Hierarchical Modeling by Recursive Unsupervised Spectral Clustering and Network Extended Importance Measures to Analyze the Reliability Characteristics of Complex Network Systems

Yi-Ping Fang, Enrico Zio

Scientific Research

Hierarchical Modeling by Recursive Unsupervised Spectral Clustering and Network Extended Importance Measures to Analyze the Reliability Characteristics of Complex Network Systems

Yiping Fang[1], Enrico Zio[1,2]

[1]Chair on Systems Science and the Energetic Challenge, Ecole Centrale Paris and Supelec, Paris, France
[2]Energy Department, Politecnico di Milano, Milano, Italy
Email: yiping.fang@ecp.fr

ABSTRACT

The complexity of large-scale network systems made of a large number of nonlinearly interconnected components is a restrictive facet for their modeling and analysis. In this paper, we propose a framework of hierarchical modeling of a complex network system, based on a recursive unsupervised spectral clustering method. The hierarchical model serves the purpose of facilitating the management of complexity in the analysis of real-world critical infrastructures. We exemplify this by referring to the reliability analysis of the 380 kV Italian Power Transmission Network (IPTN). In this work of analysis, the classical component Importance Measures (IMs) of reliability theory have been extended to render them compatible and applicable to a complex distributed network system. By utilizing these extended IMs, the reliability properties of the IPTN system can be evaluated in the framework of the hierarchical system model, with the aim of providing risk managers with information on the risk/safety significance of system structures and components.

Keywords: Complex Network System; Hierarchical Modeling; Spectral Clustering; Extended Importance Measure

1. Introduction

Critical infrastructures are engineered distributed systems which provide the fundamental support to modern Industry and society. Examples are computer and communication systems, power transmission and distribution systems, rail and road transportation systems, oil/gas systems and water distribution systems. Failures of such systems can have multiple, transnational impacts of significant size [1-3]. Hence, identifying and quantifying the reliability and vulnerability of such systems is crucial for designing the adequate protections, mitigation and emergency actions against failures [2].

These systems are exposed to multiple hazards and threats, some of which are even unexpected and emergent, and consist of a large number of elements whose interactions are not easily modeled and quantified, so that a complete analysis by exhaustive treatment cannot be pursued. As a result, the performance and reliability assessment of such "complex" systems has proved to be a non-trivial task in practice.

Recent studies suggest that many real complex network systems exhibit a modularized organization [4,5]. In many cases, these modularized structures are found to correspond to functional units within networks (ecological niches in food webs, modules in biochemical networks) [6]. Broadly speaking, clusters (also called communities or modules) are found in the network, forming groups of elements that are densely interconnected with each other but only sparsely connected with the rest of the network. Furthermore, hierarchically modularized organization, which is a central idea for the life process in biology [5,7], is also found to characterize the internal structure of many technological networks [8]. This sparks the idea of utilizing the hierarchical, modularized structure as a basis to model these complex systems, for their analysis and understanding [9].

In the analysis of systems with respect to their failure behavior, Importance Measures (IMs) are used to identify the weak points and quantify the impact of component failures [10,11]. IMs provide numerical indicators to determine which components are most important for system reliability improvement or most critical for system failure. Many different IMs have been proposed in the literature [12,13], among which classical and relevant statistics are Birnbaum [14], Fussell-Vesely [15] and Criticality Importance [16,17]. However, none of these

Figure 1. The 380 kV Italian power transmission network.

measures can be applied directly to complex network systems, because of the distributed character of functionality and service that they provide.

The purpose of this paper is twofold: firstly to propose a scheme of recursive clustering to obtain a hierarchical modeling framework associated with different varied-size grained virtual networks; then to introduce Extended Importance Measures (EIMs) which are compatible with the distributed characteristics of complex network systems, to evaluate the components importance in the framework of the hierarchical system representation.

The remainder of this paper is organized as follows:

Section 2 presents the methodology of hierarchical modeling, taking the structure of the 380 kV Italian Power Transmission Network (IPTN) as an example for illustration; in Section 3, the basic terminal-pair connection reliability problem is first introduced, based on which the traditional IMs are extended and then calculated for the IPTN system; conclusions are drawn in Section 4.

2. Hierarchical Modeling of Complex Network System

2.1. Network Representation

Graph Theory provides a framework for the mathematical representation of complex networks. A graph is an ordered pair $G(V,E)$ comprising a set of vertices (nodes) $V = \{v_1, v_2, \cdots, v_N\}$ together with a set of edges (also called arcs or links) $E = \{e_1, e_2, \cdots, e_M\}$, which are two-element subsets of V. The network structure is usually defined by the $N \times N$ *adjacency matrix*, which defines which two nodes are connected by assigning a 1 to the corresponding element of the matrix; otherwise, the value in the matrix is 0 if there is no connection between the two nodes. As described, this type of graph is unweighted and undirected. A graph is weighted if a value (weight) is assigned to each edge representing properties of the connection like cost, reliability, capacities, etc. For example, the *matrix of physical distances* is often used in conjunction with the *adjacency matrix* to describe a network also with respect to its spatial dimension [18,19].

In this paper, we take for exemplification of the analyses proposed the 380 kV Italian power transmission network (IPTN) (**Figure 1**). This network is a branch of the high-voltage-level transmission network, which can be modeled as a graph of $N = 127$ nodes ($N_G = 30$ generators and $N_D = 97$ distributors) connected by $M = 171$ links [20,21], defined by its $N \times N$ adjacency matrix A whose entries $\begin{bmatrix} a_{ij} \end{bmatrix}$ are 1 if there is an edge joining node i to node j or 0 otherwise. In **Figure 1**, the generators, *i.e.* hydro and thermal power plants, are represented by squares whereas the distribution substations are represented by circles.

