RESEARCH ARTICLE

Advanced Clustering Approach for Peer-to-Peer Local Energy Markets Considering Prosumers' Preference Vectors

GODWIN C. OKWUIBE [1,2], (Member, IEEE), AMIN SHOKRI GAZAFROUDI [3],
ESTHER MENGELKAMP [4], SARAH HAMBRIDGE[5], PETER TZSCHEUTSCHLER [1],
AND THOMAS HAMACHER[1]

[1]School of Engineering and Design, Technical University of Munich, 80333 Munich, Germany
[2]OLI Systems GmbH, 67376 Harthausen, Germany
[3]StromDAO GmbH, 69256 Mauer, Germany
[4]MK Consulting, 21365 Adendorf, Germany
[5]Grid Singularity, 10965 Berlin, Germany

Corresponding author: Godwin C. Okwuibe (godwin.okwuibe@tum.de)

ABSTRACT Local energy markets (LEMs) are utilized in a bottom-up power systems approach for reducing the complexity of the traditional, centralized power system and to enable better integration of decentralized renewable energy resources (RES). Peer-to-peer (P2P) energy trading creates opportunities for prosumers to trade their RES with other prosumers in the LEM. Although several scenarios were proposed in the literature for modelling P2P energy trading, there is still a gap in the literature considering the heterogeneous characteristics of prosumers' bidding preferences during P2P matching in the LEM. In this paper, we present heterogeneous characteristics of bidding preferences for prosumers considering energy quantity, bid/offer price, geographic location, location of agents on the local community and cluster welfare. Moreover, this paper proposes an advanced clustering model for P2P matching in the energy community considering the heterogeneous characteristics of bidding preferences for prosumers. For evaluating our proposed model performance, two German real case scenarios of a small and large communities were studied. The simulations results show that using price preference, as the criterion for clustering, offers more technical and economic benefits to energy communities compared to other clustering scenarios. On the other hand, clustering scenarios based on location of prosumers ensure that energy is traded among prosumers who are closer to each other.

INDEX TERMS Energy community, advanced clustering, local energy market, matching mechanism, peer-to-peer trading.

I. INTRODUCTION

A. MOTIVATION AND BACKGROUND

The transition from a fossil fuel-based centralized power system to a more sustainable, low-emissions and renewable-based system supported by the fast growth of distributed energy resources (DER) add higher complexity to the power system network [1]. Thus, because of the variability of the

The associate editor coordinating the review of this manuscript and approving it for publication was Salvatore Favuzza.

distributed energy resources, maintaining the reliability and stability of the the power system grid becomes more challenging and complex with the decentralization of assets. Local energy markets (LEMs) are introduced within the past two decades as a means to sustain grid balances at the distribution level and ensure that electricity is consumed closer to where it is produced [2], [3]. LEMs are platforms for trading locally sourced DERs among prosumers and consumers within a geographic and social neighborhood at the distribution level in a competitive and economic

efficient way [4], [5]. Notwithstanding the numerous research already conducted in the field of the LEM, predominant market mechanisms have not yet been developed and the LEM business model is still unclear [1]. Consequent to the numerous advantages it promised, including the opportunity for prosumers to decide whom to buy/sell their electricity from/to, peer-to-peer (P2P) energy trading has been proposed for local electricity trading by many researchers [6], [7].

B. LITERATURE REVIEW

P2P energy trading is a transactional structure which is based on prosumers and consumers directly or indirectly negotiating their electricity requirements to exchange electricity among each other [8]. In most countries, the electricity market is based on time series. Therefore, the consumers and prosumers are required to consistently bid/offer and negotiate their trade(s) every time slot to achieve a full P2P trading with direct negotiation between consumers and prosumers. This is time consuming and inefficient. Thus, modelling agents and P2P matching algorithms considering prosumers' energy requirements, choices, preferences and price are developed.

In terms of negotiation, P2P energy trading can be classified into mediated and non-mediated P2P negotiations [9]. In a mediated P2P negotiation, a neutral mediator expedites the mediation of the trade between the consumers and the prosumers [10]. This can be based on a many-to-one P2P negotiation between prosumers with an aggregator [11] or a virtual agent as a mediator [12]. Authors of [13] presented a many-to-one P2P energy trading where the distributions system operator mediates between prosumers and consumers and uses a double auction mechanism to decide the P2P winners of the auction. In a non-mediated or full P2P trading, prosumers and consumers directly negotiate with each other for electricity trading. This can be a one-to-one negotiation between a buyer and seller or many-to-many negotiation between all prosumers and consumers [10]. Ref. [14] proposed a one-to-one full P2P trading model based on bilateral trading coefficients. The authors of [15] proposed a many-to-many P2P energy trading model for local energy and flexibility trading based on decentralized negotiation.

Moreover, P2P energy trading can be classified based on the matching or decision approaches. Different approaches used for P2P energy trading are based on game-theory, algorithms, optimization models and reinforcement learning [7]. Ref. [16] proposed Stackelberg game-based energy sharing framework for P2P energy trading within in a multi-sharing region. Ref. [17] proposed a P2P multi-energy market mechanism for electricity and heat trading based on cooperative behaviors between the peers in an LEM. Authors of [18] in their work proposed a model based on Nash bargaining fair sharing of trading benefits for P2P energy trading in interconnected LEMs. Ref. [19] proposed a two-tier P2P trading model based on dual decomposition and distributed consensus mechanism for a double layer hierarchical LEM.

The authors of [20] proposed a P2P model, based on P2P coordination for energy. In recent research works, reinforcement learning is used for P2P model. Ref. [21] proposed a multi-agent deep reinforcement learning model based on the combination of the multi-agent deep deterministic policy gradient algorithm and technique of parameter sharing for P2P energy trading. Their simulation study show that the proposed model reduced daily electricity cost and demand peak.

Additionally, recent researches in the field of LEM focused on prosumers preference vectors. In order to consider and satisfy LEM participants' willingness, Ref. [22] proposed a new auction-based LEM model that takes into account the participants' preferences and willingness to pay more for a certain energy quality. The model promised to increase the coverage of local demand and local supply compared to the conventional periodic double auction mechanism. Authors of [14] proposed LEM model which provide opportunity for the prosumers to select their preferred trading partners in a P2P LEM platform. The work of Ref. [21] is based on grouping prosumers into different multiple clusters based on the type of their distributed energy resources(DERs). Ref. [23] proposed a two stage preference based merit-order market mechanism for valuation between green, local, and energy source. The model showed that different DERs supply prosumers according to their willingness to pay and market situation.

Clustering as an unsupervised learning algorithm is used in most studies to identify similarities between set of large data and group the large data into sets of smaller data with similar properties. Clustering is recently used in P2P energy trading to identify similarities in different groups of microgrid, and for coordinating trades in LEM [24]. This further results to proper decentralization of energy markets and access to renewable energy resources [24]. Ref. [25] proposed a virtual LEM that clusters prosumers daily based on their load profiles. Ref. [26] in their work proposed a game theoretic approach for clustering microgrids using particle swarm optimization in a P2P energy trading system. The concept of adaptive segmentation for P2P market clearing was proposed by Ref. [27]. The model use balanced k-means clustering algorithm to cluster energy players into different segments to negotiate energy trading separately among themselves.

The concept of clustering has been used mainly in power markets for classifying loads and forecasting the future electricity demands of the users within a cluster [28]. Ref. [29] used the known traditional K-means clustering to cluster electricity customers demand in order to make a forecast of the future customers load demand. Ref. [30] used the K-means and Fuzzy C-means clustering algorithms to model the electricity price time series patterns for forecasting future electricity price. Ref. [31] developed a semi-supervised automatic clustering algorithm based on a self-adapting metric learning process for determining the household electricity consumers demand patterns.

C. CONTRIBUTION AND ORGANIZATION

The literature contains several studies proposing different models for LEMs, P2P trading approaches, and prosumers preference vectors for LEM design. However, there is still a gap in literature concerning models considering the heterogeneous characteristics (i.e. preferences) of prosumers within the bids/offers during P2P matching. In this paper, we develop a non-mediated P2P matching model based on advanced clustering model that simultaneously takes into account the geographic location, location on the local community, bid/offer price, bid/offer quantity, and cluster welfare for prosumers in the energy community. The literature [32], [33], [34] contains several studies on clustering and application of clustering in energy systems, however, this is the first of its kind where clusters are built with developed weights instead of the distance or values of the sets. Hence, in this model, the developer can decide on a set of variables within the multi set that is of more importance to him or her and decide to manually give more weight to this variable before the start of clustering. By doing so, the model will perform the clustering while considering that a certain variable is of more importance to the model compared to others and therefore perform according to this instruction. The model was implemented on an interface and open source code-base of the Grid Singularity Exchange to simulate, and optimize energy trading in local communities [35]. Furthermore, we evaluate the model using performance indicators such as self-sufficiency, self-consumption ratio, share of market savings and traded energy quantity. Our model can be used for any multi set functions where the user wishes to cluster multi variables while attributing more importance to some variable compared to others. Thus the application of the developed novel clustering algorithm extends beyond LEM but also to other applications where similar multi sets need to be classified. The main contributions of the paper are summarized in the following:

- Proposing an advanced clustering algorithm for grouping of heterogeneous characteristics of bidding preferences for prosumers considering energy quantity, geographic location, location on the community and bid/offer price.
- Presenting novel P2P matching in energy communities based on our proposed advanced clustering algorithm.
- Implementation of the proposed model in a real case German community.
- Assessing the performance of our proposed P2P model based on performance indicators for the LEM.

The remaining sections of this work are structured as follows. Section II introduces clustering and the proposed advanced clustering algorithm. The proposed LEM model is described in Section III and sample discussion of the proposed model with example presented in Section IV. The community set-up, data and price components are presented in Section V. Section VI discusses the results of our case studies and the findings in details. Finally, the paper is concluded in Section VII.

II. PROPOSED CLUSTERING ALGORITHM

A. INTRODUCTION TO CLUSTERING

Clustering is an unsupervised learning algorithm with the objective of extracting underlying information of data samples and using the information to split the data into different groups so-called clusters [36]. Clustering is usually used for analyzing market research, pattern recognition, data analysis, image processing and categorizing genes with similar functionalities [37]. K-means clustering is one of the most popular clustering algorithms used in data mining because of its simplicity and computational efficiency [36], [38]. In K-mean clustering, the Euclidean[1] is used as a means to measure data that are nearby each other and as a means for determining the centroid for group of unlabelled data [38]. Hence, K-means clustering uses a local search to group the data sample by first; randomly selecting k points $\{\mu_1, \ldots, \mu_k\}$ as the corresponding initial centers for k clusters, then optimizing them iteratively until the objective function (Eq. 3) is minimized [39]. Supposed we have a multi-set of d-dimensional vector, X, then,

$$X = \{x_1, x_2, \ldots, x_p\}, \tag{1}$$

where p is number of observations and x_p is the p-th observation. The objective of K-means clustering is to cluster the p observations into k multi-sets of homogeneous clusters by minimizing the objective function represented by Eq. (3) [39], [40].

$$S = \{s_1, \ldots, s_k\}, \tag{2}$$

$$J = \sum_{i=1}^{k} \sum_{x_j \in S_i} ||x_j - \mu_i||_2^2, \tag{3}$$

Here, the number of clusters should be less than or equal to the number of observations ($k \leq p$). Moreover, x_j is the j-th observation belonging to s_i as i-th cluster. It is noticeable that in K-means clustering algorithm, the number of clusters must be defined initially. Hierarchical clustering is another popular clustering algorithm used recently in data mining. Hierarchical clustering algorithm consist of nested partitions in which homogeneous observations are grouped by recursively clustering only two observations at a time [41]. Unlike K-means clustering algorithm that has fixed number of clusters, the number of clusters in hierarchical clustering changes in every iteration cycle [42]. Hierarchical clustering is classified into agglomerative and divisive clustering algorithms which can simply be explained as bottom-up and top-down approaches of the clustering, respectively [42]. In an agglomerative clustering approach, the algorithm begins with a singleton observation by pairing similar clusters at each iteration. This process is repeated every iteration until all observations are included into a single cluster or a defined criterion is met [41]. On the other hand, in divisive clustering approach, clustering begins with one big cluster obser-

---

[1]Euclidean in this context is the study of solid geometry based on the work of Euclid and the corresponding elementary geometry.

vation and breaks down into clusters consisting of smaller observations in a hierarchical top-down manner. Thus, for p number of observations grouped into q sets of clusters, $\{c_1, c_2, \ldots, c_q\}$, using hierarchical clustering, it is given that

$$\{c_j \subset c_i\} \mid \{c_i \subset c_j\} \mid \{c_i \cap c_j\} = \emptyset, \quad \forall i \neq j, \quad (4)$$

hence, it is either one of clusters i and j is a subset of the other one hierarchically, or a disjoint of it [43].

In hierarchical clustering, different linkage criteria such as single linkage, complete linkage, unweighted pair group method average, the weighted pair group method average, the unweighted pair group method centroid, the weighted pair group method centroid, and the ward linkage are used to measure the distance between two clusters [44]. For the single linkage, the similarity between two clusters depends on the closest pair of members in the two clusters. Therefore, the closeness of the pairs in terms of distance are considered. However, for complete linkage, the similarity between two clusters depends on the farthest pair of members in the two clusters. Therefore, farther pairs are grouped first. For the unweighted pair group method average, the average distance between all individual clusters is calculated and clusters with the least average distance are joined to form the next hierarchy [44], [45]. For a weighted pair group method average, the hierarchical clustering are formed by joining to nearby clusters, the distance between the new formed cluster and any other cluster is the arithmetic mean of the average distances between members of the formed cluster and any other cluster [34]. The unweighted pair group method centroid is similar to unweighted pair group method average, however, for unweighted pair group method centroid, the proximity between two clusters is the proximity between their geometric centroids. Also, the weighted pair group method centroid is similar to weighted pair group method average, however, for weighted pair group method centroid, the proximity between two clusters is the proximity between their geometric centroids [33], [34]. For the wards linkage, two clusters are joined to form a hierarchy once they minimize the increase of sum of square error [34], [44], [45]. For instance in [32], the the unweighted pair average distance and ward linkage methods were used for grouping the clusters. In [33], the single linkage criterio was used. In [34], the single linkage, complete linkage, unweighted pair group method average, the weighted pair group method average, the unweighted pair group method centroid, the weighted pair group method centroid, and the ward linkage were used and compared.

