Expert Systems With Applications 196 (2022) 116557

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Identifying critical nodes in power networks: A group-driven framework

Yangyang Liu [a,b], Aibo Song [a,b,*], Xin Shan [c], Yingying Xue [a,b], Jiahui Jin [a,b]

[a] School of Computer Science and Engineering, Southeast University, Nanjing, China
[b] Key Laboratory of Computer Network and Information Integration (Southeast University), Nanjing, China
[c] NARI Group Corporation (State Grid Electric Power Research Institute), Nanjing, China

ARTICLE INFO

Keywords:
Power network
Critical nodes
Group structure
Group-driven framework

ABSTRACT

Cascading failures can easily occur and cause a major blackout in power systems when a critical devices breaks down. It is an essential problem to evaluate the importance of devices/nodes for power networks. Even though approaches for identifying a power network's critical nodes have been investigated in the past, accurately achieving critical nodes' identification has proven to remain a challenging task. Here, we propose a group-driven framework, called *GDF-ICN*, in which the potential group information is introduced for the first time, to enhance the performance of critical nodes identification. It is a novel framework performed in an iterative manner, the introduction of nodes clustering effect improves the reliability of identifying critical nodes, while the identification of critical nodes promotes the characterization of a denser group structure. Specifically, we adopt the electrical and group information of a power network simultaneously to define each node's electrical coupling that might affect the importance of nodes. To produce denser groups, we propose a fuzzy tightness metric that can be regarded as group optimization's objective function. We also provide a preliminary but systematic research on how to transform any metric of evaluating hard group structure into a metric of evaluating fuzzy group structure. Comprehensive experiments on several benchmark power networks show the necessity of considering group information to evaluate the importance of nodes and the efficiency of *GDF-ICN*.

1. Introduction

In recent years, there have been frequent large-scale power outages around the world, such as the power failure in north American (2003), the national power failure in India (2012), the national power failure in Venezuela (2019) and the California power failure in USA (2020). These outages urgently require the scholars and engineers to design more effective strategies to avoid or reduce their destructive effects (Fang et al., 2018; Wang, Gu et al., 2021). It can be seen from the follow-up investigation results and analysis reports that power outages are often a series of chain failures caused by the failure of some critical nodes/equipments that play an extremely important role in the stable operation of a power system. Therefore, how to quickly and effectively identify these critical nodes is of great significance for helping the power supply department to accurately protect critical nodes in advance and achieve differentiated management.

For the identification of critical nodes in power networks, the researchers have done some research, here we roughly divide them into two categories: the first type of research is simply based on a network's topology information, which applies degree, betweenness and other indicators known in complex network theory (Lv et al., 2016) to build static or dynamic models, and then evaluate the importance of all nodes (Guan et al., 2014). Finally, the nodes with high importance value are regraded as the desired critical nodes (Fang et al., 2018). However, the biggest difference between power network and other types of networks lies in the existence of electrical information in the former. The approaches based on pure topological information fail to capture the electrical characteristics. Therefore, most researchers at present seek to design critical nodes identification approaches more in line with the actual electrical characteristics of the power network (Li et al., 2019).

The second type of research mainly takes electrical characteristics into consideration and defines relevant indicators to represent the importance of nodes in power networks (Yang et al., 2020; Zhao et al., 2020). Meanwhile, the network's topology is taken also into account to some extent in this kind of research (Liu et al., 2018; Wang, Lv et al., 2021; Zhou et al., 2019). Although some achievements have been made in this kind of research, the group structure characteristics of the network have not been well considered. Formation of groups,

* Corresponding author at: School of Computer Science and Engineering, Southeast University, Nanjing, China.
E-mail addresses: yyliu@seu.edu.cn (Y. Liu), absong@seu.edu.cn (A. Song), shanxin@sgepri.sgcc.com.cn (X. Shan), y-xue@seu.edu.cn (Y. Xue), jjin@seu.edu.cn (J. Jin).

https://doi.org/10.1016/j.eswa.2022.116557
Received 9 February 2021; Received in revised form 20 October 2021; Accepted 15 January 2022
Available online 17 February 2022
0957-4174/© 2022 Elsevier Ltd. All rights reserved.

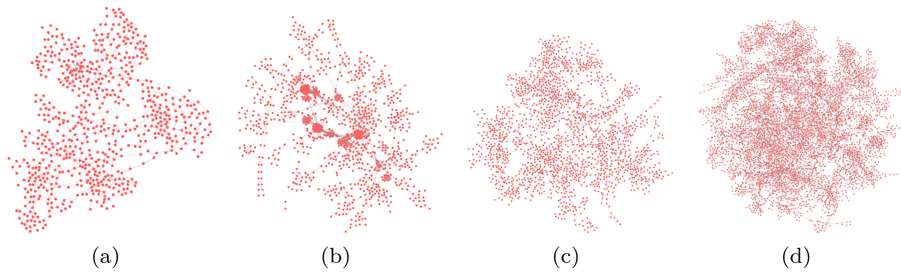


Fig. 1. Illustration of topology of four benchmark power networks[1] with varying number of nodes and edges. (a) 685 bus power network (685 nodes and 1967 edges); (b) 1176 bus power network (1176 nodes and 8687 edges); (c) western US power network in 2009 (1723 nodes and 2394 edges); (d) western US power network in 2010 (5300 nodes and 8271 edges). It can be observed that all four networks display an obvious group structure.

also referred as clusters or communities, is a phenomenon commonly observed across several domains where groups of nodes are densely connected with each other but sparsely connected to the rest of the graph. This phenomenon is also called by scholars as nodes clustering effect or agglomeration effect in a networked system. In the past two decades, a plethora of works were proposed to investigate it from different domains with networked system (Fortunato, 2010; Liu et al., 2020; Newman, 2004). In practice, group structures are usually formed in a bottom-up way where the nodes in a group centered on a "leader" node who dominates all other nodes in the realm of corresponding group. We can observe similar phenomenon in the deployment of power devices where there is a electrical hub in an administration region whose destruction will cause a major blackout in power systems. Taking inspiration from the analysis above, it is obvious that the nodes in a power system are not only affected by electrical characteristics, but also restricted by the group to which they belong. Thus, here we propose *GDF-ICN*—a Group-Driven Framework for Identifying Critical Nodes in power networks. Our motivation is mainly elicited by answering following two questions:

1. Why should group information be considered when analyzing and identifying the critical nodes in a power network?
2. How can we use group information to facilitate the identification of critical nodes in a power network?

To answer these two questions, we first analyze the topology of four representative benchmark networks shown in Fig. 1. As Fig. 1 shows, all the networks display a certain degree of agglomeration phenomenon that some nodes are densely connected. It is worth noting that power networks in reality are often in accordance with the construction principle of stratification and division. Not only are the equipments within a common administrative region geographically close, but the energy transfer among them is also frequent. This characteristic is consistent with the agglomeration phenomenon shown in Fig. 1. In power networks, when a node fails, it first affects the state of neighbor nodes, and then affects other nodes in the network according to the electrical coupling correlation path between nodes (Yang et al., 2020; Zhou et al., 2019). It can be seen that the strength of electrical coupling correlation between nodes is the key to judge critical nodes in a power network. If a node has strong electrical coupling with its neighboring nodes, its fault may cause a large-scale chain failure. It is hence regarded as vulnerable (critical) nodes of power networks. Many studies have shown that in a real-world network, the more groups two nodes coexist, the stronger the correlation between these two nodes would be (Yang & Leskovec, 2013). In power systems, there are always some nodes that belong to two or more groups simultaneously. Therefore, given a power network, the electrical coupling correlation between nodes is related to not only its electrical characteristics, but also its agglomeration phenomenon. It is of great theoretical and practical significance to introduce group information into the identification of critical nodes in power networks.

Now let us consider the second question. In reality, the power network is equipped with complicated topology and electrical configuration. Its group information captured by expert experience is unreliable, and will mislead the identification of critical nodes. Hence, group information cannot be directly captured to facilitate the identification of critical nodes. Thus here we seek to design a group-driven framework, i.e., *GDF-ICN*, to address this issue. In particular, *GDF-ICN* identifies critical nodes in a group-driven iterative manner. Initially, we capture a power network's group structure by expert experience. Then in the process of each iteration, we use group information to enhance the accuracy of critical nodes identification, while the obtained critical nodes are adopted to promote the optimization of group structure. Finally, *GDF-ICN* can obtain not only critical nodes, but also the optimal group structure related to these critical nodes, which is conducive to the power network's differentiated management.

GDF-ICN is a meaningful attempt of introducing the group structure into critical nodes identification task in power network analysis. It contains three modules: critical nodes identification, group expansion and membership update. Specifically, we introduce the group and electrical information in module 1 to identify critical nodes. Note that we have finished the critical nodes identification task of one round of iteration after the module 1 is performed. In the next two modules, we only focus on network's group structure. Then a fast and effective strategy is designed in module 2 for assigning all nodes to some groups based on critical nodes. In module 3, we propose a fuzzy tightness metric that can be regarded as the objective function of group optimization for finding more denser groups. In addition, we also conduct deeper analysis and discussion on *GDF-ICN*, which provide us some insightful idea for the researches on critical nodes and fuzzy group structure in power systems. It is worth noting that *GDF-ICN* also integrates the electrical information and topological information in an implicit way to identify critical nodes in power networks. In practice, we treat a power network as a directed weighted graph where the branch impedance and current direction of two nodes are regarded as the weight and direction of the graph, respectively.

In a nutshell, *GDF-ICN* provides a special perspective to identify the critical nodes in power networks while the electrical characteristics as well as the structural information are preserved simultaneously. Compared to the state-of-the-art approaches, it is novel and effective that firstly introduces the group structural information in an iterative manner into the construction of critical nodes identification framework. Notice that *GDF-ICN* can finally achieve two goals: critical nodes identification and group structure characterization. Namely, we can also finish the other task when any one of two tasks is finished. Specifically, *GDF-ICN* generates the group structure at the beginning using a bottom-up way and then a fuzzy tightness optimization strategy is adopted to further refine the group structure. The refined group structure in current round of iteration is then input into the next round to recompute (update) the nodes' importance.

Our contributions in this paper are summarized in four folds:

- We propose a group-driven critical nodes identification framework in power networks, i.e., *GDF-ICN*, which identifies critical nodes by considering the influence of nodes' agglomeration on nodes' electrical coupling. We see *GDF-ICN* as a powerful

approach to identify critical nodes in the context where the denser group structure itself is a consequence of the identification process.
- We design a fuzzy tightness metric to measure the goodness of a group structure in a power network, with which a metric optimization strategy is performed to accomplish better group structure characterization.
- Also, we provide a preliminary tool to soften the metrics of evaluating hard groups. It motivates us and can help researchers use a more systematic way to study network's fuzzy group structure.
- We conduct experiments on several power networks, and the experimental results show that our *GDF-ICN* can well identify the critical nodes and also reveal the fuzzy group structure associated with these critical nodes.

The remainder of this paper is organized as follows. After reviewing the related works in Section 2, we offer some preliminaries on power network and problem definition in Section 3. Details of *GDF-ICN* for the problem are given in Section 4, while in the following Section 5 we provide a further analysis and discussion on initialization, definition of repulsive force, fuzzy membership and model's computation complexity. In Section 6, comprehensive experiments results are given, followed by conclusion remarks in Section 7.

2. Related works

In this section, we first review the related works of critical nodes identification in power networks. Then, we briefly study the existing literature of network's potential group structure which motivates us to pursue the current study.

