Adaptive clustering-based hierarchical layout optimisation for large-scale integrated energy systems

Hui Guo\(^1\), Tianling Shi\(^1\), Fei Wang\(^1\), Lijun Zhang\(^1\), Zhengyu Lin\(^2\)

\(^1\)Shanghai Key Lab of Power Station Automation Technology (School of Mechatronics Engineering and Automation), Shanghai University, Baoshan District, Shanghai 200444, People’s Republic of China
\(^2\)Centre for Renewable Energy Systems Technology (School of Mechanical, Electrical and Manufacturing Engineering), Loughborough University, Loughborough, LE11, 3TU, UK

E-mail: f.wang@shu.edu.cn

Abstract: Different energy systems are generally planned and operated independently, which result in the low energy utilisation, weak self-healing ability and low system reliability. Therefore, an adaptive clustering-based hierarchical layout optimisation 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 optimisation model is formulated as to find the modified minimum spanning tree of the regional integrated energy system and multi-regional integrated energy systems, respectively, to construct an economical and reliable interconnection network. Finally, the effectiveness of the optimisation 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 characterised 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 regionally integrated energy system (RIES) \([6, 7]\).

Traditionally, different energy systems are planned and operated independently, leading to low energy utilisation, weak self-healing ability and low system reliability. Aiming to realise the coordinated optimisation and utilisation 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 realised, 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 decentralised approach is applied to solve the capacity configuration of energy production and conversion equipment of EHs, given the carbon emission constraints. Considering the uncertainity of load demands and renewable energy generation, a data-driven two-stage stochastic programming method is proposed in \([12]\) for EH capacity configuration, to minimise 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 non-linear programming model. Especially, the risk exposed uncertainties of load demands and energy prices are considered in \([13, 14]\). For a large-scale centralised non-linear model, the numerical optimisation algorithm is difficult to get the global optimal solution, therefore the heuristic optimisation algorithm is applied. Aiming at the maximisation of social welfare, genetic algorithm, differential evolution algorithm and their improved algorithms are applied and compared in \([14]\). To minimise the operation costs, carbon emission costs and reliability costs of the system, a fuzzy particle swarm optimisation algorithm is used in \([15]\). However, the heuristic optimisation algorithm is slow and easy to fall into local optimum in solving the large-scale centralised model. Therefore, a hierarchical or decentralised optimisation strategy based on spatiotemporal decoupling characteristics is proposed to solve the large-scale centralised 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 optimisation 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 optimisation of large-scale offshore wind farms, a fuzzy c-means clustering algorithm is used to allocate each wind turbine to the geographically nearest substation \([24, 25]\). However, there are two limitations, i.e. the number of substations are directly given in advance without thinking of adaptive changes, and only Euclidean distance is regarded as the similarity measurement. Following that, the layout optimisation 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 optimise the pipe network layout of district cooling systems, aiming at minimising the annual cost of pipe networks. The minimisation 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 optimise 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
Assuming that the system planning of candidate equipment associated with the operation optimisation (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 optimisation for large-scale integrated energy systems is mainly studied in this paper, and the steps are shown in Fig. 2.

To facilitate the layout optimisation of large-scale integrated energy systems, an adaptive clustering partition method is first 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 optimisation of RIESs and the looped or meshed layout optimisation of multi-RIESs are realised 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 Minimisation 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 optimisation model of clustering partition.

\[
\min \sum_{r=1}^{3} a_r P_r D_{io} \tag{1}
\]

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 \(o\), \(a_r\) represents the weight of energy form \(r\) in node \(i\), \(\eta_r\) represents the transmission efficiency of energy form \(r\), \(P_r\) represents the maximum transmission power of energy form \(r\) in node \(i\), \(x_{io}\) and \(u_{io}\) represent the \(X\)-distance of node \(i\) and cluster centre \(o\), respectively, \(x_{io}\), \(x_{io}\) and \(x_{io}\) represent the \(Y\)-distance of node \(i\) and cluster centre \(o\), respectively.

