Adaptive Clustering-Based Hierarchical Layout Optimization for Large-Scale Integrated Energy Systems

Hui Guo[1], *Member*, *IEEE*, Tianling Shi[1], *Student Member*, *IEEE*, Fei Wang[1*], *Senior Member*, *IEEE*, Lijun Zhang[1], *Member*, *IEEE*, Zhengyu Lin[2], *Senior Member*, *IEEE*

[1] Shanghai Key Lab of Power Station Automation Technology (School of Mechatronics Engineering and Automation), Shanghai University, Baoshan District, Shanghai 200444, China
[2] Centre for Renewable Energy Systems Technology (School of Mechanical, Electrical and Manufacturing Engineering), Loughborough University, Loughborough, LE11, 3TU, U.K.
[*]f.wang@shu.edu.cn (Corresponding Author)

Abstract: Different energy systems are generally planned and operated independently, which result in the low energy utilization, weak self-healing ability, and low system reliability. Therefore, an adaptive clustering-based hierarchical layout optimization method is proposed for a large-scale integrated energy system, considering energy balance, transmission losses and construction costs. First, an adaptive clustering partition method based on energy balance and load moments is proposed to determine the optimal location of energy hubs and to allocate each distributed generation and load to different energy hubs, forming multiple regional integrated energy systems adaptively. Then, the proposed hierarchical layout optimization model is formulated as to find the modified minimum spanning tree of regional integrated energy system and multi-regional integrated energy systems respectively, to construct an economical and reliable interconnection network. Finally, the effectiveness of the optimization model and strategy is verified by simulations.

1. Introduction

Facing the energy crisis and environmental pollution, the distributed renewable energy generation has received great attention [1-2]. However, the renewable energy generation is characterized by a large number, small capacity, wide distribution, and time variability. To cope with the spatiotemporal characteristics of distributed generation and solve the problem of reliable access to the utility grid, the microgrid as an effective solution has been proposed [3-5]. In order to satisfy the diversified demands and further improve the primary energy efficiency, multi-energy forms including cool, heat and gas are introduced into microgrids to constitute a regional integrated energy system (RIES) [6-7].

Traditionally, different energy systems are planned and operated independently, leading to the low energy utilization, weak self-healing ability, and low system reliability. Aiming to realize the coordinated optimization and utilization of multi-energy forms, the energy hub (EH) with functions of conversion, storage and scheduling is introduced and used for the RIES planning [8-9]. A two-stage mixed integer linear programming method is proposed for RIES planning in [10], where the type selection and connection relationship of energy production and conversion equipment of EHs are realized, considering the access of distributed renewable energy. Then, a bi-level expansion planning model of RIES and multi-RIESs is proposed in [11], where a decentralized approach is applied to solve the capacity configuration of energy production and conversion equipment of EHs, given the carbon emission constraints. Considering the uncertainty of load demands and renewable energy generation, a data-driven two-stage stochastic programming method is proposed in [12] for EH capacity configuration, to minimize the construction cost and operation cost.

Furthermore, the coordinated planning including the operation control of large-scale integrated energy systems has been proposed in [13-16], which transforms the expansion planning of power and gas systems into a mixed integer linear or nonlinear programming model. Specially, the risk exposed uncertainties of load demands and energy prices is considered in [13-14]. For a large-scale centralized nonlinear model, the numerical optimization algorithm is difficult to get the global optimal solution, therefore the heuristic optimization algorithm is applied. Aiming at the maximization of social welfare, genetic algorithm, differential evolution algorithm and their improved algorithms are applied and compared in [14]. To minimize operation costs, carbon emission costs and reliability costs of the system, a fuzzy particle swarm optimization algorithm is used in [15]. However, the heuristic optimization algorithm is slow and easy to fall into local optimum in solving the large-scale centralized model. Therefore, a hierarchical or decentralized optimization strategy based on spatiotemporal decoupling characteristics is proposed to solve the large-scale centralized problem in a decomposed fashion [17-19].

However, the above studies on system planning of integrated energy systems mainly focus on the type selection, capacity configuration and connection relationship of production and conversion equipment in EHs, as well as the system operation control. While the optimal location of EHs and the layout optimization of large-scale integrated energy systems are seldom considered in the existing research. As an effective solution, the clustering algorithm has been widely applied in industrial fields to reduce the complexity of a large-scale system [20-23]. To facilitate the layout optimization of large-scale offshore wind farms, a fuzzy c-means clustering algorithm is used to allocate each wind turbine to the geographically nearest substation [24-25]. But there are two limitations, i.e., the number of substations is directly given in advance without thinking of adaptive changes, and only Euclidean distance is regarded as the similarity measurement.

Following that, the layout optimization of energy networks is taken into consideration, which has a significant impact on the overall efficiency and economy of systems [26]. In [27], the graph theory and improved genetic algorithm are combined to optimize the pipe network layout of district cooling systems, aiming at minimizing the annual cost of pipe networks. The minimization of average production costs and total cable trenching lengths are respectively taken as objective functions in [28-29], where the minimum spanning tree (MST) algorithm of graph theory is used to optimize the cable layout of wind farms. Different from the radial topology applied commonly in windfarm systems, a looped or meshed topology is needed for multi-RIES layouts to increase the reliability of energy interaction between different RIESs, but which is less considered in existing studies [30]. Besides, the above work only focuses on the clustering and layout optimization of single energy system, without considering the reliable and economic layout optimization of integrated energy systems.

To solve the above problems, an adaptive clustering-based hierarchical layout optimization method is proposed for a large-scale integrated energy system, considering energy balance, transmission losses and construction costs. First, an adaptive clustering partition method based on energy balance and load moments is proposed to determine the optimal location of EHs and to allocate each distributed generation and load to different EHs, forming multiple RIESs adaptively. Then, the proposed hierarchical layout optimization model is formulated as to find the modified MST of integrated energy systems, to realize the dynamic radial layout optimization of RIES and the looped or meshed layout optimization of multi-RIESs respectively, constructing an economical and reliable interconnection network.

2. Example of the EH based RIES

From the view of completeness, the EH based RIES is comprised of energy input, energy conversion, energy storage, energy transmission and energy consumption, which are depicted in Fig. 2 as an example [31-33].