2.2. Construct Network Hierarchy by Successive Clustering

Modularity is ubiquitous in many networks of scientific and technological interest, ranging from the World Wide Web to biological networks [7,22]. As a result, it is often possible to identify groups of elements that are highly interconnected with each other, but have only a few links to components outside of the group to which they belong to. These communities usually combine into each other in a hierarchical manner [7], in which nodes form groups and then join the groups of groups, and so forth, starting from the lowest levels of organization (individual nodes) up to the level of the entire system. This suggests the development of a hierarchical structure to describe a complex network system at different levels of resolution, with the aim of managing the complexity of the system more effectively.

A successive Unsupervised Spectral Clustering Algorithm (USCA) [23], which is invariant to cluster shapes and densities and simple to implement, has been adopted in this study to build the hierarchical structure of the IPTN system. Cluster analysis aims at recognizing natural groups within classes of entities [24]. The problem is to assign categories to unlabelled data, encouraging the search of implicit information in the network structure encoded in its graph [25]. Consequently, modularity patterns within a complex network system can be revealed without a priori knowledge of their existence. The detailed description of different clustering methods is beyond the scope of this article. For a systematic and synthetic review, the reader is encouraged to look at [24-26].

The USCA makes use of the spectrum (eigenvalues) of the similarity matrix of the data to perform dimensionality reduction before Fuzzy c-Means (FCM)—clustering in fewer dimensions. Schematically, it is performed by the steps [23] in **Table 1**.

In the first step, the Laplacian matrix L_{sym} is calculated from the similarity (affinity) matrix as follows. The input similarity matrix S is of size $n \times n$ and its generic element s_{ij} represents the similarity between nodes i and j in the network. The diagonal components s_{ii} are set to 1 and the matrix is symmetric $\left(s_{ij} = s_{ji} \right)$. The degree matrix D is the diagonal matrix with diagonal entries $d_1, d_2, \cdots, d_n$ defined by

$$d_i = \sum_{j=1}^{N} s_{ij}, i = 1, 2, \cdots, n \tag{1}$$

Then, the normalized graph Laplacian matrix can be obtained:

$$L_{sym} = D^{-1/2} L D^{-1/2} = I - D^{-1/2} S D^{-1/2} \tag{2}$$

where $L = D - S$ and I is the identity matrix of size $n \times n$.

By recursively operating the USCA on the data of the IPTN presented in Section 2.1 above, a 5-levels hierarchical structure of the system is constructed which contains the complete system at the top and individual elements at the bottom (the top panel of **Figure 2** gives out the structure of the hierarchy, detailed in the first 3 levels).

2.3. Hierarchical Modeling of the Network

Based on the hierarchy structure resulting from the successive application of USCA, artificial networks can be

Table 1. Unsupervised spectral clustering algorithm.

Input: Similarity matrix $S \in \mathbb{R}^{n \times n}$.

Compute the normalized graph Laplacian matrix L_{sym}.

Compute the first k eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_k$ and corresponding eigenvectors $\bar{u}_1, \bar{u}_2, \cdots, \bar{u}_k$ of matrix L_{sym}. The first k eigenvalues are such that they are very small whereas λ_{k+1} is relatively large.

The number of clusters c is set equal to k, according to the eigengap heuristic theory [24].

Let $U \in \mathbb{R}^{n \times k}$ be the matrix containing the vectors $\bar{u}_1, \bar{u}_2, \cdots, \bar{u}_k$ as columns. Form the matrix $T \in \mathbb{R}^{n \times k}$ from U by normalizing the rows to norm 1, that is set $t_{ij} = u_{ij} / \left(\sum_k u_{ik}^2 \right)^{1/2}$.

For $i = 1, \cdots, n$ let $y_i \in \mathbb{R}^k$ be the vector corresponding to the i-th row of T.

Resort to the FCM algorithm [27,28] to partition the data points $(y_i)_{i=1,\cdots,n}$ into $c = k$ clusters $C_1, \cdots, C_k$.

Output: Clusters $A_1, \cdots, A_k$ with $A_i = \{j | y_j \in C_i\}$

Figure 2. The hierarchy structure of the IPTN system and associated artificial networks of the first three levels.

defined at each layer. The artificial network at level l of the hierarchy is described as a graph $G^{(l)}\left(\Lambda^{(l)}, E^{(l)}\right)$ with $1 \le l \le L$, where L is the number of levels of the hierarchy. We use $V_i^{(l)}$ to represent the artificial node i

$\left(\text{for } i = 1, 2, \cdots, \left|\Lambda^{(l)}\right|\right)$ at level l, which corresponds to a cluster of real network nodes. Artificial nodes are connected by artificial links

$$E_{ij}^{(l)}\left(\text{for } i = 1, 2, \cdots, \left|\Lambda^{(l)}\right| \text{ and } i \neq j\right),$$

composed by those actual network links connecting (in parallel) the actual nodes in the clusters forming the artificial nodes,

$$E_{ij}^{(l)} = \left\{ e_{st} \middle| v_s \in V_i^{(l)}, v_t \in V_j^{(l)} \right\}.$$

The connection pattern between artificial nodes at level l is illustrated by an *adjacency matrix* $A^{(l)}$ whose element

$$A^{(l)}\left(V_i^{(l)}, V_j^{(l)}\right) = 1 \text{ if } E_{ij}^{(l)} \neq \varnothing,$$

i.e. if in the artificial nodes $V_i^{(l)}$ and $V_j^{(l)}$ there is at least one actual link connecting two actual nodes, and 0 otherwise.

Figure 2 presents the hierarchy structure of the IPTN system and the artificial networks associated with the first 3 levels of the hierarchy. At the top of the hierarchy (*i.e.* $l = 1$), the network is a single unit, *i.e.* one artificial node $V_1^{(1)}$, which consist of all actual nodes. At the second level $(l = 2)$, we have

$$\Lambda^{(2)} = \left\{ V_1^{(2)}, V_2^{(2)}, V_3^{(2)}, V_4^{(2)} \right\}$$

and $E^{(2)} = \left\{ E_{13}^{(2)}, E_{14}^{(2)}, E_{34}^{(2)}, E_{24}^{(2)} \right\}$

with $V_1^{(2)}, V_2^{(2)}, V_3^{(2)}, V_4^{(2)} \subset V_1^{(1)}$.