B. ADVANCED CLUSTERING ALGORITHM

In this section, a novel advanced clustering algorithm is proposed. Theoretically, the proposed advance clustering algorithm is derived by combining k-means and agglomerative clustering algorithms. In this way, two different multi-sets, X and W, of d-dimensional vectors each with $\mathcal{N}$ and $\mathcal{M}$ number

of observations, are given in Eqs. (5) and (6), respectively.

$$X = \{x_1, x_2, \ldots, x_{\mathcal{N}}\} \quad (5)$$
$$W = \{w_1, w_2, \ldots, w_{\mathcal{M}}\} \quad (6)$$

1) STAGE I: INITIAL CLUSTERING

First, the sets from (5) and (6) are clustered into $\mathcal{K}$ and $\mathcal{U}$ number of homogeneous sets, by minimizing (7) and (8), respectively.

$$J^x = \sum_{k=1}^{\mathcal{K}} \sum_{x_y \in c_k^x} ||x_y - \mu_k||_2^2, \quad (7)$$

Subject to:
$$\mathcal{K} \leq \mathcal{N},$$

$$J^w = \sum_{u=1}^{\mathcal{U}} \sum_{w_g \in c_u^w} ||w_g - \mu_u||_2^2, \quad (8)$$

Subject to:
$$\mathcal{U} \leq \mathcal{M}.$$

Here, x_y and w_g are the y-th and g-th sets from the observations X and W, belonging to the k-th and u-th clusters, c_k^x and c_u^w, respectively. Hence, μ_k and μ_u are the initial cluster points for sets X and W as given in (5) and (6), respectively. Eq. (9) and (10) represent the solutions for (7) and (8), respectively.

$$C^x = \{c_1^x, c_2^x, \ldots, c_k^x\}, \quad (9)$$
$$C^w = \{c_1^w, c_2^w, \ldots, c_u^w\}, \quad (10)$$

where c_1^x to c_k^x and c_1^w to c_u^w represent the selected homogeneous clusters for sets X and W, respectively.

2) STAGE II: HIERARCHICAL CLUSTERING BASED ON LINKAGE CRITERIA

The homogeneous clusters from (9) and (10) are further clustered by determining the individual heterogeneous features of the clusters elements and pairing them one after the other, hierarchically. This is achieved by solving the optimization problem given in (11).

$$(c_{k*}^x, c_{u*}^w) = \arg\max \sum_{c_k^x \in C^x, c_u^w \in C^w} \omega_{k,u}(c_k^x, c_u^w) \quad (11)$$

where c_{k*}^x and c_{u*}^w are clusters selected from C^x and C^w, respectively, as solution of (11). Moreover, $\omega_{k,u}$ is the chosen metric which is derived based on the properties of the d-dimensional observation sets X and W. To solve the optimization problem, Eq. (11) is reduced to Eq. (12) to compute the weighted linkage ω^*, and (12), and (11) are solved iteratively until (9) and (10) are reduced to a single cluster.

$$\omega^* = \max\{\omega_{k,u}(c_k^x, c_u^w) : c_k^x \in C^x, c_u^w \in C^w\} \quad (12)$$

The pseudocode of our proposed advanced clustering is represented in Algorithm 1. The output $R_1, R_2, \ldots R_{\mathcal{N}*}$ is the result of matching two clusters c_{k*}^x and c_{u*}^w obtained as solution to Eq. (11).

Algorithm 1 Advanced Clustering Algorithm

Require: $X = \{x_1, x_2, \ldots, x_{\mathcal{N}}\}$, $W = \{w_1, w_2, \ldots, w_{\mathcal{M}}\}$.

Output: $\{R_1, R_2, \ldots R_{\mathcal{N}^*}\}$

Solve: Eq. (7) and Eq. (8), to obtain $C^x = \{c_1^x, c_2^x, \ldots c_k^x\}$ and $C^w = \{c_1^w, w_2^w, \ldots c_u^w\}$

  while (size(C^x x C^w) > 1) **do**
    Determine: ω^* from Eq. (12)
    Solve: Eq. (11) with solution from Eq. (12)
    Merge: Merge clusters of Eq. (9) and (10) based on solution from Eq. (11)
  end while

III. PROPOSED P2P MARKET MODEL

In this Section, our proposed P2P market model is presented. In the first step, each prosumer submits its offers and bids at each time slot to the energy community framework as shown in Fig. 1. Then, prosumers are clustered based on the proposed advanced clustering algorithm. P2P matching among prosumers occurs and the cluster weight matrix is updated. This process is repeated in the next time slot until end of the market as shown in Fig. 2. The motivation of clustering prosumers and consumers in P2P energy trading is to reduce mismatch during P2P matching, create cluster of markets where the members of the clusters have similar features and further reduce imbalance in the local community. By clustering prosumers, we reduce the big local market into smaller groups of markets. The participants in each of these smaller groups have similar features/properties, thereby creating opportunity for more bids and offers to be matched in the group levels which will lead to efficient LEM. Also, by creating clusters of members with the similar load profiles, the market reduce imbalance in the local community and ensure that more energy is matched at local level and that grid instability is reduced. Moreover, by creating clusters with similar bidding/offering strategies, the model ensure that there is no mismatch during the market clearing of bids and offers. Because of the establishment of previous works [6], [46], [47] in this area that P2P market for LEM trading offers economic and technical benefits to the LEM participants compared to merit order clearing mechanism, we use the P2P trading for our model. Also, P2P trading gives opportunity for consumers and prosumers to say how much they are willing to pay/receive per kW of their electricity thereby engaging both producers and consumers in the market which is the major idea of LEM unlike the merit order where only the producers gives their cost price and market is matched based on this. Also, P2P energy trading provides opportunity for more energy to be traded within the local community compared to merit order clearing mechanism.

A. STAGE I: BIDDING AND OFFERING BY PROSUMERS

For the proposed LEM model, the prosumers submit their bids and offers as a set of prosumers' preferences at each time slot to the local electricity market framework. Eq. (13)

represents the bid of buyer i at time slot t.

$$b_{i,t} = [p_{i,t}^b, q_{i,t}^b, l_i^g, l_i^d], \quad \forall i, t, \qquad (13)$$

where, $b_{i,t}$ represents the bidding vector of buyer i at time t which contains energy quantity ($q_{i,t}^b$), bid price ($p_{i,t}^b$), geographic location (l_i^g) of the buyer i in terms of latitude and longitude, and the location (l_i^d) of the buyer i in energy community. The location of the buyer, which is a prosumer in the energy community, is the area location consisting of the building block, street, zone and/or district location of the buyer. Hence, while l_i^g represents the distance location of the buyer, l_i^d represents the particular community the buyer belongs out of the energy communities located inside the LEM. Similar to Eq. (13), Eq. (14) represents the offer of seller j at time slot t.

$$s_{j,t} = [p_{j,t}^s, q_{j,t}^s, l_j^g, l_j^d], \quad \forall j, t, \qquad (14)$$

where, $s_{j,t}$ represents the offering vector of seller j which contains the energy quantity ($q_{j,t}^s$), the offer price ($p_{j,t}^s$), the geographic location (l_j^g) of the seller, and the location (l_j^d) of the seller on the energy community. Similar to buyers, l_i^g represents the distance location of the seller, and l_i^d represents the particular community the seller belongs out of the energy communities located inside the LEM. Thus, the bidding/offering vectors submitted by buyers /sellers to the LEM frame work at time slot t are represented by (15) and (16), respectively.

$$B_t = \{b_{1,t}, b_{2,t}, \ldots, b_{\mathcal{N},t}\}, \quad \forall t, \qquad (15)$$

$$S_t = \{s_{1,t}, s_{2,t}, \ldots, s_{\mathcal{M},t}\}, \quad \forall t, \qquad (16)$$

where B_t and S_t represent bidding and offering vectors, respectively. Moreover, $\mathcal{N}$ and $\mathcal{M}$ express number of buyers and sellers, respectively.

B. STAGE II: ADVANCED CLUSTERING OF BIDS AND OFFERS

1) STEP I: INITIAL CLUSTERING

After bidding and offering vectors are submitted to the LEM framework, they are clustered based on the proposed advanced clustering algorithm. This way, similar to K-means clustering algorithm, bidding and offering vectors are defined as multi-sets and are clustered independently based on the defined criterion by minimizing Eqs. (17) and (18), respectively:

$$J_t^b = \sum_{k=1}^{\mathcal{K}} \sum_{b_{f,t} \in c_{k,t}^b} ||b_{f,t} - \mu_{k,t}||_2^2, \qquad (17)$$

Subject to:
$$\mathcal{K} \leq \mathcal{N}.$$

$$J_t^s = \sum_{u=1}^{\mathcal{U}} \sum_{s_{g,t} \in c_{u,t}^s} ||s_{g,t} - \mu_{u,t}||_2^2, \qquad (18)$$
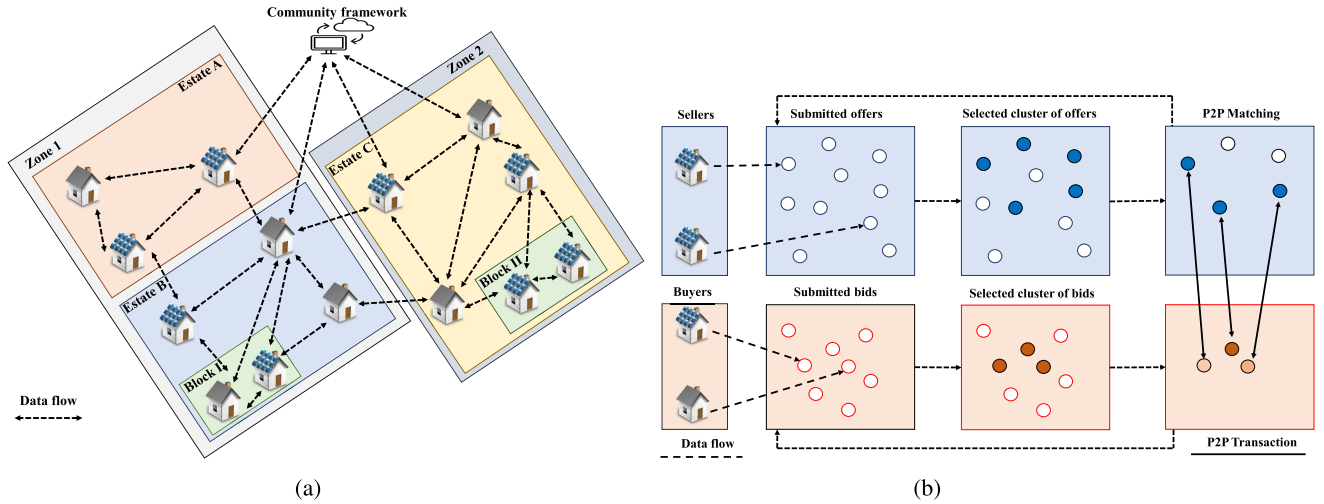
Subject to:
$$\mathcal{U} \leq \mathcal{M}.$$

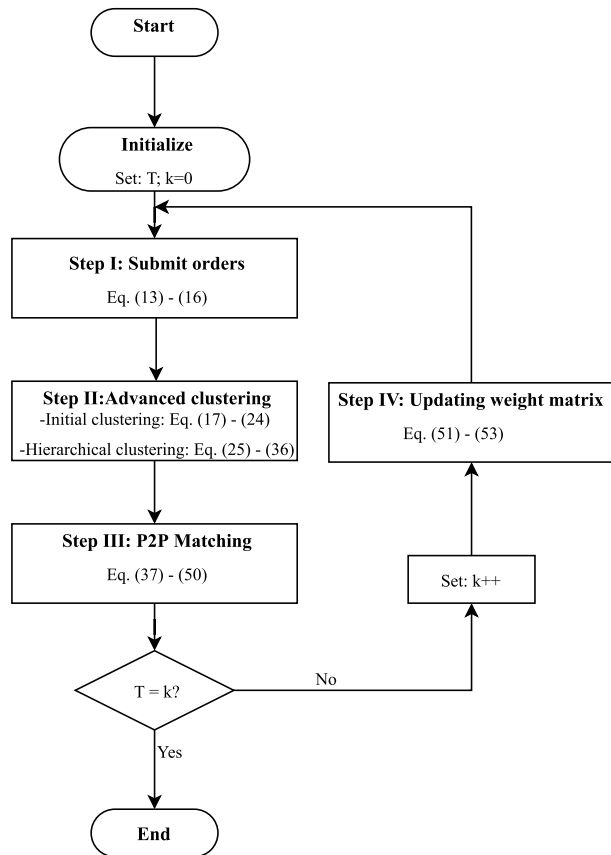**FIGURE 1.** Proposed (a) LEM design and (b) P2P market model.

FIGURE 2. Flowchart of the proposed P2P market model.