The relationship between different energy systems (electricity, gas and heat) is determined by the conversion and storage equipment of EHs. The type selection and

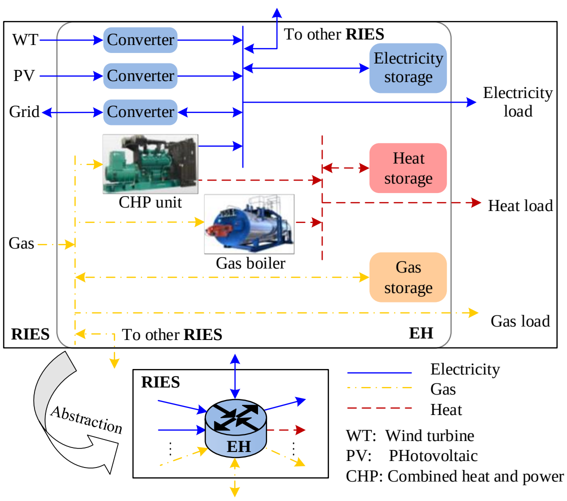


Fig. 2. Example diagram of the EH based RIES

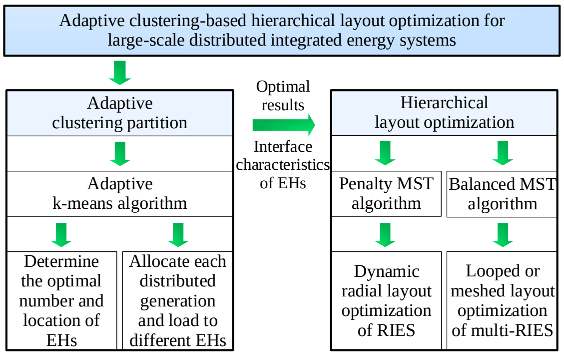


Fig. 1. *Steps of the proposed adaptive clustering-based hierarchical layout optimization*

capacity configuration of candidate equipment in EHs associated with the system operation optimization are really important, which have been introduced in [10-12,31] and partly studied in our previous work [19]. While the optimal location of EHs and the layout optimization of large-scale integrated energy systems are ignored or seldom considered in the existing research, which are vital for the construction of an economical and reliable interconnection network.

Assumed that the system planning of candidate equipment associated with the operation optimization (including system power balance constraints) in the EH is completed, the EH can be abstracted as a multi-port energy conversion node. Based on this assumption, the adaptive clustering-based hierarchical layout optimization for large-scale integrated energy systems is mainly studied in this paper, and the steps are shown in Fig. 1.

To facilitate the layout optimization of large-scale integrated energy systems, an adaptive clustering partition method is firstly proposed to determine the optimal location of EHs and to allocate each distributed generation and load to different EH-based RIESs. Then, according to the interface characteristics of EHs (including interface number and capacity), the dynamic radial layout optimization of RIES and the looped or meshed layout optimization of multi-RIESs are realized by applying the modified MST algorithm.

3. Adaptive clustering partition

3.1. Main principles of EH-based clustering partition

Two main principles of the EH-based clustering partition are as follows.

3.1.1 Minimization of transmission loss and construction cost: As a vital indicator of multi-energy transmission losses and construction costs, the integrated load moment of electricity, gas and heat shown in (1)-(3) should be introduced into the optimization model of clustering partition.

$$\min \sum_{r=1}^{3} \alpha_{ir} P_{ir} D_{io} \tag{1}$$

$$\alpha_{ir} = \frac{P_{ir}/\eta_r}{\sum_{r=1}^{3} (P_{ir}/\eta_r)} \tag{2}$$

$$D_{io}=\sqrt{(x_{i1}-u_{o1})^2+(x_{i2}-u_{o2})^2} \tag{3}$$

where r represents the type of energy forms, r=1 represents electricity, r=2 represents gas, and r=3 represents heat. D_{io} represents the distance between node i and cluster centre o, α_{ir} represents the weight of energy form r in node i, η_r represents the transmission efficiency of energy form r, P_{ir} represents the maximum transmission power of energy form r in node i, x_{i1} and u_{o1} represent the X-distance of node i and cluster center o respectively, x_{i2} and u_{o2} represent the Y-distance of node i and cluster center o respectively.

When the cluster center o is selected as the optimal location of EHs, the integrated load moment of electricity, gas and heat shown in (1) should be minimized. Therefore, the distributed nodes with smaller load moments can be divided into the same partition as far as possible.

3.1.2 Energy balance of distributed generation and load: In the process of clustering partition, in order to improve the local utilization of distributed energy and reduce the long-distance transmission loss, the minimization of (4) means that the total capacity of generation nodes should be close to the total capacity of load nodes in each partition C_o.

$$\min \sum_{r=1}^{3} | x_{i3r}+u_{o3r} | \tag{4}$$

$$u_{o3r}=\frac{1}{|C_o|}\sum_{x_i\in C_o} x_{i3r} \tag{5}$$

where $x_{i3r}(r=1,2,3)$ represents the capacity of energy form r in the generation node i ($x_{i3r}>0$) or load node i ($x_{i3r}<0$), $|C_o|$ represents the total number of generation nodes and load nodes in the partition C_o, u_{o3r} represents the average capacity of energy form r for the total generation nodes and load nodes in the partition C_o.

When $u_{o3r}<0$, the total capacity of generation nodes is less than the total capacity of load nodes for energy form r in the partition C_o. Therefore, the generation node i ($x_{i3r}>0$) should be divided into the partition C_o to minimize the (4).

When $u_{o3r}>0$, the total capacity of generation nodes is greater than the total capacity of load nodes for energy form r in the partition C_o. Therefore, the load node i ($x_{i3r}<0$) should be divided into the partition C_o to minimize the (4).

In order to realize the local utilization of large-scale distributed energy and reduce the long-distance transmission loss, a clustering partition optimization model based on load moments and energy balance is proposed.

3.2. EH-based clustering partition optimization model

According to the above analysis, the optimization goal of clustering partition should reflect the load moment from distributed nodes to the cluster center, as well as the energy balance of distributed generations and loads in each partition C_o.

3.2.1 Equivalent load moment: According to the load moment and energy balance, the equivalent load moment L_{io} between node i and cluster center o is defined, as shown in (6).

$$\begin{aligned} L_{io} &= \alpha_1 \sum_{r=1}^{3}\alpha_{ir}P_{ir}D_{io}+\alpha_2\sum_{r=1}^{3}| x_{i3r}+u_{o3r} | \\ P_{ir} &= | x_{i3r} | \end{aligned} \tag{6}$$

where the first term of (6) represents the integrated load moment of electricity, gas and heat between node i and cluster center o as well as the weighting coefficient α_1. The second term represents the energy balance of electricity, gas and heat in each partition C_o as well as the weight coefficient α_2.

3.2.2 Objective function: As shown in (7), the objective function J of clustering partition based on equivalent load moments, can be expressed as to minimize total equivalent load moments between node i and cluster center o.

$$\begin{aligned} &\min: J=\sum_{o=1}^{k}\sum_{x_i\in C_o} L_{io} \\ &\text{s.t.: } 1\le i\le m \\ &\qquad 1\le o\le k \end{aligned} \tag{7}$$

If there is a common connection point 0, the load moment from the cluster center o to the common connection point 0 should also be considered into the objective function, as shown in (8).

$$\begin{aligned} &\min: \sum_{o=1}^{k}\sum_{x_i\in C_o} L_{io}+\sum_{o=1}^{k}P_0 D_{0o} \\ &\text{s.t.: } 1\le i\le m \\ &\qquad 1\le o\le k \end{aligned} \tag{8}$$

Where P_0 represents the maximum transmission power of the common connection point 0, and D_{0o} represents the distance from the common connection point 0 to the cluster center o.

