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], Consolación 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)*

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 more sparse. 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 the networks representing power grids, it is discussed the implications of these results from the viewpoint of the engineering task, and how they could be used to analyse the vulnerability risk of power grids to avoid large-scale cascading failures.

Keywords: Electric power system, power grid, high voltage transmission networks, complex networks, community detection, genetic algorithms.

Email addresses: mgl220@fm.ual.es (Manuel Guerrero), pagilm@ual.es (Francisco G. Montoya), rbanos@ual.es (Raúl

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 [3], including engineering tasks [10, 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 maximised is an NP-hard problem [25, 9]. 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 [4, 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. detect any number 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 briefly describes 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. Section 5 discuss the results and the implications of the work on the engineering domain, with emphasis in the possibility of using the proposed methods for contingency

Baños), aalcayde@ual.es (Alfredo Alcayde), cgilm@ual.es (Consolación Gil)

analysis. Finally, The conclusions of this work are provided 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 [7]. 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]. Usually, the nodes of the network represent the power plants, distribution substations and transmission 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 analysed the vulnerability of power grids to blackout using graph topological indexes [17].

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.

2.1. Community detection in power grids

Some studies have applied community detection to power grids. In [4], a method was proposed that used a node similarity index to assign each node to the community sharing maximum similarity. These approaches exhibited good performance in a set of experiments performed on several IEEE standard power grids. Some authors have analyzed the optimal phasor measurement unit placement problem and have used algorithms for community detection to identify coherent groups based on an equivalent graph of generator nodes [11]. 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 analysed 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 [8] 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) [13] 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 [5, 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 [18] as a local search strategy.
- The Generational Genetic Algorithm (GGA+) [12] 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. Additional

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 undirectional 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 functions in community detection due to its simplicity and ease of calculation. Modularity provides a numerical value that represents the quality of the solution, such that the greater the value is, the more accurate the community structure. Therefore, the aim of the algorithms is to maximize the value of Modularity (Q), which is defined as:

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

where M represents the total number of edges in the network; the sub-indices i and j indicate two nodes (vertices) of the network; K_i and K_j are the degree of the i-th and j-th nodes, respectively; the parameter a_{ij} is the element of the i-th row and the j-th column of the adjacency matrix; and $\delta(i,j)$ represents the relationship between the i-th node and the j-th node, such that if node i and node j are in the same community, $\delta(i,j) = 1$; otherwise, $\delta(i,j) = 0$.

Therefore, taking into account the previous definition, the community detection problem consist of finding a network partition that maximizes modularity, Q. This problem has been proven to be NP-hard [9, 25]. This is indeed the reason why heuristics and metaheuristics are used [32].

4.2. Test cases

To analyse the performance of the genetic algorithms, several case studies have been considered. On the one hand, it is used the IEEE 118 bus test case, which represents a portion of the American Electric Power System (in the Midwestern US) as of December, 1962. This network, which contains 118 buses, 186 branches, 54 thermal generators and 91 load sides, has been selected to compare the performance of GGA+ with the node similarity index proposed in [4]. On the other hand, the graph models of a five national-scale power grids are considered. Four are European power grids: Italy, including Sardinia and Sicily, Germany, the continental territory of France, and the Iberian peninsula, including the Balearic islands. The graph models were obtained from the Transmission System Map, which includes information about the transmission system network operated by members of the European Network of Transmission System Operators for Electricity (ENTSO-E) [6]. The ENTSO-E data is taken and displayed using Gephi [20]. Then, Gephi filters and transformations are used to extract only the nodes that belong to each country. The result, for each country, is exported from Gephi to a spreadsheet file (.csv) that includes the nodes and the relationships between nodes (edges). These graphs include transmission lines designed for 220kV voltage and higher and generation stations with a net generation capacity of more than 100MW. Furthermore, it is also analyzed the graph model of the Texas power grid, the second largest state in the United States (US) by both area and population. It is well-known that America's 48 contiguous states (and most of Canada's population) receive the bulk of their electricity from three separate electric grids: the huge Eastern Interconnection, the Western Interconnection, and the relatively small Texas grid [23], which is almost entirely managed by the Electric Reliability Council of Texas (ERCOT). This network consists of a 2000 bus power system with electrical transmission lines of 345 kV, 115 kV and 13.8 kV. The synthetic network model was built based on the statistical analysis of real power systems and public information obtained from the Illinois Center for a Smarter Electric Grid (ICSEG) [15].

Table 1 describes some graph characteristics of these five high voltage power grids. The surface covered by these networks range between about 300,000 and 700,000 km^2, and the number of nodes and edges are almost proportional to the surface area. The dimensions of these networks are significantly larger than those of other power grids considered in recent studies (see e.g. [4]).

4.3. Parameter configuration

There is no a consensus for establishing an optimal population size, although a larger population size is often desirable to obtain high quality solutions. The selection of probabilities of applying crossover and mutation operators also depend of the problem at hand. However, an high crossover operator is often selected. To perform a fair comparison between the two genetic algorithms, the parameters were adjusted by means of a sensitivity analysis that consists in carrying out multiple runs of the algorithms with different

Feature - Power grid	IEEE 118	Italy	Germany	France	Iberian peninsula	Texas
Approx. surface (km^2)	N/A	301,338	357,376	551,695	582,918	695,662
Nodes	118	352	438	904	1104	2007
Edges	177	462	662	1163	1416	2607
Average degree	3.03	2.63	3.03	2.57	2.57	2.60
Network diameter	14	39	21	28	40	39
Average path length	6.31	12.67	9.22	12.09	13.17	15.50
Average clustering coefficient	0.18	0.04	0.23	0.05	0.09	0.02

Table 1: Graph description of the power grids used.

population size and probabilities of applying the evolutionary operators. In this paper, they are analysed the results obtained by different combinations of three parameters: population size (50, 100, and 200 individuals), probability of crossover (50%, 80%, and 95%), and three values of probability of mutation (10%, 20%, and 50%). To accomplish the sensitivity analysis, 10 independent runs of each of these 27 scenarios are executed. Table 2 shows the parameters in common and the model-specific parameters used by MIGA and GGA+, having into account the results of the sensitivity analysis. The experiments were performed on a personal computer with an Intel Core i7 3630Q processor (2.4 GHz, 8 GB DDR3 RAM), which executes the application developed in C# .Net Framework 4. The computational time required to perform each independent run ranged from a few seconds for GGA+ to a few minutes in the case of the MIGA algorithm, since the latter uses Simulated Annealing in the search process.