The integer that is indicated in the Figure in proximity of the generic i-th artificial node $V_i^{(2)}$ indicates the number of actual nodes which compose it, e.g. $V_1^{(2)}$ is representative of a group of 38 actual network nodes. Note that at the bottom of the hierarchy, we find the original network, *i.e.* each artificial node is an actual node and each artificial edge corresponds to an actual link.

The hierarchical model offers different levels of resolution at the different levels of the hierarchy. The artificial networks at the top of the hierarchy contain limited detail information of the local connectivity patterns (in the limit, only one node represents the whole network at the first level of the hierarchy); as we move down the hierarchy, more local information enters the model, at the expense of an increase in the dimension of the network. These characteristics can be leveraged efficiently to manage the complexity of a complex network system.
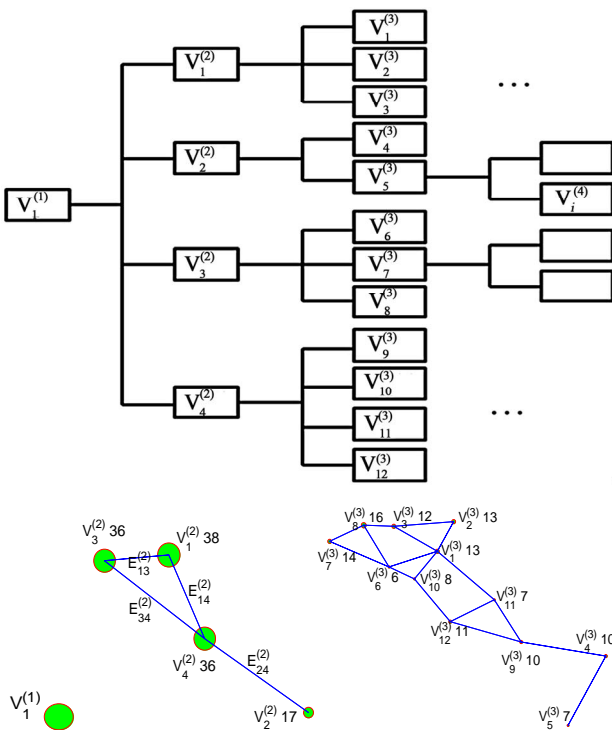
3. Reliability Analysis Based on the Hierarchical Model

It is known that most network reliability problems are *NP*-hard and therefore there is a significant gap between theoretical analysis and the ability to compute different

---

reliability parameters for large or even moderate network systems [11]. In this respect, hierarchical modeling sets up a framework based on which reliability and vulnerability characteristics of complex network systems can be computed efficiently, thanks to the multi-scaled information representation scheme.

3.1. Terminal-Pairs Reliability Assessment

The terminal-pair or node-pair reliability (TPR) problem amounts to determining the probability of successful communication between a specified source node and a terminal node in a network, given the probability of success of each link and node in the network. Let us introduce a binary vector $S_k = \{x_1, \cdots, x_M, y_1, \cdots, y_N\}$ to represent the state of the network, *i.e.* the state x of each of its M edges and the state y of each of its N nodes, where $x_i = 1$ if edge e_i is operating and 0 otherwise (y for node). For simplicity of illustration, we assume that nodes cannot fail, while edges can (thus y is no longer considered hereafter). The state of the network is defined as being *non-failure* if the specified terminal-pair is connected by at least one path of operating edges; otherwise it is *failure*. All possible *failure* states are included in the subset Ω_F of the set Ω containing all possible scenarios (*failure* and *non-failure*). An inclusive TPR analysis requires considering all elements in Ω_F. We then define the TPR as:

$$R_{sd} = \Pr(\phi_{sd}(S_k) = 1), S_k \in \Omega \qquad (3)$$

where ϕ_{sd} is a binary function which indicates the connection availability between node-pair s and d (1 = connection; 0 = no-connection). Let us assume that each edge e_i has associated a probability p_i of being operating and a probability $q_i = 1 - p_i$ of being failed; then, the TPR of the network can be calculated as:

$$R_{sd} = 1 - \sum_{S_k \in \Omega_F} \left(\prod_{x_i \notin X_f} (1 - p_i) \prod_{x_i \in X_f} p_i \right) \qquad (4)$$

where x_i represents the state of network edge e_i and X_f is the set of failed edges for a given state $S_k \in \Omega_F$. Note that the implicit assumption underpinning Equation (4) is that the network edges are independent.

When the computational cost of the network is high (it grows exponentially with the number of network components), then, the artificial network at a suitable level of the hierarchy can be leveraged to carry out the TPR. At the generic level of the hierarchy, the artificial link $E_{ij}^{(l)}$ connecting nodes (clusters) $V_i^{(l)}$ and $V_j^{(l)}$ is composed by actual network links in parallel,

$$E_{ij}^{(l)} = \left\{ e_{st} \mid v_s \in V_i^{(l)}, v_t \in V_j^{(l)} \right\};$$

then, the reliability of the artificial edge $E_{ij}^{(l)}$ at level l can be calculated by:

$$p\left(E_{ij}^{(l)}\right) = 1 - \prod_{e_{st} \in E_{ij}^{(l)}} q(e_{st}); \ v_s \in V_i^{(l)}, v_t \in V_j^{(l)} \qquad (5)$$

where $q(e_{st})$ indicates the failure probability of the actual link e_{st} that in the real network connects nodes v_s and v_t.