3.3. Adaptive k-means algorithm

Due to simple and efficient clustering characteristics, the k-means algorithm has been widely used [34 - 35]. However, there is also a certain limitation in the clustering process of traditional k-means algorithm [36 - 37]. For example, the number of cluster centers is directly given in advance without thinking of adaptive changes, and only Euclidean distance is regarded as the similarity measurement. Therefore, an adaptive k-means algorithm based on cluster center optimization and attribute weights is proposed to realize the adaptive clustering partition of large-scale distributed generations and loads, as well as the optimal location of EHs, as shown in Fig. 4.

For large-scale distributed generations and loads, the clustering partition optimization model involves different indicators, such as physical location and power capacity, which are unable to compare and synthesize directly, due to different attributes or dimensions of the indicators. Therefore, it is necessary to eliminate the impact of different indicators through data standardization, in order to obtain reliable and rational results.

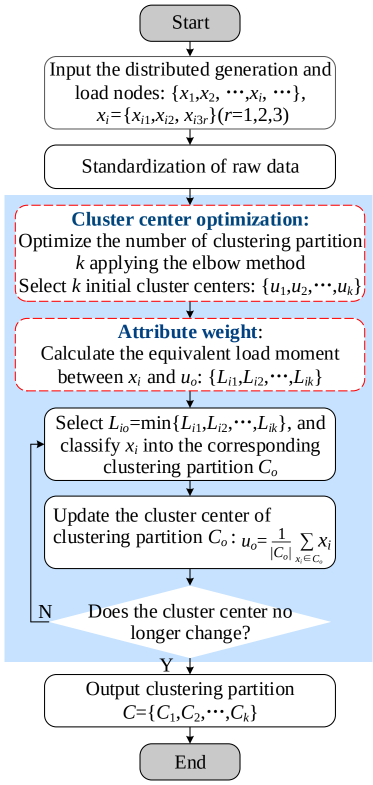


Fig. 4. *Adaptive k-means algorithm for clustering partition*

3.3.1 Cluster center optimization: The idea of the elbow method is to run k-means clustering for a given range of cluster centers *k* and calculate the objective function *J* for each value of *k*. Then, the *J* is plotted for each value of *k*, and the “elbow” (the inflection point on the curve) is the best value of *k*. Following that, *k* distributed nodes are selected as the initial cluster centers.

3.3.2 Attribute weight: Considering the physical location, energy form and power capacity, the equivalent load moment is obtained by weighting the load moment and energy balance. To improve the efficiency and quality of clustering partition, the distributed nodes with slightly larger load moments but favorable for energy balance, can be divided into the same partition. The distributed nodes with slightly smaller load moments but unfavorable for energy balance, can be divided into different partitions.

Based on the adaptive clustering partition, the distributed generation and load nodes are divided into *k* RIESs, and the cluster center near distributed nodes in each RIES is selected as the optimal location of EHs. Then, the hierarchical layout optimization of RIES and multi-RIESs are carried out respectively, to construct an economical and reliable interconnection network.

4. Hierarchical layout optimization

The proposed hierarchical layout optimization of integrated energy systems includes the dynamic radial layout optimization of RIES and the looped or meshed layout optimization of multi-RIESs.

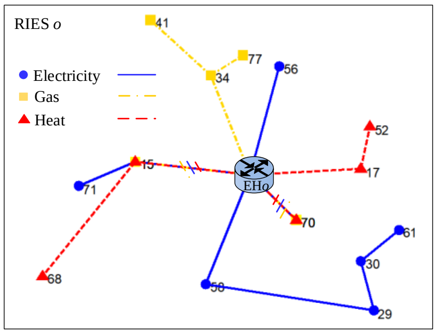


Fig. 3. *Layout optimization of distributed nodes in RIES o*

4.1. Dynamic radial layout optimization of RIES based on penalized MST

To optimize the network layout of each RIES, not only the transmission loss but also the conversion loss of different energy forms is necessary to be considered. Within a RIES, the dynamic radial layout optimization is implemented independently for different energy forms, but the coordinated conversion of different energy forms can be realized by the EH located in cluster centers.

As shown in Fig. 3, multiple distributed generations and loads in each RIES *o* can be abstracted into a weighted graph $G_o=\{V_o, E_o\}$, where the cluster center o $(o=1,2,\ldots,k)$ is the optimal location of EH *o*. The distributed generation and load are abstracted into nodes and constitute a node set $V_o=\{1,2,\ldots,|C_o|\}$ ($|C_o|$ represents the total number of distributed nodes in the partition C_o, i.e. RIES *o*). The interconnection between distributed generation and load is abstracted into node edge and constitutes an edge set E_o.

4.1.1 Penalty weight of edge $w_{ij,r}(r=1,2,3)$: When the dynamic radial layout optimization is performed for a RIES, the load moment of energy form *r* shown in (9) is taken as the penalty weight of edge (*i*,*j*), which reflects the transmission loss and construction cost of the energy form *r*.

$$w_{ij,r}=\begin{cases}\max(P_{ir},P_{jr})D_{ij} & i\neq o \text{ and } j\neq o,(i,j)\in E_o\\ \beta_{or}\cdot P_{jr}D_{ij} & i=o \text{ and } j\neq o,(i,j)\in E_o\\ \beta_{or}\cdot P_{ir}D_{ij} & i\neq o \text{ and } j=o,(i,j)\in E_o\\ \infty & (i,j)\notin E_o\end{cases} \tag{9}$$

where $w_{ij,r}$ represents the load moment of energy form *r* related to the edge (*i*,*j*), D_{ij} represents the distance of the edge (*i*,*j*), β_{or} represents the penalty factor of each node directly connected to the EH *o* in energy form *r*.

4.1.2 Objective function: To minimize the transmission loss and construction cost, the dynamic radial layout optimization of energy form *r* between distributed nodes in each RIES *o* is transformed into (10), which is used to find the penalized MST in the graph G_o according to the penalty edge weight. For energy form *r*, the internal connection degree of EH *o* is not greater than the internal interface number of EH *o* connecting distributed generation and load.

$$
\begin{aligned}
&\min\text{: } J_{or} = \sum_{(i,j)\in T_{or}} w_{ij,r} \\
&\text{s.t.: } d_r^{\text{in}}(o) \le n_{or}^{\text{in}} \\
&\qquad 1 \le r \le 3 \\
&\qquad 1 \le i, j \le |C_o|
\end{aligned} \tag{10}
$$

where J_{or} represents the total load moments of energy form r of the penalized MST for the RIES o, T_{or} represents the penalized MST of energy form r in the graph G_o, $d_r^{\text{in}}(o)$ represents the internal connection degree of energy form r of EH o, n_{or}^{in} represents the internal interface number of energy form r of EH o connecting distributed generation and load.