General parameters	
Population size	200
Iterations	200
Maximum iterations without improvement	50
Crossover probability	0.8
Mutation probability	0.2
Specific parameters (algorithm)	
MIGA	Initial temperature = 800000 Cooling rate = 0.99 loop count l= 10
GGA+	Reproduction ratio = 0.2

Table 2: Parameter settings.

4.4. Results and discussion

The accuracy of MIGA and GGA+ is evaluated according to the modularity (Q) values. Table 3 shows the mean and standard deviation values obtained by both algorithms after 50 independent runs when detecting from 2 to 10 communities. Based on these results, GGA+ achieves the best mean and standard deviation in all test instances independently of the number of communities to be detected. These results also indicate that the larger the problem instance, the greater the advantage of GGA+ with respect to MIGA. Moreover, the standard deviation (shown between parentheses) obtained from the results of these 50 independent runs is lower in GGA+ than in MIGA, which denotes the robustness of the results obtained by GGA+.

		Number of communities								
Test case	Method	2	3	4	5	6	7	8	9	10
Italy	MIGA	0,482	0,626	0,701	0,736	0,757	0,774	0,784	0,788	0,788
		(0,002)	(0,005)	(0,005)	(0,005)	(0,007)	(0,005)	(0,004)	(0,004)	(0,003)
	GGA+	0,491	0,638	0,715	0,743	0,770	0,787	0,798	0,806	0,810
		(0,000)	(0,000)	(0,001)	(0,002)	(0,005)	(0,001)	(0,001)	(0,002)	(0,002)
Germany	MIGA	0,478	0,637	0,707	0,737	0,759	0,769	0,781	0,789	0,793
		(0,003)	(0,005)	(0,004)	(0,005)	(0,007)	(0,006)	(0,006)	(0,005)	(0,004)
	GGA+	0,485	0,643	0,714	0,754	0,784	0,800	0,810	0,814	0,817
		(0,001)	(0,001)	(0,001)	(0,003)	(0,002)	(0,001)	(0,001)	(0,003)	(0,002)
France	MIGA	0,477	0,636	0,709	0,750	0,776	0,795	0,804	0,811	0,818
		(0,002)	(0,003)	(0,004)	(0,003)	(0,005)	(0,005)	(0,004)	(0,003)	(0,003)
	GGA+	0,486	0,645	0,720	0,765	0,794	0,818	0,830	0,838	0,842
		(0,001)	(0,001)	(0,001)	(0,002)	(0,003)	(0,001)	(0,001)	(0,002)	(0,002)
Iberian peninsula	MIGA	0,480	0,637	0,712	0,758	0,787	0,812	0,823	0,828	0,832
		(0,002)	(0,006)	(0,003)	(0,005)	(0,005)	(0,003)	(0,003)	(0,002)	(0,003)
	GGA+	0,489	0,650	0,727	0,773	0,804	0,829	0,836	0,843	0,851
		(0,001)	(0,001)	(0,002)	(0,001)	(0,001)	(0,001)	(0,002)	(0,002)	(0,002)
Texas	MIGA	0,485	0,635	0,704	0,748	0,772	0,789	0,804	0,814	0,826
		(0,002)	(0,005)	(0,004)	(0,004)	(0,004)	(0,003)	(0,004)	(0,003)	(0,003)
	GGA+	0,493	0,654	0,731	0,776	0,804	0,823	0,838	0,851	0,858
		(0,001)	(0,001)	(0,003)	(0,002)	(0,003)	(0,002)	(0,004)	(0,002)	(0,003)

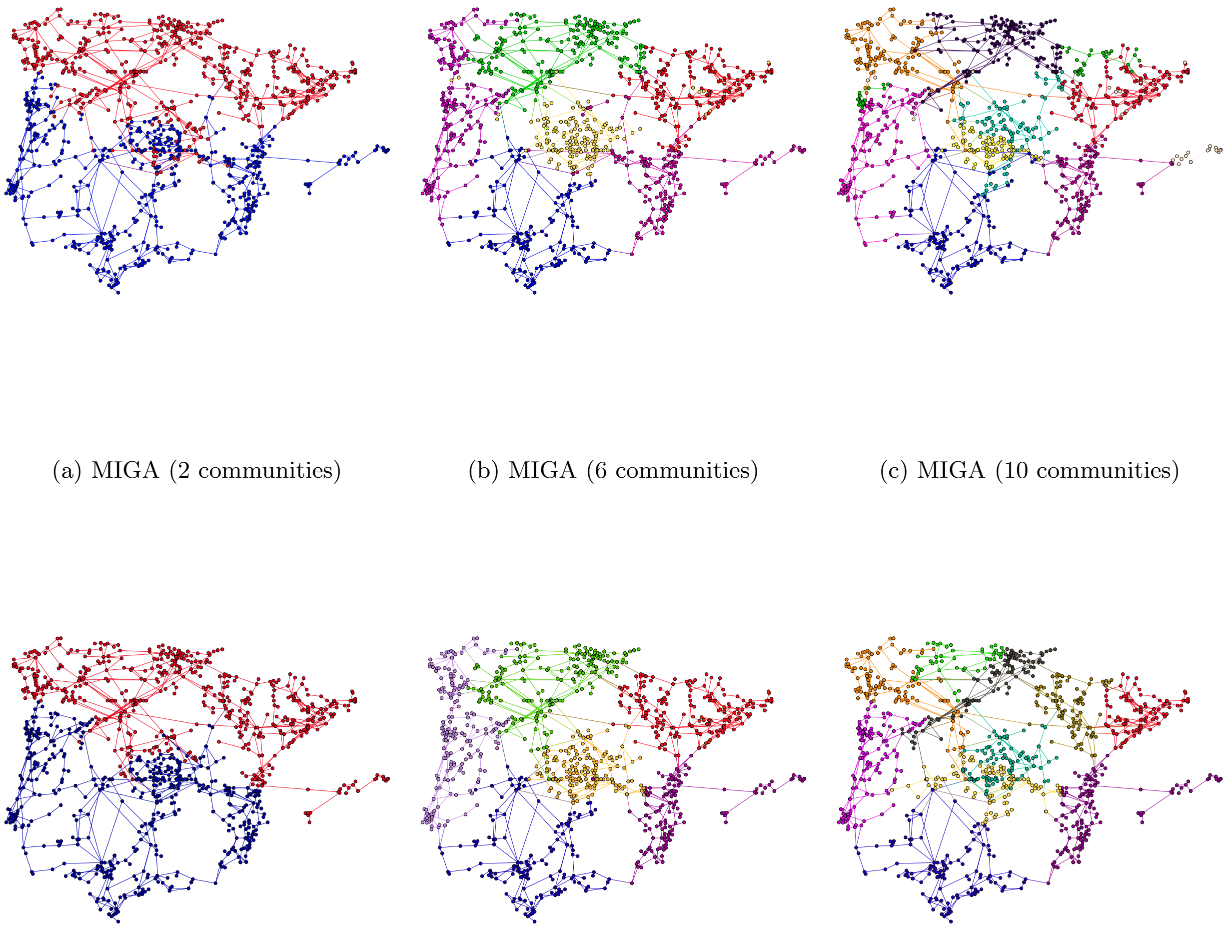
Table 3: Mean (and standard deviation) of the modularity obtained by MIGA and GGA+ after 50 independent runs.