Various algorithms to solve the classic TPR problem have been reported in literature, with various computational efficiencies [29-31]. A so-called Modified Dotson algorithm [30], which has been claimed and tested to subdue others in computational time, is used here for the TPR assessment based on the hierarchical modeling. The failure probability of the transmission lines in the IPTN system is computed based on outage statistics provided in [32], by assuming that the edge failure probability is proportional to its length with an average failure rate $\lambda = 1.380635$ occ/100 mile-year, and average outage duration time $t = 64.81$ hours/occ.

In **Figure 3** right-panel, the connection reliability between nodes 1 and 127 in the IPTN network system (left panel in **Figure 3**) is shown as resulting from evaluations at each of the five levels of the hierarchical model described in the previous section. The right panel of **Figure 3** gives the probabilities of connectivity failure between nodes 1 and 127 from level 2 to level 5 (top) and
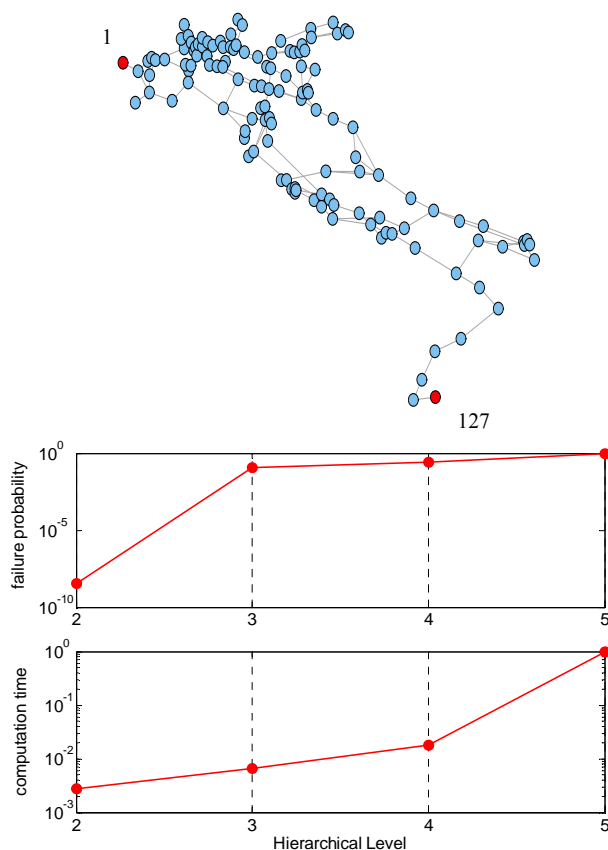

Figure 3. Illustrative example of terminal pair reliability assessment of IPTN system.

the computational time needed for the analysis (bottom); the values have been normalized with respect to the maximum values of connectivity failure probability and computational time, which occur at the bottom of the hierarchy (level 5) corresponding to the whole network. The result at the first level has not been shown since its value is simply 0, *i.e.*, node 1 and 127 are in a single unit and will not disconnect. One can see that the difference between the actual and estimated failure probabilities decreases as the assessment moves downs to the bottom of the hierarchy, balanced by the computation time which instead increases significantly. The decision maker can obtain satisfying estimations of the failure probability at a hierarchical level of lower complexity, e.g. level 3, thus saving significantly in computation time.

3.2. Component Extended Importance Measures

Component importance measures are widely used in system engineering to identify components within the system that most significantly influence the system behavior with respect to reliability, risk and/or safety. The indications drawn are valuable for establishing direction and prioritization of actions, related to reliability improvement during system design and optimization of operation and maintenance.

A well known IM is the so called Birnbaum IM defined as (with reference to system reliability R_s, as the system performance indicator) [14]:

$$I_i^B = \frac{\partial R_s}{\partial R_i} = R_s\left(R_i = 1\right) - R_s\left(R_i = 0\right) \tag{6}$$

where I_i^B is the Birnbaum Importance (BI) of component i; R_s represents the reliability of the system; R_i is the reliability of component i; $R_s\left(R_i = 1\right)$ is the system reliability calculated assuming that component i is perfectly operating and $R_s\left(R_i = 0\right)$ the system reliability in the opposite case of component i failed. The BI measures the significance of component i to system reliability by the rate at which system reliability improves with the reliability of component i. As shown in Equation (6), the BI of component i does not depend on R_i itself, so that two components i and j may have a similar value I^B although they have different reliability values R_i and R_j, respectively; this could be seen as a limitation of BI.

The Criticality Importance (CI) measure overcomes the above limitation by considering component unreliability [17]. It is defined as:

$$I_i^C = I_i^{\text{BI}} \frac{F_i}{F_s} = \left[R_s\left(R_i = 1\right) - R_s\left(R_i = 0\right)\right]\frac{1 - R_i}{F_s} \tag{7}$$

where F_i is the unreliability of component i and F_s is the system unreliability. Now, a less reliable component is more critical than another one with same value of BI.

Fuessell & Vesely [15] proposed an alternative impor-

tance measure according to which the importance of a component in the system depends on the number and on the order of the cut sets in which it appears [17]. Most commonly used as a risk reduction indicator, Fuessell & Vesely Importance (FVI) quantifies the maximum decrement in system reliability caused by a particular component being failed $\left(R_i = 0\right)$:

$$I_i^{FV} = \frac{R_s - R_s\left(R_i = 0\right)}{R_s} \tag{8}$$

The previously proposed IMs (BI, CI and FVI) are functionally different. They evaluate subtly different properties of the system behavior, and therefore, are often used in a complementary fashion to infer different information. To apply the IMs for analyzing a network system such as the IPTN, it is necessary to extend the definition of the IMs to account for the multiple terminal or node pairs (e.g. generator-distributor pairs) where connectivity defines the network functionality.