4.1.3 Penalty MST algorithm: Prim algorithm and Kruskal algorithm are both classic algorithms for finding MST from connected graphs. Strategically speaking, Kruskal algorithm can find the MST only by sorting the edge weight once, which has higher efficiency. To minimize the transmission loss and construction cost, a penalty MST algorithm based on the Kruskal algorithm is proposed to realize the dynamic radial layout optimization of RIES, where the penalty factor is dynamically adjusted and determined according to the internal interface number of EHs connecting distributed generation and load, as shown in Fig. 5.

Initially $\beta_{or} = 1$. When $\beta_{or} > 1$, the distributed nodes directly connected to the cluster center o in energy form r will decrease as the increase of β_{or}. When $0 < \beta_{or} < 1$, the distributed nodes directly connected to the cluster center o in energy form r will increase as the decrease of β_{or}.

4.2. Looped or meshed layout optimization of multi-RIESs based on balanced MST

The layout optimization of multi-RIESs is considered based on the optimal location of EHs. According to the estimation from the U.S. Energy Information Administration, the average loss rate of electricity transmission is about 5% [38]. The loss rate of gas transported through pipeline systems is about 8% [39]. However, a high level of heat losses during transportation is up to 30 % of generation [40]. Based on the above estimated transmission loss of different energy systems, the long-distance transmission loss of heat is larger, compared with electricity and gas. Therefore, the interconnection of multi-RIESs only includes electricity and gas networks in this paper.

The long-distance transmission of gas requires a booster station, and some booster stations are electrically driven. Considering the reuse and sharing of certain infrastructures during the transmission process of electricity and gas, the layout optimization of different energy forms between EHs is considered together, by minimizing the integrated load moment of electricity and gas.

As shown in Fig. 6, multi-RIESs based on the EH are abstracted into a weighted graph $G_{\text{EH}} = \{V, E\}$. The EHs are abstracted into nodes and constitute a node set $V = \{1, 2, \ldots, k\}$ (k represents the number of partitions, that is, the total number of EHs). The interconnection between EHs is abstracted into node edges and constitutes an edge set E.

4.2.1 Integrated weight of edge w_{mn}: As shown in (11)-(12), the integrated load moment of electricity and gas is

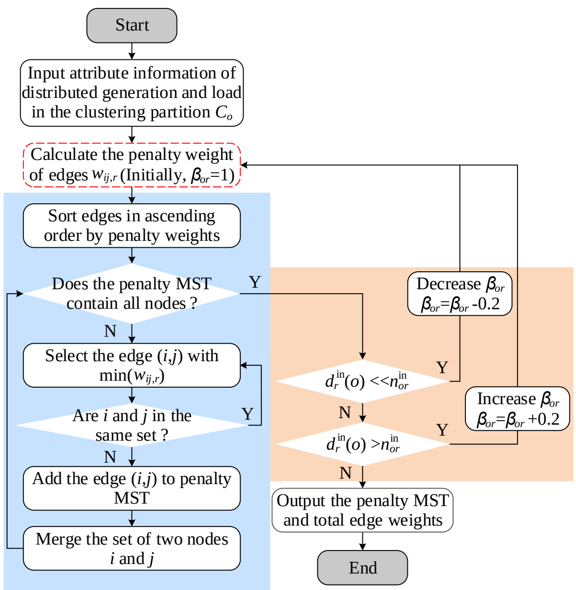


Fig. 5. *Dynamic radial layout optimization based on the penalty MST*

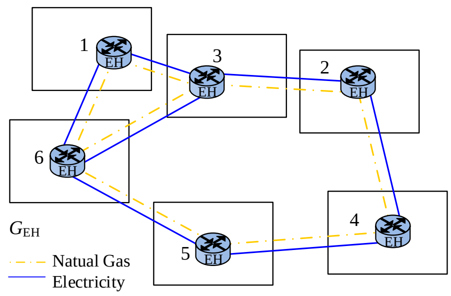


Fig. 6. *Layout optimization of multi-RIESs based on the EH*

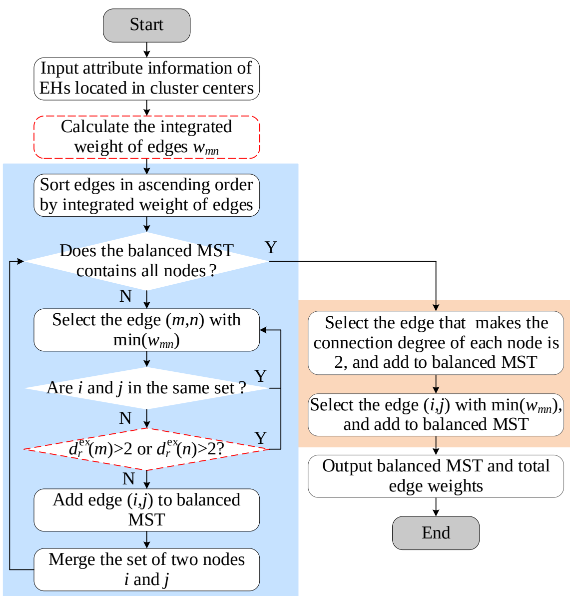


Fig. 7. *Looped or meshed layout optimization based on the balanced MST*

taken as the integrated weight of edge (m,n) connecting EH m and n, which reflects the transmission loss and construction cost of the electricity and gas network.

$$w_{mn} = \sum_{r=1}^{2} w_{mn,r} \tag{11}$$

$$w_{mn,r} = \begin{cases} \alpha_{mr} P_{mr} D_{mn} & P_{mr} = \max(P_{mr}, P_{nr}), (m,n) \in E \\ \alpha_{nr} P_{nr} D_{mn} & P_{nr} = \max(P_{mr}, P_{nr}), (m,n) \in E \\ \infty & (m,n) \notin E \end{cases} \tag{12}$$

where w_{mn} represents the integrated load moment of electricity and gas related to the edge (m,n), $w_{mn,r}$ represents the load moment of energy form r related to the edge (m,n), D_{mn} represents the distance of edge (m,n), α_{mr} represents the weight of energy form r in the EH m, P_{mr} represents the maximum transmission power of energy form r in the EH m.

4.2.2 Objective function: To minimize the transmission loss and improve the energy supply reliability, the layout optimization of multi-RIESs based on the EH is transformed into (13), which is used to find a balanced MST of electricity and gas in the weighted graph G_{EH} , where the external connection degree of energy form r of EH m is not greater than 2.

$$\begin{aligned} &\min\text{: } J_{\mathrm{EH}} = \sum_{(m,n)\in T} w_{mn} \\ &\text{s.t.: } d_r^{\mathrm{ex}}(m) \le 2 \\ &\qquad d_r^{\mathrm{ex}}(n) \le 2 \\ &\qquad 1 \le r \le 2 \\ &\qquad 1 \le m,n \le k \end{aligned} \tag{13}$$

where J_{EH} represents the total integrated load moment of electricity and gas of the balanced MST for multi-RIESs, T represents the balanced MST of electricity and gas in the graph G_{EH} , $d_r^{\mathrm{ex}}(m)$ represents the external connection degree of energy form r of EH m connecting the other EHs.