Obviously, since power network is one type of complex networks, scholars initially adopted classical methods proposed for identifying key nodes in universal complex networks to identify critical nodes in power networks. For example, authors in Guan et al. (2014) combined betweenness and degree, and presented a new index to identify the critical nodes. Similarly, scholars also try to rank vulnerability of lines by adopting the complex network theory. In literature (Fang et al., 2018), authors proposed a maximum flow-based complex network approach to identify the critical lines in a power network. Specifically, this algorithm was implemented to rank vulnerable lines in a western Danish power network. However, the biggest difference between power network and other types of networks lies in the existence of electrical information in former. Therefore, scholars seek to combine the complex network theory with the electrical information (Yang et al., 2020; Zhao et al., 2020) to identify critical nodes and achieve a lot of convincing research results. For instance, Zhou et al. (2019) developed an approach for analyzing nodes' vulnerability by outlining the electrical coupling of dynamic region. Wang, Lv et al. (2021) proposed a novel approach for evaluating the power network's robustness, in which critical nodes are firstly identified with the combination of topological structures and electrical characteristics. Liu et al. (2018) proposed an index of node electrical centrality to describe nodes' transfer ability. However, the reliability of identification still can be promoted in the complex power networks, by making better use of network's structural information, such as the group structure. In addition, as discussed in Section 1, network's group structure, i.e., all nodes' group membership, also influence the identification of critical nodes in a power system. Therefore, in this paper we try to introduce the clustering characteristic to the process of critical nodes identification.

Group structure is ubiquitous in networked systems. In general, we assume that the edges between nodes in common groups are densely distributed, while the edges between nodes in different groups are sparsely distributed (Radicchi et al., 2004). Research in this phenomenon has attracted a lot of attention from researchers in several disciplines, such as the computer science (call it cluster) (Priebe et al., 2019), social science (call it community) (Waniek et al., 2018) and biological science (call it protein complex) (Martinet et al., 2020). In the past two decades, a plethora of approaches were proposed aiming to fully uncover a network's group structure, such as Louvain (Blondel et al., 2008) and BigClam (Yang & Leskovec, 2013). Here we briefly provide some representative approaches and readers who are interested in this research topic can review related survey literatures (Fortunato, 2010; Javed et al., 2018). Heuristic method (Chakraborty et al., 2017) is a kind of classic but alive strategy to uncover network's group structure, whose goal is to optimize a specific objective function of group structure, such as the best-known modularity (Newman, 2004) and permanence (Chakraborty et al., 2014). Another interesting line of studies is based on formulating each node as a agent in a multiagent system under a specific protocol (Cao et al., 2019) or as a player in a game under a specific rule (Jonnalagadda & Kuppusamy, 2016). The goal of this kind of method is to find the final group structure by model the interactions between a target node and her "neighbors". The final stable state is regarded as the required group structure (Chen et al., 2010). For example, Cao et al. (2019) proposed a multiagent method to find prosumer-community groups in smart grid. Bu et al. (2020) developed a game theory based graph k-means to find group structure in social networks, where each node is regarded as a self-centered player and the interaction among them is formulated as a non-cooperative game. Recently, there has been an increasing interest in representative learning based approach (Cai et al., 2018) where each node is transformed into a dense low-dimensional Euclidean space. Although we can use off-the-shelf powerful machine learning tools to produce final group structure (Liu et al., 2020), many scholars from the field of electricity take a principled position that there is always a tradeoff among accuracy and interpretability, reliably identifying network's group structure is still be a challenging task. All in all, group structure exists in many real networks, and it is vital for scholars when their research object is a networked system. However, as far as we know, although there are some works aiming to uncover power network's group structure (Chen et al., 2015; Wei et al., 2020; Zhao et al., 2019), we do not find any work applying group information to the task of identifying critical nodes in a power network, while the electrical characteristics are also preserved. Thus our group-driven framework for critical nodes identification, i.e., *GDF-ICN*, suggests a promising perspective to understand the nodes and their organizing characteristics in complex power systems.

3. Preliminaries and problem definition

Let us now provide some definitions about basic complex network theory and power grid that are instrumental in the derivation of our framework, starting with the notions of network.

Definition 1 (*Network*)**.** We denote by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ a plain network, $\mathcal{V} = \{v_i\}_{i\in[1,n]}$ is a vertex set with n elements and $\mathcal{E} = \{e_{ij}\}_{\forall i,j\in[1,n]}$ is an edge set with $m = |\mathcal{E}|$ elements, $|\cdot|$ denotes the cardinality of a set.

Here we can customize a network by adopting a vertex-type mapping function $f_v : \mathcal{V} \to \mathcal{T}^v$ and an edge-type mapping function $f_e : \mathcal{E} \to \mathcal{T}^e$. $\mathcal{T}^v$ and $\mathcal{T}^e$ are the set of vertex type and edge type, respectively. Any vertex $v_i \in \mathcal{V}$ or edge $e_{ij} \in \mathcal{E}$ is associated with a specific type, i.e., $f_v(v_i) \in \mathcal{T}^v$, $f_e(e_{ij}) \in \mathcal{T}^e$.

Definition 2 (*Directed Heterogeneous Network, DHN*)**.** We denote by $\mathcal{G}_{dhn} = (\mathcal{V}, \mathcal{E})$ a directed heterogeneous network. Its all nodes belong to one single type, i.e., $|\mathcal{T}^v| = 1$; its edges have multiple types (different weight), i.e., $|\mathcal{T}^e| > 1, e_{ij} \neq e_{ji}, \exists e_{ij}, e_{ji} \in \mathcal{E}, i \neq j$.

As a kind of complex network, power grid is a typical system that can be naturally transformed into a *DHN*. In the early stage of complex network modeling for real power grid, it was often regarded as an undirected and unweighted topology model. In the model, generators, substations and loads were regarded as nodes; lines between nodes

Table 1
Correspondence between complex network and power system.

Complex network	Power system
Node	Generator, substation, load
Edge	High voltage, transformer branch, circuit
Weight of edge	Branch impedance
Direction of edge	Direction of current

were regarded as undifferentiated edges, double or multiple lines were combined into one edge. Although the conclusions obtained on this basis have some guiding significance, they are incapable of performing the quantitative calculation and analysis of the actual power grid. The fundamental reason is that the essential physical relations existing in the actual power grid are ignored in these models. Therefore, here we introduce the physical characteristics of a power grid into the complex network model and using a *DHN* to approximate a real power grid. The correspondence between power grid and complex network is offered in Table 1. Based on the corresponding relationship, we provide a definition of directed heterogeneous power network (*DHPN*) as follows.

Definition 3 (*Directed Heterogeneous Power Network, DHPN*)**.** Given a power grid, we transform it into a special DHN $\mathcal{G}_{dhn} = (\mathcal{V}, \mathcal{E})$, called DHPN $\mathcal{G}_{dhpn} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$, where $\mathcal{W}$ stores edges' weight information. Specifically, each node in DHPN represents a device (i.e., generator, substation and load) in a power grid; the impedance of an branch is regarded as the edge's weight; the direction of electric energy (that only can flow from the generation node through the transmission node to the load node) is regarded as edges' direction.

In the research field of complex network, meso-structure/ community/ group/ cluster attracts many attentions. As a kind of complex network, power network also shows the nodes clustering effect where the electrical equipments within a common administrative region geographically close, but the energy transfer among them is also frequent. Let $C = \{C_1, C_2, \cdots, C_k, \cdots, C_K\}$ denote a certain group structure with K groups, C_k is a set of nodes who reside in the kth group. Any network's group structure (community in social network), describing the clustering effect of nodes, helps us understand, predict and optimize the behavior of networked systems.

Definition 4 (*Power Group Structure*)**.** Given a DHPN $\mathcal{G}_{dhpn} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ and the number of groups K ($K \ll n$), the power group structure is denoted by a set of groups, i.e., $C = \{C_1, C_2, \cdots, C_K\}$, $\bigcup_{k=1}^{K} C_k = \mathcal{V}$, in which nodes within common groups are densely connected, while nodes distributed in different groups are sparsely connected. In particular, a power group represents a administrative region for a real power system, in which devices exchange power and information more frequently.

Traditionally, according to the number of groups to which each node belongs, groups can be categorized into two types: non-overlapping and overlapping. Non-overlapping groups are $\mathcal{V}$'s non-empty, mutually exclusive subsets, i.e., $\forall i, j \in \{1, 2, \ldots, K\}, i \neq j, C_i \cap C_j = \emptyset$; while in overlapping groups, each vertex can be assigned to more than one group, i.e., $\exists i, j \in \{1, 2, \ldots, K\}, i \neq j, C_i \cap C_j \neq \emptyset$. However, this classification has some limitations, it ignores the network connectivity. In fact, in a real networked system, the nodes undoubtedly belong to any group. Specifically, the probability of belonging to a small part of groups is large, while the probability of belonging to most of groups is small. We call this group structure as soft group structure denoted by nodes' membership matrix F in which F_{ik} (ith row, kth column) denotes node v_i's preference toward C_k. For consistency, we call the another group structure as hard group structure where any one of all nodes can only belong to a certain group or not belong to this group. Besides, we assume that the sum of the probabilities of any node belonging to each group is equal to 1, i.e., $\forall v_i \in \mathcal{V}, \sum_k F_{ik} = 1$. Here, let $|C_k| = \sum_i F_{ik}$ represent the size of C_k, i.e., the number of nodes reside in C_k.

Problem 1 (*Group-Driven Critical Nodes Identification in Power Networks, GDCNI*)**.** Given a power grid that can be transformed into a *DHPN* $\mathcal{G}_{dhpn} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$, the problem of identifying group-driven critical nodes aims to find the vulnerable nodes based on the network's topology and electrical information, while the potential group structure $C = \{C_1, C_2, \ldots, C_K\}$ is also preserved.

It is worth noting that, with the accomplishment of task of identifying group-driven critical nodes, the derived potential fuzzy groups also describes the meso-structure in a complex network with electrical information.

4. Framework setup

In this section, we will first introduce the framework and then detail it in following three subsections.

4.1. Main framework

Our group-driven critical nodes identification framework (*GDF-ICN*) is depicted in Fig. 2. *GDF-ICN* consists of three main modules: critical nodes identification, group expansion and membership update. In Fig. 2, rectangle with orange background and number 1 represents the structural information of a *DHPN*; rectangle with green background and number 2 represents the electrical information of a *DHPN*. In addition, the solid line represents *GDF-ICN*'s specific operation flow, and the dashed line represents the relevant information required by a module. As Fig. 2 shows, both structural information and electrical information of the *DHPN* are required in all three modules.

In a real power system, devices within a common administrative region are geographically close, and power and information exchange are also closer. Therefore, at the beginning of iteration, we can set the number of groups K according to the administrative areas of the power network and expert experience. However, due to the diversity of power network operation mode, the method based on the division of administrative areas and expert experience only consider the natural property of network and ignore its electric characteristics, is extremely unfavorable to safe operation of power networks. Thus in the algorithm iteration process, module 1 identifies critical nodes according to structural information, operating conditions and node group memberships of power systems, then module 2 and 3 redivides and optimizes the groups centering on the critical nodes. The three modules are carried out in a coupled iterative way. The introduction of nodes clustering effect improves reliability of identifying critical nodes, and the identification of critical nodes promotes the formation of denser group structure. Note that a key point of *GDF-ICN* is the construction of correspondence between network and power system as shown in Table 1. This makes *GDF-ICN* take the electrical information and topological information into consideration simultaneously. Note also that the group structure can be regarded as the topological/structural information because it is derived from the original topological information.