To compare the two algorithms, it is necessary to determine if there are significant differences between the results obtained by the different methods. With this aim, a one-way ANOVA has been applied, and the results indicate that the $p\text{-value}<0.05$ in all cases, i.e., the null hypothesis is always rejected, indicating that there is a significant variation between at least some of the means of the different groups. Therefore, the results obtained by GGA+ are significantly different from those obtained by MIGA, supporting the conclusion based on the mean values displayed in Table 3.

Figure 1 shows the communities detected by MIGA and GGA+ in the Iberian peninsula power grid. These results reveal the significant differences between the two algorithms, especially when the number of communities increases. Even when detecting two communities, MIGA has some difficulty in assigning communities in the central part of the graph, while GGA+ obtains clearly differentiated communities. Based on the data included in these figures and the results in Table 3, GGA+ not only outperforms MIGA, but it also exhibits good performance in these large networks.

The results obtained by GGA+ are analysed in more detail. Figure 2(a) displays the communities detected by GGA+ in the Iberian peninsula power grid with different levels of detail. These data reveal that this algorithm is able to obtain quality solutions even when the number of communities increases. Moreover, Figure 2(b) provides a different layout based on the Force Atlas2 [16] plugin in Gephi. This visualisation method builds a force directed layout by simulating a physical system in order to accommodate nodes and links in a spatial network. Nodes repel each other like charged particles, while edges attract their nodes like springs. The aim of this method is to help construct a balanced state network that facilitates data interpretation. The analysis of Figure 2 demonstrates the good performance of GGA+ independently of the degree of abstraction.

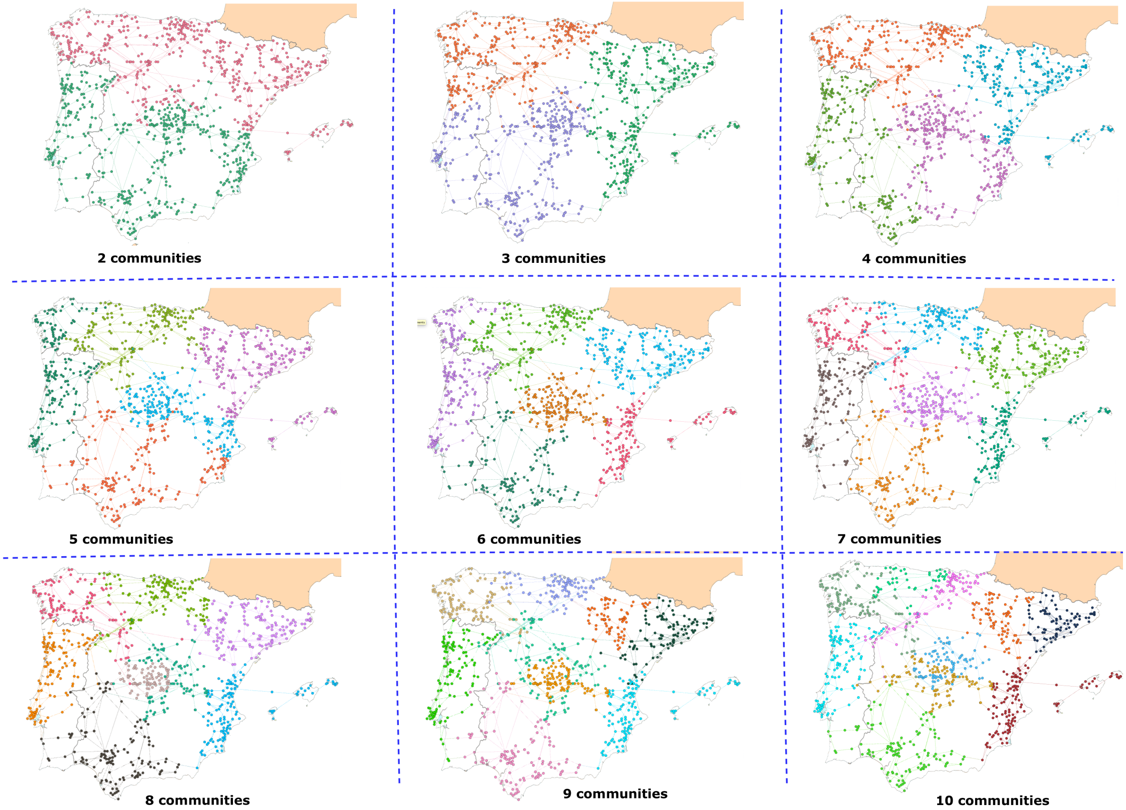
Figure 3 displays the communities detected by GGA+ in the networks representing the power grids of Italy, Germany and France with 3 and 8 communities, respectively. These results clearly show that independently of the number of communities, GGA+ detects clearly differentiated communities. In the case of the Italian network (Figure 3(a)), the solution provided by GGA+ clearly divides the Italian power grid into three communities: north, central and south. The analysis of the communities detected in the German power grid (Figure 3(b)) also provides interesting results. In particular, the structure of three communities obtained by GGA+ is coincident with the historical division of Germany during the Cold War, with one community covering a large area in East Germany and two communities in West Germany covering the north-central area (British and French sectors) and the south area (American sector). These results coincide with previous studies that have highlighted that the evolution of electric systems in Europe was largely affected