When the cluster centre \(o\) is selected as the optimal location of EHs, the integrated load moment of electricity, gas and heat shown in (1) should be minimised. 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 utilisation of distributed energy and reduce the long-distance transmission loss, the minimisation of (4) means that the total capacity of generation nodes should be close to the total capacity of load nodes in each partition \(C_p\).

\[
\min \sum_{r=1}^{3} [x_{io} + u_{io}] \tag{4}
\]

\[
u_{io} = \frac{1}{|C_p|} \sum_{r=1}^{3} x_{io} \tag{5}
\]

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

When $u_{io} < 0$ 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 $i$ ($x_{io} > 0$) should be divided into the partition $C_o$ to minimise the (4).

When $u_{io} > 0$ 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 $i$ ($x_{io} < 0$) should be divided into the partition $C_o$ to minimise the (4).

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

### 3.2 EH-based clustering partition optimisation model

According to the above analysis, the optimisation goal of clustering partition should reflect the load moment from distributed nodes to the cluster centre, 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 centre $o$ is defined as:

$$L_{io} = a_i \sum_{r=1}^{3} a_{io} P_{o} D_{io} + a_i \sum_{r=1}^{3} |x_{io} + u_{io}|$$  \hspace{1cm} (6)

$$P_{o} = |x_{io}|$$

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

#### 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 minimise total equivalent load moments between node $i$ and cluster centre $o$:

$$\min : J = \sum_{o=1}^{k} \sum_{i \in C_o} L_{io}$$  \hspace{1cm} (7)

$$\text{s.t. : } 1 \leq i \leq m$$

$$1 \leq o \leq k$$

If there is a common connection point 0, the load moment from the cluster centre $o$ to the common connection point 0 should also be considered into the objective function, as:

$$\min : \sum_{o=1}^{k} \sum_{i \in C_o} L_{io} + \sum_{o=1}^{k} P_{o} D_{io}$$  \hspace{1cm} (8)

$$\text{s.t. : } 1 \leq i \leq m$$

$$1 \leq o \leq k$$

where $P_{o}$ represents the maximum transmission power of the common connection point 0 and $D_{io}$ represents the distance from the common connection point 0 to the cluster centre $o$.

### 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 the traditional $k$-means algorithm \[36, 37\]. For example, the number of cluster centres 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 centre optimisation and attribute weights is proposed to realise the adaptive clustering partition of large-scale distributed generations and loads, as well as the optimal location of EHs, as shown in Fig. 3.

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

#### 3.3.1 Cluster centre optimisation

The idea of the elbow method is to run $k$-means clustering for a given range of cluster centres $k$ 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 centres.

#### 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 favourable for energy balance can be divided into the same partition. The distributed nodes with slightly smaller load moments but unfavourable 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 centre near distributed nodes in each RIES is selected as the optimal location of EHs. Then, the hierarchical layout optimisation of RIES and multi-RIESs are carried out, respectively, to construct an economical and reliable interconnection network.
Hierarchical layout optimisation

4 Hierarchical layout optimisation

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

4.1 Dynamic radial layout optimisation of RIES based on penalised MST

To optimise 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 optimisation is implemented independently for different energy forms, but the coordinated conversion of different energy forms can be realised by the EH located in cluster centres.

As shown in Fig. 4, multiple distributed generations and loads in each RIES can be abstracted into a weighted graph $G_o = (V_o, E_o)$, where the cluster centre $o$ ($o = 1, 2, \ldots, k$) 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 the edge $w_{ij,r}$ ($r = 1, 2, 3$): When the dynamic radial layout optimisation is performed for a RIES, the load moment of energy form $r$ in (9) is taken as the penalty weight of edge $(ij)$, which reflects the transmission loss and construction cost of the energy form $r$.