The pseudocode for *GDF-ICN*, is given in Algorithm 1. We take a directed heterogeneous power network $\mathcal{G}_{dhpn}$ as an input which is generated by a real power network. In addition, we also input the number of groups K and the maximum number of iterations $\mathcal{O}$. Output of the framework is the top K nodes with high value of $\mathcal{J}$, i.e., $\mathcal{I}$, and the nodes membership matrix F. After initialize F, we perform framework's three main parts in an iterative manner. The framework stops when the number of iterations is large than a threshold parameter, i.e., $\mathcal{O}$ which is set to 100 in our setting. Initially, module 1 is applied to compute each node's vulnerability (line 3 in Algorithm 1), i.e., $\mathcal{J}_i$. Meanwhile, the top K nodes with high value of $\mathcal{J}$ are selected as the

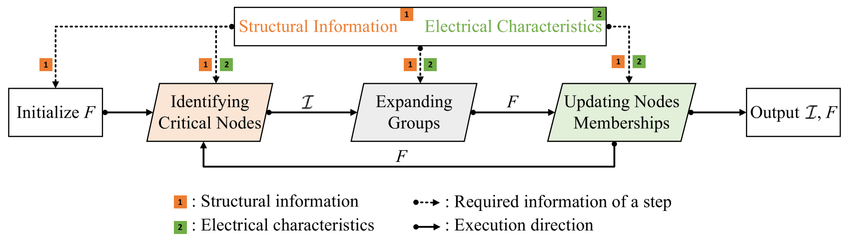


Fig. 2. Illustration of our *GDF-ICN* framework for identifying critical nodes in power networks.

critical nodes in each iteration. Then, group expansion module (line 6 in Algorithm 1) is applied to expand the "region" of these critical nodes allowing us to generate an initial nodes membership. Next, the membership update module (line 7 in Algorithm 1) is performed to promote each group's tightness. After $\mathcal{O}$ iterations, we obtain not only the critical nodes under the help of potential groups, but also the nodes membership in a complex network with specific electrical information.

Algorithm 1: *GDF-ICN*

Input: Directed heterogeneous power network $\mathcal{G}_{dhpn}$, number of groups K and the maximum number of iterations $\mathcal{O}$.
Output: Critical nodes set $\mathcal{I}$, membership matrix F.

```
1 Initialize F;
2 for (NumIter = 0; NumIter < O; NumIter++) do
3     I ← CNI(G_dhpn, F); // Identifying critical nodes.
4     F ← GE(G_dhpn, I); // Expanding groups.
5     F ← NMU(G_dhpn, K, F, O); // Updating membership.
6 end
7 return I, F;
```

4.2. Identifying critical nodes

Due to the existence of electrical coupling characteristics between nodes in the power system (varying degrees of electrical coupling correlation exist between grid nodes), when a node fails, it first affects the state of neighbor nodes, and then affects other nodes in the network according to the electrical coupling correlation path between nodes. It can be seen that the electrical coupling correlation between nodes is the key to identify critical nodes in a power system. Therefore, if a node has strong electrical coupling with its neighboring nodes, its fault will cause a large-scale chain failure. We thus regard these nodes that have strong coupling with its neighboring nodes as system's critical (vulnerable) nodes.

Here we first take a power system's structural information, electric parameters and the agglomeration effect into consideration, define the electrical coupling correlation among nodes. Then, we analyze a node's electrical coupling with its neighboring nodes by using the power system's present operation condition, to reflect the influence degree of this node on system operation stability when it fails. And we sort all nodes' electrical coupling in descending order, select top-K nodes as critical nodes.

Now we first define electrical distance that is an important index reflecting the electrical closeness between nodes.

Definition 5 (*Electrical Distance, ED*)**.** Electrical distance between nodes v_i and v_j is denoted by the equivalent impedance $Z_{ij,euq}$ between them, which is equal to the voltage U_{ij} between nodes v_i and v_j after injecting unit current I_i from vertex v_i, namely:

$$Z_{ij,euq} = \frac{U_{ij}}{I_i} = U_{ij}. \tag{1}$$

Also, according to the superposition principle, $Z_{ij,euq}$ can be expressed by the system node impedance matrix, as follows

$$Z_{ij,euq} = (Z_{ii} - Z_{ij}) - (Z_{ij} - Z_{jj}). \tag{2}$$

where Z_{ij} is an entry (ith row, jth column) of the system node impedance matrix. $Z_{ij,euq}$ expresses the degree of difficulty in changing the state of node v_j caused by the failure of node v_i. The larger the $Z_{ij,euq}$ is, the weaker the influence between different nodes will be, and vice versa (Poudel et al., 2018).

As a kind of complex network, power network has node agglomeration effect. In power systems, there are always some nodes that belong to two or more groups simultaneously. Many studies have shown that in a real-world network system, the more groups two nodes coexist, the stronger the correlation between the two nodes would be (Yang & Leskovec, 2013). Therefore, the electrical coupling correlation between nodes is not only related to the structural and operation characteristics of the power grid, but also related to its agglomeration phenomenon. In order to give consideration to the clustering effect of grid nodes, we define the electrical coupling correlation between nodes v_i and v_j as follows:

$$\mathcal{T}_{ij} = \frac{1}{Z_{ij,euq}} \times \frac{1}{1 + r(F_i, F_j)}. \tag{3}$$

where $r(\cdot,\cdot)$ is the repulsive force between two nodes' membership. F is all nodes' group membership matrix with size $n \times K$, n and K are the number of nodes and groups, respectively. $F_i = (F_{i1}, F_{i2}, \ldots, F_{ik}, \ldots, F_{iK})$, F_{ik} denotes the probability that node v_i belongs to the kth group C_k, also denotes node v_i's preference toward C_k. We introduce the node group membership matrix F into the calculation process of node vulnerability, so as to acquire both agglomeration effect and electrical characteristics. Notice that in Eq. (3) that the more similar the group membership vectors of the two nodes are, the greater the correlation degree will be. The smaller the electrical distance between two nodes, the greater the correlation degree.

According to the direction of power flow, when fault information is propagated in the network, a node can be affected or affect other nodes. For a target node, we divide its neighbors into two types: the neighbor source node and the neighbor end node. The neighbor source node supplies power to it, and the neighbor end node extracts power from it. Considering the nature of this, we define the electrical coupling Ω_i between node v_i and its neighbor nodes, to reflect v_i's ability to influence all its neighbor nodes' status, as follows:

$$\Omega_i = \sum_{j \in N_i} \mathcal{T}_{ij} = \sum_{p \in N_i^{in}} \mathcal{T}_{pi} + \sum_{q \in N_i^{out}} \mathcal{T}_{iq}, \tag{4}$$

where N_i^{in} and N_i^{out} are the sets of source nodes and end nodes for the target node v_i, respectively.

In the present operating condition of a power system, the interaction among nodes is directly related to the power transmission between them. The apparent power S_{ij} from node v_i to node v_j is calculated as

$$S_{ij} = \sqrt{P_{ij}^2 + Q_{ij}^2}, \tag{5}$$

where P_{ij} and Q_{ij} are the active power and reactive power from node v_i to node v_j, respectively. For a target node v_j, we define its influence on its neighbor source node v_i as follows:

$$w_{ij} = \frac{S_{ij} / \sum_{h \in N_j^{in}} S_{hj}}{\sum_{j \in N_i^{out}} (S_{ij} / \sum_{h \in N_j^{in}} S_{hj})} \cdot \Omega_j, \tag{6}$$

where the numerator denotes the ratio of the power absorbed by node v_j from node v_i to the total power absorbed by node v_j; the denominator represents the ratio of the output power of node v_i to node v_j to the total output power of node v_i. We perform normalization as follows:

$$p_{ij} = \frac{w_{ij}}{\sum_{i\in N_j^{in}} w_{ij}}, \tag{7}$$

Similarly, for the target node v_j, we define its influence on neighbor end node v_k as follows:

$$w_{jk} = \frac{S_{jk}}{\sum_{h\in N_k^{in}} S_{hk}} \cdot \Omega_j, \tag{8}$$

Normalizing w_{jk}, we obtain

$$p_{jk} = \frac{w_{jk}}{\sum_{k\in N_j^{out}} w_{jk}}, \tag{9}$$

Combining Eqs. (7) and (9), we redefine the electrical coupling $\mathcal{J}_i$ between nodes v_i and its neighbor nodes as

$$\mathcal{J}_i = \frac{1}{\alpha(1-\sum p_{ij}\ln p_{ij})\sum w_{ij} + (1-\alpha)(1-\sum p_{jk}\ln p_{jk})\sum w_{jk}}. \tag{10}$$

where $\alpha \in [0,1]$ is a parameter determining the degree of influence. When $\alpha > 0.5$, node v_i's state is more dependent on its end nodes. According to the value of $\mathcal{J}_i$, we sort them in a descending order, and select the top-K nodes with the biggest electrical coupling as the grid's critical nodes.

The pseudocode of Critical Nodes Identification (CNI) module is given in Algorithm 2.

Algorithm 2: *Critical Nodes Identification (CNI)*

Input: Directed heterogeneous power network $\mathcal{G}_{dhpn}$, membership matrix F.
Output: Critical node set $\mathcal{I}$.

Calculate the power flow of this power system, to obtain apparent power on each line;
for $v_i \in \mathcal{V}$ **do**
  for $v_j \in N_i$ **do**
    Calculate $\mathcal{T}_{ij}$ between nodes v_i and v_j using Eq. (3);
  end
  Calculate Ω_i using Eq. (4);
end
for $v_i \in \mathcal{V}$ **do**
  for $v_j \in N_i$ **do**
    if $v_j \in N_i^{in}$ **then**
      Calculate w_{ji} using Eq. (6);
      Normalize w_{ji} using Eq. (7);
    end
    if $v_j \in N_i^{out}$ **then**
      Calculate w_{ik} using Eq. (8);
      Normalize w_{ik} using Eq. (9);
    end
  end
  Calculate $\mathcal{J}_i$ using Eq. (10);
end
Sort $\mathcal{J}$ in descending order;
Select the top-K nodes with the biggest electrical coupling as critical nodes, and build critical nodes set $\mathcal{I}$;
return $\mathcal{I}$;

4.3. Expanding groups

The goal of group extension is to update the nodes membership matrix F. In each round of iteration, module 1 is used to determine the critical nodes based on the group structure in the last round of iteration. It is worth noting that the critical nodes identification task is accomplished when we finish the module 1. The goal of follow-up two modules, i.e., modules 2 and 3, is to extend and refine the group structure respectively for the next round of critical nodes identification task. Intuitively, a critical node in a power system locates at the "center" of a vulnerable region. Hence, we choose the critical node identified by module 1 as the centroid of each group, and initially classify the vulnerable group related to each critical node. Overall, we can perform module 2 in two steps:

Compared with the nodes adjacent to the topology structure, the electrical distance is more closely related to the nodes adjacent to it, while the electrical distance and the topological distance in the grid are not completely consistent. Therefore, for the rest unassigned nodes, we calculate the electrical distance to the critical nodes, and derive the preference toward all groups after which the value is normalized to guarantee the summation of each node's preference toward all groups equals 1.

Step 1: We choose K critical nodes as the centroids of K groups. Then, each centroid absorbs its neighbor nodes to the group it dominates. In particular, if a node v_i is the neighbor of c centroids simultaneously, then the probability that it belongs to any one of the c groups is equal to $1/c$, and the probabilities that it belongs to other groups are all 0. This ensures that the sum of the probabilities that any node belongs to all the groups always equals 1.