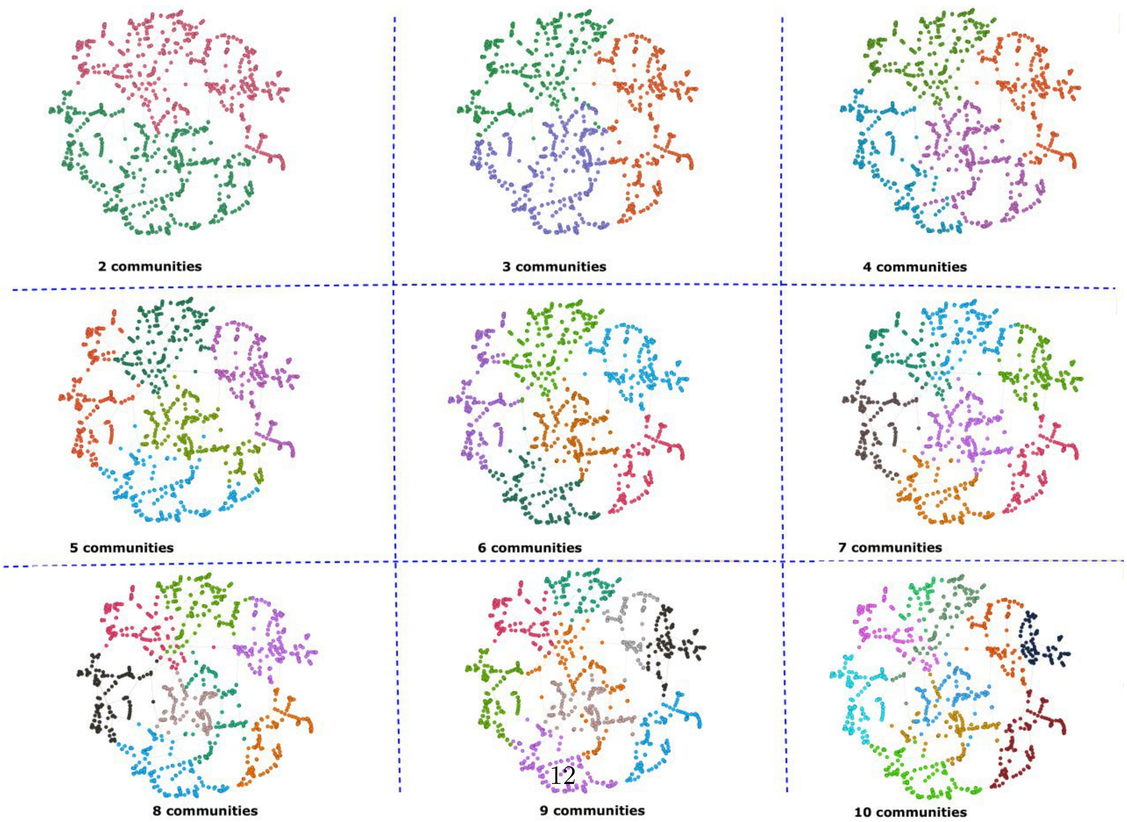
(a) MIGA (2 communities) (b) MIGA (6 communities) (c) MIGA (10 communities)

(d) GGA+ (2 communities) (e) GGA+ (6 communities) (f) GGA+ (10 communities)

Figure 1: Results obtained by MIGA and GGA+ in the Iberian peninsula network considering 2, 6 and 10 communities.