According to Eqs. (17) and (18), $\mathcal{K}$ and $\mathcal{U}$ are the number of clusters for the bids and offers, respectively. The number of defined bidding clusters $\mathcal{K}$, must be less than or equal to the total number of bids at time step t. Also, the number of the defined offering clusters $\mathcal{U}$, must be less than or equal to the total numbers of offers at time step t. Besides, $b_{f,t}$ and $s_{g,t}$

are the f-th bid and g-th offer, belonging to the k-th and u-th clusters, $c_{k,t}^b$ and $c_{u,t}^s$, respectively. Moreover, μ_k and μ_u are the initial cluster points for bids and offers, respectively. The initial cluster points are selected randomly and minimization of Eqs. (17) and (18) is performed iteratively until the optimal cluster points are determined. Thus, for initial k-th cluster of bids and u-th cluster of offers, it is given:

$$c_{k,t}^b = \{b_{1,t}^k, \ldots, b_{n^k,t}^k\}, \quad \forall t, \tag{19}$$

$$c_{u,t}^s = \{s_{1,t}^u, \ldots, s_{m^u,t}^u\}, \quad \forall t, \tag{20}$$

where n^k is the number of bids belonging to k-th cluster of bids, $c_{k,t}^b$, which have similar homogeneous features based on the clustering criterion at time step t. Similarly, m^u is the number of offers belonging to u-th cluster of offers, $c_{u,t}^s$, which have similar homogeneous features based on the clustering criterion at time step t. In our proposed model, different criterion are defined for clustering bids and offers based on the elements of bids and offers submitted by prosumers, as represented by Eqs. (13) and (14), respectively, and cluster welfare. For a buyer i in $c_{k,t}^b$ and seller j in $c_{u,t}^s$ at time slot t, with bid and offer as represented by Eqs. (13) and (14), respectively, the welfare (π) of the pair is represented in Eq. (21),

$$\pi = \begin{cases} (p_{i,t}^b - p_{j,t}^s) \times q_{j,t}^s: & q_{i,t}^b = q_{j,t}^s \\ (p_{i,t}^b - p_{j,t}^s) \times q_{j,t}^s: & q_{i,t}^b > q_{j,t}^s \\ (p_{i,t}^b - p_{j,t}^s) \times q_{i,t}^b: & q_{i,t}^b < q_{j,t}^s \end{cases}, \quad p_{i,t}^b \geq p_{j,t}^s, \ \forall i, j, t.$$
$$\tag{21}$$

For pairs with $p_{i,t}^b < p_{j,t}^s$, the welfare is not considered since there will not be matching of such pairs even after clustering. The social welfare of the clusters in Eqs. (19) and (20) termed cluster welfare is represented in Eq. (22). Hence, for our model, cluster welfare means social welfare of the clusters and not the social welfare of the local community. Hence, hierarchical clustering with cluster welfare as the chosen

metric criterion means clustering pairs with higher $\pi*$ using the hierarchical clustering algorithm/method in the next step.

$$\pi* = \sum_{i=1}^{n^k} \sum_{j=1}^{m^u} \pi_{i,j}, \qquad (22)$$

Additionally, Eqs. (23) and (24) represent the the set of all initial clusters for bids and offers:

$$C_t^b = \{c_{1,t}^b, \ldots, c_{k,t}^b\}, \quad \forall t, \qquad (23)$$
$$C_t^s = \{c_{1,t}^s, \ldots, c_{u,t}^s\}, \quad \forall t, \qquad (24)$$

where $c_{1,t}{}^b$ to $c_{k,t}{}^b$ and $c_{1,t}{}^s$ to $c_{u,t}{}^s$ represent the selected homogeneous initial clusters of bids and offers at time step t resulting from Eqs. (17) and (18), respectively. Eqs. (25) and (26) show that the selected initial clusters in Eqs. (23) and (24) are from the bids and offering vectors of buyers and seller represented in Eqs. (15) and (16), respectively.

$$B_t = \{c_{1,t}^b \cup \ldots \cup c_{k,t}^b\}, \quad \forall t, \qquad (25)$$
$$S_t = \{c_{1,t}^s \cup \ldots \cup c_{u,t}^s\}, \quad \forall t. \qquad (26)$$

2) STEP II: HIERARCHICAL CLUSTERING BASED ON LINKAGE CRITERIA

After initial clusters obtained by Eqs. (23) and (24), bids and offers are clustered hierarchically. Thus, initial clusters are inputs for the hierarchical clustering stage and its optimization problem represented by (27):

$$(c_{k*,t}^b, c_{u*,t}^s) = \arg\max_{c_{k,t}^b \in C_t^b, c_{u,t}^s \in C_t^s} \sum \omega_{k,u,t}(c_{k,t}^b, c_{u,t}^s), \quad \forall t, \qquad (27)$$

where $c_{k*,t}^b$ and $c_{u*,t}^s$ are clusters selected from C_t^b and C_t^s, respectively. Moreover, $\omega_{k,u,t}$ is the chosen metric and is defined as the weight for clustering pair of bidding cluster k and offering cluster u into a single cluster at time slot t. Accordingly, Ω is defined as a chosen weight matrix as represented in Eq. (28).

$$\Omega_{k \times u,t} = \begin{bmatrix} \omega_{1,1,t} & \cdots & \omega_{1,u,t} \\ \vdots & \ddots & \vdots \\ \omega_{k,1,t} & \cdots & \omega_{k,u,t} \end{bmatrix}, \quad \forall t. \qquad (28)$$

For reducing the computational burden, Eq. (27) can be restated as given in (29) and (30):

$$\omega_t^* = \max\{\omega_{k,u,t}(c_{k,t}^b, c_{u,t}^s) : c_{k,t}^b \in C_t^b, c_{u,t}^s \in C_t^s\}, \quad \forall t, \qquad (29)$$

$$(c_{k*,t}^b, c_{u*,t}^s) = \arg(\omega_t^*), \quad \forall t, \qquad (30)$$

where ω_t^* is the linkage criteria and is defined as the maximum value of the weighted linkage from the chosen weight matrix represented in Eq. (28). From (28), the chosen metric criteria, $\omega_{k*,u*,t}$, is defined as sum of the average weights of bidding and offering elements including bid and offer prices, energy quantity, location of the prosumers on the local

community for bids and offers clusters combination of $c_{k*,t}^b$ and $c_{u*,t}^s$, respectively, which are given in (31):

$$\omega_{k*,u*,t} = \overline{\omega}_{k*,u*,t}^p + \overline{\omega}_{k*,u*,t}^q + \overline{\omega}_{k*,u*}^l. \qquad (31)$$

From (31), $\overline{\omega}_{k*,u*,t}^p$ is the average price weight for possible matching bidding cluster, $c_{k*,t}^b$, and offering cluster, $c_{u*,t}^s$, as represented in the price weight matrix of (32) and given in (38).

$$\omega_{k*,u*,t}^p = \begin{bmatrix} w_{1,1,t}^p & \cdots & w_{1,m^{u*},t}^p \\ \vdots & \ddots & \vdots \\ w_{n^{k*},1,t}^p & \cdots & w_{n^{k*},m^{u*},t}^p \end{bmatrix}, \quad \forall t, \qquad (32)$$

where n^{k*} and m^{u*} are number of bids and offers in bidding cluster, $c_{k*,t}^b$, and offering cluster, $c_{u*,t}^s$, respectively. Moreover, $w_{i,j,t}^p$ is the price weight for matching bid $b_{i,t}$ belonging to $k*$-th bidding cluster and offer $s_{j,t}$ belonging to $u*$-th offering cluster at time slot t as represented by (33),

$$w_{i,j,t}^p = \begin{cases} 2: & p_{i,t}^b = p_{j,t}^s \\ 1: & p_{i,t}^b > p_{j,t}^s \\ -1: & \text{else} \end{cases}, \quad \forall i, j, t. \qquad (33)$$

From Eq. (33), a bid and an offer pair with equal bidding and offering price has the maximum weight which is 2, because, this pair provide the maximum satisfaction to both prosumers considering their bid and offer preference price. This pair also provide the maximum welfare for the both pairs while considering the bid/offer price of the prosumers. Therefore, giving the maximum price weight to prosumers with the same bid/offer price will ensure that more energy is traded within the community. For a bidding price greater than the offering price, the price weight is 1, because, this bid and offer pair provide less price preference satisfactions to both prosumers compared to the former. For a bidding price less than the offering price, the price weight is -1 because the bid and offer pair does not satisfy the sellers requirements and validation requirements and consequently, this bid and offer pair cannot match. This negative weight reduce the possibility of matching the bid and offer pair with sellers offer price greater than the buyers bid price.

Moreover, $\overline{\omega}_{k*,u*,t}^q$ and $\overline{\omega}_{k*,u*}^l$ represent weights of average energy quantity and location for possible matching of bidding cluster, $c_{k*,t}^b$, and offering cluster, $c_{u*,t}^s$, as represented by (34) and (36), respectively.

$$\omega_{k*,u*,t}^q = \begin{bmatrix} w_{1,1,t}^q & \cdots & w_{1,m^{u*},t}^q \\ \vdots & \ddots & \vdots \\ w_{n^{k*},1,t}^q & \cdots & w_{n^{k*},m^{u*},t}^q \end{bmatrix}, \quad \forall t, \qquad (34)$$

where $w_{i,j,t}^q$ represents the energy quantity weight of a buyer i and seller j belonging to clusters $c_{k*,t}^b$ and $c_{u*,t}^s$, respectively, as given by (35):

$$w_{i,j,t}^q = \begin{cases} 2: & q_{i,t}^b = q_{j,t}^s \\ 1: & \text{else} \end{cases}, \quad \forall i, j, t. \qquad (35)$$

From Eq. (35), bid and offer pair with equal bidding and offering energy quantity has the maximum weight which is 2, because, this pair provides the maximum energy quantity preference satisfaction to both prosumers considering their bid and offer energy quantity. Thus, there is high probability that the buyer and seller will trade all their energy requirements with each other without requesting for another prosumer to take care of their energy requirements that is not satisfied by their pair. For pair with bidding energy quantity greater than the offering energy quantity or offering energy quantity greater than the bidding energy quantity, the price weight is 1, because, this bid and offer pair provide less energy quantity preference satisfactions to one of the prosumers unlike the former that provide maximum energy quantity preference satisfaction to both prosumers. Also, the prosumer with higher energy quantity requirements will still be paired in the next iteration meaning that all the energy requirements cannot be satisfied with the pair and hence, the reason for lesser energy quantity weight.

Additionally, Eq. (37) represents the location weight ($w_{i,j}^l$) of a buyer agent i and seller agent j belonging to clusters $c_{k*,t}^b$ and $c_{u*,t}^s$, respectively.

$$\omega_{k*,u*}^l = \begin{bmatrix} w_{1,1}^l & \cdots & w_{1,m^{u*}}^l \\ \vdots & \ddots & \vdots \\ w_{n^{k*},1}^l & \cdots & w_{n^{k*},m^{u*}}^l \end{bmatrix}, \quad (36)$$

$$w_{i,j}^l = \begin{cases} 4: & \{l_i \cup l_j\} \subset l^b \\ 3: & \{l_i \cup l_j\} \subset l^e \\ 2: & \{l_i \cup l_j\} \subset l^z \\ 1: & \{l_i \cup l_j\} \subset l^c \\ 0: & \text{else} \end{cases}, \quad (37)$$

where l^b, l^e, l^z and l^c represent apartment block, estate, zone and local community, respectively. In other words, Eq. (37) expresses that the closer the seller and buyer are on the local community, the higher their location weight.

Thus, the average weight for all bidding/offering elements is determined by Eq. (38):

$$\overline{\omega}_{k*,u*}^{(.)} = \frac{1}{n^{k*} \times m^{u*}} \sum_{i=1}^{n^{k*}} \sum_{j=1}^{m^{u*}} w_{i,j}^{(.)}, \quad (38)$$

where $w_{i,j}^{(.)}$ can be weight for price (p), energy quantity (q) and location (l). In this way, the average weights calculated by Eq. (38) will be the inputs of (31).

3) STEP III: P2P MATCHING OF CLUSTERED BIDS AND OFFERS

After prosumers are clustered in bid and offer clusters separately, P2P matching occur among prosumers which their bids and offers belong to the selected clusters. This way, considering $c_{k*,t}^b$ and $c_{u*,t}^s$ are selected as optimum bidding and offering clusters from Eq. (30) which are represented by

Eqs. (39) and (40), respectively.

$$c_{k*,t}^b = (b_{1,t}^k, \ldots, b_{n^{k*},t}^k), \quad \forall t, \quad (39)$$

$$c_{u*,t}^s = (s_{1,t}^u, \ldots, s_{m^{u*},t}^u), \quad \forall t, \quad (40)$$

where $c_{k*,t}^b$ contains n^{k*} number of bids and $c_{u*,t}^s$ contains n^{u*} number of offers at time t. Besides, bid i and offer j are vectors consisting of their corresponding bidding and offering elements as represented by Eqs. (13) and (14). Thus, prosumer i as a potential buyer and prosumer j as a potential seller are eligible to negotiate with each other directly for making P2P transaction because they belong to clusters $c_{k*,t}^b$ and $c_{u*,t}^s$, respectively. Hence, the matching among prosumers i and j can occur if Eqs. (41) and (42) as P2P validation criteria are satisfied.

$$p_{i,t}^b \geq p_{j,t}^s + g_{i,j,t}^l, \quad \forall i, j, t, \quad (41)$$

$$(l_i^d \subset l) \wedge (l_j^d \subset l) = 1, \quad \forall i, j, \quad (42)$$

where $g_{i,j,t}^l$ represents grid fee of P2P energy transaction among prosumers i and j in the local community l where l_j and l_i exist. In other words, the grid fee is amount of money in cent/kWh which the paired prosumers i and j pay for using the local grid for energy exchange in the energy community l. Eq. (43) represents the community grid fees,

$$g_{i,j,t}^l = \begin{cases} g_t^{lb}: & \{l_i \cup l_j\} \subset l^b \\ g_t^{le}: & \{l_i \cup l_j\} \subset l^e \\ g_t^{lz}: & \{l_i \cup l_j\} \subset l^z \\ g_t^{lc}: & \{l_i \cup l_j\} \subset l^c \\ g_t^{lu}: & \text{else} \end{cases}, \quad \forall i, j, t, \quad (43)$$

where $g_t^{lb}, g_t^{le}, g_t^{lz}, g_t^{lc}$ and g_t^{lu} represent grid fees for apartment blocks, estates, zones, local community and the upstream grid, respectively. Each grid fee is the combination of metering fee (g^m), local grid fee ($g_{i,j,t}$), upstream grid fee (g_t^u) and 19% value-added-tax (VAT) as represented in Eq. (44)

$$g_t^{l*} = g^m + g_{i,j,t} + g_t^u + 0.19 \times (p_{i,j,t}^{p2p} + g^m + g_{i,j,t} + g_t^u); \quad \forall i, t, \quad (44)$$

If Eqs. (41) and (42) are satisfied among prosumers i and j, the selected pair of bid i and offer j are matched. Thus, $q_{i,j,t}$ and $p_{i,j,t}^{p2p}$ represent matched P2P energy quantity and price between buyer i and seller j at time slot t which are given in (45) and (46), respectively.

$$q_{i,j,t} = \begin{cases} q_{i,t}^b: & q_{j,t}^s > q_{i,t}^b \\ q_{j,t}^s: & \text{else} \end{cases}, \quad \forall i, j, t, \quad (45)$$

$$p_{i,j,t}^{p2p} = \frac{p_{j,t}^s + p_{i,t}^b}{2}, \quad \forall i, j, t, \quad (46)$$

$$p_{i,j,t}^b = p_{i,j,t}^{p2p} + \frac{g_{i,j,t}^l}{2}, \quad \forall i, j, t, \quad (47)$$

$$p_{j,i,t}^s = p_{i,j,t}^{p2p} - \frac{g_{i,j,t}^l}{2}, \quad \forall i, j, t, \quad (48)$$

where $p_{i,j,t}^b$ is the bought price for buyer i to trade P2P energy with seller j, and $p_{i,j,t}^s$ is the sold price for seller j to trade P2P energy with buyer i at time slot t as represent by Eqs. (47) and (48), respectively. After P2P matching between seller i and buyer j, the bid and offer are updated by subtracting the matched energy quantity from the submitted bid/offer quantity as represented in (49) and (50), respectively.

$$q_{i,t}^{b'} = q_{i,t}^b - q_{i,j,t}; \quad \forall i, t, \quad (49)$$

$$q_{j,t}^{s'} = q_{j,t}^s - q_{i,j,t}; \quad \forall j, t, \quad (50)$$

where $q_{i,t}^{b'}$ and $q_{j,t}^{s'}$ is the deficit/excess bid and offer energy quantity for buyer i and seller j respectively, after P2P matching between them. In case, the selected bid and offer does not satisfy constraints (41) and (42), another bid and offer pair ($b_{i^*,t}$ and $s_{j^*,t}$) are selected according to (51) and (52), respectively.

$$b_{i^*,t} = [p_{i^*,t}^b, q_{i^*,t}^b, l_{i^*}^g, l_{i^*}^d]; \{b_{i^*,t} \subset c_{k^*,t}^b, i^* \neq i\}, \quad \forall t, \quad (51)$$

$$s_{j^*,t} = [p_{j^*,t}^s, q_{j^*,t}^s, l_{j^*}^g, l_{j^*}^d]; \{s_{j^*,t} \subset c_{u^*,t}^s, j^* \neq j\}, \quad \forall t. \quad (52)$$

Eqs. (51) and (52) imply that bids and offers paired earlier cannot be paired again in the same time slot, t. The proposed matching process from Eqs. (39) to (52) is repeated until all the bids and offers are selected.