Step 2: For any remaining node v_i after Step 1, we first calculate the shortest electrical path between it and each group. Then the average shortest electrical distance from this node to any group k is calculated as the gravity of group k to node v_i, denoted by g_{iC_k}.

$$g_{iC_k} = \ln \frac{\sum_{j\in C_k} d_{ij} + |C_k|}{\sum_{j\in C_k} d_{ij}}, \tag{11}$$

where d_{ij} denotes the number of edges in the shortest electrical path between nodes v_i and v_j. The shortest electrical path denotes the path with the shortest electrical distance from nodes v_i to v_j. Compared to the shortest path, the shortest electrical path can offer reference for power flow distribution, makes power flow distribute through the path with the least loss. Then, we can derive the probability that node v_i belongs to group k as follows:

$$F_{ik} = g_{iC_k} / \sum_k g_{iC_k}. \tag{12}$$

Here we provide a toy example in Fig. 3 to illustrate the process of group expansion. We construct a small network with 9 nodes and 2 groups allowing us to more clearly describe how groups expand from K centroids. For simplicity, we set the weight of all edges as 1. At the beginning, as shown in Fig. 3(1), each node is denoted by a white circle, i.e., they all belong to no groups. Moreover, we associate each node with a value pair to represent the probabilities that node belongs to the two groups. Next, based on the calculation of $\mathcal{J}$ in module 1, we identify two critical nodes (node 1 and node 6), as Fig. 3(2) shows. Specifically, we can observe that these two nodes are highlighted by two different colors (blue and orange), and the probabilities that each node belongs to two groups also change. Then, we perform Step 1 of module 2, as shown in Fig. 3(3). We can see that node 1 and node 6 absorb their neighbor nodes to their groups, respectively. In particular, we highlight these nodes using corresponding color. We also note that, node 3 is highlighted by a type of different color because she is the neighbor of both node 1 and node 6. Meanwhile, the probability of these nodes that belong to two groups changes. After Step 1 of module 2, only node 5 and node 8 are unassigned. We now perform Step 2 to assign them to two groups. Details are shown in Table 2. It is worth noting that we have to calculate the electrical distance between each unassigned node and assigned nodes in corresponding groups. After we finish the calculation of probability that unassigned nodes belong to each groups (see Table 2), we complete the module 2, i.e., group extension, as shown in Fig. 3(4).

The pseudocode of module 2 is shown in Algorithm 3. This module is completed in a fast manner. The computation cost is mainly on the step 2 for the assignment of remaining nodes.

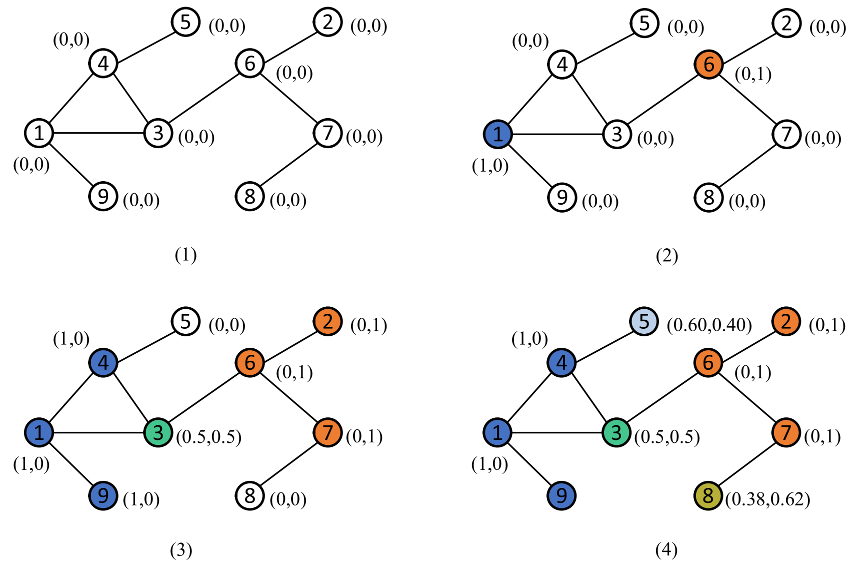


Fig. 3. Toy example of group expansion process.

Table 2
Membership analysis of node-5 and node-8 in the expansion process of two groups.

NodeId	GroupId	NPair	Distance	Gravity	Probability
node-5	group-1	(5,1)	2	ln(3/2)	ln(3/2)/ln(51/26) ≈ 0.6018
		(5,3)	2		
		(5,4)	1		
		(5,9)	3		
	group-2	(5,2)	4	ln(17/13)	ln(17/13)/ln(51/26) ≈ 0.3982
		(5,3)	2		
		(5,6)	3		
		(5,7)	4		
node-8	group-1	(8,1)	4	ln(5/4)	ln(5/4)/ln(65/36) ≈ 0.3776
		(8,3)	3		
		(8,4)	4		
		(8,9)	5		
	group-2	(8,2)	3	ln(13/9)	ln(13/9)/ln(65/36) ≈ 0.6224
		(8,3)	3		
		(8,6)	2		
		(8,7)	1		

Algorithm 3: *Group Expansion (GE)*

Input: Directed heterogeneous power network $\mathcal{G}_{dhpn}$, critical nodes set $\mathcal{I}$.
Output: Membership matrix F.
Select K critical nodes as the centroids of K groups;
Assign all critical nodes' neighbors to corresponding groups and update F;
Calculate the attraction of each group to the remaining nodes using Eq. (11);
Calculate the membership of the remaining nodes using Eq. (12) and update F;
return F;

4.4. Updating memberships

In the last subsection (module 2), we assign nodes to K groups according to the neighborhood relationship and electrical distance to the K centroid nodes. In this section, we proceed by promoting groups' agglomeration phenomenon by optimizing a novel metric, called *Fuzzy Tightness*. *Tightness* (Bu et al., 2017) is a classic metric can be used to evaluate the significance of group structure in a complex network.

The tightness of a group in a given *DHPN* is approximately derived by the difference between the number of intra-edges and inter-edges, as follows

$$T(C_k) = \frac{2M_k^{intra}}{|C_k|^2} - \frac{M_k^{inter}}{|C_k|(n-|C_k|)}, \tag{13}$$

where M_k^{intra} represents the number of edges that two end-nodes are all in group C_k; M_k^{inter} represents the number of edges that one end-node is in group C_k and the another one node is in other groups except C_k. n is the number of nodes, $|C_k|$ is the number of nodes belong to kth group C_k.

In particular, to alleviate the impact derived by too small and too large groups, we can define *Adjusted Tightness*, as follows:

$$\begin{aligned} T_A(C_k) &= |C_k|(n-|C_k|)\left(\frac{2M_k^{intra}}{|C_k|^2} - \frac{M_k^{inter}}{|C_k|(n-|C_k|)}\right) \\ &= \frac{2(n-|C_k|)}{|C_k|}M_k^{intra} - M_k^{inter}. \end{aligned} \tag{14}$$

Compared to *Tightness*, *Adjusted Tightness* penalizes too small and too large groups and produces better solutions (well evaluates the significance of a *DHPN*'s group structure) as an objective function. However, due to the nature of network's connectivity, each node must belong to any one of all groups. In other words, evaluating group structure based on hard memberships is not the best choice. Thus, here we define $\mathcal{T}_{C_k}^{fe}$—the *fuzzy electrical tightness* of a group C_k using nodes group membership matrix F (with size of $n \times K$) as follows:

$$\mathcal{T}_{C_k}^{fe}(F) = \frac{2(n-|C_k|)}{|C_k|}m_k^{intra} - m_k^{inter}, \tag{15}$$

where m_k^{intra} denotes the fuzzy number of edges that two end-nodes are all in C_k, as

$$m_k^{intra} = \sum_i \sum_{j \in N_i} F_{ik}F_{jk}, \tag{16}$$

and, m_k^{inter} denotes the fuzzy number of edges that one end-node is in C_k and another one end-node is in a different groups except C_k, as

$$m_k^{inter} = \sum_i \sum_{j \in N_i \wedge (\mathcal{N} \setminus C_k)} \sum_{\hat{k} \in C \setminus C_k} F_{ik}F_{j\hat{k}}. \tag{17}$$

where $\mathcal{T}_{ij}$ is the electrical coupling correlation degree from node v_i to v_j, as defined in Eq. (3). Obviously, using Eq. (15) we can obtain a tightness measure over the whole group structure, given by

$$\mathcal{T}_C^{fe}(F) = \sum_k \frac{2(n-|C_k|)}{|C_k|}m_k^{intra} - m_k^{inter}. \tag{18}$$

Like the modularity optimization methods (Blondel et al., 2008), we can optimize this metric function to gain a better group structure with higher fuzzy tightness score. But, this equation is hard to solve. Now, for any one node v_i, her contribution to the fuzzy tightness score can be derived as

$$\mathcal{T}_i^{fe}(F) = \frac{2(n-|C_k|)}{|C_k|}\sum_{j\in N_i} F_{ik}F_{jk} - \sum_{j\in N_i \wedge(\mathcal{N}\setminus C_k)}\sum_{\hat{k}\in C\setminus C_k} F_{ik}F_{j\hat{k}}, \tag{19}$$

where N_i is a set of v_i's out-neighbor nodes. Thus, for the node v_i, we adopt the gradient ascent method (Yang & Leskovec, 2013) to update her group membership, and the gradient of F_i of v_i at kth group is presented as

$$\frac{\partial \mathcal{T}_i^{fe}(F)}{\partial F_{ik}} = \frac{2(n-|C_k|)}{|C_k|}\sum_{j\in N_i} F_{jk} - \sum_{j\in N_i \wedge(\mathcal{N}\setminus C_k)}\sum_{\hat{k}\in C\setminus C_k} F_{j\hat{k}}, \tag{20}$$

Then the update process of group membership of vertex v_i toward C_k is given as

$$F_{ik}^{new} = \max(F_{ik}^{old} + \beta\frac{\partial \mathcal{T}_i^{fe}(F)}{\partial F_{ik}}, 0). \tag{21}$$

where β is a step size.

This module's pseudocode is given in Algorithm 4. It must be noted that we can run this module in a parallel manner. Specifically, we can employ multi-threads to compute F's gradient of all nodes. This module is also performed in an iterative way and terminates when the change of value of F_{ik} for any node v_i at group k is less than 0.0001 or the number of iterations exceeds the predefined maximum number of iterations $\mathcal{O}$.

Algorithm 4: *Nodes Membership Update (NMU)*

Input: Directed heterogeneous power network $\mathcal{G}_{dhpn}$, number of groups K, nodes' group membership matrix F and the maximum number of iterations $\mathcal{O}$.
Output: Membership matrix F.

$flag \leftarrow 0$, $Iter \leftarrow 0$;
while ($flag == 0 \parallel Iter > \mathcal{O}$) **do**
  $flag \leftarrow 1$;
  // **parallel computing**
  for $v_i \in \mathcal{V}$ **do**
    for ($k = 1; k <= K; ++k$) **do**
      Calculate $\frac{\partial \mathcal{T}_i^{fe}(F)}{\partial F_{ik}}$ using Eq. (20);
      Calculate new F using Eq. (21);
      $F_{ik}^{old} \leftarrow F_{ik}^{new}$;
      if $F_{ik}^{new} - F_{ik}^{old} > 0.0001$ **then**
        $flag \leftarrow 0$;
      end
    end
  end
  $Iter \leftarrow Iter + 1$;
end

5. Model analysis and discussion

We proceed by analyzing and discussing some interesting as well as valuable details about our critical nodes identification framework *GDF-ICN*.

5.1. Initialization

At the beginning of our framework, in addition to transform a given power grid into a *DHPN*, we have to initialize the nodes group membership matrix F, as well as the number of groups K. There are many strategies can be used to initialize the initial group structure, i.e., F. However, many authors favor regarding neighborhood based groups as the initial nodes membership. The reason behind it is that the neighborhood groups can give us a simple and powerful heuristic for speeding up whole nodes partitioning approaches (Gleich & Seshadhri, 2012). For example in Yang and Leskovec (2013), Jaewon and Jure proposed a fast NMF based method for uncovering overlapping social communities in scale, where initially they sort all nodes by their conductance value in ascending order, then each node is regarded as a centroid and orderly absorbs their neighbors to the group they reside. Specifically, each assigned node (centroids or her neighbors) will lose the chance to be a centroid. The assignment terminates if the number of groups exceeds the predefined number of groups K or all nodes are assigned. Moreover, if there exist unassigned nodes when the number of groups exceeds K, they perform a random strategy to initialize the remaining nodes' membership.