4.2.3 Balanced MST algorithm: As shown in Fig. 7, to increase the reliability and minimize the transmission loss and construction cost, a balanced MST algorithm is proposed to realize the looped or meshed layout optimization of multi-RIESs, where the external connection degree of each EH is kept as balanced as possible. Considering the system reliability, economy and equivalence, the edge that makes T a looped layout (the external connection degree of electricity and gas of each EH is 2) is selected to join T. Then, the edge with the lowest weight is selected to join T, forming a meshed layout of multi-RIESs.

Table 1 System parameters of 80-distributed nodes

Node	X (m)	Y (m)	Capacity (kW)			Node	X (m)	Y (m)	Capacity (kW)			Node	X (m)	Y(m)	Capacity (kW)		
			r=1	r=2	r=3				r=1	r=2	r=3				r=1	r=2	r=3
1	16.6	16.2	200	0	0	28	30	57.9	-150	0	0	55	34.8	42.1	0	0	-180
2	85.2	17.7	0	0	-150	29	63.9	2	-120	0	0	56	54.2	30.5	-160	0	0
3	97.8	56.9	100	0	0	30	62.6	7.8	150	0	0	57	27.5	80.7	0	-180	0
4	81.8	76.1	0	-160	0	31	24.8	82.3	0	0	-150	58	46.7	5.1	100	0	0
5	44.9	79.5	0	-150	0	32	80.3	58.5	0	-240	0	59	91.2	91.3	0	0	-90
6	96	85.2	-110	0	0	33	33.8	47	0	0	-150	60	5.2	37.5	0	-160	0
7	95.1	78.6	0	0	-100	34	47.2	29.4	0	-260	0	61	66.5	11.4	-180	0	0
8	27.1	85.3	200	300	250	35	82.9	77.6	300	450	300	62	49	77.7	0	-240	0
9	92.1	6.3	-280	0	0	36	36.9	98.8	-200	0	0	63	8	53.7	300	400	300
10	19.7	63.5	-140	0	0	37	17.9	61.2	-160	0	0	64	31.5	45.9	0	0	-160
11	74.7	57.1	-100	0	0	38	3.5	55.7	-80	0	0	65	32.9	98.8	-160	0	0
12	93.6	41.1	300	0	0	39	28.7	75	0	-200	0	66	49.8	100.3	-100	0	0
13	33.4	88.2	-180	0	0	40	69.3	73.3	-200	0	0	67	91	77.8	0	0	-60
14	56.8	91.1	-250	0	0	41	41	35.9	0	-180	0	68	29.9	6	0	0	-160
15	39.4	19.4	150	300	200	42	15.9	100.7	0	0	-160	69	25.1	52.8	-150	0	0
16	34.4	98	150	0	0	43	90.3	87.9	0	0	-50	70	56	12.6	300	400	300
17	62.6	18.5	0	0	-140	44	99.3	30.4	0	-120	0	71	33.6	16.5	-130	0	0
18	35.1	61.4	100	0	0	45	69.7	72	-180	0	0	72	24.9	59.9	160	0	0
19	55.3	74.9	-180	0	0	46	62.7	63.5	0	0	-80	73	22.2	48.5	200	300	400
20	2	46.3	0	-100	0	47	80.7	33	0	-180	0	74	75.6	69	200	350	0
21	85.4	19.8	0	0	-140	48	49.3	59.6	0	0	-60	75	92.9	20	-260	0	0
22	8.2	18.9	0	-180	0	49	7.3	34.7	0	-130	0	76	31.3	70.8	0	-220	0
23	4.6	37.1	0	-120	0	50	86.3	57.3	0	-180	0	77	50.5	31.7	0	-240	0
24	6.1	71.3	200	0	0	51	31	83.3	0	-150	0	78	8.4	79.3	0	0	-130
25	27.9	89.7	300	460	200	52	63.5	23.4	0	0	-200	79	1.8	62.7	-180	0	0
26	74.8	96.7	0	-210	0	53	45	69.5	350	400	250	80	92.2	25.3	250	300	300
27	19.6	38.8	0	0	-200	54	67.9	71.5	0	0	-100						

Note that the capacity value of electricity/gas/heat generation is positive, and the capacity value of electricity/gas/heat consumer is negative. r=1 represents electricity, r=2 represents gas, and r=3 represents heat.

5. Simulation results and analysis

An integrated energy system with 80-distributed nodes is used as a simulation case to be planned. According to the system plan design, the physical location of 80 distributed nodes and their energy form, power capacity can be obtained. Table 1 shows the X-distance, Y-distance, type and capacity of distributed nodes.

5.1. Adaptive clustering partition

5.1.1 Optimal k (the number of EHs or partitions): For the integrated energy system with 80-distributed nodes, the elbow method is used to optimize and determine the value of *k*. As shown in Fig. 8, the objective function *J* is plotted for each value of *k* from 2 to 10. As the *k* increases, the decrease of *J* tends to be smooth, creating the inflection point on the curve. Therefore, *k*=6 is selected as the optimal value in this paper.

5.1.2 Optimal clustering partition: After the optimal number of EHs *k*=6 is determined, the optimal clustering partition of the integrated energy system with 80-distributed nodes is shown in Fig. 9. In each partition, the cluster centre is selected as the optimal location of EHs, close to the distributed generation and load, which will bring benefits to realize the local utilization of distributed energy and reduce the long-distance transmission loss.

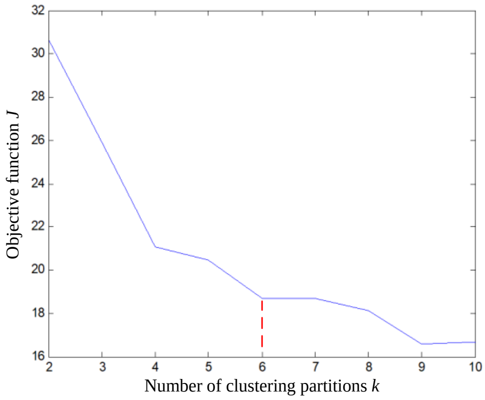


Fig. 8. *Elbow method showing the optimal k*

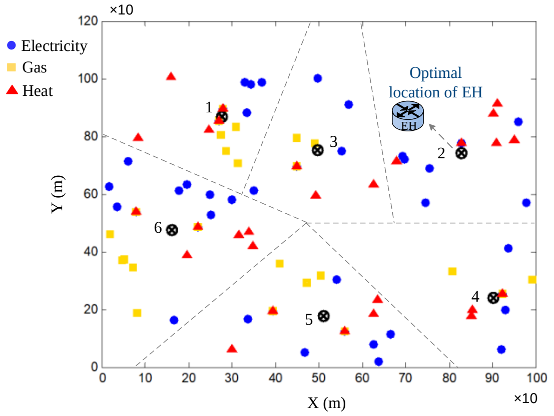


Fig. 9. *Optimal clustering partition of the integrated energy system with 80-distributed nodes*