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

(9)

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

4.1.2 Objective function: To minimise the transmission loss and construction cost, the dynamic radial layout optimisation of energy form $r$ between distributed nodes in each RIES $o$ is transformed into (10), which is used to find the penalised 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

$$ \min J_{o,r} = \sum_{(i, j) \in E_o} w_{ij,r} $$

s.t.:

$$ d_{o,r}^{P}(o) \leq n_o^{\text{inter}, r} $$

$$ 1 \leq r \leq 3 $$

$$ 1 \leq i, j \leq |C_o| $$

where $J_{o,r}$ represents the total load moments of energy form $r$ of the penalised MST for the RIES $o$, $n_o^{\text{inter}, r}$ represents the penalised MST of energy form $r$ in the graph $G_o$, $d_{o,r}^{P}(o)$ represents the internal connection degree of energy form $r$ of EH $o$, $n_o^{\text{inter}, r}$ 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, the Kruskal algorithm can find the MST only by sorting the edge weight once, which has higher efficiency. To minimise the transmission loss and construction cost, a penalty MST algorithm based on the Kruskal algorithm is proposed to realise the dynamic radial layout optimisation 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, $\beta_{o,r} = 1$, when $\beta_{o,r} > 1$, the distributed nodes directly connected to the cluster centre $o$ in energy form $r$ will decrease as the increase of $\beta_{o,r}$. When $0 < \beta_{o,r} < 1$ the distributed nodes directly connected to the cluster centre $o$ in energy form $r$ will increase as the decrease of $\beta_{o,r}$.
The layout optimisation of multi-RIESs is considered based on the balanced MST algorithm.

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

The layout optimisation of multi-RIESs is considered based on the optimal location of EHs. According to the estimation from the US Energy Information Administration, the average loss rate of electricity transmission is ≈5% [38]. The loss rate of gas transported through pipeline systems is ≈8% [39]. However, a high level of heat losses during the 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 optimisation of different energy forms between EHs is considered together, by minimising the integrated load moment of energy form and gas.

As shown in Fig. 6, multi-RIESs based on the EH are abstracted into a weighted graph \( G_{EH} = (V, E) \). The EHs are abstracted into nodes and constitute a node set \( V = \{1, 2, \ldots, k\} \), where \( k \) 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 \( w_{int} \): As shown in (11) and (12), the integrated load moment of electricity and gas is 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.

\[
\begin{align*}
\text{Fig. 6} & \quad \text{Layout optimisation of multi-RIESs based on the EH} \\
\text{Fig. 7} & \quad \text{Looped or meshed layout optimisation based on the balanced MST}
\end{align*}
\]

\[
w_{int} = \sum_{r=1}^{T} w_{int,r} \quad (11)
\]

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

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

4.2.2 Objective function: To minimise the transmission loss and improve the energy supply reliability, the layout optimisation 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 \( >2 \).

\[
\min \quad J_{EH} = \sum_{(m,n) \in T} w_{int} \\
\text{s. t.} : \quad d_{ex}^r(m) \leq 2 \\
\quad 1 \leq r \leq 2 \\
\quad 1 \leq m,n \leq k
\]

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 \( G_{EH} \), and \( d_{ex}^r(m) \) 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 minimise the transmission loss and construction cost, a balanced MST algorithm is proposed to realise the looped or meshed layout optimisation 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 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.

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 optimise 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 realise the local utilisation of distributed energy and reduce the long-distance transmission loss.

5.2 Hierarchical layout optimisation

According to optimisation results of the adaptive clustering partition, the partition 5 is selected as a case to analyse the dynamic radial layout optimisation 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 centre of partition 5 is the optimal location of EH 5.

The cluster centres of partitions 1–6 are obtained from optimisation 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 optimisation of multi-RIESs is carried out.