Similarly, Cao et al. (2019) also provided a heuristic method based on the "centroid" to initialize prosumer-community groups in smart grids. The difference between Cao's method and Jaewon's method for initialization is in three folds: (1) Application field. Jaewon aims to uncover community structure in large scale social networks, while Cao aims to detect prosumer-community groups in smart grids; (2) Order of nodes. Jaewon assumes that nodes with small conductance are more important, while Cao assumes that the nodes with high energy density (ED) are more important; (3) Assignment of remaining nodes. Jaewon randomly initialize these nodes' community membership, while Cao assumes that these nodes belongs to no community initially. Specifically, based on the discussion above, we regard the degree (sum of in degree and out degree) as the importance of a node in our framework. In contrast to Cao's assumption, we believe that the assignment of remaining nodes is also important. Here, we assign each remaining node to the group where she has shortest path to this group's centroid.

5.2. Repulsive force

Repulsive force function in Eq. (3) measures the mutual exclusivity of two nodes regarding to their group memberships. For any two nodes, if both of them reside in several common groups, they will have strong relationship. In contrast, if they are in different groups (for example, two users in a social platform have no common interests), we think that they have weak relationship. Here we provide the definition of repulsive force as follows.

Definition 6 (*Repulsive Force*)**.** We denote by $f(\cdot) : \mathbb{R}^K \to \mathbb{R}_0^+$ a differentiable continuously non-negative convex function, and the repulsive force between nodes v_i and v_j is derived by the Bregman divergence as follows

$$r(v_i, v_j) = \varpi(F_i, F_j) = f(F_i) - f(F_j) - (F_i - F_j)'\nabla f(F_j). \tag{22}$$

where $r(\cdot,\cdot) : \mathbb{R}^K \times \mathbb{R}^K \to \mathbb{R}_0^+$ is non-negative, continuous and differentiable, which satisfies $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^K$: (1) $r(\mathbf{x}, \mathbf{y}) \geq 0$; (2) $r(\mathbf{x}, \mathbf{x}) = 0$; and (3) $r_{y_q}(\mathbf{x}, \mathbf{y})$ is continuous and differentiable on $\mathbf{x}_p$, $p, q \in [1, K]$.

Based on the definition of repulsive force above, we can assign $f(\cdot)$ different convex properties, so as to generate different types of repulsive force. In particular, if we set $f^1(\mathbf{x}) = \|\mathbf{x}\|^2$, $\mathbf{x} \in \mathbb{R}^K$, we can derive first repulsive force $r_f^1(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2$; if we set $f^2(\mathbf{x}) = \sum_{i=1}^K x_i \log_2 x_j$, $\mathbf{x} \in \Delta$, $\Delta = \{\mathbf{x} \in \mathbb{R}^K : \sum_{k=1}^K x_i = 1, \mathbf{x} \geq 0\}$ is the standard K-simplex of $\mathbb{R}^K$, we can derive second repulsive force $r_f^2(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^K x_k \log_2(x_k/y_k)$; and if we set $f^3(\mathbf{x}) = \|\mathbf{x}\|$, $\mathbf{x} \in \mathbb{R}^K$, we can derive third repulsive force $r_f^3(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\| - \frac{\mathbf{x}\otimes\mathbf{y}}{\|\mathbf{y}\|}$, where $\otimes$ is an inner product operator. In the following experiments section, we will test our framework under different choices of repulsive force.

5.3. Fuzzy membership

In Section 4.4, we define a fuzzy tightness metric $\mathcal{T}_C^{fe}(F)$ to measure *DHPN*s' group structure. $\mathcal{T}_C^{fe}(F)$ can measure not only network's fuzzy group structure and also network's hard group structure.

Moreover, electrical information is also introduced to this metric where we regard the electrical quantity $\mathcal{T}_{ij}$ as the weight of edge from node v_i to node v_j. Regarding the $\mathcal{T}_C^{fe}(F)$, we ponder whether we can transform the other traditional metrics of evaluating hard group structure into fuzzy metrics of depicting network's fuzzy group structure. Motivated by this, here we take classic modularity, its alternative and variants as example, transform them into a new metric for evaluating fuzzy group structure. For simplicity, we will not introduce the electrical information, i.e., $\mathcal{T}_{ij}$, to these metrics.

- **Modularity** (Newman, 2004):

$$Q = \frac{1}{2m}\sum_{i,j}\left[A_{ij} - \frac{|N_i \parallel N_j|}{2m}\right]\delta_{i,j}, \tag{23}$$

where $\delta_{i,j}$ is the Kronecker delta function which equals 1 if v_i and v_j are in a common group, 0 otherwise.

- **Fuzzy Modularity:**

$$Q^f = \frac{1}{2m}\sum_{i,j}\left[A_{ij} - \frac{|N_i \parallel N_j|}{2m}\right]\sum_k F_{ik}F_{jk}. \tag{24}$$

Notice that the key of implementing transformation is replacing $\delta_{i,j} \in \{0,1\}$ by $\sum_k F_{ik}F_{jk} \in \mathbb{R}^+$. In particular, $\sum_k F_{ik}F_{jk}$ here denotes the number of groups of v_i and v_j reside simultaneously. We can also adopt different strategies to replace this value, such as $\sum_k (F_{ik}+F_{jk})/2$, $\sum_k \max(F_{ik},F_{jk})$ and $\sum_k \min(F_{ik},F_{jk})$. Now let us look at a variant of Modularity—Similarity-based Modularity.

- **Similarity-based Modularity (group C_k)** (Feng et al., 2007):

$$SQ_{C_k} = \frac{\sum_i\sum_j s(i,j)\varphi_{ik}\varphi_{jk}}{\sum_i\sum_j s(i,j)} - \left(\frac{\sum_i\sum_j s(i,j)\varphi_{ik}}{\sum_i\sum_j s(i,j)}\right)^2, \tag{25}$$

where SQ_{C_k} is the contribution of kth group C_k to the whole modularity. Thus we can easily use $\sum_k SQ_{C_k}$ to calculate the whole modularity. $\varphi_{ik} = 1$ if $v_i \in C_k$, 0 otherwise. $s(i,j)$ denotes the topological similarity between nodes v_i and v_j. There are many similarity metrics (Wang, Bu et al., 2021) can be used to calculate the similarity-based modularity. For example, the Jaccard coefficient, Salton index, Sorensen index, Hub promoted index and Hub depressed index (Wang, Bu et al., 2021).

- **Fuzzy Similarity-based Modularity (group C_k):**

$$SQ_{C_k} = \frac{\sum_i\sum_j s(i,j)F_{ik}F_{jk}}{\sum_i\sum_j s(i,j)} - \left(\frac{\sum_i\sum_j s(i,j)F_{ik}}{\sum_i\sum_j s(i,j)}\right)^2. \tag{26}$$

Note that here we finish the transformation by replacing $\varphi_{ik} \in \{0,1\}$ with $F_{ik} \in (0,1\})$. There are many variants of Modularity proposed for evaluating groups with hard membership can be simply transformed by the replacement, such as the Dist-modularity, Diffusion-based modularity, Influence-based modularity and Motif modularity (Chakraborty et al., 2017).

Under these two situations above, we successfully soften some metrics of evaluating hard group structure by some simple replacement operations. Here let us turn to an alternative of Modularity.

- **Modularity (group C_k)** (Newman, 2004):

$$\hat{Q}_{C_k} = \frac{M_k^{intra}}{m} - \left(\frac{(M_k^{intra}+M_k^{inter})}{2m}\right)^2, \tag{27}$$

where $\hat{Q}_{C_k}$ is the contribution of kth group C_k to the whole modularity. M_k^{intra} and M_k^{inter} the number of intra-edges and inter-edges of C_k, respectively, which can be counted directly in a network with hard group structure.

- **Fuzzy modularity (group C_k):**

$$\hat{Q}^f_{C_k} = \frac{m_k^{intra}}{m} - \left(\frac{(m_k^{intra}+m_k^{inter})}{2m}\right)^2, \tag{28}$$

Table 3
Correspondence of notations between metrics of evaluating hard group structure and fuzzy group structure. Meaning of notations can be found in the corresponding equations.

Position	Hard	Domain	Fuzzy	Domain
Eqs. (25)–(26)	φ_{ik}	{0,1}	F_{ik}	(0,1]
Eqs. (23)–(24)	$\delta_{i,j}$	{0,1}	$\sum_k F_{ik}F_{jk}$	$\mathbb{R}^+$
Eqs. (28)–(29)	M_k^{intra}	$\mathbb{N}^+$	$\sum_i\sum_{j\in N_i} F_{ik}F_{jk}$	$\mathbb{R}^+$
Eqs. (28)–(29)	M_k^{inter}	$\mathbb{N}^+$	$\sum_i\sum_{j\in N_i\wedge(\mathcal{N}\setminus C_k)}\sum_{\hat{k}\in C\setminus C_k} F_{ik}F_{j\hat{k}}$	$\mathbb{R}^+$

Table 4
Power network statistics. n: number of nodes, m: number of edges, d_{avg}: average degree, d: network diameter, c: global clustering coefficient, q: network density, I_e: existence of required electrical information.

Network	n	m	d_{avg}	d	c	q	I_e
IEEE 39-Bus	39	46	2.359	10	0.038	0.062	✓
IEEE 118-Bus	118	177	2.983	16	0.186	0.025	✓
Eris-1176	1176	9864	14.776	53	0.432	0.013	✗
Western-US-09	1723	4117	2.779	37	0.076	0.002	✗
US-Grid	4941	6594	2.669	46	0.080	0.001	✗
Western-US-10	5300	13 571	3.121	49	0.088	0.001	✗

where m_k^{intra} and m_k^{inter} are the fuzzy number of intra-edges and inter-edges of C_k, respectively, given as

$$\begin{cases} m_k^{intra} = \sum_i \sum_{j\in N_i} F_{ik}F_{jk}, \\ m_k^{inter} = \sum_i \sum_{j\in N_i\wedge(\mathcal{N}\setminus C_k)} \sum_{\hat{k}\in C\setminus C_k} F_{ik}F_{j\hat{k}}. \end{cases} \tag{29}$$

Note that this transformation becomes complex, and the derived fuzzy modularity function is similar to our fuzzy electrical tightness metric $\mathcal{T}_{C_k}^{fe}$, this is because they are both defined by approximately calculating the difference between the number of intra-edges and inter-edges. Now we summarize the corresponding relationship of notations between metrics for evaluating group structure with hard and fuzzy membership in Table 3. It is also worth noting that Table 3 just provide a list of notations used for our discussion in this paper. However, the analysis provides a preliminary tool on softening the metrics of evaluating hard group structure. It motivates us and can also help researchers use a systematic way to study network's fuzzy group structure.

6. Experiments

In this section, we aim to carry out comprehensive experiments for verifying *GDF-ICN*'s performance from two perspectives on several benchmark power networks. Since *GDF-ICN*'s core contribution is the introduction of group structure that is influenced by the setting of different repulsive power functions in module 3, we initially test the effect of different repulsive power function on the performance of group structure's characterization. Details are shown in Section 6.2. Then, to verify *GDF-ICN*'s performance in terms of critical nodes identification, i.e., our core task in this paper, we further perform a series of tests on two power networks with electrical information. Section 6.3 shows the details. Now we first provide the experiment setup.

6.1. Experiment setup

6.1.1. Datasets

We use six benchmark power networks for evaluation, and summarize their statistics in Table 4. As shown in Table 4, most networks are with sparse links except Eris-1176. In addition, compared to Western-US-09, Western-US-10 has three times the number of nodes and edges, but have similar other statistics.