Specializing such extension for the analysis of the importance of components of the IPTN system, we introduce the Extended Birnbaum Importance (EBI) measure as the average of all BI values obtained considering all possible Generator-Distributor pairs reliabilities in the network system:

$$\begin{aligned} I_i^{E-B} \\ = \frac{1}{N_G N_D} \sum_{s \in V_G, d \in V_D} \frac{\partial R_{sd}}{\partial R_i} \\ = \frac{1}{N_G N_D} \sum_{s \in V_G, d \in V_D} \left(R_{sd}\left(R_i = 1\right) - R_{sd}\left(R_i = 0\right)\right) \end{aligned} \tag{9}$$

where N_G and N_D are the number of generators and distributors in the network respectively; V_G and V_D are sets of node generators and distributors respectively; R_{sd} is the TPR between node s and node d; $R_{sd}\left(R_i = 1\right)$ and $R_{sd}\left(R_i = 0\right)$ represent the terminal pair reliabilities between node s and node d, in the condition that component i is perfectly operating and completely failed, respectively.

Similarly, we can define Extended Criticality Importance (ECI) and Extended Fussell & Vesely Importance (EFVI) measures:

$$I_i^{E-C} = \frac{1}{N_G N_D} \sum_{s \in V_G, d \in V_D} \left[R_{sd}\left(R_i = 1\right) - R_{sd}\left(R_i = 0\right)\right]\frac{1 - R_i}{1 - R_{sd}} \tag{10}$$

$$I_i^{E-FV} = \frac{1}{N_G N_D} \sum_{s \in V_G, d \in V_D} \frac{R_{sd} - R_{sd}\left(R_i = 0\right)}{R_{sd}} \tag{11}$$

where I_i^{E-C} is the Extended Criticality Importance (ECI) measure of component i and I_i^{E-FV} is the Extended Fussell & Vesely Importance measure.

3.3. Numerical Example: Results and Discussions

The EIMs introduced have been calculated for the IPTN system at different levels of the hierarchical model of the system developed. For the evaluation, an artificial node functions as a generator as long as there is at least one actual generator node within it; otherwise it is simply a distributor.

Tables **2** and **3** report the results of the importance assessment (EBI, EFVI are given in **Table 2** and ECI in **Table 3**) for the artificial edges of the network at level 2 of the hierarchy. For EBI and EFVI, all components in the artificial network have the same importance rank, but with slight differences between EBI and EFVI values, and the artificial edge {2-4} is the most important in the artificial network (see the bottom panel of **Figure 2**). This is due to the fact that this artificial edge is the only possible link between a generator in artificial node $V_2^{(2)}$ and the distributors in other artificial nodes, and thus its disconnection would cause a large-scale generator-distributor connectivity failure. The rank based on the ECI is different from that of EBI and EFVI, and the most important artificial edge is {3-4}; the difference lies in the definition, as discussed before: EBI depends only on the structure of the system and not on the reliability of the considered component, whereas ECI takes the unreliability of the component into consideration, and in fact, the artificial edge {3-4} is made of only one actual edge with relatively high probability of failure, which leads to the highest ECI value.

By combining the indications of EBI and ECI, it is advisable to offer advices to the decision maker for the purpose of system maintenance and operation optimization [10]. When EBI & EFVI is high and ECI is low like in the case of artificial edge {2-4}, the system safety can be improved by protecting against failure of each component, e.g., adding alternative edges between artificial node $V_2^{(2)}$ and node $V_1^{(2)}$ (or $V_3^{(2)}$). For the case of low EBI & EFVI and high ECI (artificial edge {3-4}), the decision maker should invest in improvements of the component itself, to decrease the failure probability.

Tables **4** and **5** report the evaluation results at level 3 of the hierarchy. Fictitious edge {4-9}, composed by actual edges {110-111, 112-114, 107-109}, has highest EBI and EFVI values but relatively low ECI value (ranked 15th among all 17 artificial edges), indicating that the system reliability is highly sensitive to its failure, whereas the component itself is relatively reliable. On the contrary, the artificial edge {1-10} composed by only one actual edge {64-78} is highly unreliable itself, and its EBI and EFVI values are both ranked 8th among all 17 edges. It is important to pay attention to these artificial edges with both relatively high EBI & EFVI ranks and ECI ranks, which means not only that their failures cause a significant deterioration of the system reliability but also that they are vulnerable themselves. In this respect, by combining **Tables 4** and **5**, we find that artificial edges {1-11} (whose actual network link is {71-83}), {6-10} (which is composed by actual link {76-79}), and {10-12} (which is composed by actual links {75-88, 80-95}) are the three artificial edges most critical for the system reliability.

The bold edges in **Figure 4** represent the edges of the actual network system which have resulted most critical based on the extended importance measure evaluation carried out at level 3 of the hierarchy model. These edges should be paid special attention. For links {110-111, 112-114, 107-109}, improving the defense in depth against

Table 2. EBI and EFVI at level 2 of the hierarchical model.

Artificial Edge	EBI		EFVI		Associated Actual Edges
	Rank	Value	Rank	Value	
{2-4}	1	0.3750	1	0.3750	{107-109, 112-114, 110-111}
{1-4}	2	1.9606E−03	2	1.9605E−03	{64-78, 71-83}
{1-3}	3	1.4817E−03	3	1.4817E−03	{59-60, 61-62, 30-34, 30-31}
{3-4}	4	1.5100E−05	4	1.4900E−05	{76-79}

Table 3. ECI at level 2 of the hierarchical model.

Artificial Edges	Rank	ECI	Associated Actual Edges
{2-4}	4	0.37	{107-109, 112-114, 110-111}
{1-4}	2	7699812.62	{64-78, 71-83}
{1-3}	3	16.55	{59-60, 61-62, 30-34, 30-31}
{3-4}	1	7699828.67	{76-79}

Table 4. EBI and EFVI at level 3 of the hierarchical model.