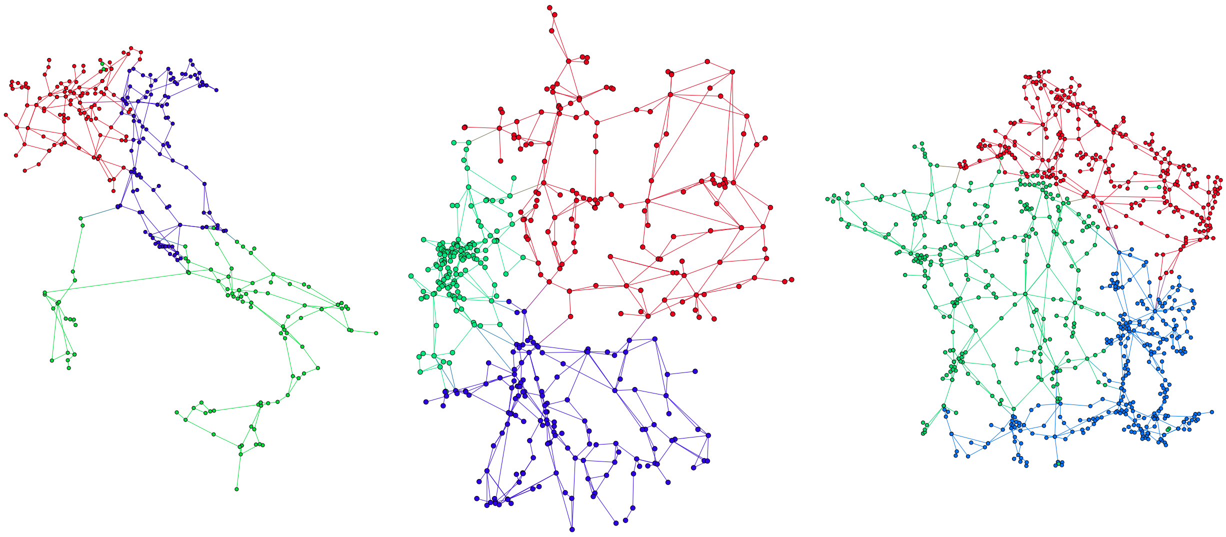


(a) Physical layout

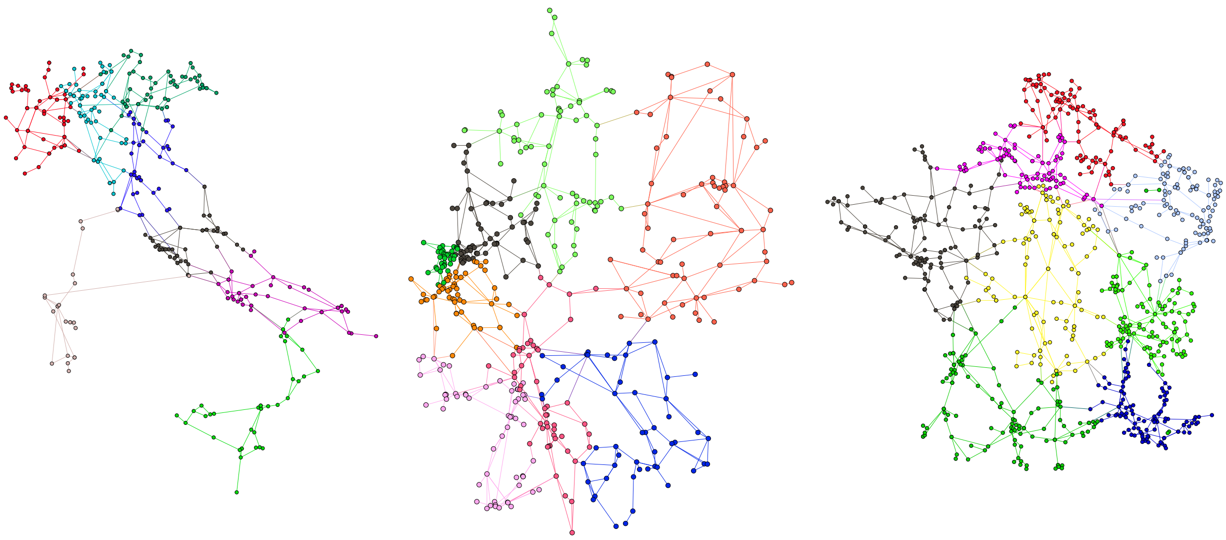


(b) Force Atlas 2

Figure 2: Results obtained by GGA+ in the Iberian peninsula and Balearic islands considering from 2 to 10 communities.



(a) GGA+ (3 communities): Italy (b) GGA+ (3 communities): Germany (c) GGA+ (3 communities): France



(d) GGA+ (8 communities): Italy (e) GGA+ (8 communities): Germany (f) GGA+ (8 communities): France

Figure 3: Results obtained by GAA+ in the Italian, German and French power grids considering 3 and 8 communities.

by political conflicts such as the two world wars, the Cold War and the political separation of Eastern and Western Europe [31]. The network representing the French power grid exhibits a homogeneous geographical distribution of high voltage transmission lines, except for the central zone, where the density is slightly lower than in the south-east and north. When the network is divided into three communities (Figure 3(c)), they are of similar size, although the north-east and southeast communities cover a smaller area than the central community. When these three networks are divided into eight communities (Figures 3(d), 3(e), and 3(f)) the connected regions representing each community are of similar size.

Figure 4 presents some solutions obtained by GGA+ in the Texas network with different levels of detail (3, 4, 5, and 6 communities). A significant concentration of substations and electrical lines is observed around the major cities in Texas, i.e., Houston, San Antonio, Dallas, and Austin. In fact, Figure 4 shows that when GGA+ obtains a set of five communities, four communities include these important metropolitan cities, while the fifth region covers the sparsely populated west region of Texas.

In addition to the comparison using the five previous networks, Figure 5 compares the results obtained by GGA+ with those obtained by the method recently proposed by [4] in the IEEE 118 bus network. Figure 5(a), shows the one-line diagram of IEEE 118-bus test system including the communities detected by [4], while Figure 5(b) shows the communities detected by GGA+. Figure 5(c) displays the graph with the communities detected by [4], which obtains a modularity value of 0.689, while the modularity obtained by GGA+ (Figure 5(d)) is 0.726. Figure 5(e) shows the difference between both solutions, such that those nodes (buses) vary from community between the solutions are highlighted with red color with a bigger size. Therefore, it has also been proved that the genetic algorithm outperforms the method proposed by [4].

5. Implications for power grid design and operation

The results obtained in the previous section show how the genetic algorithms are able to determine communities in the networks representing national-scale power grids. In addition to the information about the topological characteristics of these systems, understanding their community structure can be a valuable information to be considered in the planning, design and operation of power systems.