### 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 |
|------|--------|--------|--------------|------|--------|--------|--------------|------|--------|--------|--------------|
| 1    | 16.6   | 16.2   | 200          | 2    | 85.2   | 17.7   | 0            | 29   | 63.9   | 2      | -150        | 0      | 0      | -180        |
| 3    | 97.8   | 56.9   | 100          | 4    | 81.8   | 76.1   | -160         | 31   | 24.8   | 82.3   | 0            | -150   | 58     | 46.7         | 5.1    | 100        | 0      |
| 5    | 44.9   | 79.5   | -150         | 6    | 96     | 85.2   | -110         | 33   | 33.8   | 47     | 0            | -150   | 56     | 5.2          | 37.5   | 0          | -160   |
| 7    | 95.1   | 78.6   | 0            | 8    | 27.1   | 85.3   | 200          | 300  | 250    | 35     | 82.9         | 77.6   | 300    | 450          | 300    | 400        | 300    |
| 9    | 92.1   | 6.3    | -280         | 10   | 19.7   | 63.5   | -140         | 0    | 0      | 0      | -200        | 0      | 63     | 5.3          | 53.7   | 300        | 400    |
| 11   | 74.7   | 57.1   | -100         | 12   | 93.6   | 41.1   | 300          | 0    | 0      | 0      | -200        | 0      | 66     | 49.8         | 100.3  | 0          | -160   |
| 13   | 33.4   | 88.2   | -180         | 15   | 56.8   | 91.1   | -250         | 0    | 0      | 0      | -180        | 0      | 68     | 29.9         | 6      | 0          | -160   |
| 17   | 39.4   | 19.4   | 150          | 16   | 34.4   | 98     | 150          | 0    | 0      | 0      | -160        | 69     | 25.1   | 52.8         | -150   | 0          | 0      |
| 18   | 62.6   | 18.5   | 0            | 17   | 35.1   | 61.4   | 100          | 0    | 0      | -120   | 0            | 71     | 33.6   | 16.5         | -130   | 0          | 0      |
| 22   | 8.2    | 18.9   | 0            | 23   | 4.6    | 37.1   | 0            | -120 | 0      | 50     | 86.3         | 57.3   | -180   | 77           | 50.5   | 31.7        | -240   |
| 24   | 6.1    | 71.3   | 200          | 26   | 27.9   | 89.7   | 300          | 460  | 200    | 52     | 63.5         | 23.4   | 0      | -200        | 79     | 1.8         | 62.7    | -180   | 0          | 0      |
| 27   | 19.6   | 38.8   | 0            | 28   | 27.4   | 96.7   | -210         | 0    | 53     | 45     | 69.5         | 350    | 400    | 250          | 80     | 92.2        | 25.3    | 250    | 300        | 300    |
| 21   | 85.4   | 19.8   | 0            | 25   | 27.4   | 96.7   | 300          | 460  | 200    | 52     | 63.5         | 23.4   | 0      | -200        | 79     | 1.8         | 62.7    | -180   | 0          | 0      |
| 24   | 19.6   | 38.8   | 0            | 27   | 19.6   | 38.8   | 0            | -200 | 0      | 54     | 67.9         | 71.5   | 0      | 0            | -100   | 0          | 0      |

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.

Fig. 8 Elbow method showing the optimal k

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 optimise 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 realise the local utilisation of distributed energy and reduce the long-distance transmission loss.

5.2 Hierarchical layout optimisation

According to optimisation results of the adaptive clustering partition, the partition 5 is selected as a case to analyse the dynamic radial layout optimisation 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 centre of partition 5 is the optimal location of EH 5.

The cluster centres of partitions 1–6 are obtained from optimisation 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 optimisation of multi-RIESs is carried out.
5.2.1 Dynamic radial layout optimisation of RIES: The dynamic radial layout optimisation of RIES 5 based on the penalty MST is shown in Fig. 10. When $\beta_5 = 1$, as shown in Fig. 10a, the distributed nodes 15, 56, 58 and 70 containing energy form, electricity, and with smaller equivalent load moment to cluster centre 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 centre 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 centre 5 are directly connected to the EH 5 through heat pipes.

As shown in Figs. 10b and c, when $\beta_5 > 1$, the distributed nodes in RIES 5 tend to construct a penalty MST in a serial manner. As $\beta_5$ increases, the distributed nodes directly connected to the EH located in the cluster centre 5 gradually decreases. As shown in Figs. 10d and e, when $0 < \beta_5 < 1$, the distributed nodes in RIES tend to directly connect the EH and construct a penalty MST. As $\beta_5$ decreases, the distributed nodes directly connected to the EH located in the cluster centre 5 gradually increases, but which cannot be greater than the internal interface number of the EH 5. Therefore, the layout optimisation of RIES influenced by penalty factors, can be adjusted according to the internal interface number of EHS connecting distributed generation and load.