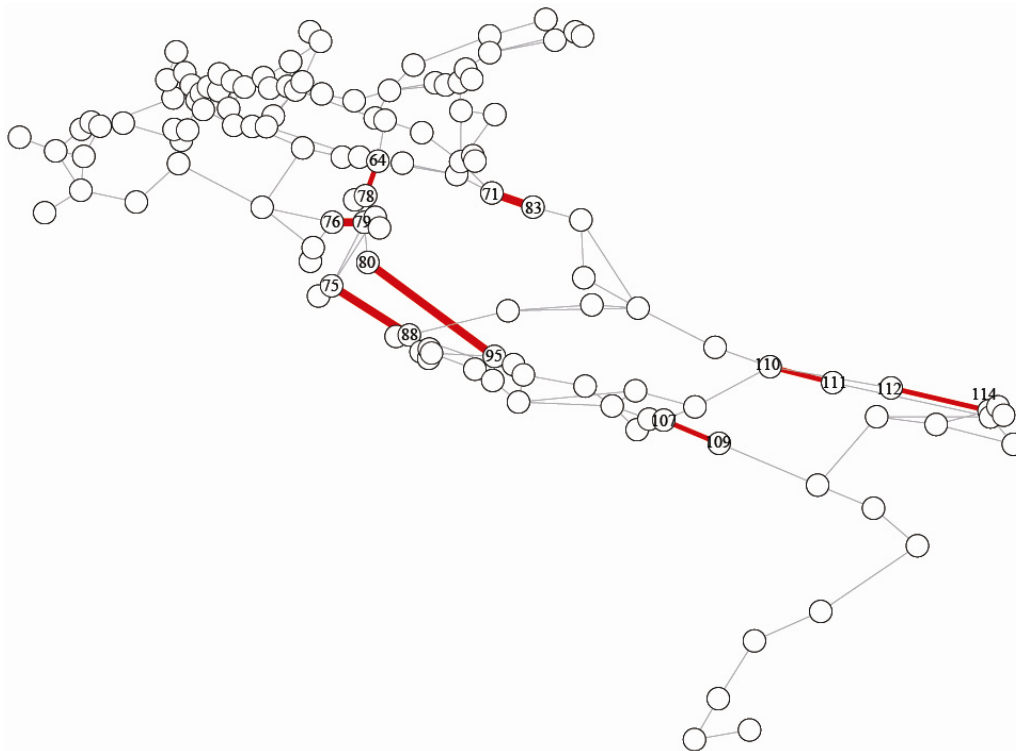| Artificial Edges	EBI		EFVI		Associated Actual Edges
	Rank	Value	Rank	Value	
{4-9}	1	0.2867	1	0.2879	{110-111, 112-114, 107-109}
{4-5}	2	0.1591	2	0.1591	{119-122}
{9-12}	3	0.0030	3	0.0030	{98-99, 94-97, 97-98}
{10-12}	4	0.0028	4	0.0028	{75-88, 80-95}
{2-3}	5	0.0007	5	0.0007	{42-43, 40-43}
{1-11}	6	0.0002	6	0.0002	{71-83}
{6-10}	7	1.55E−05	7	1.54E−05	{76-79}
{1-10}	8	1.17E−05	8	1.15E−05	{64-78}
{3-8}	9	8.04E−06	9	8.05E−06	{30-31, 30-34, 59-60}
{9-11}	10	7.52E−06	10	7.38E−06	{102-110}
{11-12}	11	4.82E−06	11	4.65E−06	{86-88}
{7-8}	12	4.11E−06	12	4.11E−06	{10-16, 10-21, 20-21}
{1-2}	13	3.00E−06	13	2.98E−06	{47-48}
{1-3}	14	8.43E−08	14	8.40E−08	{40-41, 60-63}
{1-6}	15	7.58E−08	16	5.56E−08	{61-62}
{6-7}	16	5.58E−08	15	4.96E−08	{11-12, 12-13}
{6-8}	17	1.43E−08	17	3.92E−08	{59-61}

Figure 4. Most critical edges at level 3 of the hierarchical model.

Table 5. The results of ECI assessment at level 3 of the hierarchical model.

Artificial Edges	Rank	ECI	Associated Actual Edges
{1-10}	1	3029896	{64-78}
{6-10}	2	2975998	{76-79}
{1-11}	3	2763614	{71-83}
{10-12}	4	139883.50	{75-88, 80-95}
{11-12}	5	45071.41	{86-88}
{6-8}	6	24763.84	{59-61}
{1-6}	7	20374.07	{61-62}
{1-2}	8	13626.99	{47-48}
{1-3}	9	212.10	{40-41, 60-63}
{6-7}	10	196.24	{11-12, 12-13}
{2-3}	11	57.85	{42-43, 40-43}
{3-8}	12	10.65	{30-31, 30-34, 59-60}
{7-8}	13	0.38	{10-16, 10-21, 20-21}
{4-5}	14	0.16	{119-122}
{4-9}	15	0.07	{110-111, 112-114, 107-109}
{9-11}	16	0.05	{102-110}
{9-12}	17	0.02	{98-99, 94-97, 97-98}

their failures is advisable to improve the reliability of the system, while for links {64-78, 71-83, 76-79, 80-95, 75-88}, the edge unreliability should also be mitigated.

Tables 6 and **7** report the results of the EIMs evaluation at level 4 of the IPTN hierarchical model. It turns out that artificial edge {7-11} (corresponding to actual link {119-122}) has the highest EBI and EFVI values and artificial edge {1-22} (corresponding to actual link {64-78}) has the highest ECI rank and relatively high EBI and EFVI ranks, indicating its criticality to system reliability.

Finally, **Table 8** reports the computation times required for the calculations of the EIMs at different levels in the hierarchy: as expected, the more we go down in the hierarchy the higher the computation time.

4. Conclusions

The modeling and analysis of complex network systems is a non-trivial task. Related decision-making regarding reliability and vulnerability is limited by computational resources.