Table 2 System parameters of partition 5-RIES 5

Nodes	X (m)	Y (m)	Capacity (kW)		
			r=1	*r*=2	*r*=3
EH 5	51.3	17.9	/	/	/
15	39.4	19.4	150.0	300.0	200.0
17	62.6	18.5	0	0	-140
29	63.90	2	-120.0	0	0
30	62.6	7.8	150	0	0
34	47.2	29.4	0	-260	0
41	41.0	35.9	0	-180.0	0
52	63.5	23.4	0	0	-200
56	54.2	30.5	-160.0	0	0
58	46.7	5.1	100	0	0
61	66.5	11.4	-180.0	0	0
68	29.9	6.0	0	0	-160.0
70	56.0	12.6	300.0	400.0	300.0
71	33.6	16.5	-130.0	0	0
77	50.5	31.7	0	-240.0	0

Table 3 Interface capacity of EH 1-6

Nodes	X (m)	Y (m)	Interface capacity (kW)		
			r=1	*r*=2	*r*=3
EH 1	27.7	87.0	300	460	250
EH 2	82.8	74.2	300	450	300
EH 3	49.8	75.3	350	400	250
EH 4	90.2	24.2	300	300	300
EH 5	51.3	17.9	300	400	300
EH 6	16.1	47.6	300	400	400

5.2. Hierarchical layout optimization

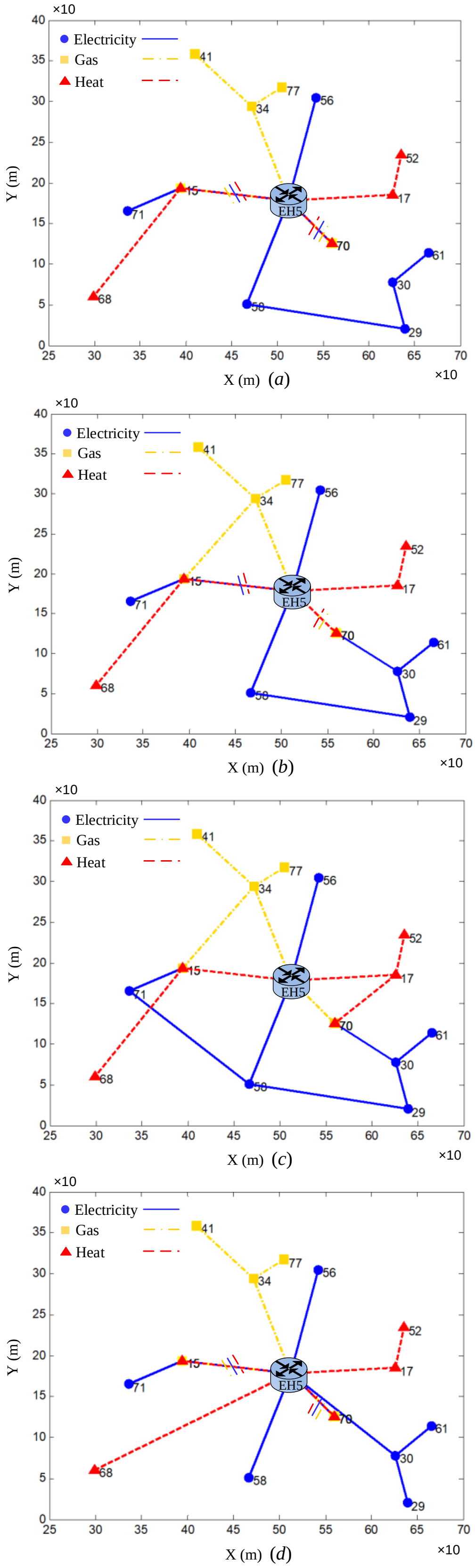
According to optimization results of the adaptive clustering partition, the partition 5 is selected as a case to analyze the dynamic radial layout optimization of RIES. The attributes (X-distance, Y-distance, type and capacity) of each distributed node in partition 5 are shown in Table 2, where the cluster center of partition 5 is the optimal location of EH 5.

The cluster centers of partition 1-6 are obtained from optimization results of the adaptive clustering partition and used as the optimal locations of EH 1-6. Besides, the interface capacity of electricity, gas and heat of EH 1-6 is shown in Table 3. Based on the optimal location and interface capacity of EHs, the looped or meshed layout optimization of multi-RIESs is carried out.

5.2.1 Dynamic radial layout optimization of RIES: The dynamic radial layout optimization of RIES 5 based on the penalty MST is shown in Fig. 10.

When $\beta_{or}=1$, as shown in Fig. 10 (a), the distributed nodes 15, 56, 58 and 70 containing energy form-electricity and with smaller equivalent load moment to cluster center 5 are directly connected to the EH 5 through transmission lines. The distributed nodes 15, 34 and 70 containing energy form-gas and with smaller equivalent load moment to cluster center 5 are directly connected to the EH 5 through gas pipes. The distributed nodes 15, 17 and 70 containing energy form-heat and with smaller equivalent load moment to cluster center 5 are directly connected to the EH 5 through heat pipes.

As shown in Fig. 10 (b) and Fig. 10 (c), when $\beta_{or} > 1$, the distributed nodes in RIES 5 tend to construct a penalty MST in a serial manner. As β_{or} increases, the distributed nodes directly connected to the EH located in the cluster center 5 gradually decreases. As shown in Fig. 10 (d) and Fig. 10 (e), when $0 < \beta_{or} < 1$, the distributed nodes in RIES tend to directly connect the EH and construct a penalty MST. As β_{or} decreases, the distributed nodes directly connected to the EH located in the cluster center 5 gradually increases, but which cannot be greater than the internal interface number of the EH 5. Therefore, the layout optimization of RIES influenced by penalty factors, can be adjusted according to the internal interface number of EHs connecting distributed generation and load.



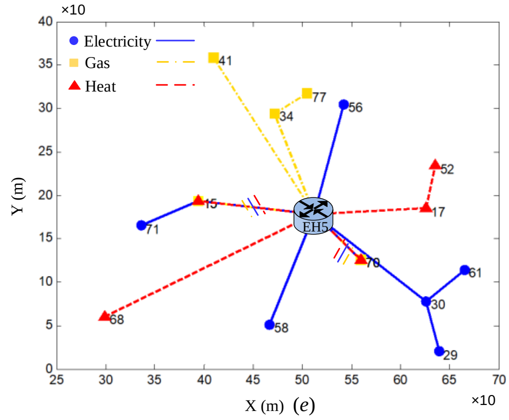


Fig. 10. *Dynamic radial layout optimization of RIES 5 under different penalty factors* β_{or}

(a) $\beta_{or} = 1$, ***(b)*** $\beta_{or} = 1.2$, ***(c)*** $\beta_{or} = 1.4$, ***(d)*** $\beta_{or} = 0.8$, ***(e)*** $\beta_{or} = 0.6$

5.2.2 Looped or meshed layout optimization of multi-RIESs: After determining the optimal location of EHs, the traditional MST of multi-RIESs based on the integrated load moment of electricity and gas is shown in Fig. 11 (a). The RIESs 1, 2 and 4 corresponding to the EHs 1, 2 and 4 only have one connection degree, which are vulnerable in the energy supply. The RIES 3 corresponding to the EH 3 has a high connection degree, which is more reliable in the energy supply.