4) STEP IV: UPDATING THE CHOSEN WEIGHT MATRIX

After matching bidding and offering clusters $c_{k^*,t}^b$ and $c_{u^*,t}^s$, respectively, according to Section III-B3, the chosen weight matrix will be updated. In this way, if all energy quantity of both bid i and offer j are fully matched, the chosen weight matrix is updated as given in Eq. (53) by deleting the corresponding bid's column and offer's row stated in Eq. (28). In other words, Eq. (28) is reduced to Eq. (53) by deleting matched bidding cluster, $c_{k^*,t}^b$, and offering cluster, $c_{u^*,t}^s$, from Eqs. (23) and (24), before solving (27) in the next iteration.

$$\Omega_{k \times u, t} = \begin{bmatrix} \omega_{1,1,t} & \cdots & \omega_{1,u-1,t} \\ \vdots & \ddots & \vdots \\ \omega_{k-1,1,t} & \cdots & \omega_{k-1,u-1,t} \end{bmatrix}, \quad \forall t. \quad (53)$$

However, if only the offered energy quantity of the cluster $c_{u^*,t}^s$ is fully cleared while the bid energy quantity of $c_{k^*,t}^b$ is not cleared completely, then Eq. (28) is updated to (54) by deleting the corresponding offer's row stated in Eq. (28).

$$\Omega_{k \times u, t} = \begin{bmatrix} \omega_{1,1,t} & \cdots & \omega_{1,u-1,t} \\ \vdots & \ddots & \vdots \\ \omega_{k,1,t} & \cdots & \omega_{k,u-1,t} \end{bmatrix}, \quad \forall t. \quad (54)$$

On the other hand, if only the energy quantity of the bid cluster $c_{k^*,t}^b$ is cleared completely, Eq. (28) is updated to Eq. (55) by deleting the corresponding bid's column stated

in Eq. (28).

$$\Omega_{k \times u, t} = \begin{bmatrix} \omega_{1,1,t} & \cdots & \omega_{1,u,t} \\ \vdots & \ddots & \vdots \\ \omega_{k-1,1,t} & \cdots & \omega_{k-1,u,t} \end{bmatrix}, \quad \forall t. \quad (55)$$

Finally, if there are unmatched energy quantities of bid and offer within both ($c_{k^*,t}^b$ and $c_{u^*,t}^s$) clusters, the size of Eq. (28) remains unchanged. However, matched bids and offers are deleted leaving unmatched ones. Then the chosen weight matrix from Eq. (28) is computed again based on Eqs. (32), (34) and (36), and used to solve Eqs. (29) and (27). The process is repeated until the chosen weight matrix, represented in Eq. (28), forms a single matrix in which the input clusters cannot satisfy the matching criterion described in Eqs. (41) and (42), or all the bids and offers are fully matched. All deficit/excess bids and offers which are not cleared in the local energy market, will be traded with the utility using the upstream-grid price. The pseudocode and flowchart of our proposed advanced clustering approach for a single time slot t is shown in Algorithm 2 and Fig. 3, respectively. From Algorithm 2, $R_1, R_2, \ldots R_{\mathcal{N}^*}$ is the output of the algorithm which is the result of matching two clusters $c_{k^*,t}^b$ and $c_{u^*,t}^s$ based on Eqs. (39) to (52).

Algorithm 2 Advanced Clustering Algorithm

Require: $B_t = \{b_{1,t}, b_{2,t}, \ldots, b_{\mathcal{N},t}\}$, $S_t = \{s_{1,t}, s_{2,t}, \ldots, s_{\mathcal{M},t}\}$.   ▷ Bids and offers
Ensure: $M_{criterion} = True$ ▷ All bids and offers contain no zero or empty set
Output: $\{R_1^{b,s}, R_2^{b,s}, \ldots R_{\mathcal{N}^*}^{b,s}\}$
Solve: Eq. (17) and Eq. (18), to obtain $C_t^b = \{c_{1,t}^b, \ldots, c_{k,t}^b\}$ and $C_t^s = \{c_{1,t}^s, \ldots, c_{u,t}^s\}$

while ($M_{criterion} = True \mid$ (size($C_t^b \times C_t^s$) > 1)) **do**
   Calculate: Eq. (28)
   Determine: ω^* and ($c_{k^*,t}^b, c_{u^*,t}^s$) from Eqs. (29) and (30), respectively.
   Matching: Match the clusters ($c_{k^*,t}^b, c_{u^*,t}^s$) based on Eq. (39) to (52).
   if : All ($c_{k^*,t}^b$ & $c_{u^*,t}^s$) are matched **then**
     Reduce Eq. (28) to Eq. (53);
   else if : All ($c_{u^*,t}^s$) are matched **then**
     Reduce Eq. (28) to Eq. (54);
   else if : All ($c_{k^*,t}^b$) are matched **then**
     Reduce Eq. (28) to Eq. (55);
   else:
     Delete individual matched bids and offers;
   end if
end while

IV. PROPOSED P2P MARKET MODEL DISCUSSION
In this Section, our proposed P2P market model is described with a simple example to illustrate how bids and offers are clustered and matched according to prosumers' preferences.
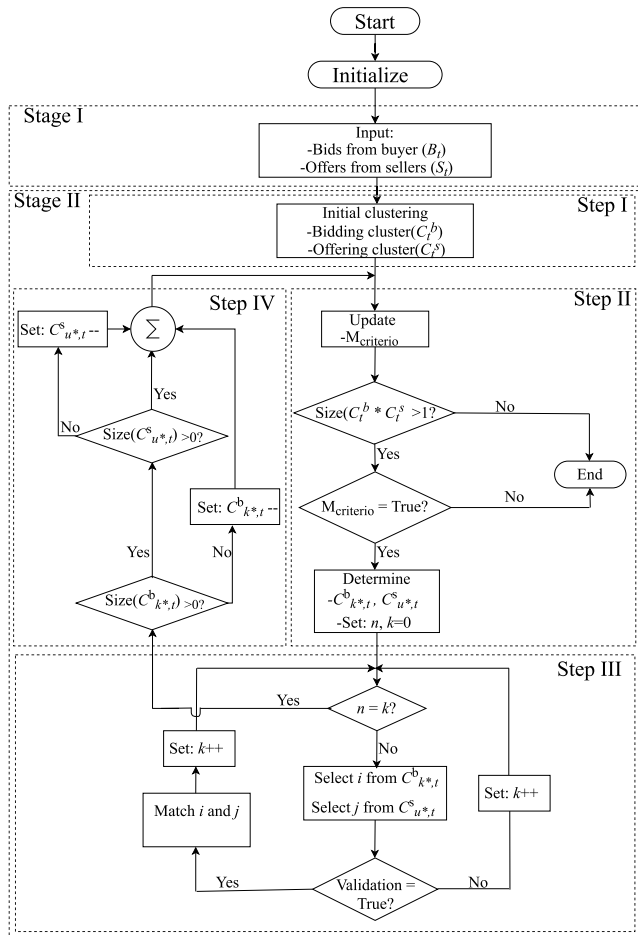
FIGURE 3. Flowchart of the proposed advanced clustering model.

Table 1 displays the sample of prosumers' bids and offers containing the prosumers preference vector for a single time step. The preference vector contains the prosumer's identification (ID), bid/offer price, energy quantity, geographic location and location of the prosumer on the energy community. The geographic location is the location of the prosumer in terms of latitude (La) and longitude (Lo). On the other hand, location of the prosumer in the energy community contains the its zone, estate and building block. According to Table 1, there are two zones consisting of zone A (ZA) and zone B (ZB). Moreover, there are three estates namely yellow (YE), ash (AS), and orange (OR), and six building blocks B1 to B6. In this way, as shown in first row of Table 1, ZA-YE-B1 means that the prosumer 1 belong to zone A, yellow estate and B1.

Fig. 4 displays the clustering and matching of the sample bids and offers of Table 1 based on the proposed P2P market model using advanced clustering algorithm. The bids and offers are arranged according to the bidding/offering price and energy quantity as shown on the lowest level of Fig. 4. The bids are b_1 to b_5, and the offers are s_1 to s_4. In the next level, bids and offers are clustered based on the initial clustering algorithm described in Section III-B1 with cluster welfare

as the clustering criterion. The number of clusters for bids and offers are set as two for each of them. The initial clusters is shown on the second level from the bottom layer of Fig. 4. For the bids, the initial clusters are represented by c_1^b and c_2^b, while the initial clusters of offers are represented by c_1^s and c_2^s. Afterward, first two optimal clusters are selected from the four initial clusters by solving Eq. (27). In this sample, the first selected optimal clusters are c_1^s and c_2^b. This way, bids and offers within these selected clusters are matched using the matching algorithm described in Section III-B3 and forms a single cluster $c_{2,1}^s$ after matching. During matching, the first negotiation is between prosumers b_4 and s_4 because of their close price and energy quantity preference. Consequently, 2.85kWh is traded between both prosumers at 23.75 ct./kWh. Hence, b_4 is fully matched while s_4 is partially matched with a left over of 0.16kWh. In the same way, b_5 and s_3 negotiate to exchange 3.12kWh at 24.15ct./kWh. This results in s_3 been fully matched and b_5 been partially matched with a left over of 0.08kWh. Then, b_5 and s_4 negotiate to trade their left over. This results in 0.08kWh been traded among b_5 and s_4 at 25.5 ct./kWh. Hence, b_5 is now fully matched while s_4 is partially matched with a left over of 0.08kWh. Consequently, the formed single cluster contains two offers s_2 and s_4. While s_4 is partially matched, s_2 is not matched at all. Here, $c_{2,1}^s$ represents a combined cluster of bidding cluster c_2^b represented by first subscript 2, and offering cluster c_1^s represented by second subscript 1 and they both formed an offering cluster represented by superscript s. The fist hierarchy now contains three clusters namely $c_{2,1}^s$, c_2^s and c_1^b. For the next iteration, two clusters are selected from three clusters in the first hierarchy by solving Eq. (27) and are termed second selected optimal clusters. The second selected optimal clusters are $c_{2,1}^s$ and c_1^b. Bids and offers within these selected clusters are matched using the same matching algorithm described in Section III-B3 and forms a single cluster $c_{1,(2,1)}^b$ after matching. $c_{1,(2,1)}^b$ contains only one bid, b_1, which is partially matched. At this stage, only two clusters consisting of c_2^s and $c_{1,(2,1)}^b$ are left and form the second hierarchy. These two clusters are selected and matched to form the single cluster $c_{(1,(2,1)),2}^s$ which contains a single offer s_1 that is partially matched. In this way, the final cluster, $c_{(1,(2,1)),2}^s$, will trade with the upstream grid using the upstream grid price as it was not able to match in the local community.

Thus, our model considers similarities based on bidding/offering prices and load profiles. This way, two prosumers in an LEM with similar profiles will be able to trade P2P energy with each other. If the bidding price of the buyer is less than the offering price of the seller, there will be no matching notwithstanding the similarities in the load profiles. This shows the importance of bidding/offering price in an LEM. Bidding/offering price just likes prosumer profiles has effect in LEM trading which if not considered will lead to mismatch in trading and further result in trading with the upstream grid thereby making the market uneconomical and even cause grid instability. As our trading time step

TABLE 1. Prosumers bid/offer preferences.

S/N	ID	Price (ct./kWh)	Quantity (kWh)	Type (Bid/Offer)	Geographic location (La, Lo)	Energy Community
1	s_1	20.05	1.1	Offer	48.717548, 9.26708901	ZA-YE-B1
2	s_2	22.20	2.25	Offer	48.71766159, 9.26803016	ZA-AS-B2
3	s_3	20.80	3.12	Offer	48.71665553, 9.26887395	ZB-OR-B4
4	s_4	23.50	3.01	Offer	48.72076089, 9.26598559	ZB-OR-B5
5	b_1	25.45	0.42	Bid	48.72327604, 9.26718637	ZA-YE-B3
6	b_2	25.00	1.20	Bid	48.72433078, 9.27088607	ZA-YE-B1
7	b_3	27.00	1.00	Bid	48.72397379, 9.27069135	ZA-YE-B6
8	b_4	24.00	2.85	Bid	48.72325981, 9.27184345	ZB-OR-B4
9	b_5	27.50	3.20	Bid	48.71775895, 9.26848451	ZB-OR-B5

FIGURE 4. Clustering of bids/offers based on our proposed advanced clustering.

economic goal of social welfare which is mainly creating additional savings for the local participants. This additional savings results in increase in the share of market savings of the local prosumers participating in the local energy trading. Moreover, clustering is based on considering the social welfare of clusters as defined in Eqs. (21) and (22) and not the social welfare of the local community. In comparison to existing clustering algorithms, the proposed clustering algorithm provides opportunity for the LEM operator managing the local market to give more importance to certain prosumers preferences that is more importance to the prosumers. Furthermore, the algorithm converges faster in lesser time when compared to the traditional clustering algorithms like K-means clustering.