In this work, we have introduced a framework for hierarchical modeling of complex network systems, which leads to the definition of different varied-size grained artificial networks. The construction of the hierarchical model is obtained by a recursive unsupervised spectral clustering method. The hierarchical model thereby obtained provides a multi-scaled representation of the original network system, with more detailed information but high complexity at the lower levels of the hierarchy, and simplified structure but relatively low complexity at the higher levels. The availability of different scales of modeling resolution allows a flexible management of the analysis, at the level of details desired for its purposes. The 380 kV Italian Power Transmission Network (IPTN) has been taken as an illustration.

Furthermore, Importance Measures (IMs) such as Birnbaum, Fuessell & Vesely and Criticality, have been extended for application to the terminal-pair reliability problem in complex distributed network systems.

The calculation of the extended IMs at different levels of the hierarchical system modeling has demonstrated the effectiveness of the proposed hierarchical modeling, with the IM-ranking of the IPTN elements offering insights on how to improve the system against failures of most critical elements.

5. Acknowledgements

Y. P. FANG, E. ZIO

Table 6. EBI and EFVI at level 4 of the hierarchical model (only the top 20 elements are reported).

Artificial Edges	EBI		EFVI		Associated Actual Edges
	Rank	Value	Rank	Value	
{7-11}	1	0.1504	1	0.1511	{119-122}
{3-4}	2	0.0787	2	0.0788	{47-49, 51-54}
{10-11}	3	0.0782	3	0.0788	{125-126}
{22-23}	4	4.2717E−4	4	4.2606E−4	{78-81}
{24-25}	5	3.5490E−4	5	3.5551E−4	{84-101, 85-101}
{12-13}	6	3.3570E−4	6	3.3605E−4	{14-73, 14-76}
{1-22}	7	3.0044E−4	7	2.9915E−4	{64-78}
{21-28}	8	2.1515E−4	8	2.1436E−4	{94-97}
{26-28}	9	1.7038E−4	9	1.6954E−4	{92-93}
{2-25}	10	1.6962E−4	10	1.6906E−4	{71-83}
{17-19}	11	1.0216E−4	11	1.0206E−4	{17-18}
{14-19}	12	7.53E−05	12	7.51E−05	{10-16}
{23-29}	13	6.50E−05	13	6.43E−05	{75-88}
{7-21}	14	5.10E−05	14	5.09E−05	{107-109}
{9-20}	15	4.24E−05	16	4.22E−05	{110-111}
{23-27}	16	3.74E−05	15	3.66E−05	{80-95}
{13-23}	17	3.36E−05	17	3.35E−05	{76-79}
{21-27}	18	3.22E−05	18	3.23E−05	{97-98, 98-99}
{7-8}	19	3.07E−05	19	3.07E−05	{113-120}
{8-20}	20	2.64E−05	20	2.61E−05	{112-114}

Table 7. ECI at level 4 of the hierarchical model (only the top 20 elements are reported).

Artificial Edges	Rank	ECI	Associated Actual Edges
{1-22}	1	868094.790	{64-78}
{2-4}	2	750781.848	{47-48}
{1-12}	3	737490.646	{61-62}
{13-23}	4	645088.015	{76-79}
{22-23}	5	602356.820	{78-81}
{12-14}	6	44554.9988	{12-13}
{14-18}	7	43748.7434	{10-21}
{14-15}	8	40914.4150	{7-9}
{14-19}	9	23137.9590	{10-16}
{17-19}	10	17031.1229	{17-18}
{12-13}	11	14138.4808	{14-73,14-76}
{12-18}	12	8829.8833	{59-61}
{1-5}	13	6285.1115	{40-41,60-63}
{5-16}	14	6013.8315	{30-31}
{16-18}	15	5235.8073	{27-59}
{6-16}	16	5051.5230	{30-34}
{5-18}	17	4665.6252	{59-60}
{15-18}	18	2481.7982	{20-21}
{12-15}	19	1666.1408	{11-12}
{4-5}	20	325.9829	{40-43}

Table 8. EIMs evaluation time at each level of the hierarchical model.

EIMs	Computation time (seconds on a computer with 2 CPU 3.06 G 3.07 G)		
	Level 2	Level 3	Level 4
EBI	0.3856	108.5	31763.58
EFVI	0.2086	112.2	32179.50
ECI	0.5152	175.0	47621.58

vant information on the network structure used as reference system and to Dr. Yanfu Li of Supelec for fruitful discussions.

REFERENCES

[1] W. Kröger, "Critical Infrastructures at Risk: A Need for a New Conceptual Approach and Extended Analytical Tools," *Reliability Engineering & System Safety*, Vol. 93, No. 12, 2008, pp. 1781-1787. doi:10.1016/j.ress.2008.03.005

[2] E. Zio, "Reliability Engineering: Old Problems and New Challenges," *Reliability Engineering & System Safety*, Vol. 94, No. 2, 2009, pp. 125-141. doi:10.1016/j.ress.2008.06.002

[3] W. Kröger and E. Zio, "Vulnerable Systems," Springer, Berlin, 2011. doi:10.1007/978-0-85729-655-9

[4] A. Mason, J. Onnela and P. Mucha, "Communities in Networks," *Notices of the American Mathematical Society*, Vol. 56, No. 9, 2009, pp. 1082-1166.

[5] S. Fortunato, "Community Detection in Graphs," *Physics Reports*, Vol. 486, No. 3, 2010, pp. 75-174. doi:10.1016/j.physrep.2009.11.002

[6] B. Karrer, E. Levina and M. Newman, "Robustness of Community Structure in Networks," *Physical Review E*, Vol. 77, No. 4, 2008, Article ID: 046119. doi:10.1103/PhysRevE.77.046119

[7] A. Clauset, C. Moore and M. E. J. Newman, "Hierarchical Structure and the Prediction of Missing Links in Networks," *Nature*, Vol. 453, No. 7191, 2008, pp. 98-101. doi:10.1038/nature06830

[8] M. Sales-Pardo, R. Guimerá, A. Moreira, A. Moreira and L. A. N. Amaral, "Extracting the Hierarchical Organization of Complex Systems," *Proceedings of the National Academy of Sciences*, Vol. 104, No. 39, 2007, pp. 15224-15229. doi:10.1073/pnas.0703740104

[9] C. Gómez, M. Sánchez-Silva and L.Duenas-Osorio, "Clustering Methods for Risk Assessment of Infrastructure Network Systems," In: Faber, Köhler and Nishijima, Eds., *Applications of Statistics and Probability in Civil Engineering*, CRC Press, Boca Raton, 2011.