Considering the system reliability, economy and equivalence, the MST based on the integrated load moment of electricity and gas for multi-RIESs is optimized. In the optimized network layout shown in Fig. 11 (b), the external connection degree of each EH is as balanced as possible, that is, the external connection degree of each node is not greater than 2.

To improve the reliability of the energy supply, a looped or meshed layout is preferable. Firstly, the edge (2,4) that makes the balanced MST form a looped layout (the external connection degree of each EH is 2) is selected to join the balanced MST. Then, the smallest weighted edge (3,6) is selected from all the remaining edges to join the balanced MST, as shown in Fig. 11 (c).

6. Conclusions

In order to construct an economical and reliable interconnection network of integrated energy system, an adaptive clustering-based hierarchical layout optimization method is proposed for a large-scale integrated energy system, considering energy balance, transmission losses and construction costs.

First, the proposed adaptive clustering partition method based on energy balance and load moments realizes the optimal location of EHs and allocates each distributed generation and load to different EHs, forming multiple RIESs adaptively.

Then, the proposed hierarchical layout optimization method realizes the dynamic radial layout optimization of RIES and the looped or meshed layout optimization of multi-RIESs, constructing an economical and reliable interconnection network.

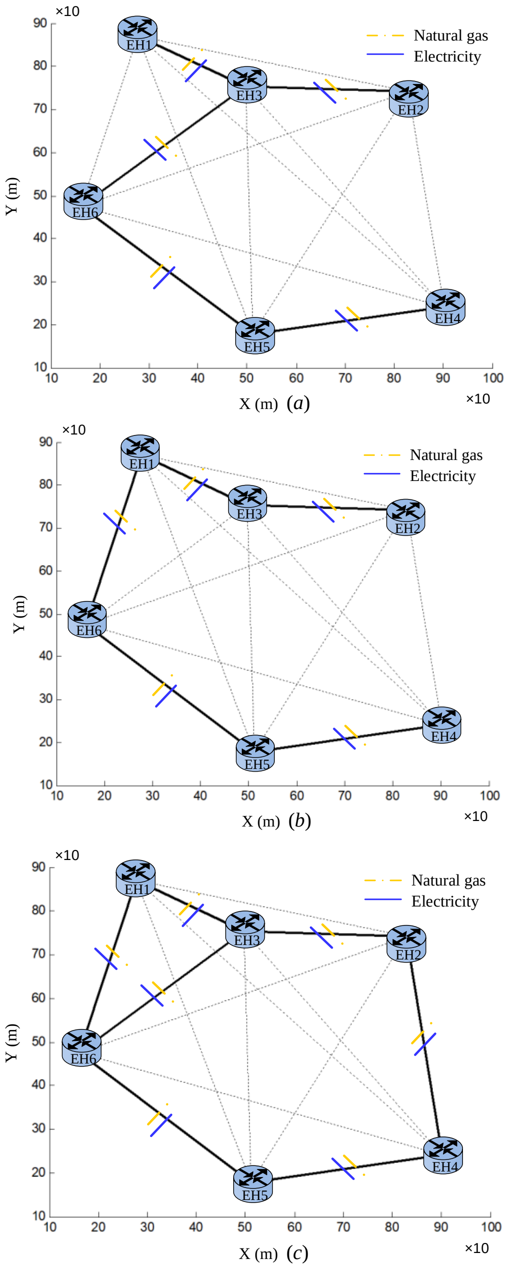


Fig. 11. Looped or meshed layout optimization of multi-RIESs based on the EH
(a) Traditional MST, (b) Balanced MST, (c) Balanced MST-based looped or meshed layout

7. Acknowledgments

The authors are grateful for the financial and technical support of the National Key R&D Program of China (Project no.2017YFE0112400) and the European Union's Horizon 2020 research and innovation programme (Project no.734796).

8. Reference

[1] Pursiheimo, E., Holttinen, H., Koljonen, T.: 'Path toward 100% renewable energy future and feasibility of power-to-gas technology in Nordic countries', IET Renewable Power Generation, 2017, 11, (13), pp. 1695-1706

[2] Wang, T., Nian, H., Zhu, Z.Q.: 'Using inverter-based renewable generators to improve the grid power quality—A review', Chinese Journal of Electrical Engineering, 2018, 4, (4), pp. 16-25

[3] Harmon, E., Ozgur, U., Cintuglu, M.H., et al.: 'The internet of microgrids: a cloud-based framework for wide-area networked microgrids', IEEE Trans. Ind. Informatics, 2018, 14, (3), pp. 1262-1274

[4] Li, Y., Zhang, P., Luh, P.B.: 'Formal analysis of networked microgrids dynamics', IEEE Trans. Power Syst., 2018, 33, (3), pp. 3418-3427

[5] Faisal R.B., Purnima D., Subrata K.S., Sajal K. D.: 'A survey on control issues in renewable energy integration and microgrid', Protection and Control of Modern Power Systems, 2019, 4, (1), pp. 87-113

[6] Massrur, H.R., Niknam, T., Fotuhi-Firuzabad, M., Nikoobakht, A.: 'Hourly electricity and heat demand response in the OEF of the integrated electricity-heat-natural gas system', IET Renewable Power Generation, 2019, 13, (15), pp. 2853-2863

[7] Zhang, X.H., Liu, X.Y., Zhong, J.Q., Gao, W.B.: 'Electricity-gas-integrated energy planning based on reward and penalty ladder-type carbon trading cost', IET Generation, Transmission & Distribution, 2019, 13, (23), pp. 5263-5270

[8] Alipour, M., Zare, K., Abapour, M.: 'MINLP probabilistic scheduling model for demand response programs integrated energy hubs', IEEE Trans. Ind. Informatics, 2018,14, (1), pp. 79-88

[9] Dolatabadi, A., Mohammadi-Ivatloo, B., Abapour, M., Tohidi, S.: 'Optimal stochastic design of wind integrated energy hub', IEEE Trans. Ind. Informatics, 2017, 13, (5), pp. 2379-2388

[10] Huang, W.J., Zhang, N., Yang, J.W., et al.: 'Optimal configuration planning of multi-energy systems considering distributed renewable energy', IEEE Trans. Smart Grid, 2019, 10, (2), pp. 1452-1464