These networks are equipped with different topological or electrical configurations. To test the contribution to the whole framework of module 3, four larger networks (Eris-1176, Western-US-09, US-Grid and Western-US-10) are adopted to verify the effectiveness of proposed fuzzy tightness (i.e., module 3).[1] In particular, we adopt two classical benchmark networks (IEEE 39-Bus and IEEE 118-Bus systems) for testing performance on critical nodes identification, to verify overall performance of $GDF-ICN$ framework. The MATLAB MATPOWER Toolbox is used to calculate the electrical information in the two test systems.

6.1.2. Baseline methods

To test the performance of fuzzy tightness on group detection, we compare it with three baseline methods, i.e., Louvain (Blondel et al., 2008), Bigclam (Yang & Leskovec, 2013) and Node2vec (Grover & Leskovec, 2016). Louvain (Blondel et al., 2008) is a classical method for detecting communities based on modularity optimization. Bigclam (Yang & Leskovec, 2013) is a NMF-based method proposed by Jaewon and Jure that can process large scale social networks. Node2vec (Grover & Leskovec, 2016) is a node representation learning based method which adjusts the strategy of random walk to achieve a higher performance on several tasks, such as nodes classification, link prediction and visualization.

To evaluate the results of *GDF-ICN* on critical nodes identification, we adopt six baseline methods as follows: Li et al. (2019) developed an identification method of key nodes in power system based on an improved PageRank algorithm. Specifically, they use Jacob matrix to obtain sensitivity matrix of voltage reactive power and the sensitivity matrix of phase angle active power, then define the link matrices under different indices of system. Zhu et al. (2016) proposed a method for identifying vital nodes in power grid based on an importance evaluation matrix. Zhu et al. (2018) constructed an original power Google matrix based on the topology information and direction of current, and proposed a modified PageRank algorithm to detect critical nodes in power networks. Yang et al. (2020) proposed an ECCI index to rank nodes by using electric cactus structure. Bompard et al. (2011) incorporated several specific features of power systems such as electrical distance, power transfer distribution factors and line flow limits, and developed a topological approach for identifying critical nodes. Wei (2020) proposed an evaluation method of key nodes of power grid based on the intermediate weight entropy. For the sake of fairness, we compare our results calculated by *GDF-ICN* with the original results shown in these literatures. In particular, we perform and compare *GDF-ICN* and the first three baseline methods in IEEE 39-Bus; while perform and compare *GDF-ICN* and the last three baseline methods in IEEE 118-Bus.

6.1.3. Evaluation metrics

We have to evaluate methods' performance on tasks of critical nodes identification and group detection. However, because there is no ground-truth information of group structure, we naturally adopt three representative metrics for evaluating methods' performance on group detection, i.e., Conductance (Gleich & Seshadhri, 2012), Average of association (Jianbo Shi & Malik, 2000) and Permanence (Chakraborty et al., 2014) as follows:

$$Cond(C) = \sum_{k=1}^{|C|} \frac{cut(C_k)}{\min(vol(C_k), vol(\overline{C_k}))}, \tag{30}$$

$$Asso(C) = \sum_{k=1}^{|C|} \frac{edges(C_k)}{|C_k|}, \tag{31}$$

$$Perm(C) = \sum_{i=1}^{|\mathcal{V}|} [\frac{I(v_i)}{E_{max}(v_i)D(v_i)}] - [1 - c_{in}(v_i)]. \tag{32}$$

[1] http://networkrepository.com/.

where $\overline{C_k} = \mathcal{V}\backslash C_k$ is the complement set of kth group C_k, $vol(C_k)$ denotes the sum of degrees of nodes in C_k. $cut(C_k) = vol(C_k) - edges(C_k)$, $edges(C_k)$ is twice the number of links among nodes in C_k. In Eq. (32), for a node v_i, $I(v_i)$ denotes its internal connections, $D(v_i)$ is its total degree, $E_{max}(v_i)$ is the maximum number of links to each other group and $c_{in}(v_i)$ denotes the internal clustering coefficient.

6.2. The effectiveness of module 3

We begin by singly taking our module 3, i.e., fuzzy tightness optimization, to detect groups on four large scale benchmark power networks. Firstly, we examine how the different choices of repulsive force functions affect the performance of module 3 on group detection. The results are shown in Fig. 4. The x axis denotes the number of iterations in optimizing Eq. (18), whereas the y axis denotes the normalized value of fuzzy tightness. The tight group structure is more significant when the value of fuzzy tightness is higher. It is graphically obvious that the value of fuzzy tightness increases with the increase of the number of iterations. As we can see in Fig. 4, the normalized value of fuzzy tightness increases initially with a big slope, however, the rate of increase gradually slows down over the number of iterations. In particular, with the increase of the number of iterations from 1 to 10, the dynamics of module 3 under these three repulsive power functions are similar, then fuzzy tightness with type 2 increases significantly slower than the other two types of functions. In addition, when we adopt third repulsive power function (Type-3), we can see that the stable group structure forms in a faster speed in the reminding three power networks except Western-US-09. However, stable group structure forms fastest speed when the second repulsive power function (Type-2) is adopted in Western-US-09. Based on the above analysis on Fig. 4, the third type of repulsive force function is more better to be used to find stable group structure at a faster speed. Thus, we will fix the repulsive power function as type-3, i.e., $r_f^3(\mathbf{x},\mathbf{y}) = \|\mathbf{x}\| - \frac{\mathbf{x}\otimes\mathbf{y}}{\|\mathbf{y}\|}$, in the follow-up experiments.

Based on four benchmark power networks, we proceed by evaluating the performance of module 3 and three baseline approaches on detecting groups w.r.t three representative metrics. Details are shown in Fig. 5. In network Eris-1176 (see Fig. 5(a)), it can be observed that our module 3 obtains better performance in terms of Conductance and Permanence. We find our module 3 is also second in another metric, i.e., Association, only to Louvain. In addition, counterintuitive as it might seem, Node2vec displays the worst performance on this network w.r.t all metrics compared to the other three approaches. As shown in network Western-US-09 (see Fig. 5(b)), our module 3 can also obtain the best performance w.r.t Conductance; and the second in Association and Permanence, to Louvain and Node2vec, respectively. In network US-Grid (see Fig. 5(c)), we can see that our module 3 outperforms the other three approaches w.r.t Conductance and Association. In network Western-US-10 (see Fig. 5(d)), our module 3 also can obtain comparable performance w.r.t all three metrics, compared to other approaches. Overall, our module 3 has obvious advantage on the task of group detection compared to the other three baseline methods w.r.t three metrics.

To more clearly demonstrate the advantages of our module 3, we accumulate the normalized performance of each method in terms of all three metrics separately on any data set. Fig. 6 shows the results. It can be seen that our module 3 outperforms all the other three baseline methods w.r.t the sum of normalized three metrics. It is also worth noting that Bigclam also behaves well in network Western-US-10 because it is developed to detect overlapping groups in networks at scale.

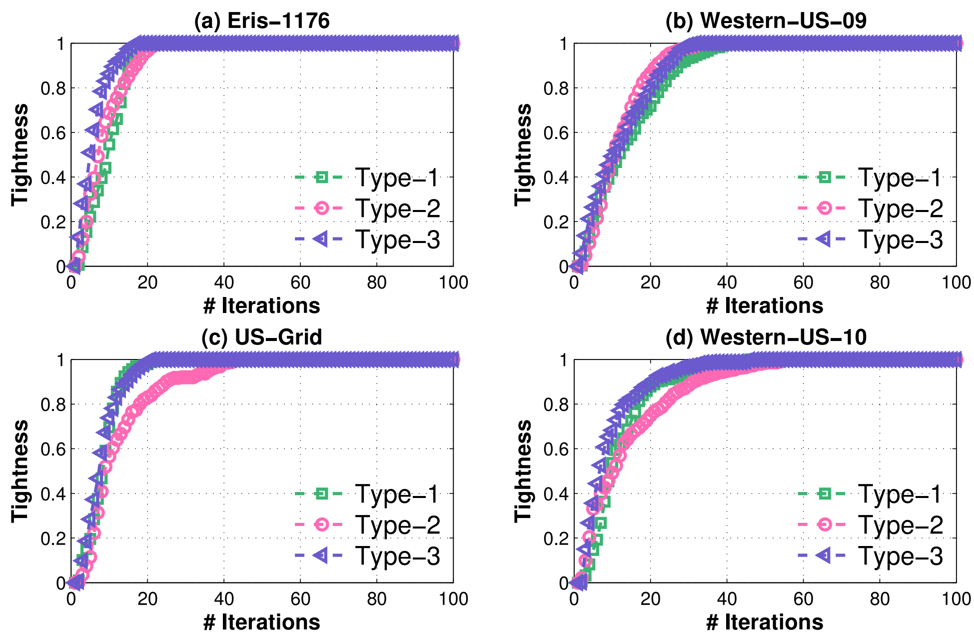


Fig. 4. Dynamics of fuzzy tightness optimization on four benchmark power networks under three different repulsive power functions (Type-1: $r_f^1(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^2$, Type-2: $r_f^2(\mathbf{x},\mathbf{y}) = \sum_{k=1}^{K} x_k \log_2(x_k/y_k)$, Type-3: $r_f^3(\mathbf{x},\mathbf{y}) = \|\mathbf{x}\| - \frac{\mathbf{x}\otimes\mathbf{y}}{\|\mathbf{y}\|}$).

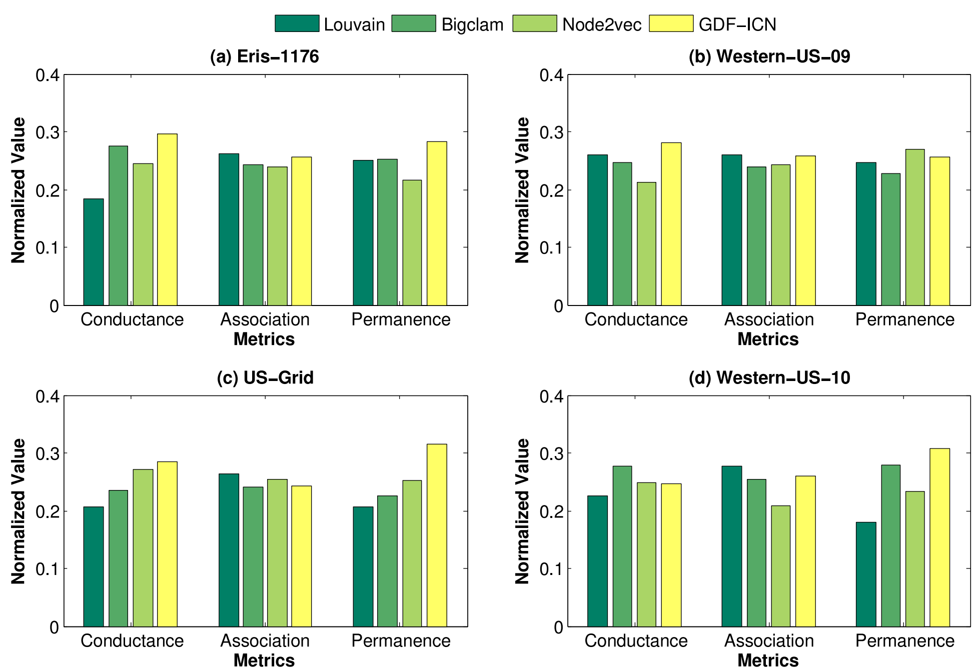


Fig. 5. Performance evaluation of different methods for group detection on four benchmark power networks w.r.t three representative metrics (Conductance (Gleich & Seshadhri, 2012), Association (Jianbo Shi & Malik, 2000) and Permanence (Chakraborty et al., 2014)).