[10] M. Van der Borst and H. Schoonakker, "An Overview of PSA Importance Measures," *Reliability Engineering & System Safety*, Vol. 72, No. 3, 2001, pp. 241-245. doi:10.1016/S0951-8320(01)00007-2

[11] I. Gertsbakh and Y. Shpungin, "Network Reliability Importance Measures: Combinatorics and Monte Carlo Based Computations," *WSEAS Transactions on Computers*, Vol. 7, No. 4, 2008, pp. 216-227.

[12] E. Zio, "Computational Methods for Reliability and Risk Analysis," World Scientific Publishing, Singapore, 2009. doi:10.1142/7190

[13] E. Zio, "Risk Importance Measures," *Safety and Risk Modeling and Its Applications*, Springer, Berlin, 2011.

[14] Z. W. Birnbaum, "On the Importance of Different Components in a Multicomponent System," In: P. R. Krishnaiah, Ed., *Multivariate Analysis II*, Academic Press, New York, 1969.

[15] J. B. Fussell, "How to Hand-Calculate System Reliability and Safety Characteristics," *IEEE Transactions on Reliability*, Vol. 24, No. 3, 1975, pp. 169-174. doi:10.1109/TR.1975.5215142

[16] M. C. Cheok, W. P. Gareth and R. Sherry, "Use of Importance Measures in Risk-informed Regulatory Applications," *Reliability Engineering & System Safety*, Vol. 60, No. 3, 1998, pp. 213-226. doi:10.1016/S0951-8320(97)00144-0

[17] J. F. Espiritu, D. W. Coit and U. Prakash, "Component Criticality Importance Measures for the Power Industry," *Electric Power Systems Research*, Vol. 77, No. 5, 2007, pp. 407-420. doi:10.1016/j.epsr.2006.04.003

[18] V. Latora and M. Marchiori, "Economic Small-World Behavior in Weighted Networks," *The European Physical Journal B—Condensed Matter and Complex Systems*, Vol. 32, No. 2, 2003, pp. 249-263. doi:10.1140/epjb/e2003-00095-5

[19] U. Brandes and T. Erlebach, "Network Analysis: Methodological Foundations," Springer, Berlin, 2005.

[20] E. Zio and G. Sansavini, "Component Criticality in Failure Cascade Processes of Network Systems," *Risk Analysis*, Vol. 31, No. 8, 2011, pp. 1196-1210. doi:10.1111/j.1539-6924.2011.01584.x

[21] V. Rosato, S. Bologna and F. Tiriticco "Topological Properties of High-Voltage Electrical Transmission Networks," *Electric Power Systems Research*, Vol. 77, No. 2, 2007, pp. 99-105. doi:10.1016/j.epsr.2005.05.013

[22] E. Ravasz and B. Albert-László, "Hierarchical Organization in Complex Networks," *Physical Review E*, Vol. 67, No. 2, 2003, Article ID: 026112. doi:10.1103/PhysRevE.67.026112

[23] U. Von Luxburg, "A Tutorial on Spectral Clustering," *Statistics and Computing*, Vol. 17, No. 4, 2007, pp. 395-416. doi:10.1007/s11222-007-9033-z

[24] A. K. Jain, M. Narasimha Murty and P. J. Flynn, "Data Clustering: A Review," *ACM Computing Surveys*, Vol. 31, No. 3, 1999, pp. 264-323.

[25] M. Filippone, F. Camastra, F. Masulli and S. Rovetta, "A Survey of Kernel and Spectral Methods for Clustering," *Pattern Recognition*, Vol. 41, No. 1, 2008, pp. 176-190. doi:10.1016/j.patcog.2007.05.018

[26] S. E. Schaeffer, "Graph Clustering," *Computer Science Review*, Vol. 1, No. 1, 2007, pp. 27-64. doi:10.1016/j.cosrev.2007.05.001

AJOR

[27] S. Leguizemón, H. Pelgrum and S. Azzali, "Unsupervised Fuzzy C-means Classification for the Determination of Dynamically Homogeneous Areas," *Revista SELPER*, Vol. 12, No. 12, 1996, pp. 20-24.

[28] M. Alata, M. Molhim and A. Ramini, "Optimizing of Fuzzy C-Means Clustering Algorithm Using GA," *World Academy of Science*, *Engineering and Technology*, Vol. 1, No. 5, 2008, pp. 224-229.

[29] W. Dotson and J. Gobien, "A New Analysis Technique for Probabilistic Graphs," *IEEE Transactions on Circuits and Systems*, Vol. 26, No. 10, 1979, pp. 855-865. doi:10.1109/TCS.1979.1084573

[30] Y. B. Yoo and N. Deo, "A Comparison of Algorithms for Terminal-pair Reliability," *IEEE Transactions on Reliability*, Vol. 37, No. 2, 1988, pp. 210-215.

[31] C.-C. Jane, W.-H. Shen and Y.-W. Laih, "Practical Sequential Bounds for Approximating Two-Terminal Reliability," *European Journal of Operational Research*, Vol. 195, No. 2, 2009, pp. 427-441. doi:10.1016/j.ejor.2008.02.022

[32] A. A. Chowdhury and D. O. Koval, "High Voltage Transmission Equipment Forced Outage Statistics Including Different Fault Types," *Proceedings of the* 10*th International Conference on Probabilistic Methods Applied to Power Systems*, 25-29 May 2008, pp. 1-8, 25-29.