V. SIMULATION SET-UP

A. SIMULATION FRAMEWORK

The proposed LEM advanced clustering model is developed as a Python code and implemented by integrating with the open-source Grid Singularity Exchange (GSy-E) [35], [48], [49] and bidding agent application programming interface (API) as represented in Fig. 5. Each consumer and prosumer is represented by a bidding/offering agent which communicates their bids and offers/bids, respectively, to the exchange engine. The exchange engine is responsible for (i) receiving the bids/offers from the bidding/offering agents of the prosumers, (ii) storing them in the ordered book, and (iii) sending them for onward clearing by the matching API. The matching API is responsible for matching the bids/offers among prosumer agents using the proposed advanced clustering model developed in this work. All unmatched bids and offers are transacted with the up-stream grid agent using the upstream grid price. In other words, any offer and bid of prosumers which is not matched with other prosumers in the energy community is traded with the up-stream grid. In our simulation model, each prosumer communicates their bids or offers individually to the exchange engine every 15 minutes time slot before the energy exchange time. Afterward, the results from the matching API is sent back to individual prosumers.

B. COMMUNITY SET-UP AND DATA

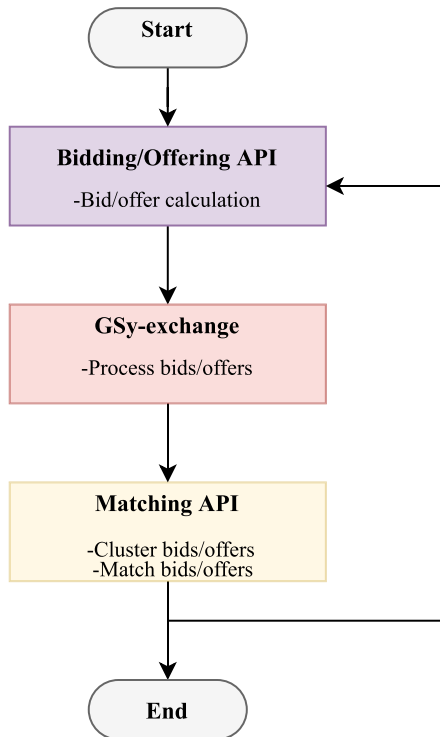The proposed P2P LEM approach is verified in two simulation case studies for a period of one day. Case study I is a

(15 minutes) is very short, uncertainty was not considered as it does not have high impact on the LEM. In the LEM, prosumers are allowed to take strategic actions that will benefit them and the market. However, they are not allowed to take actions that will result in inequality.

Clustering brings together prosumers into pools of their fellow prosumers where they all posses similar characteristics. This results in reducing mismatch in the LEM as clusters contain pairs with similar bidding/offering preferences which is used for the P2P matching. Clustering further ensures that more energy is traded at the local levels and creates more local economy for energy community. Hence, considering heterogeneous characteristics of prosumers is relevant to the

```
┌─────────────┐
│    Start    │
└─────────────┘
       │
       ▼
┌────────────────────────┐
│  Bidding/Offering API  │
│                        │◄──┐
│  -Bid/offer calculation│   │
└────────────────────────┘   │
       │                     │
       ▼                     │
┌────────────────────────┐   │
│     GSy-exchange       │   │
│                        │   │
│  -Process bids/offers  │   │
└────────────────────────┘   │
       │                     │
       ▼                     │
┌────────────────────────┐   │
│     Matching API       │   │
│                        │   │
│   -Cluster bids/offers │   │
│   -Match bids/offers   │───┘
└────────────────────────┘
       │
       ▼
┌─────────────┐
│     End     │
└─────────────┘
```

FIGURE 5. Flowchart of simulation framework for proposed model.

community with 25 prosumers consisting of 15 households with only consumption devices and 10 households with consumption and production devices. Case study II is a community with 120 prosumers consisting of 68 households consumers, 4 commercial consumers, 7 commercial prosumers, 6 industrial prosumers and 35 household prosumers. For both case studies, the load profiles are combinations of profiles from [50], LoadProfileGenerator [51], [52] and standard load profiles [53], [54]. These two case studies are chosen to show an in-depth and clear analysis of the advantage/benefits of our proposed model using case study I and further prove its application in a real-case LEM using case study II. Besides, the PV production profiles are from Renewables Ninja [55], [56] using Stuttgart region as a community. The PV system losses are also varied between 5% and 15% with a constant tilt angle of 35°. For prosumers with storage systems, the storage capacities are between 7.5 kWh to 13.5 kWh with maximum absolute power between 3.5 kW to 5.25 kW. Additionally, two community storage systems of capacities 140 kWh and 120 kWh, with maximum power of 40 kW are included in the local community of case study II. The minimum allowed state of charge for all storage systems is 10%. The profiles are randomly fitted into a community (Scharnhausen) in Stuttgart region as shown in Fig. 6(a) and (b) for case studies I and II, respectively, using the OpenStreet map and QGIS. In this study, prosumers in the same building like P1, C1 and P2, are considered to be in the same block. Moreover, prosumers in the same neighbourhood (street) like P1, P4 and C2 are considered to be in the same estate, and prosumers in the same

post code (e.g P1, C6, C9 and P3) are considered to be on the same zone.

The price components of the simulation case studies are presented in Table 2. The price components of the local electricity market consist of energy price of the electricity producer, metering fee, local grid fee and 19% Value-Added-Tax (VAT) [57]. The energy price is the minimum amount of money (cents/kWh) a seller prosumer is willing to receive per kWh of electricity produced and traded with other prosumers. The metering fee is the surcharge paid by the consumer for maintaining the metering infrastructure while the local grid fees are for maintaining the local distribution grid infrastructure. The VAT is the tax paid for transacting electricity between prosumers. It is noted that the VAT is 19% of the sum of energy price, metering, local and up-stream grid fees. Additionally, the up-stream grid fee is added if the electricity is traded with the up-stream grid. Because of the average cost of electricity in Germany, the maximum cost of electricity which is the cost of buying electricity from the grid is capped at 31.5 cents/kWh [58]. Moreover, the minimum cost of electricity is 11.00 ct/kWh which is the feed-in tariff of PV. Hence, consumers (also prosumers during bidding) set their preferred bid price between 11.38 to 29.54 ct/kWh, that is the range of total cost for buying energy within a block. However, prosumers (during offering) set their preferred offering price between 11 to 24.5 ct/kWh. This is the range of energy price for selling energy within a block as shown in Table 2. The range of energy price is between the feed-in tariff price and the maximum price that prosumers can sell their energy within the community. Each prosumer within the LEM has an agent which is a software code that submit their bidding/offering vectors on their behalf to the LEM framework. The bidding price is chosen randomly by the agent within the range provided by the prosumer while the energy quantity is selected from the consumption/production profile of the prosumer. The agents on behalf of the prosumers and consumers can only send one bid or offer per time slot. Furthermore, electricity which is not traded within the LEM is exchanged with the up-stream grid at the grid price which is less profitable compared to trading within the LEM.

VI. RESULTS AND DISCUSSION
A. CASE STUDY I
The analysis and discussion of the simulations results for case study I are presented in this section.

1) GENERAL ANALYSIS OF SIMULATION RESULTS
To analyze the traded electricity volume and the electricity exchanged with the upstream grid, the geographic location and location of the agent in the local community are used as criteria for initial cluster grouping and hierarchical clustering chosen metric, respectively. The simulation is analyzed and presented in this subsection.

Fig. 7 shows the net electricity demand for (a) consumers, (b) prosumers and (c) the prosumers' electricity supply for

(a)

(b)

FIGURE 6. Geospatial information for (a) case study I and (b) II.

TABLE 2. Price components for sub-communities and with the up-stream grid.

Components	Block Trades	Estate Trades	Zonal Trades	Community Trades	Up-stream Grid Trades
Energy price(ct./kWh)	[11 24.50]	[11 23.75]	[11 23.25]	[11 22.50]	[11 17.75]
Metering fee(ct./kWh)	0.32	0.32	0.32	0.32	0.32
Local grid fee(ct./kWh)	-	1.31	1.81	3.33	3.33
Up-stream fee(ct./kWh)	-	-	-	-	4.75
VAT(ct./kWh)	[0.06 4.72]	[0.31 4.82]	[0.40 4.83]	[0.85 4.96]	[1.60 5.35]
Total (ct./kWh)	[11.38 29.54]	[12.94 30.20]	[13.53 30.21]	[15.50 31.11]	[21.00 31.50]

the simulation period. As shown in Figs. 7a and 7b, the peak demand happens in the morning (around 6:00am) and late evening (around 8:00pm). This is typical of households as most people always prepare for their daily activities in the morning and consequently will need to use their electricity appliances. Moreover, at late in the evening, most people will be cooking their dinner, listening to local news, etc. Consequently, the peak electricity demand is obvious. The early morning peak electricity demand of the prosumer households is not noticeable in Fig. 7b. This is because the simulation is conducted for a summer day, and most of the prosumers' electricity demand in the morning of the simulation day is covered by the PV production and hence cutting off the morning peak demand of the prosumers. In addition, the households show an almost constant electricity demand from about 11:30pm until 5:00am. This is because most electricity

consumers at this time are asleep and therefore only use their constant electricity demand appliances such as refrigerators. From Fig. 7c, net electricity supply starts early in the morning at around 5:30am with peak generation between 10:00am to 12:30pm, ending at around 4:00pm. This is a typical summer day, prior to and after this specified time, electricity generated by the prosumers is used to satisfy the internal demand of the prosumer households.

Fig. 8 displays the electricity traded within the LEM for selected (a) consumers (b) prosumers and (c) energy exchange with the upstream grid for the simulation period. As shown in Fig. 8, it is evident that electricity trading happens within the LEM mainly during the net PV production time. Since the batteries are not internally controlled, the prosumers that own batteries sell their charged electricity a few time steps after the PV generation within the LEM
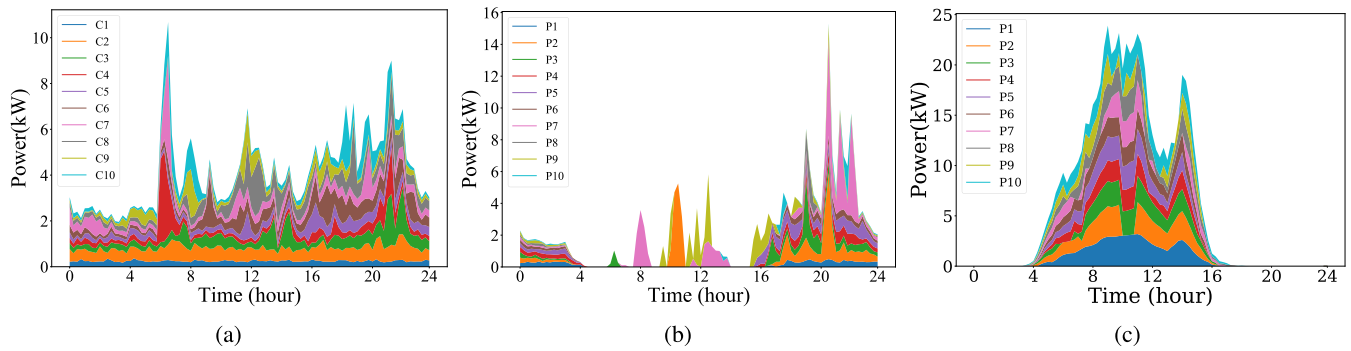
FIGURE 7. Power demand for (a) Consumers (b) Prosumers and (c) Prosumers' supply.

has stopped. The stacked area of the prosumers' (Fig. 8b) traded energy is bigger compared to the stacked area of the consumers' (Fig. 8a) traded energy. This is because prosumers can act as consumers in some time steps. At this instance, the traded energy between prosumers is recorded for both prosumers (i.e. prosumers that produce the energy and those that also consume the energy) thereby increasing the area of Fig. 8b compared to Fig. 8a where two consumers cannot exchange energy. Fig. 8c displays the energy imported/exported from/to the upstream grid to/from the LEM. From Fig. 8c, there is no energy export to the upstream grid from 12:00am until about 6:00am and from about 4:00pm until 12:00am. This is because there is no net PV production at this period. The major electricity import from the upstream grid also happens at this period. Since all prosumers within the LEM may not have the same price preference, some prosumers may not be able to bid the price that will enable them buy electricity from the LEM. At same time, some prosumers many not be able to offer the price that will enable them sell all their electricity at the LEM. This is the cause of the simultaneous imports and exports witnessed at some periods around 6:00am until about 4:00pm in in Fig. 8c. Comparing Figs. 8c and 7a, it is evident that the peak demand witnessed in Fig. 7a at around 6:30am which is cut off by the LEM (as shown in Fig. 8c) from 10kW to about 4kW showing about 60% cut off. The peak import (Fig. 8c) witnessed from around 9:00pm is because of the high demand from both consumers and prosumers at late evening and there is no PV production at this time to reduce it.

2) ANALYSIS OF CLUSTERING SCENARIOS
The analysis and discussion of the simulation results for the different clustering scenarios are presented in this section.

a: CONSTANT CLUSTERING CRITERIA FOR INITIAL CLUSTER AND HIERARCHICAL CLUSTER CHOSEN METRIC
In this subsection, we analyze the simulation for the constant clustering scenarios. The same approach is used for both initial cluster grouping and hierarchical clustering chosen metric. In this way, four scenarios are introduced based on location, offering/bidding price, energy quantity, and cluster

welfare. These scenarios are analyzed in terms of how they affect electricity trading, and the economic and technical benefits of the LEM. Fig. 9 shows (a) the traded energy, (b) average trade rate and (c) number of trade per slot for four clustering scenarios. According to our study, the four clustering scenarios show similar behaviour for traded energy, average trade rate and number of trades per slot. In the proposed scenarios, from 12:00am until about 5:00am, the energy traded within the LEM is zero and consequently, the average trade rate and the number of trades per slot are 31.5ct/kWh and 0, respectively. At this time, there is no PV generation and all the electricity is bought from the upstream grid at a higher price. In the morning, the PVs start generating and consequently, the traded volume and the number of trades per slot increase gradually which further results in gradual decrease in the average trade rate. During the day, as the PVs are generating electricity, higher volume of energy is traded within the LEM resulting in higher value of number of trades per slot and lower value of the average trade rate. On the other hand, during the evening around 8:00pm, as the sun sets, the traded volume and the number of trades per slot reduce gradually to zero. The average trade rate increases in the same way to 31.5ct/kWh.