6.3. Testing GDF-ICN

Now we evaluate the performance of our *GDF-ICN* and six baseline methods on the task of critical nodes identification. We first perform comparisons on network IEEE-39-Bus, and details are summarized in Table 5.

As shown in Table 5, eight nodes with top high value of index is displayed for each methods. Compared to the critical nodes selected by literature (Li et al., 2019), which is a PageRank based method, our *GDF-ICN* selects three nodes with same ID, i.e., node-16, node-22 and node-19. Compared to the critical nodes selected by literature (Zhu et al., 2016), our *GDF-ICN* selects four nodes with same ID, i.e., node-16, node-4, node-26 and node-3. Compared to the critical nodes selected by literature (Zhu et al., 2018), our *GDF-ICN* selects three nodes with same ID, i.e., node-6, node-16 and node-4. We can see that node-16 is always ranked as one of the top three nodes, this is because it is located in the power transmission channel of the generators 33–36. In addition, as a heavy load node, node-16 undertakes important power transmission and load tasks. Aiming to analysis the selected critical nodes by *GDF-ICN* in a more comprehensive manner, we now provide more details shown in Fig. 7 viewed from different perspective.

Fig. 7(a) denotes the final stable value of fuzzy tightness under different number of groups in IEEE-39-Bus. The x axis denotes the

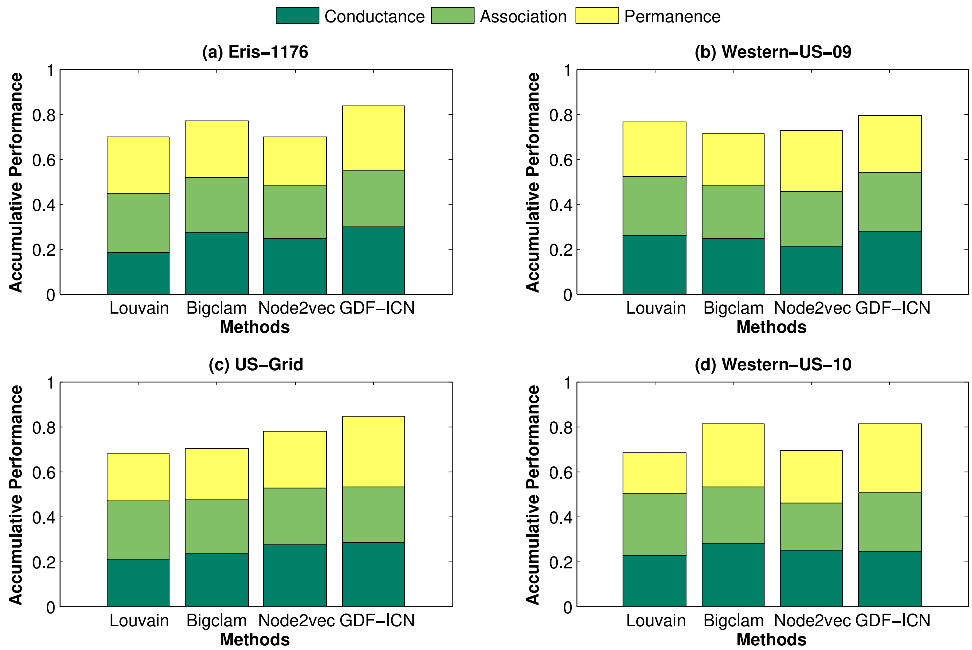


Fig. 6. Accumulative performance evaluation of different methods for group detection on four benchmark power networks w.r.t the sum of normalized three representative metrics (Conductance (Gleich & Seshadhri, 2012), Association (Jianbo Shi & Malik, 2000) and Permanence (Chakraborty et al., 2014)).

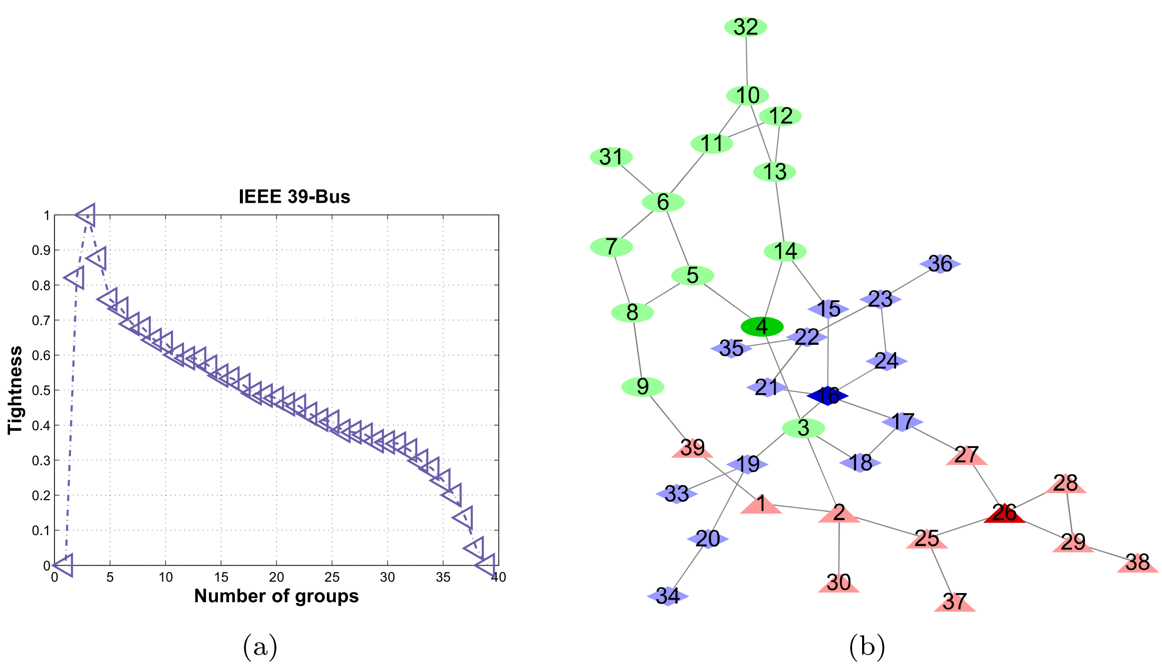
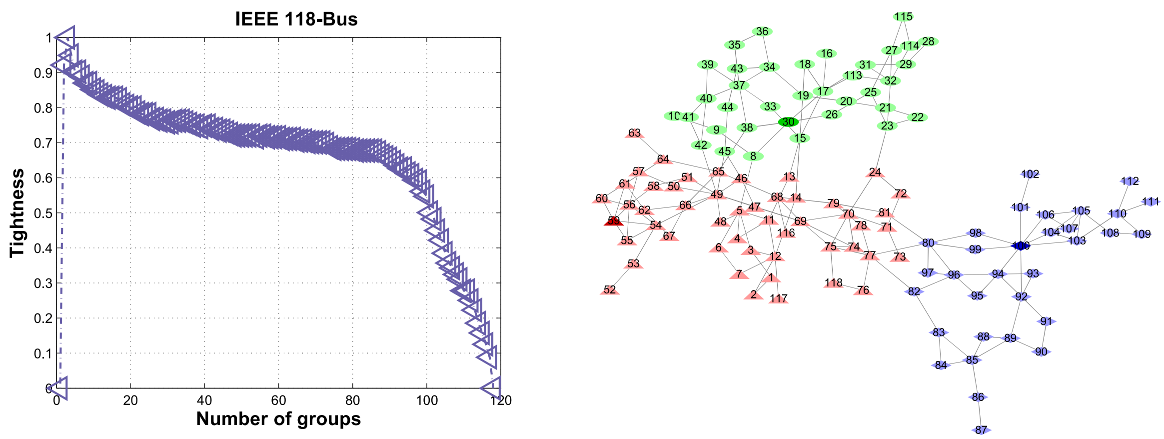


Fig. 7. (a) The final value of fuzzy tightness with different number of groups in IEEE 39-Bus. (b) The obtained group structure in IEEE 39-Bus when the number of groups is fixed as 3.

number of groups from 1 (whole network) to 39 (number of nodes); whereas the y axis denotes the value of fuzzy tightness. It can be observed from Fig. 7(a) that the biggest value of tightness is obtained when the number of groups is prefixed as 3. After the number of groups exceeds 3 and as it increases, the tightness gradually decreases. It is also worth noting that as the number of iterations approaches the total number of nodes, the curve drops to 0 with a larger slope. Fig. 7(b) shows the topology of IEEE-39-Bus, in which each color/shape represents a type of group detected by *GDF-ICN* when the number of groups is fixed as 3. According to the results shown in Table 5, we also highlight the three centroid nodes (node-16, node-26 and node-4) in darker color. Fig. 7(b) shows a tight group structure and the three centroid nodes also play an important role from the perspective of topology. All nodes are distributed in the three groups dominated by the three centroid nodes, respectively. It is worth noting that node-25 is selected only by our *GDF-ICN* as one of the top eight critical nodes. From the perspective of group structure, node-25 is in the group dominated by node-26, which is an important hub connecting node-26 and node-2.

Similarly, we perform our *GDF-ICN* and other three baseline methods in another power network—IEEE-118-Bus. Compared to IEEE-39-Bus, IEEE-118-Bus has complex topology and electrical interactions. The top 10 critical nodes identified by the four methods are shown in Table 6. Compared to the critical nodes selected by literature (Yang



(a) (b)

Fig. 8. (a) The value of fuzzy tightness with different number of groups in IEEE 118-Bus. (b) The obtained group structure in IEEE 118-Bus when the number of groups is fixed as 3.

Table 5
Critical nodes identified on IEEE-39 by three baselines and *GDF-ICN*. The importance of these nodes is also displayed for clear comparison.

Rank	Li et al. (2019)		Zhu et al. (2016)		Zhu et al. (2018)		*GDF-ICN*	
	NId	Index	NId	Index	NId	Index	NId	Index
1	16	0.0329	16	1	39	1.59	16	0.8385
2	29	0.0327	4	0.9114	6	1.35	26	0.8344
3	38	0.0308	12	0.8757	16	1.32	4	0.8290
4	10	0.0295	26	0.8667	38	1.15	22	0.8287
5	20	0.0288	3	0.8015	29	1.13	3	0.7142
6	22	0.0280	11	0.7865	8	1.04	6	0.4843
7	8	0.0277	15	0.7372	20	1.02	25	0.2451
8	19	0.0276	5	0.7222	4	0.99	19	0.2336

Table 6
Critical nodes identified on IEEE-118-Bus by three baselines and *GDF-ICN*. The importance of these nodes is also displayed for clear comparison.

Rank	Yang et al. (2020)		Bompard et al. (2011)		Wei (2020)		*GDF-ICN*	
	NId	Index	NId	Index	NId	Index	NId	Index
1	42	5.15	65	1	47	3.6	30	0.8306
2	80	3.61	68	0.8803	12	3.07	59	0.6809
3	15	2.92	80	0.7917	100	3.01	100	0.6478
4	55	2.55	38	0.7051	80	2.7	49	0.6262
5	84	2.55	30	0.6786	18	2.6	70	0.6082
6	95	2.55	81	0.6367	38	2.24	105	0.4931
7	118	2.55	100	0.6354	59	2.17	80	0.3885
8	5	2.24	49	0.5973	65	2.15	68	0.3513
9	18	2.24	77	0.5902	73	2.13	85	0.2847
10	19	2.24	69	0.5506	92	2.12	18	0.2612

et al., 2020), our *GDF-ICN* selects three nodes with same ID, i.e., node-80 and node-18. Compared to the critical nodes selected by literature (Bompard et al., 2011), our *GDF-ICN* selects four nodes with same ID, i.e., node-68, node-80, node-100 and node-49. Compared to the critical nodes selected by literature (Wei, 2020), our *GDF-ICN* selects four nodes with same ID, i.e., node-100, node-80, node-18 and node-59. It can be seen that node-80 is always ranked as one of the top 10 critical nodes. Also, node-100 is ranked as one of the top 10 critical nodes except the method in literature (Yang et al., 2020). Aiming to analysis the selected critical nodes by *GDF-ICN* in a more comprehensive manner, we also provide more details shown in Fig. 8 viewed from different perspective.

