methods. Some recent investigations have proposed the analysis of power grid infrastructures using graph-based complex network analysis techniques [29], such that the nodes of the network often represent the power plants, distribution substations and trans- mission substations, while the edges correspond to transmission lines. Advances in computer science have allowed for the efficient representation, management and processing of large amounts of data, including graph-based networks representing real systems. The application of graph-based analysis techniques has allowed for the analysis of the topological structure of networks representing power grids [2]. For example, some studies have analyzed the vulnerability of power grids to blackouts using graph topological indexes [18].

All complex systems share a common characteristic: community structures [26]. Communities consist of groups of nodes inside a network that are more densely connected with each other than with the remaining nodes of the network. As the nodes belonging to the same community have a higher likelihood of interaction, detecting those communities can reveal characteristics or functional relationships in a given network. Therefore, the community detection problem consists of partitioning the nodes in a network into groups such that there are many edges connecting nodes within the same group, and comparatively few edges connecting nodes in different groups. In the case of power grids, communities represent high-voltage lines that are densely connected.

Some studies have applied community detection to power grids. In [5], it was introduced a node similarity index to assign each node to the community sharing maximum similarity, which exhibited a good performance in a set of experiments performed on several IEEE standard power grids. Other authors have analyzed the optimal phasor measurement unit placement problem using algorithms for community detection to identify coherent groups based on an equivalent graph of generator nodes [12]. Another study presented a hierarchical spectral clustering method that reveals the internal connectivity structure of the power transmission capability of islanding systems using a network with nodes and links representing buses and electrical transmission lines, respectively [37]. Some investigations have analyzed the possibility of using community detection for islanding power systems as an emergency response to isolate failures that might propagate and lead to major disturbances [30]. Community detection has also been applied to analyze the vulnerability of the power systems under terrorist attacks [40], among other applications.

3. Evolutionary algorithms

Evolutionary computation [9] is a research field closely related to computational intelligence that is focused on designing algorithms to solve complex global optimisation problems. Evolutionary algorithms are problem-solving procedures that include evolutionary processes as the key design elements. In particular, an evolutionary algorithm consists of a population of individuals that continually and selectively evolve until a termination criterion is fulfilled.

Among evolutionary techniques, Genetic Algorithms (GAs) [14] are likely the most widely used. A genetic algorithm mimics natural selection by evolving a population of individual solutions to the problem at hand over time until a termination condition is fulfilled and the best individual is taken as an acceptable solution. Two of the most important characteristics of GAs are the representation used (e.g., binary or real) and the genetic operators employed (e.g., mutation and crossover).

As the community detection problem is highly complex, researchers have applied heuristics and meta-heuristics to obtain high quality solutions with a reduced runtime. In particular, GAs are selected because they have been used to solve many electrical problems [6,33,34,39,42]. In this study, two genetic algorithms have been adapted to solve community detection problems in power grids. These recently proposed algorithms (MIGA and GGA+) have been shown to be more effective than other approaches to community detection based on benchmarks typically used to compare algorithms for this problem.

- The Modularity and Improved Genetic Algorithm (MIGA) [36] takes the modularity (*Q*) as the objective function, and uses the number of community structures as prior information to improve the stability and accuracy of community detection. MIGA also uses Simulated Annealing [19] as a local search strategy.
- The Generational Genetic Algorithm (GGA+) [13] includes efficient and safe initialisation methods in which a maximum node size is assigned to each community. Several operators are applied to migrate or exchange nodes between communities while using the modularity function as the objective function. An important feature of GGA+ is that it is able to rapidly obtain community partitions with different degrees of abstraction.

4. Empirical study

This section analyses the performance of the MIGA and GGA+ algorithms in detecting communities in several national-scale high voltage transmission networks with different characteristics. Neglecting complex electrical properties, the nodes of the network represent the power plants, distribution and transmission substations, while the edges correspond to transmission lines. In this way, the power grid is simplified as an unidirectional and unweighted network.

4.1. Modularity

Most optimisation methods apply *modularity* to detect communities in networks. *Modularity* [27] may be the most extensively applied objective function in community detection problems due to its simplicity and ease of calculation. Modularity provides a numerical value that represents the quality of the solution, such that the higher the value, the more accurate the community structure. Therefore, the aim of the algorithms is to maximize the value of *modularity* (Q), defined as:

$$Q = \frac{1}{2M} \sum \left(a_{ij} - \frac{K_i K_j}{2M} \right) \delta(i, j) \tag{1}$$