Fig. 10 displays the internal traded energy within the LEM, total energy import and export from/to the upstream grid for all clustering scenarios. The internal traded energy is the total energy traded between consumers and prosumers within the LEM. The energy imported/exported from/to the upstream grid are known as external energy exchange. As shown in Fig. 10, the price clustering scenario shows the best performance for the internal and external energy exchange compared to other clustering scenarios with location showing the least. This is evident with the high internal and lower external energy exchange of the price clustering scenario and the lower internal and higher external energy exchange of the location clustering scenario. Higher internal energy exchange means that more energy is traded within the LEM and hence, more benefits is expected to be created by this higher trade. On the other hand, higher external energy exchange means that most of the trades happens between the prosumers and the upstream grid, this is expected to create less benefits to the
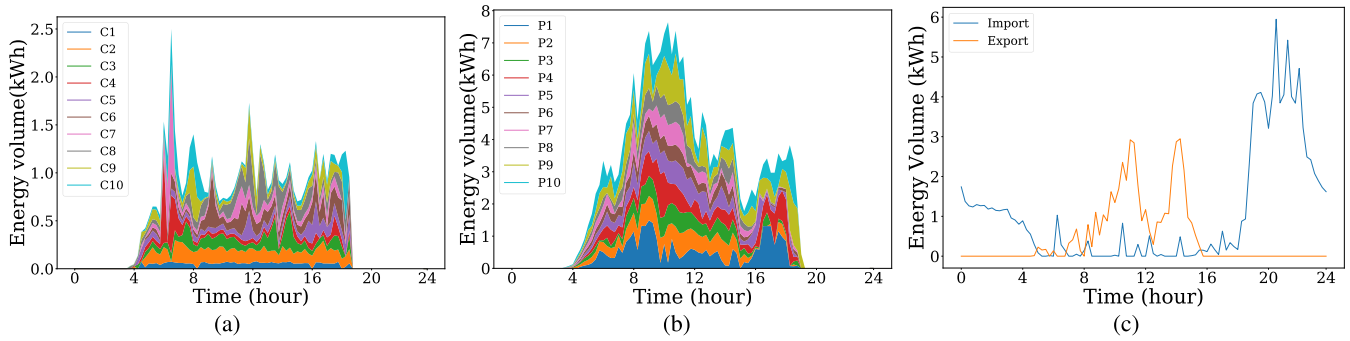
FIGURE 8. Electricity traded within the LEM for selected (a) Consumers, (b) Prosumers and (c) exchange with upstream grid.
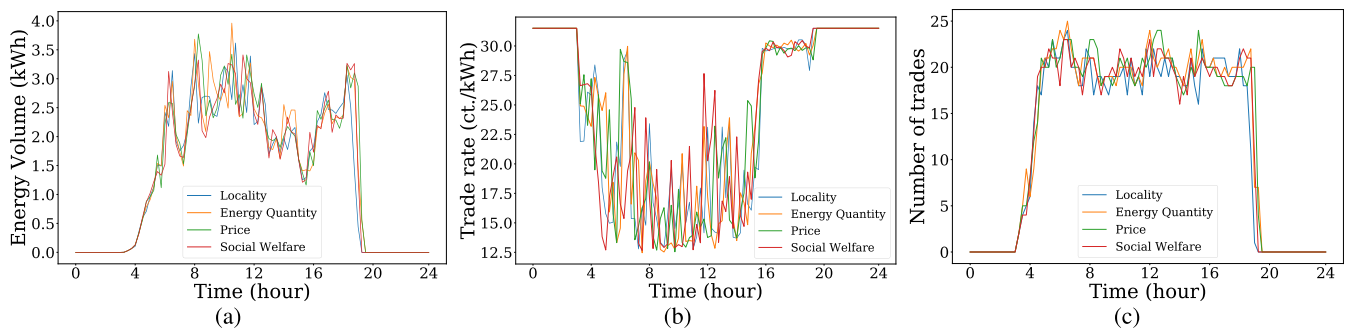

FIGURE 9. Comparison of (a) traded energy (b) average trade rate and (c)number of trades per slot for constant criteria for initial cluster grouping and chosen metric.
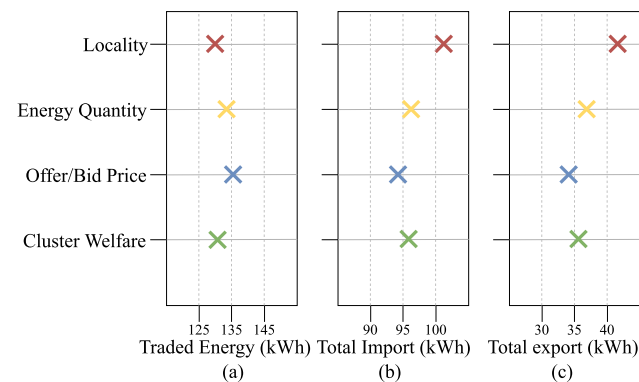

FIGURE 10. Comparison of internal and external energy exchanged for constant clustering criteria scenarios.

LEM. The least performance of location clustering scenario compared to the price clustering scenario is because, although some consumers and prosumers may be close to each other and may be clustered to trade electricity, but they will not trade if the bidding price is less than the offering price. This way, clustering prosumers and consumers who have the same price preference range will result in higher matching of bids and offers within the LEM, thereby creating more internal energy trade and less energy exchange with the upstream grid. The self-sufficiency (SS), self consumption ratio (SC) and the share of market savings (SMS) for the four clustering scenarios are shown in Fig. 11. The SMS is the share of

profit made by the local consumers and prosumers for trading within the LEM compared to when there is no LEM [59]. The SS, SC and SM of the four scenarios share a close range with each other. However, for all the performance indicators (SS, SC and SMS), the price based clustering scenario has the best performance and location has the least. This is because, the price based clustering scenario provides more internal traded energy and less external traded energy compared to other clustering scenarios. Higher internal traded energy and lower external energy exchange results in higher performance of the LEM as the LEM does not depend much on the upstream grid for its energy production/consumption. Hence, clustering prosumers and consumers, who have the same price preference range will be more beneficial to the LEM as most trades will match in this cluster scenario, thereby increasing the LEM'S performance indicators.

Figs. 12 (a) and (b) show box plots of the net cost of trading electricity for consumers and prosumers, respectively, in twelve clustering scenarios (L1 to L12). For the consumers, Fig. 12a, the mean cost for all the clustering scenarios is between 2.6 to 2.8 ct./kWh. It is evident that the consumers are household consumers as this cost is within the range of the daily electricity cost for a households in Germany and this correspond to the household data used for the simulation of this case scenario. The minimum cost is 1.5 ct./kWh and all the clustering scenarios show the same minimum cost. However, the maximum cost varies for the different scenarios with L1, L3,L4, L6, L9 and L10 showing a lesser maximum
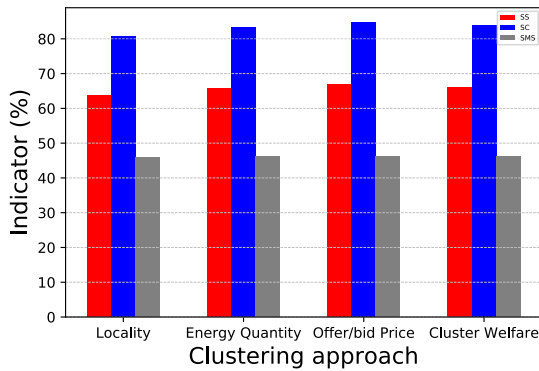
FIGURE 11. Comparison of SS, SC and SMS for constant clustering criteria scenarios.

TABLE 3. Varied clustering scenarios.

S/N	Initial cluster criteria	chosen metric criteria	Symbol
1	location	Offer/bid price + location	L1
2	location	Energy quantity + location	L2
3	location	Cluster welfare + location	L3
4	Offer/bid price	Offer/bid price + location	L4
5	Offer/bid price	Energy quantity + location	L5
6	Offer/bid price	Cluster welfare + location	L6
7	Energy quantity	Offer/bid price + location	L7
8	Energy quantity	Energy quantity + location	L8
9	Energy quantity	Cluster welfare + location	L9
10	Cluster welfare	Offer/bid price + location	L10
11	Cluster welfare	Energy quantity + location	L11
12	Cluster welfare	Cluster welfare + location	L12

cost value compared to L3, L5, L7, L8, L11 and L12 clustering scenarios. For the prosumers, Fig. 12b, the mean cost is between −1.8 to −1.6 ct./kWh for all the clustering scenarios. Negative mean cost means that the prosumers gained some monetary value for exchanging their electricity with the consumers. The maximum net cost of the prosumers is higher for L3, L4, and L8 compared to other clustering scenarios. On the other hand, the minimum net cost is lower for L4, L5, L6 and L7 compared to other clustering scenarios. L6, L7 and L9 show the best performance because of their lower mean, maximum, and minimum net cost.

Fig. 13 shows the internal traded energy within the LEM, total energy import and export from/to the upstream grid in twelve clustering scenarios (L1 to L12). As seen in Fig. 13, L4, L5 and L6 scenarios present the best performance for the internal and external energy exchange compared to other clustering scenarios with L1, L2 and L10 having the worst performance. This is evident with the high internal and lower external energy exchange of the L4, L5 and L6 clustering scenarios and the lower internal and higher external energy exchange of the L1, L2 and L10 clustering scenarios. L4, L5 and L6 have similar cluster criterion for initial cluster grouping which is the offering/bidding price. This similarity shows why the three scenarios perform better than the others in terms of energy exchange within the LEM and upstream grid. It is also evident that clustering prosumers and consumers who have the same price range based on initial cluster
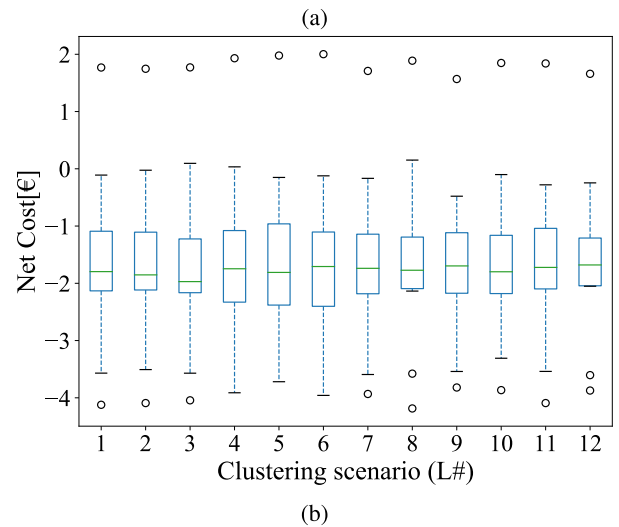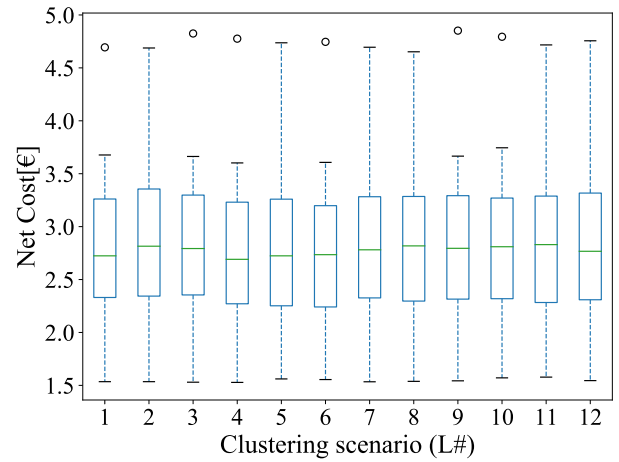

(a)

(b)

FIGURE 12. Net cost for (a) Consumers and (b) Prosumers based on variable clustering criteria scenarios.

grouping will result in higher matching of bids and offers within the LEM which create more internal energy trade and less energy exchange with the upstream grid.

Fig. 14 shows the SS, SC and SMS for the clustering scenarios L1 to L12. The SS, SC and SMS of the twelve scenarios L1 to L12 show close range value to each other. However, for all the performance indicators, the price based initial cluster criterion's (L4, L5 and L6) clustering scenarios show the outstanding performance compared to other scenarios. The most outstanding clustering scenario is the L6 scenario which is based on initial grouping of bids and offers based on bids/offer price and advanced grouping based on location and cluster welfare. L6 has the highest SS, SC and SM compared to other clustering scenarios. This is because, the price based initial cluster grouping scenario provides more internal traded energy and less external traded energy compared to other clustering scenarios. Furthermore, the location and cluster welfare based criteria for chosen metric ensures that electricity is traded close to where they are consumed and further create additional welfare to the community. This provides additional SMS to the LEM.
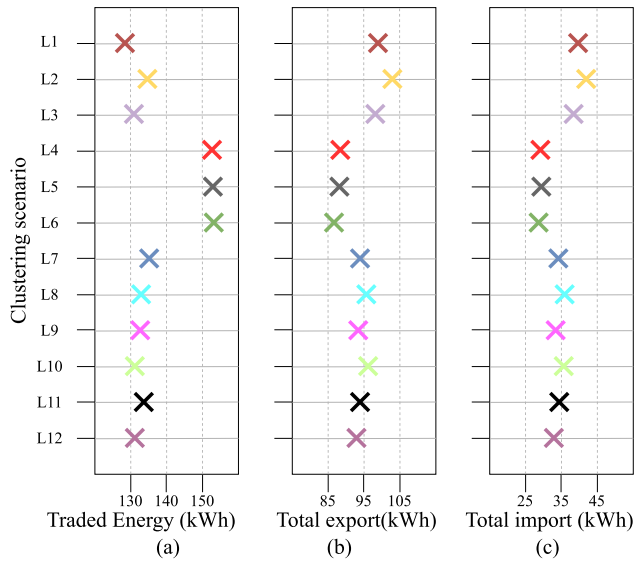
FIGURE 13. Comparison of internal and external energy exchanged for variable clustering criteria scenarios.
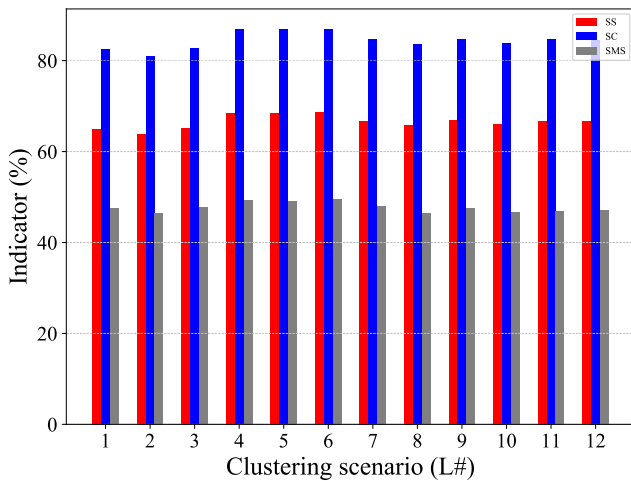

FIGURE 14. Comparison of SS, SC and SMS for variable clustering criteria scenarios.

b: P2P ENERGY TRADE
Fig. 15 shows the P2P energy trade between consumers and prosumers for the L6 clustering approach. Since the prosumers can act as consumers in some time slots, and consequently buy energy from other prosumers, the prosumers are further added to the horizontal axis. Comparing Fig. 15 with Fig. 6a shows that P2P energy is traded within the LEM according to the closest cluster that can be formed from Fig. 6a. Regions on Fig. 15 where higher quantity of energy is exchanged represent a particular region or cluster of household in Fig. 6a where households are close to each other. An example is the upper most center of Fig. 15 where higher quantity of energy is exchanged between P8, P9 and P10 as prosumers (producing) and; C13, C14 and C15 as consumers. This is evident in Fig. 6a where P8, P9, P10, C13, C14 and C15 are located in the same estate known as AS-EST. The higher quantity of energy exchanged between P2 and

P13 is because they belong to the same block, therefore, the cluster model matches them easily and ensure that electricity is traded close to where it is produced. In the same way, the higher quantity of energy exchanged between P10 and C15 is also because they are located in the same block, therefore, exchanging energy among them will create additional welfare to the LEM.