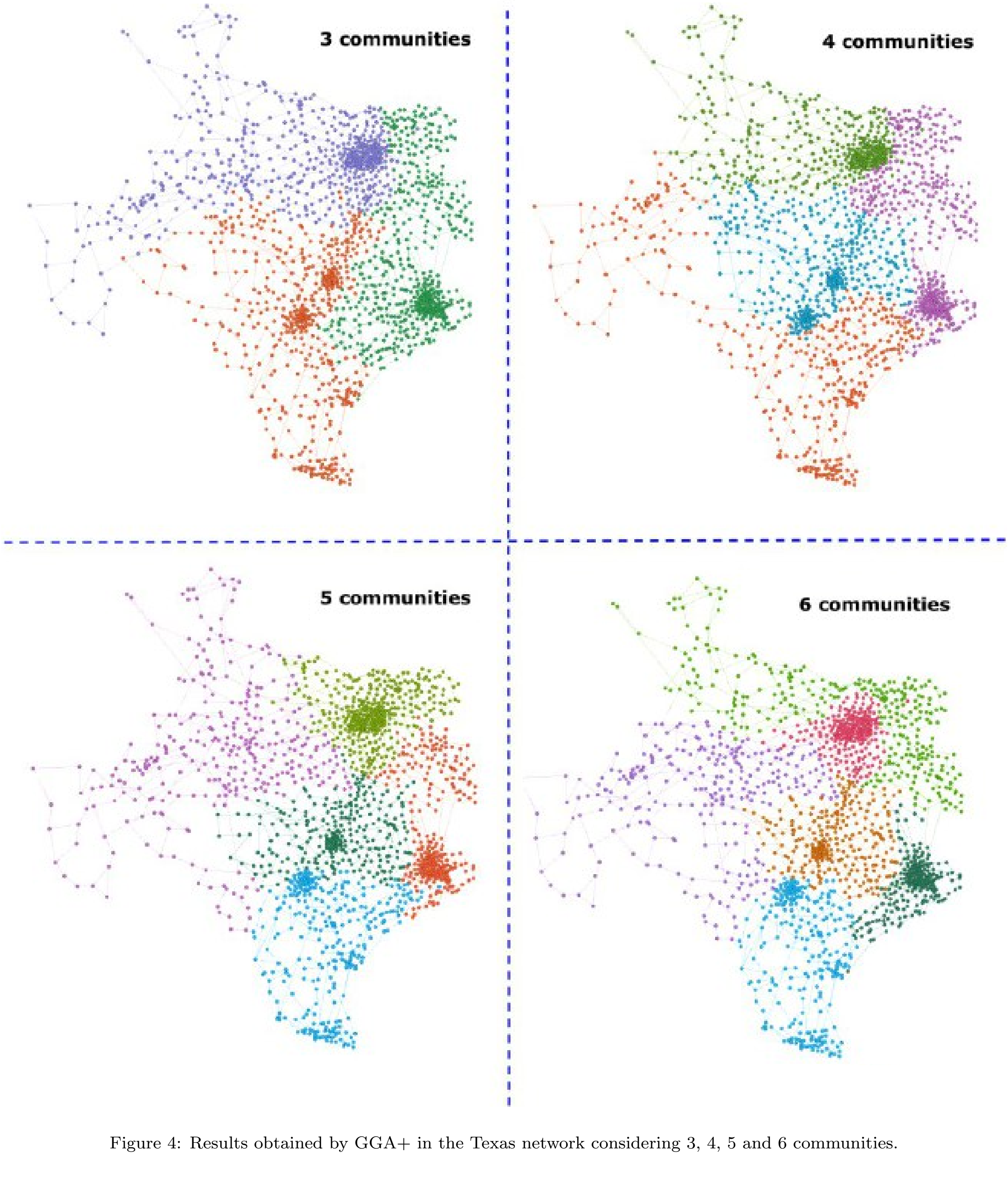


Figure 4: Results obtained by GGA+ in the Texas network considering 3, 4, 5 and 6 communities.

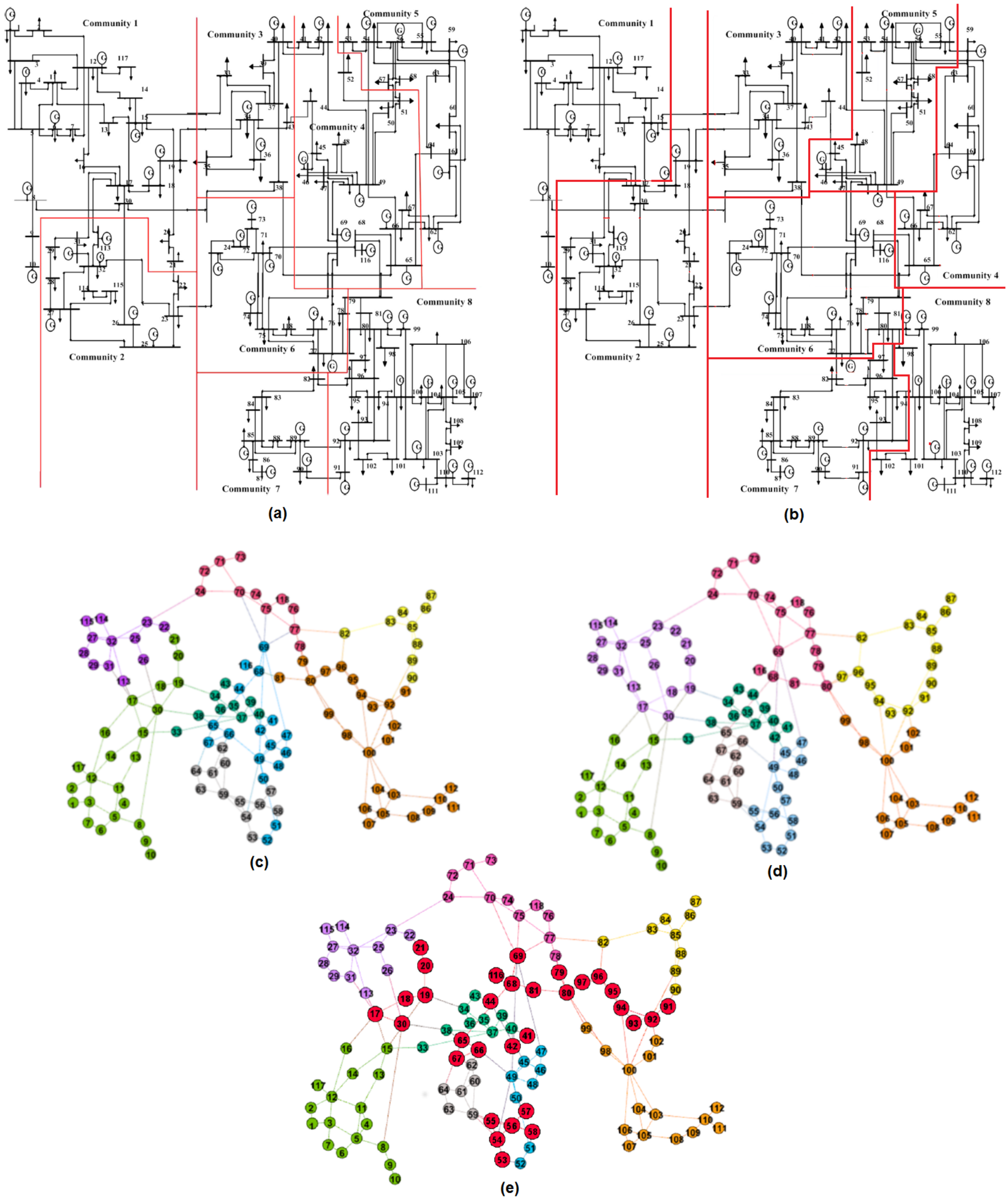


Figure 5: One-line diagram of the IEEE 118-bus test system with eight communities detected by [4] (a) and GGA+ (b). Graph with the nodes of each community detected by [4] (c) and GGA+ (d), and its differences (e).

POWER FLOW ANALYSIS (n-1)-security criterion					COMMUNITY DETECTION ANALYSIS
Branch (failure)	From bus	To bus	Ploss (MW)	Qloss (MVAr)	Is boundary line in the community detected?
8	8	5	197,03	1136,16	YES (k={10,11,12,13,14,15,16,17,18,19,20})
51	38	37	169,57	919,93	YES (k={18,19,20})
96	38	65	167,84	906,67	YES (k={2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20})
38	26	30	165,90	936,08	YES (k={2,4})
36	30	17	150,14	859,82	YES (k={18,19,20})

Table 4: Critical transmission lines -sorted from highest to lowest power losses (Ploss) due to failure in one transmission line (branch)- that are boundary edges between different communities -graph obtained by GGA+ considering from k=2 to k=20 communities-.