[11] Cheng, Y.H., Zhang, N., Lu, Z.X., et al.: 'Planning multiple energy systems toward low-carbon society: A decentralized approach', IEEE Trans. Smart Grid, 2019, 10, (5), pp. 4859-4869

[12] Cao, Y., Wei, W., Wang, J.H., et al.: 'Capacity planning of energy hub in multi-carrier energy networks: A data-driven robust stochastic programming approach', IEEE Trans. Sustain. Energy, 2020, 11, (1), pp. 3-14

[13] Ding, T., Hu, Y., Bie, Z.H.: 'Multi-stage stochastic programming with nonanticipativity constraints for

expansion of combined power and natural gas systems', IEEE Trans. Power Syst., 2018, 33, (1), pp. 317-328
[14] Qiu, J., Dong, Z.Y., Zhao, J.H., et al.: 'Multi-stage flexible expansion co-planning under uncertainties in a combined electricity and gas market', IEEE Trans. Power Syst., 2015, 30, (4), pp. 2119-2129
[15] Qiu, J., Dong, Z.Y., Zhao, J.H., et al.: 'Low carbon oriented expansion planning of integrated gas and power systems', IEEE Trans. on Power Syst., 2015, 30, (2), pp. 1035-1046
[16] Saldarriaga, C.A., Hincapie, R.A., Salazar, H.: 'A holistic approach for planning natural gas and electricity distribution networks', IEEE Trans. on Power Syst., 2013, 28, (4), pp. 4052-4063
[17] Massrur, H.R., Niknam, T., Aghaei, J., Shafie-Khah, M., et al.: 'Fast decomposed energy flow in large-scale integrated electricity-gas-heat energy systems', IEEE Trans. Sustain. Energy, 2018,9, (4), pp. 1565-1577
[18] Moeini-Aghtaie, M., Abbaspour, A., Fotuhi-Firuzabad, M., et al.: 'A decomposed solution to multiple-energy carriers optimal power flow', IEEE Trans. Power Syst., 2014, 29, (2), pp. 707-716
[19] Guo, H., Wang, F., Zhang, L., et al.: 'A hierarchical optimization strategy of the energy router-based energy internet', IEEE Trans. Power Syst., 2019, 34, (6), pp. 4177-4185
[20] Ouchitachen, H., Hair, A., Idrissi, N.: 'Improved multi-objective weighted clustering algorithm in wireless sensor network', Egyptian Informatics Journal, 2017, 18, (1), pp. 45-54
[21] Al-Wakeel, A., Wu, J.Z., Jenkins, N.: 'K-means based load estimation of domestic smart meter measurements', Applied Energy, 2017, 194, (MAY15), pp. 333-342
[22] Huang, L.L., Chen, N., Zhang, H.Y., et al.: 'Optimization of large-scale offshore wind farm electrical collection systems based on improved FCM'. IEEE IET Int. Conf. on Sustainable Power Generation and Supply, Hangzhou, China, September 2012, pp. 1-6
[23] Li, F., Huang, Y., Wu, F., Liu, Y., Zhang, X.: 'Research on clustering equivalent modeling of large-scale photovoltaic power plants', Chinese Journal of Electrical Engineering, 2018, 4, (4), pp. 80-85
[24] Chen, Y.Y., Dong, Z.Y., Meng, K., et al.: 'A novel technique for the optimal design of offshore wind farm electrical layout', Journal of Modern Power Systems and Clean Energy, 2013, 1, (3), pp. 254-259
[25] Chen, Y.Y., Dong, Z.Y., Meng, K., et al.: 'Collector system layout optimization framework for large-scale offshore wind farms', IEEE Trans. on Sustain. Energy, 2016, 7, (4), pp. 1398-1407
[26] Adefarati, T., Bansal, R.C.: 'Integration of renewable distributed generators into the distribution system: a review', IET Renewable Power Generation, 2016, 10, (7), pp. 873-884
[27] Feng, X.P., Long, W.D.: 'Applying single parent genetic algorithm to optimize piping network layout of district cooling system'. IEEE 2008 Fourth Int. Conf. on Natural Computation, Jinan, China, October 2008, pp. 176-180
[28] Hou, P., Hu, W.H., Chen, C., et al.: 'Optimisation of offshore wind farm cable connection layout considering levelised production cost using dynamic minimum spanning tree algorithm', IET Renewable Power Generation, 2016, 10, (2), pp. 175-183
[29] Dutta, S., Overbye, T.J.: 'Optimal wind farm collector system topology design considering total trenching length', IEEE Trans. Sustain. Energy, 2012, 3, (3), pp. 339-348
[30] Guo, H., Wang, F., James, G., et al.: 'Graph theory based topology design and energy routing control of the energy internet', IET Generation, Transmission & Distribution, 2018, 12, (20), pp. 4507-4514
[31] Zhang, X.P., Che, L., Shahidehpour, M., et al.: 'Reliability-based optimal planning of electricity and natural gas interconnections for multiple energy hubs', IEEE Trans. Smart Grid, 2017, 8, (4), pp. 1658-1667
[32] Barmayoon, M. H., Fotuhi-Firuzabad, M., Rajabi-Ghahnavieh, A., et al.: 'Energy storage in renewable-based residential energy hubs', IET Generation Transmission & Distribution, 2016, 10, (13), pp. 3127-3134
[33] Zarif, M., Khaleghi, S., Javidi, M.H.: 'Assessment of electricity price uncertainty impact on the operation of multi-carrier energy systems', IET Generation Transmission & Distribution, 2015, 9, (16), pp. 2586-2592
[34] Zhang, L.B., Lu, F.M., Liu, A., et al.: 'Application of k-means clustering algorithm for classification of NBA guards', International Journal of Science and Engineering Applications, 2016, 5, (1), pp. 1-6
[35] Arora, P., Varshney, S., et al.: 'Analysis of k-means and k-medoids algorithm for big data', Procedia Computer Science, 2016, 78, pp. 507-512
[36] Manoharan, J.J., Ganesh, S.H.: 'Initialization of optimized k-means centroids using divide-and-conquer method', ARPN Journal of Engineering and Applied Sciences, 2016, 11, (2), pp. 1086-1091
[37] Elbasiony, R.M., Sallam, E.A., Eltobely, T.E., et al.: 'A hybrid network intrusion detection framework based on random forests and weighted k-means', Ain Shams Engineering Journal, 2013, 4, (4), pp. 753-762
[38] 'How much electricity is lost in electricity transmission and distribution in the United States?', https://www.eia.gov/tools/faqs/faq.php?id=105&t=3, accessed 31 December 2019
[39] 'Natural Gas is Energy Efficient', https://www.socalgas.com/smart-energy/benefits-of-natural-gas/energy-efficient, accessed 2016
[40] Syranov, D.V., Kovalnogov, V.N., Zolotov, A.N.: 'Modeling, research and optimization of heat losses during transport in energy systems'. 2016 2nd Int. Conf. on Industrial Engineering, Applications and Manufacturing, Chelyabinsk, Russia, May 2016, pp. 1-4