B. CASE STUDY II
In this Section, we analyze the performance of the L6 clustering scenario with 120 households as described in Section V-B and varying number of clusters for initial cluster grouping and compare it with P2P matching where there is no clustering.

1) ENERGY EXCHANGE FOR VARYING NUMBER OF CLUSTERS
Fig. 16 displays the internal and external energy exchange for varying number of clusters. The external energy exchange are the energy export and import to/from the upstream grid. As shown in Fig. 16, the total energy exchange, that is the sum of the internal and external energy exchange, is constant for all the simulation scenarios. This is evident that the total energy requirements of the local community is constant throughout the simulations. Furthermore, the P2P matching model without advanced clustering scenario has higher external and lower internal energy exchange compared to all other scenarios. Higher external and lower internal energy exchange results in LEM depending totally on the upstream grid and impacts the LEM participants negatively. Moreover, increasing the number of clusters for initial cluster grouping increases the internal energy exchange and reduces the external energy exchange of the LEM. For the case scenario studied in this work, the simulation achieved its optimum when the number of clusters is four, therefore, increasing the numbers of clusters beyond four has no impact on the performance of the LEM. Based on the L6 advanced clustering approach, the number of initial cluster groups where the LEM reaches its optimum depends on the individual bidding strategies and location of the LEM participants on the local grid.

2) ECONOMIC AND TECHNICAL BENEFITS FOR VARYING NUMBER OF CLUSTERS
Fig. 17 displays the SS, SC and SMS for varying number of initial cluster groups. The P2P matching model without advanced clustering scenario has lower SS, SC and SMS compared to all other scenarios. This is because, the P2P matching model without clustering scenario has more external energy exchange and lower internal energy exchange compared to all other scenarios, making the LEM less beneficial to the local markets participants. Also, increasing the number of clusters for initial cluster groups increases the SS, SC and the SMS of the local community. This is because, increasing the number of clusters for initial cluster groups reduces the external energy and increases the internal energy exchange of the LEM and thereby creating additional benefits to the
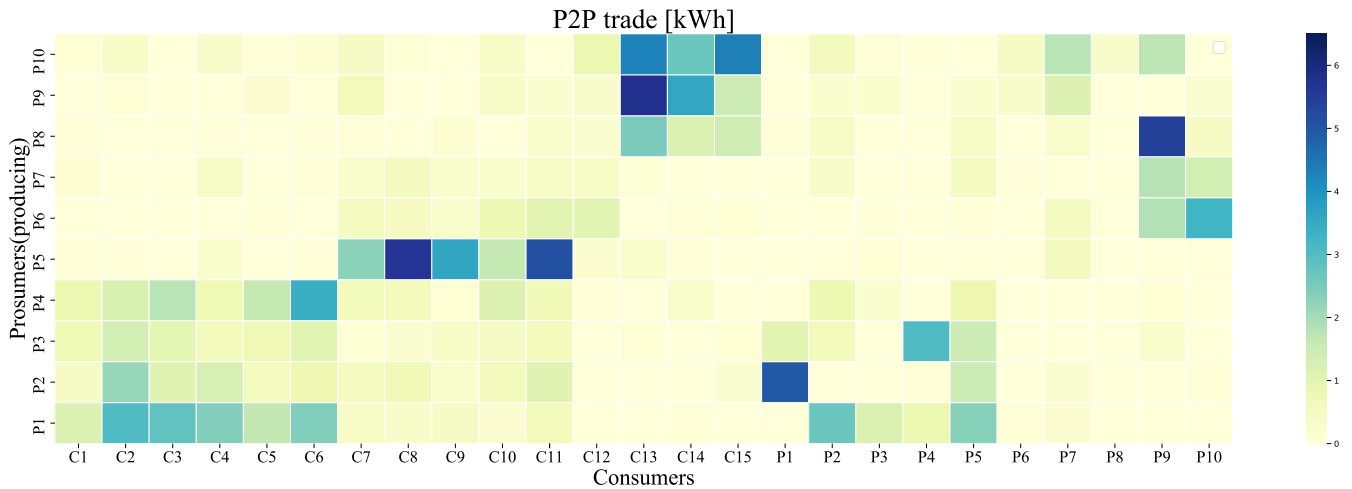
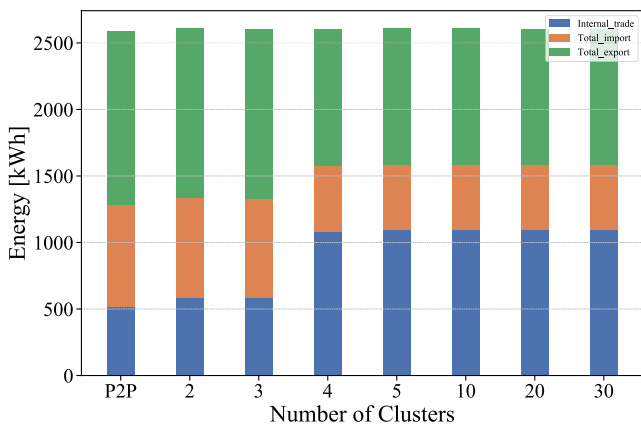**FIGURE 15.** P2P energy trade for L6 clustering approach.

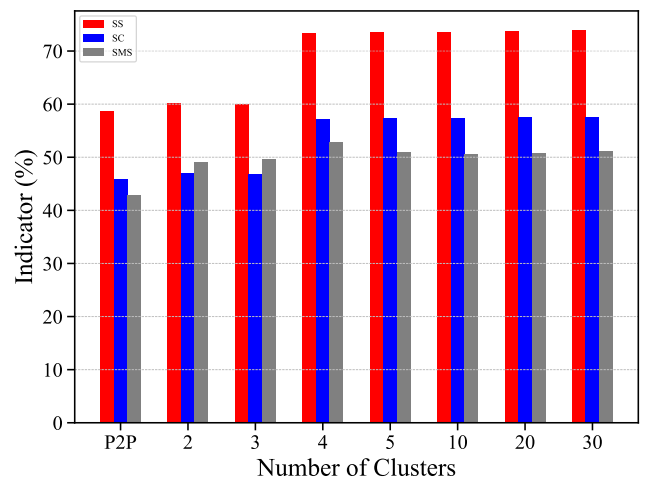**FIGURE 16.** Comparison of energy exchange for variable number of clusters.

local community. The optimum SS, SC and SMS is achieved when the number of initial cluster groups is four. This is the same for internal and external energy exchange of the LEM. This is evidence that the technical and economic benefits of the local community depend on the energy exchange between the community and the upstream grid.

3) TIME COMPLEXITY
The proposed clustering algorithm takes less than 1 seconds to solve a given problem in an Ubuntu operating system computer. This means running the model only, without a market framework such as GSy-exchange. Running it with a market framework takes more time as simulation data needs to be stored and read from different files. Table 4 shows the average simulation computational time per slot to complete the matching in the GSy-exchange matching API for varying number of clusters. The P2P matching mechanism without clustering has the least computational time per slot compared to all other scenarios. This is because the P2P matching mechanism without clustering is a simple random matching and has no clustering model. With the smallest number of

FIGURE 17. Comparison of SS, SC and SMS for variable number of clusters.

TABLE 4. Average computational time per time slot for varying number of clusters.

Number of cluster	Average computation time per slot (s)
P2P	15.5
2	22.55
3	28.36
4	32.1
5	42.02
10	60.08
20	73.05
30	82.5

clusters which is 2, the run time is 22.55 seconds per slot for a community of 120 prosumers. Increasing the number of initial cluster groups increases the computational time. This is because increasing the number of clusters will require more time for k-means clustering to determine the individual cluster centroids, additionally, it will require more time for advanced clustering to build up the hierarchical model as there will be a greater number of branches from the initial

cluster groups. Also, comparing the average computational time of our model with that of a multi-layer LEM model proposed in [48], where a similar community was simulated with the Grid Singularity platform shows that our proposed model has the capability of reducing the average computational time by half when the number of clusters is less than 5. However, with large number of clusters, it may take more time for the algorithm to complete its calculation and matching. For our simulation case scenario, our model achieved its optimum when the number of clusters is four as discussed in Sections VI-B1 and VI-B2, this is evidence that the model is computational efficient when compared to previous works already published in the same area.

VII. CONCLUSION

In this paper, an advanced clustering algorithm was proposed for P2P matching in the energy community considering the heterogeneous characteristics of bidding preferences for prosumers consisting of geographic location, location on the local community, bid/offer price, bid/offer quantity, and cluster welfare. The proposed model was tested in a German real-case scenario and simulated for 120 German households. The simulations results show that the model was able to leverage the preference opportunity to ensure that electricity was traded and consumed closer to where it is produced in a P2P trading structure. Thereby, reducing the peak demand of the local community, increasing the traded energy at the local community, and reducing the energy exchange of the local community and the upstream grid. Also, the model benefit the local prosumers as it can help to increase the share of market saving, self consumption and self sufficiency of the prosumers. By ensuring that energy is traded closer to where is is produced, the model further increase the local economy of the local prosumers.

Furthermore, in the proposed advanced clustering model, using price preference as the criterion for initial cluster grouping, and location of the prosumers in the community and cluster welfare as the criteria for the chosen weight metric, offered more technical and economic benefits to the local community compared to other clustering scenarios. Also, comparing our proposed model with a P2P model without clustering shows that the computational time of both models are closer in range and hence our model can be easily adopted without increasing the market computational time. In future work, we will investigate how reinforcement learning can be used by the prosumers to intelligently decide their bidding/offering preference and how the model will be implemented in a distributed blockchain platform to ensure efficient preservation of prosumers' privacy and conformation to data protection laws.

REFERENCES

[1] C. Weinhardt, E. Mengelkamp, W. Cramer, S. Hambridge, A. Hobert, E. Kremers, W. Otter, P. Pinson, V. Tiefenbeck, and M. Zade, "How far along are local energy markets in the DACH+ region?" in *Proc. 10th ACM Int. Conf. Future Energy Syst.*, New York, NY, USA, Jun. 2019, pp. 544–549, doi: 10.1145/3307772.3335318.

[2] E. Mengelkamp, J. Diesing, and C. Weinhardt, "Tracing local energy markets: A literature review," *it-Inf. Technol.*, vol. 61, nos. 2–3, pp. 101–110, Apr. 2019, doi: 10.1515/itit-2019-0016.

[3] T. Chmielak, T. Glinka, and G. Kosman, "Local market of electric energy as an aid for restructuring of Silesia region," *Energetyka*, vol. 50, no. 7, pp. 407–415, 1996.

[4] P. Siano, G. De Marco, A. Rolan, and V. Loia, "A survey and evaluation of the potentials of distributed ledger technology for peer-to-peer transactive energy exchanges in local energy markets," *IEEE Syst. J.*, vol. 13, no. 3, pp. 3454–3466, Sep. 2019.

[5] E. Mengelkamp, S. Bose, E. Kremers, J. Eberbach, B. Hoffmann, and C. Weinhardt, "Increasing the efficiency of local energy markets through residential demand response," *Energy Inform.*, vol. 1, pp. 1–18, Aug. 2018, doi: 10.1186/s42162-018-0017-3.

[6] T. Capper, A. Gorbatcheva, M. A. Mustafa, M. Bahloul, J. M. Schwidtal, R. Chitchyan, M. Andoni, V. Robu, M. Montakhabi, I. J. Scott, C. Francis, T. Mbavarira, J. M. Espana, and L. Kiesling, "Peer-to-peer, community self-consumption, and transactive energy: A systematic literature review of local energy market models," *Renew. Sustain. Energy Rev.*, vol. 162, Jul. 2022, Art. no. 112403, doi: 10.1016/j.rser.2022.112403.

[7] E. A. Soto, L. B. Bosman, E. Wollega, and W. D. Leon-Salas, "Peer-to-peer energy trading: A review of the literature," *Appl. Energy*, vol. 283, Feb. 2021, Art. no. 116268, doi: 10.1016/j.apenergy.2020.116268.

[8] T. Sousa, T. Soares, P. Pinson, F. Moret, T. Baroche, and E. Sorin, "Peer-to-peer and community-based markets: A comprehensive review," *Renew. Sustain. Energy Rev.*, vol. 104, pp. 367–378, Apr. 2019.

[9] R. Zheng, T. Dai, K. Sycara, and N. Chakraborty, "Automated multilateral negotiation on multiple issues with private information," *INFORMS J. Comput.*, vol. 28, no. 4, pp. 612–628, Nov. 2016.

[10] M. Khorasany, A. Paudel, R. Razzaghi, and P. Siano, "A new method for peer matching and negotiation of prosumers in peer-to-peer energy markets," *IEEE Trans. Smart Grid*, vol. 12, no. 3, pp. 2472–2483, May 2021, doi: 10.1109/TSG.2020.3048397.