Contingency analysis is often considered in the planning and design stages of power systems to evaluate their security. A contingency can be defined as a unplanned outage due to the loss or failure of a small part of the power system (e.g. a transmission line). Since many years ago, power systems engineers perform contingency analyses on computer models of the power system to evaluate the effects (power losses, overloads, etc.) resulting from each outage event. Planning studies aim to guarantee the system is able to withstand sudden disturbances such as electric short circuits or unanticipated loss of system elements [41]. The (n-1)-security criterion is verified by analyzing power flows considering one component failure (e.g.: a transmission line is out of service) then determining how this contingency would affect the network operation.

Table 4 shows some results obtained from a n-1 security contingency analysis in the network IEEE 118. Since this system has 186 branches, the (n-1)-security contingency analysis requires to calculate 186 power flows. With this aim, a Matlab power system simulation package (Matpower) is used. Table 4 shows the five most critical transmission lines (branches) in terms of power losses (Ploss). When these results are compared with the communities detected by GGA+, it is observed an interesting pattern: the most critical branches in power networks according to the contingency analysis are edges that connect nodes of different communities obtained by the genetic algorithm. Therefore, as future work, it is planned to perform an exhaustive study to confirm that the detection of communities in power grids would provide useful information to those power systems engineers that perform contingency analyses.

On the other hand, it is important to remark that the operation of power systems has changed significantly in the last years. Traditional electric power distribution systems have been designed on the assumption that the sources of power are the primary substations close to conventional power plants [28]. This assumption is invalidated by the entrance of distributed generation, which allows to generate electricity in plants that are connected to a distribution network rather than the transmission network. The application of distributed generation has considerably increased in last years and it is expected to be higher in the near future due to the emerging utilization of renewable energy resources, including solar panels, micro wind turbines, and combined heat and power units. This tendency has led to a demand for a new electricity distribution paradigm, such that there is an increasing interest in the concepts of smart grids and microgrids. A microgrid is a cluster of both distributed generators and loads which act to cooperate with the main grid or autonomously from it [14]. In case of an upstream network disturbance, the microgrid is able to disconnect from the main grid then operating as a self-controlled entity. With this aim, several authors have proposed methods for island partitioning as a solution to rapidly restore the energy supply of important loads, and to reduce the outage time [28]. But it is not an easy task, since large power imbalance in the island can lead to frequency instability, then resulting in a de-energization of part or all of the system [19]. Therefore, as future work, it will be considered the use of community detection as a partitioning strategy to implement islanding strategies in power systems.

6. Conclusions

Power systems design is an engineering task consisting in supplying power from power plants to various load centers with high reliability, while the systems are operated economically at maximum efficiency. However, these systems are often affected by contingencies.

This paper analyses the performance of two efficient genetic algorithms (MIGA and GGA+) for solving community detection problems in national-scale high voltage transmission networks. These approaches use powerful initialisation methods and evolutionary search operators under the guidance of modularity, and are able to obtain good quality solutions in different networks. These algorithms enable a flexible and adaptive analysis of the characteristics of power grids with different levels of detail (number of communities).

The empirical study considers six test cases representing the national-scale high voltage power grids of

Italy, Germany, the continental territory of France, the Iberian peninsula (Portugal and Spain, including the Balearic islands), Texas (US), and the IEEE 118 bus test case that represents a portion of the American Electric Power System (in the Midwestern). Despite several studies have previously applied community detection to networks representing power grids, to our knowledge, it is the first study that has empirically analyzed the performance of evolutionary methods in network representing national-scale power grids. Moreover, this study is particularly interesting if we take into account that the development of the power grid in the U.S. was organized according to federal and state regulation, while the power grid in Europe has traditionally organized country by country.

These results are useful in showing that genetic algorithms are fast and powerful tools to detect communities in national-scale high voltage transmission networks, and they also provide interesting topological information about the physical distribution and concentration of these grids. In addition, this new knowledge is useful for engineering decision makers since it provides information about the critical elements (lines in our case) which damage may trigger cascades of failures across the grid and lead to a large blackout. Future work will be focused to study with more detail how partitions obtained with community detection algorithms could provide useful information for contingency analysis and islanding strategies. Moreover, it is expected to apply parallel and multi-objective optimisation methods to solve this problem in national and continental-scale power grids.

7. Acknowledgements

This research has been supported by the Spanish Ministry of Economy and Competitiveness and FEDER funds (project TIN2015-67020-P), and by the Regional Government of Andalusia (ceiA3 project) at the University of Almeria.

References

[1] Agency, I. E. (2016). Large-scale electricity interconnection. technology and prospects for cross-regional networks. https://www.iea.org/publications/freepublications/publication/Interconnection.pdf. Accessed: 2017-11-30.

[2] Barthelemy, M. (2011). Spatial networks. *Physics Reports*, *499*, 1–101.

[3] Bornholdt, S., & Schuster, H. G. (2003). *Handbook of Graphs and Networks: From the Genome to the Internet*. New York, NY, USA: John Wiley & Sons, Inc.

[4] Chen, Z., Xie, Z., & Zhang, Q. (2015). Community detection based on local topological information and its application in power grid. *Neurocomputing*, *170*, 384–392.

[5] Clarke, J., McLay, L., & McLeskey, J. T. (2014). Comparison of genetic algorithm to particle swarm for constrained simulation-based optimization of a geothermal power plant. *Advanced Engineering Informatics*, *28*, 81–90.

[6] ENTSO-E (2017). Entso-e transmission system map. https://www.entsoe.eu/map/Pages/default.aspx. Accessed: 2017-11-30.

[7] Etxegarai, A., Eguia, P., Torres, E., A., I., & Valverde, V. (2015). Review of grid connection requirements for generation assets in weak power grids. *Renewable and Sustainable Energy Reviews*, *31*, 1501–1514.

[8] Fogel, D. B. (2006). *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. IEEE Press.

[9] Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*, 75–174.

[10] Fortunato, S., & Hric, D. (2016). Community detection in networks: A user guide. *Physics Reports*, *659*.