Fig. 8(a) denotes the final stable value of fuzzy tightness under different number of groups in IEEE-118-Bus. The x axis denotes the number of groups from 1 (whole network) to 118 (number of nodes); whereas the y axis denotes the value of fuzzy tightness. It can be observed from Fig. 8(a) that the biggest value of tightness is obtained when the number of groups is prefixed as 3. After the number of groups exceeds 3 and as it increases, the tightness gradually decreases to 0. Fig. 8(b) shows the topology of IEEE-118-Bus. In Fig. 8(b), each color/shape represents a type of group detected by *GDF-ICN* when the number of groups is fixed as 3. According to the results shown in Table 6, we also highlight the three centroid nodes (node-30, node-59 and node-100) in darker color. Fig. 8(b) shows an obvious group structure and the three centroid nodes also play an important role from the perspective of topology. Different from the other three baseline methods, the top 3 nodes ranked by *GDF-ICN* are distributed in three "distant" space/group, respectively. The identification of these critical nodes also promotes the detection of tight group structure.

7. Conclusion

In this paper, we developed *GDF-ICN*—a group-driven framework for identifying critical nodes in power networks, behind which the motivation is based on a reasonable assumption that nodes play different roles of keeping system's stability in power networks with different group structure. As far as we know, no one has identified critical nodes in a power network by means of potential group structure. *GDF-ICN* skillfully combines the electrical information of the power grid with its structural group information, which can not only finally identify the critical nodes, but also produce a better group structure. Also, we proposed a new metric called fuzzy tightness for evaluating the goodness of a specific group structure. In particular, we also provided a preliminary tool to softening the metrics of evaluating hard groups. It motivates us and can help researchers use a more systematic way to study network's fuzzy group structure. Comprehensive experiments on several benchmark power networks were conducted to examine the performance of *GDF-ICN*. The results show that *GDF-ICN* can well identify the critical nodes and also reveal the fuzzy group structure associated with these critical nodes. Overall, *GDF-ICN* provides an interesting and promising direction for us to investigate the roles played by nodes in a power network with group structure. However, there are some issues or research directions that need our further study. For example, the theoretical relationship between structural group information and electrical information. The research on it can help us further understand the impact of group structure on a power system. Besides, *GDF-ICN* can not only help us find the critical entities in a complex system, but also uncover the potential tightly linked entities through the lens of network. Since power system is one of typical complex networked system with informative entities and relationship among them, in the future we seek to apply the idea of *GDF-ICN* to more broader application fields, such as social network, biological network and other engineering networks.

CRediT authorship contribution statement

Yangyang Liu: Discussed the results, Commented on the manuscript, Analyzed the data, Developed the models and methods, Wrote the manuscript. **Aibo Song:** Discussed the results, Commented on the manuscript, Analyzed the data, Developed the models and methods, Directed the research, Wrote the manuscript. **Xin Shan:** Discussed the results, Commented on the manuscript, Directed the research, Wrote the manuscript. **Yingying Xue:** Discussed the results, Commented on the manuscript, Analyzed the data, Developed the methods. **Jiahui Jin:** Discussed the results, Commented on the manuscript, Analyzed the data, Developed the methods.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work is supported by the National Key Research and Development Program of China under grant 2018YFB08034, the National Natural Science Foundation of China under grant 62061146001, SGCC Science and Technology Program "Research on Key Technologies of Reactive Voltage Coordinated Control Including New Energy, Energy Storage and Flexible Load".

References

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008.

Bompard, E., Wu, D., & Xue, F. (2011). Structural vulnerability of power systems: A topological approach. *Electric Power Systems Research*, *81*(7), 1334–1340.

Bu, Z., Gao, G., Li, H. J., & Cao, J. (2017). CAMAS: A cluster-aware multiagent system for attributed graph clustering. *Information Fusion*, *37*, 10–21.

Bu, Z., Li, H.-J., Zhang, C., Cao, J., Li, A., & Shi, Y. (2020). Graph k-means based on leader identification, dynamic game and opinion dynamics. *IEEE Transactions on Knowledge and Data Engineering*, *32*(7), 1348–1361.

Cai, H., Zheng, V. W., & Chang, K. C.-C. (2018). A comprehensive survey of graph embedding: problems, techniques, and applications. *IEEE Transactions on Knowledge and Data Engineering*, *30*(9), 1616–1637.

Cao, J., Bu, Z., Wang, Y., Yang, H., Jiang, J., & Li, H.-J. (2019). Detecting prosumer-community groups in smart grids from the multiagent perspective. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *49*(8), 1652–1664.

Chakraborty, T., Dalmia, A., Mukherjee, A., & Ganguly, N. (2017). Metrics for community analysis: A survey. *ACM Computing Surveys*, *54*, 1–37.

Chakraborty, T., Srinivasan, S., Ganguly, N., Mukherjee, A., & Bhowmick, S. (2014). On the permanence of vertices in network communities. In *Proceedings of the 20th acm sigkdd international conference on knowledge discovery and data mining* (pp. 1396–1405).

Chen, W., Liu, Z., Sun, X., & Wang, Y. (2010). A game-theoretic framework to identify overlapping communities in social networks. *Data Mining and Knowledge Discovery*, *21*(2), 224–240.

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

Fang, J., Su, C., Chen, Z., Sun, H., & Lund, P. (2018). Power system structural vulnerability assessment based on an improved maximum flow approach. *IEEE Transactions on Smart Grid*, *9*(2), 777–785.

Feng, Z., Xu, X., Yuruk, N., & Schweiger, T. A. J. (2007). A novel similarity-based modularity function for graph partitioning. In *Proceedings of the 9th international conference on data warehousing and knowledge discovery* (pp. 385–396).

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3–5), 75–174.

Gleich, D. F., & Seshadhri, C. (2012). Vertex neighborhoods, low conductance cuts, and good seeds for local community methods. In *Proceedings of the 18th acm sigkdd international conference on knowledge discovery and data dining* (pp. 597–605).

Grover, A., & Leskovec, J. (2016). node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*.

Guan, X., Liu, J., Gao, Z., Yu, D., & Cai, M. (2014). Power grids vulnerability analysis based on combination of degree and betweenness. In *The 26th chinese control and decision conference (2014 ccdc)* (pp. 4829–4833).

Javed, M. A., Younis, M. S., Latif, S., Qadir, J., & Baig, A. (2018). Community detection in networks: A multidisciplinary review. *Journal of Network and Computer Applications*, *108*, 87–111.

Jianbo Shi, & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(8), 888–905.

Jonnalagadda, A., & Kuppusamy, L. (2016). A survey on game theoretic models for community detection in social networks. *Social Network Analysis and Mining*, *6*(1), 1–24.

Li, C., Kang, Z., Yu, H., Li, X., & Zhao, B. (2019). Identification method of key nodes in power system based on improved PageRank algorithm. *Transactions of China Electrotechnical Society*, *39*(9), 1952–1959.

Liu, B., Li, Z., Chen, X., Huang, Y., & Liu, X. (2018). Recognition and vulnerability analysis of key nodes in power grid based on complex network centrality. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *65*(3), 346–350.

Liu, F., Xue, S., Wu, J., Zhou, C., Hu, W., Paris, C., Nepal, S., Yang, J., & Yu, P. (2020). Deep learning for community detection: Progress, challenges and opportunities. In *Proceedings of the twenty-ninth international joint conference on artificial intelligence* (pp. 4981–4987).

Lv, L., Chen, D., Ren, X. L., Zhang, Q. M., Zhang, Y. C., & Zhou, T. (2016). Vital nodes identification in complex networks. *Physics Reports*, *650*, 1–63.

Martinet, L. E., Kramer, M. A., Viles, W., Perkins, L. N., & Kolaczyk, E. D. (2020). Robust dynamic community detection with applications to human brain functional networks. *Nature Communications*, *11*(1), 2785.

Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E*, *69*, Article 066133.

Poudel, S., Ni, Z., & Sun, W. (2018). Electrical distance approach for searching vulnerable branches during contingencies. *IEEE Transactions on Smart Grid*, *9*(4), 3373–3382.

Priebe, C. E., Park, Y., Vogelstein, J. T., Conroy, J. M., Yzinski, V., Tang, M., Athreya, A., Cape, J., & Bridgeford, E. (2019). On a two-truths phenomenon in spectral graph clustering. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(13), 5995–6000.

Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., & Parisi, D. (2004). Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(9), 2658–2663.

Wang, Y., Bu, Z., Yang, H., Li, H.-J., & Cao, J. (2021). An effective and scalable overlapping community detection approach: Integrating social identity model and game theory. *Applied Mathematics and Computation*, *390*, Article 125601.

Wang, S., Gu, X., Luan, S., & Zhao, M. (2021). Resilience analysis of interdependent critical infrastructure systems considering deep learning and network theory. *International Journal of Critical Infrastructure Protection*, *35*, Article 100459.

Wang, S., Lv, W., Zhang, J., Luan, S., Chen, C., & Gu, X. (2021). Method of power network critical nodes identification and robustness enhancement based on a cooperative framework. *Reliability Engineering & System Safety*, *207*, Article 107313.

Waniek, M., Michalak, T., Rahwan, T., & Wooldridge, M. (2018). Hiding individuals and communities in a social network. *Nature Human Behaviour*, *2*, 139–147.

Wei, G. (2020). Evaluation method of key nodes of power grid based on the intermediate weight entropy. *Computer Applications and Research*, *37*(S1), 27–30.

Wei, Z., Guan, X., & Liu, L. (2020). Overview of power community structure discovery algorithms and its application in power grid analysis. *Power System Technology*, *44*(7), 2600–2609.

Yang, J., & Leskovec, J. (2013). Overlapping community detection at scale: a non-negative matrix factorization approach. In *Proceedings of the sixth acm international conference on web search and data mining* (pp. 587–596).

Yang, D. S., Sun, Y. H., Zhou, B. W., Gao, X. T., & Zhang, H. G. (2020). Critical nodes identification of complex power systems based on electric cactus structure. *IEEE Systems Journal*, *14*(3), 4477–4488.

Zhao, Y., Liu, S., Lin, Z., Yang, L., Gao, Q., & Chen, Y. (2020). Identification of critical lines for enhancing disaster resilience of power systems with renewables based on complex network theory. *IET Generation, Transmission and Distribution*, *14*(20), 4459–4467.

Zhao, C., Zhao, J., Wu, C., Wang, X., & Lu, S. (2019). Power grid partitioning based on functional community structure. *IEEE Access*, *7*, 152624–152634.

Zhou, M., Li, J., Wu, S., Liu, S., Li, G., & Liu, J. (2019). Vulnerability analysis of power system based on dynamic regional electrical coupling. *International Transactions on Electrical Energy Systems*, *29*(1), e2671.1–e2671.12.

Zhu, G., Wang, X., He, R., Tian, M., Dai, D., & Zhang, Q. (2016). Identification of vital node in power grid based on importance evaluation matrix. *High Voltage Engineering*, *42*(10), 3347–3353.

Zhu, G., Wang, X., He, R., Tian, M., Dai, D., & Zhang, Q. (2018). Identification of critical node in power gird based on modified PageRank algorithm. *Electric Power Construction*, *39*(11), 34–41.