5.2.2 Looped or meshed layout optimisation 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. 11a. The RIESs 1, 2 and 4 corresponding to the EHs 1, 2 and 4 only have one connection degree, which is 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 optimised. In the optimised network layout

| Table 2 | System parameters of partition 5-RIES 5 |
|---------|----------------------------------------|
| Nodes  | $X$, m | $Y$, m  | $r = 1$ | $r = 2$ | $r = 3$ |
| 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 | $r = 1$ | $r = 2$ | $r = 3$ |
| 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     |

© The Institution of Engineering and Technology 2020
shown in Fig. 11b, the external connection degree of each EH is as balanced as possible, that is, the external connection degree of each node is not >2.

To improve the reliability of the energy supply, a looped or meshed layout is preferable. First, 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. 11c.

6 Conclusions

In order to construct an economical and reliable interconnection network of the integrated energy system, an adaptive clustering-based hierarchical layout optimisation 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 realises the optimal location of EHs and allocates each distributed generation and load to different EHs, forming multiple RIESs adaptively.

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

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).
Fig. 11 Looped or meshed layout optimisation of multi-RIESs based on the EH

(a) Traditional MST, (b) Balanced MST, (c) Balanced MST-based looped or meshed layout

8 References

[1] Pursiheimo, E., Holthinren, H., Koljonen, T.: ‘Path toward 100% renewable energy future and feasibility of power-to-gas technology in nordic countries’, IET Renew. Power Gener., 2017, 11, (13), pp. 1695–1705

[2] Wang, T., Nian, H., Zha, Z.Q.: ‘Using inverter-based renewable generators to improve the grid power quality—a review’, Chinese J. Electr. Eng., 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. Inf., 2018, 14, (3), pp. 3418–3427

[4] Li, Y., Zhang, P., Luh, P.B.: ‘Formal analysis of networked microgrids planning natural gas and electricity distribution networks’, IEEE Trans. Power Syst., 2013, 28, (4), pp. 4052–4063

[5] Massrur, H.R., Niknam, T., 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

[6] Alipour, M., Zare, K., Abapour, M., et al.: ‘A novel technique for the optimal design of large-scale photovoltaic power plants’, IEEE Trans. Power Syst., 2018, 33, (1), pp. 317–328

[7] Huang, L.L., Chen, N., Zhang, H.Y., et al.: ‘Planning multiple energy systems toward low-carbon society: a decentralized approach’, IEEE Trans. Smart Grid, 2019, 10, (5), pp. 4859–4869

[8] 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

[9] Qin, J., Dong, Z.Y., Zhao, J.H., et al.: ‘Multi-stage flexible expansion planning under uncertainties in a combined electricity and gas market’, IEEE Trans. Power Syst., 2015, 30, (4), pp. 2119–2129

[10] Qiu, J., Dong, Z.Y., Zhao, J.H., et al.: ‘Low carbon oriented expansion planning of integrated gas and power systems’, IEEE Trans. Power Syst., 2015, 30, (2), pp. 1035–1046

[11] Saldarriaga, C.A., Hincapie, R.A., Salazar, H.: ‘A holistic approach for planning natural gas and electricity distribution networks’, IEEE Trans. Power Syst., 2013, 28, (4), pp. 4052–4063

[12] Massrur, H.R., Niknam, T., Aghaei, J., et al.: ‘Collector system layout optimization framework for large-scale offshore wind farms’, IEEE Trans. Sust. Energy, 2016, 7, (4), pp. 1398–1407

[13] 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

[14] 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 Renew. Power Gener.*, 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 Gener. Transm. Distrib.*, 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 Gener. Transm. Distrib.*, 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 Gener. Transm. Distrib.*, 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’, *Int. J. Sci. Eng. Appl.*, 2016, **5**, (1), pp. 1–6

[35] Arora, P., Deepali, , Varshney, S., et al.: ‘Analysis of k-means and k-medoids algorithm for big data’, *Procedia Comput. Sci.*, 2016, **78**, pp. 507–512

[36] Manoharan, J.J., Ganesh, S.H.: ‘Initialization of optimized k-means centroids using divide-and-conquer method’, *ARPN J. Eng. Appl. Sci.*, 2016, **11**, (2), pp. 1086–1091

[37] Elbashony, 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 Eng. J.*, 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 Second Int. Conf. on Industrial Engineering, Applications and Manufacturing, Chelyabinsk Russia, May 2016, pp. 1–4