[11] Gómez, O., & Rios, M. A. (2013). Ilp-based multistage placement of pmus with dynamic monitoring constraints. *International Journal of Electrical Power & Energy Systems*, *53*, 95–105.

[12] Guerrero, M., Montoya, F. G., Baños, R., Alcayde, A., & Gil, C. (2017). Adaptive community detection in complex networks using genetic algorithms. *Neurocomputing*, *266*, 101–113.

[13] Holland, J. H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence.*. U Michigan Press.

[14] Hosseinnezhad, V., Rafiee, M., Ahmadian, M., & Siano, P. (2018). Optimal island partitioning of smart distribution systems to improve system restoration under emergency conditions. *Electrical Power and Energy Systems*, *97*, 155–164.

[15] ICSEG (2016). Texas synthetic power grid test case. http://icseg.iti.illinois.edu/synthetic-power-cases/texas2000-june2016/. Accessed: 2017-11-30.

[16] Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PloS One*, *9*, e98679.

[17] Kim, C. J., & Obah, O. B. (2007). Vulnerability metrics and analysis for communities in complex networks. *International Journal of Emerging Electric Power Systems*, *8*, 1–15.

[18] Kirkpatrick, S., Vecchi, M. et al. (1983). Optimization by simulated annealing. *Science*, *220*, 671–680.

[19] Li, B., & Sansavini, G. (2016). Effective multi-objective selection of inter-subnetwork power shifts to mitigate cascading failures. *Electric Power Systems Research*, *134*, 114–125.

[20] M., B., S., H., & M., J. (). Gephi: an open source software for exploring and manipulating networks. In [20].

[21] Mahapatra, S., & Dasappa, S. (2012). Rural electrification: Optimising the choice between decentralised renewable energy sources and grid extension. *Energy for Sustainable Development*, *16*, 146–154.

[22] Mahmoud, M., Hussain, S., & Abido, M. (2014). Modeling and control of microgrid: An overview. *Journal of the Franklin Institute*, *351*, 2822–2859.

[23] Mazur, A., & Metcalfe, T. (2012). America's three electric grids: Are efficiency and reliability functions of grid size? *Electric Power Systems Research*, *89*, 191–195.

[24] Morton, D. L. (2002). Reviewing the history of electric power and electrification. *Endeavour*, *26*, 60 – 63.

[25] Nascimento, M. C., & Pitsoulis, L. (2013). Community detection by modularity maximization using grasp with path relinking. *Computers & Operations Research*, *40*, 3121–3131.

[26] Newman, M. E. (2012). Communities, modules and large-scale structure in networks. *Nature Physics*, *8*, 25–31.

[27] Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, *69*, 026113.

[28] Oboudi, M., Hooshmand, R., & Karamad, A. (2016). Feasible method for making controlled intentional islanding of microgrids based on the modified shuffled frog leap algorithm. *Electrical Power and Energy Systems*, *78*, 745–754.

[29] Pagani, G. A., & Aiello, M. (2013). The power grid as a complex network: A survey. *Physica A: Statistical Mechanics and its Application*, *392*, 2688–2700.

[30] Pahwa, S., Youssef, M., Schumm, P., Scoglio, C., & Schulz, N. (2013). Optimal intentional islanding to enhance the robustness of power grid networks. *Physica A: Statistical Mechanics and its Applications*, *392*, 3741 – 3754.

[31] Ponce-Jara, M., Ruiz, E., Gil, R., E., S., Pérez-Molina, C., & Castro, M. (2017). Smart grid: Assessment of the past and present in developed and developing countries. *Energy Strategy Reviews*, *18*, 38–52.

[32] Rocco, C. M., & Ramirez-Marquez, J. E. (2011). Vulnerability metrics and analysis for communities in complex networks. *Reliability Engineering and System Safety*, *96*, 1360–1366.

[33] Saleh, A. I., Rabie, A. H., & Abo-Al-Ez, K. M. (2014). A data mining based load forecasting strategy for smart electrical grids. *Advanced Engineering Informatics*, *30*, 422–448.

[34] Sánchez, P., Montoya, F. G., Manzano-Agugliaro, F., & Gil, C. (2013). Genetic algorithm for s-transform optimisation in the analysis and classification of electrical signal perturbations. *Expert Systems with Applications*, *40*, 6766–6777.

[35] Schlachtberger, D., Brown, T., Schramm, S., & Greiner, M. (2017). The benefits of cooperation in a highly renewable european electricity network. *Energy*, *134*, 469–481.

[36] Shang, R., Bai, J., Jiao, L., & Jinm, C. (2013). Community detection based on modularity and an improved genetic algorithm. *Physica A*, *392*, 1215–1231.

[37] Sánchez-García, R. J., Fennelly, M., Norris, S., Wright, N., Niblo, G., Brodzki, J., & Bialek, J. W. (2014). Hierarchical spectral clustering of power grids. *IEEE Transactions on Power Systems*, *29*, 2229–2237.

[38] Tang, W., Allman, A., Pourkargar, D. B., & Daoutidis, P. (2018). Optimal decomposition for distributed optimization in nonlinear model predictive control through community detection. *Computers & Chemical Engineering*, *111*, 43–54.

[39] Thakur, M., & Kumar, A. (2016). Optimal coordination of directional over current relays using a modified real coded

genetic algorithm: A comparative study. *International Journal of Electrical Power & Energy Systems*, *82*, 484–495.

[40] Wang, S., Zhang, J., Zhao, M., & Min, X. (2017). Vulnerability analysis and critical areas identification of the power systems under terrorist attacks. *Physica A: Statistical Mechanics and its Applications*, *473*, 156–165.

[41] Wiest, P., Rudion, K., & Probst, A. (2018). Efficient integration of (n - 1)-security into probabilistic network expansion planning. *Electrical Power and Energy Systems*, *94*, 151–159.

[42] Wu, Y., Tang, Y., Han, B., & Ni, M. (2015). A topology analysis and genetic algorithm combined approach for power network intentional islanding. *International Journal of Electrical Power & Energy Systems*, *71*, 174–183.

[43] Zhou, C., Ding, L., Zhou, Y., & Luo, H. (2018). Topological mapping and assessment of multiple settlement time series in deep excavation: A complex network perspective. *Advanced Engineering Informatics*, *36*, 1–19.