[11] T. Morstyn, A. Teytelboym, and M. D. Mcculloch, "Bilateral contract networks for peer-to-peer energy trading," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2026–2035, Mar. 2019, doi: 10.1109/TSG.2017.2786668.

[12] A. G. Azar, H. Nazaripouya, B. Khaki, C.-C. Chu, R. Gadh, and R. H. Jacobsen, "A non-cooperative framework for coordinating a neighborhood of distributed prosumers," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 2523–2534, May 2019, doi: 10.1109/TII.2018.2867748.

[13] H. Haggi and W. Sun, "Multi-round double auction-enabled peer-to-peer energy exchange in active distribution networks," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 4403–4414, Sep. 2021, doi: 10.1109/TSG.2021.3088309.

[14] T. AlSkaif, J. L. Crespo-Vazquez, M. Sekuloski, G. van Leeuwen, and J. P. S. Catalao, "Blockchain-based fully peer-to-peer energy trading strategies for residential energy systems," *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 231–241, Jan. 2022, doi: 10.1109/TII.2021.3077008.

[15] M. Khorasany, A. Shokri Gazafroudi, R. Razzaghi, T. Morstyn, and M. Shafie-khah, "A framework for participation of prosumers in peer-to-peer energy trading and flexibility markets," *Appl. Energy*, vol. 314, May 2022, Art. no. 118907, doi: 10.1016/j.apenergy.2022.118907.

[16] L. Chen, N. Liu, and J. Wang, "Peer-to-peer energy sharing in distribution networks with multiple sharing regions," *IEEE Trans. Ind. Informat.*, vol. 16, no. 11, pp. 6760–6771, Nov. 2020, doi: 10.1109/TII.2020.2974023.

[17] N. Wang, Z. Liu, P. Heijnen, and M. Warnier, "A peer-to-peer market mechanism incorporating multi-energy coupling and cooperative behaviors," *Appl. Energy*, vol. 311, Apr. 2022, Art. no. 118572, doi: 10.1016/j.apenergy.2022.118572.

[18] D. Xu, B. Zhou, N. Liu, Q. Wu, N. Voropai, C. Li, and E. Barakhtenko, "Peer-to-peer multienergy and communication resource trading for interconnected microgrids," *IEEE Trans. Ind. Informat.*, vol. 17, no. 4, pp. 2522–2533, Apr. 2021, doi: 10.1109/TII.2020.3000906.

[19] M. H. Ullah and J.-D. Park, "A two-tier distributed market clearing scheme for peer-to-peer energy sharing in smart grid," *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 66–76, Jan. 2022, doi: 10.1109/TII.2021.3058511.

[20] A. S. Gazafroudi, M. Khorasany, R. Razzaghi, H. Laaksonen, and M. Shafie-Khah, "Hierarchical approach for coordinating energy and flexibility trading in local energy markets," *Appl. Energy*, vol. 302, Nov. 2021, Art. no. 117575, doi: 10.1016/j.apenergy.2021.117575.

[21] D. Qiu, Y. Ye, D. Papadaskalopoulos, and G. Strbac, "Scalable coordinated management of peer-to-peer energy trading: A multi-cluster deep reinforcement learning approach," *Appl. Energy*, vol. 292, Jun. 2021, Art. no. 116940, doi: 10.1016/j.apenergy.2021.116940.

[22] M. Zade, S. D. Lumpp, P. Tzscheutschler, and U. Wagner, "Satisfying user preferences in community-based local energy markets—Auction-based clearing approaches," *Appl. Energy*, vol. 306, Jan. 2022, Art. no. 118004, doi: 10.1016/j.apenergy.2021.118004.

[23] B. Richter, E. Mengelkamp, and C. Weinhardt, "Vote for your energy: A market mechanism for local energy markets based on the consumers' preferences," in *Proc. 16th Int. Conf. Eur. Energy Market (EEM)*, Sep. 2019, pp. 1–6, doi: 10.1109/EEM.2019.8916544.

[24] I. Petri, A. Alzahrani, J. Reynolds, and Y. Rezgui, "Federating smart cluster energy grids for peer-to-peer energy sharing and trading," *IEEE Access*, vol. 8, pp. 102419–102435, 2020, doi: 10.1109/ACCESS.2020.2998747.

[25] N. Hashemipour, P. Crespo del Granado, and J. Aghaei, "Dynamic allocation of peer-to-peer clusters in virtual local electricity markets: A marketplace for EV flexibility," *Energy*, vol. 236, Dec. 2021, Art. no. 121428, doi: 10.1016/j.energy.2021.121428.

[26] L. Ali, S. Muyeen, H. Bizhani, and A. Ghosh, "A peer-to-peer energy trading for a clustered microgrid—Game theoretical approach," *Int. J. Elect. Power Energy Syst.*, vol. 133, Dec. 2021, Art. no. 107307, doi: 10.1016/j.ijepes.2021.107307.

[27] M. Khorasany, Y. Mishra, B. Babaki, and G. Ledwich, "Enhancing scalability of peer-to-peer energy markets using adaptive segmentation method," *J. Mod. Power Syst. Clean Energy*, vol. 7, no. 4, pp. 791–801, Jul. 2019, doi: 10.1007/s40565-019-0510-0.

[28] P.-Y. Zhou and K. C. C. Chan, "A model-based multivariate time series clustering algorithm," in *Trends and Applications in Knowledge Discovery and Data Mining*, W.-C. Peng, H. Wang, J. Bailey, V. S. Tseng, T. B. Ho, Z.-H. Zhou, and A. L. Chen, Eds. Cham, Switzerland: Springer, 2014, pp. 805–817.

[29] Y. Wang, L. Li, and Q. Yang, "Application of clustering technique to electricity customer classification for load forecasting," in *Proc. IEEE Int. Conf. Inf. Autom.*, Aug. 2015, pp. 1425–1430, doi: 10.1109/ICInfA.2015.7279510.

[30] F. Martinez Álvarez, A. Troncoso, J. M. R. Santos, and J. C. R. Santos, "Discovering patterns in electricity price using clustering techniques," *Renew. Energy Power Qual.*, vol. 1, no. 05, pp. 174–181, Mar. 2007, doi: 10.24084/repqj05.245.

[31] X. Zhang, J. L. Ramírez-Mendiola, M. Li, and L. Guo, "Electricity consumption pattern analysis beyond traditional clustering methods: A novel self-adapting semi-supervised clustering method and application case study," *Appl. Energy*, vol. 308, Feb. 2022, Art. no. 118335, doi: 10.1016/j.apenergy.2021.118335.

[32] G. Chicco, R. Napoli, and F. Piglione, "Comparisons among clustering techniques for electricity customer classification," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 933–940, May 2006, doi: 10.1109/TPWRS.2006.873122.

[33] R. Li, C. Gu, F. Li, G. Shaddick, and M. Dale, "Development of low voltage network templates—Part I: Substation clustering and classification," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3036–3044, Nov. 2015, doi: 10.1109/TPWRS.2014.2371474.

[34] G. J. Tsekouras, N. D. Hatziargyriou, and E. N. Dialynas, "Two-stage pattern recognition of load curves for classification of electricity customers," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1120–1128, Aug. 2007, doi: 10.1109/TPWRS.2007.901287.

[35] A. Trbovich, S. Hambridge, D. van den Biggelaar, E. Hesse, and F. Sioshansi, "D3A energy exchange for a transactive grid," in *Behind and Beyond the Meter*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 267–284.

[36] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, Berkeley, CA, USA, 1965, pp. 281–297.

[37] A. E. Ezugwu, A. M. Ikotun, O. O. Oyelade, L. Abualigah, J. O. Agushaka, C. I. Eke, and A. A. Akinyelu, "A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects," *Eng. Appl. Artif. Intell.*, vol. 110, Apr. 2022, Art. no. 104743, doi: 10.1016/j.engappai.2022.104743.

[38] B. Wang, Y. Li, and W. K. Härdle, "K-expectiles clustering," *J. Multivariate Anal.*, vol. 189, May 2022, Art. no. 104869, doi: 10.1016/j.jmva.2021.104869.

[39] X. Zhang, Y. He, Y. Jin, H. Qin, M. Azhar, and J. Z. Huang, "A robust k-means clustering algorithm based on observation point mechanism," *Complexity*, vol. 2020, pp. 1–11, Mar. 2020, doi: 10.1155/2020/3650926.

[40] Q.-Z. Liao, L. Xue, G. Lei, X. Liu, S.-Y. Sun, and S. Patil, "Statistical prediction of waterflooding performance by K-means clustering and empirical modeling," *Petroleum Sci.*, vol. 19, no. 3, pp. 1139–1152, Jun. 2022, doi: 10.1016/j.petsci.2021.12.032.

[41] O. Akman, T. Comar, D. Hrozencik, and J. Gonzales, "Data clustering and self-organizing maps in biology," in *Algebraic and Combinatorial Computational Biology* (MSE/Mathematics in Science and Engineering), R. Robeva and M. Macauley, Eds. New York, NY, USA: Academic, 2019, ch. 11, pp. 351–374, doi: 10.1016/B978-0-12-814066-6.00011-8.

[42] P. Wittek, "Unsupervised learning," *Quantum Machine Learning*, P. Wittek, Ed. Boston, MA, USA: Academic, 2014, ch. 5, pp. 57–62, doi: 10.1016/B978-0-12-800953-6.00005-0.

[43] R. Xu and D. Wunsch, *Clustering* (Computational Intelligence). Hoboken, NJ, USA: Wiley, 2009, pp. 221–238.

[44] Z. Li and M. de Rijke, "The impact of linkage methods in hierarchical clustering for active learning to rank," in *Proc. 40th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, New York, NY, USA, Aug. 2017, pp. 941–944, doi: 10.1145/3077136.3080684.

[45] L. Ramos Emmendorfer and A. M. de Paula Canuto, "A generalized average linkage criterion for hierarchical agglomerative clustering," *Appl. Soft Comput.*, vol. 100, Mar. 2021, Art. no. 106990, doi: 10.1016/j.asoc.2020.106990.

[46] J. Lin, M. Pipattanasomporn, and S. Rahman, "Comparative analysis of auction mechanisms and bidding strategies for P2P solar transactive energy markets," *Appl. Energy*, vol. 255, Dec. 2019, Art. no. 113687, doi: 10.1016/j.apenergy.2019.113687.

[47] A. Esmat, M. de Vos, Y. Ghiassi-Farrokhfal, P. Palensky, and D. Epema, "A novel decentralized platform for peer-to-peer energy trading market with blockchain technology," *Appl. Energy*, vol. 282, Jan. 2021, Art. no. 116123, doi: 10.1016/j.apenergy.2020.116123.

[48] G. C. Okwuibe, A. S. Gazafroudi, S. Hambridge, C. Dietrich, A. Trbovich, M. Shafie-khah, P. Tzscheutschler, and T. Hamacher, "Evaluation of hierarchical, multi-agent, community-based, local energy markets based on key performance indicators," *Energies*, vol. 15, no. 10, p. 3575, May 2022, doi: 10.3390/en15103575.

[49] Grid Singularity. (2022). *Grid Singularity Technical Approach*. Accessed: Aug. 19, 2022. [Online]. Available: https://gridsingularity.github.io/gsye/technical-approach/

[50] T. Tjaden, J. Bergner, J. Weniger, and V. Quaschning, "Representative electrical load profiles of residential buildings in Germany with a temporal resolution of one second," Dataset: HTW Berlin Uni. Appl. Sci., Berlin, Germany. Accessed: Feb. 14, 2022. [Online]. Available: https://solar.htw-berlin.de/elektrische-lastprofile-fuer-wohngebaeude/

[51] N. Pflugradt. *Loadprofilegenerator*. Accessed: Jul. 18, 2022. [Online]. Available: https://www.loadprofilegenerator.de/

[52] N. Pflugradt, "Modellierung von wasser und energieverbräuchen in haushalten," Ph.D. dissertation, Dept. Mech. Eng., TU Chemnitz, Chemnitz, Germany, 2016.

[53] Stromnetz-Berlin. *Netznutzer–Standardlastprofile*. Accessed: Jul. 2, 2022. [Online]. Available: https://www.stromnetz.berlin/netz-nutzen/netznutzer

[54] C. Fünfgeld and R. Tiedemann. *Anwendung der Repräsentativen Vdew-Lastprofile: Step-by-Step*. Accessed: Jul. 15, 2022. [Online]. Available: https://www.bdew.de/media/documents/2000131_Anwendung-repraesentativen_Lastprofile-Step-by-step.pdf

[55] S. Pfenninger and I. Staffell. *Renewables Ninja*. Accessed: Jul. 28, 2022. [Online]. Available: https://www.renewables.ninja/

[56] S. Pfenninger and I. Staffell, "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, pp. 1251–1265, Nov. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544216311744, doi: 10.1016/j.energy.2016.08.060.

[57] *German Vat Refund*. Accessed: Apr. 12, 2022. [Online]. Available: https://www.germany.info/us-en/ser-vice/09-Taxes/vat-refund/906296

[58] B. Wehrmann. (2023). What German Households Pay for Electricity, in Journalism for the Energy Transition. Clean Energy Wire. Accessed: Apr. 1, 2023. [Online]. Available: https://www.cleanenergywire.org/factsheets/what-german-households-pay-electricity

[59] G. C. Okwuibe, M. Wadhwa, T. Brenner, P. Tzscheutschler, and T. Hamacher, "Intelligent bidding strategies in local electricity markets: A simulation-based analysis," in *Proc. IEEE Electr. Power Energy Conf. (EPEC)*, Nov. 2020, pp. 1–7, doi: 10.1109/EPEC48502.2020.9320067.

GODWIN C. OKWUIBE (Member, IEEE) received the B.Eng. degree in electrical engineering from the University of Nigeria, Nsukka, Nigeria, in 2013, and the M.Sc. degree in power engineering from the Technical University of Munich, Munich, Germany, in 2019, where he is currently pursuing the Ph.D. degree. He is currently a Researcher with OLI Systems GmbH, Stuttgart, Germany. His research interests include energy markets, game theory, integration of renewable energy to power grids, distributed generation, and the application of blockchain technology to energy markets.

AMIN SHOKRI GAZAFROUDI received the B.Sc. and M.Sc. degrees in power electrical engineering, in 2011 and 2013, respectively, and the Ph.D. degree in computer science from the University of Salamanca, Salamanca, Spain, in 2019. Afterward, he was a Postdoctoral Researcher, involved in the CoNDyNet2 Project, with the Energy System Modeling Research Group, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, from October 2019 to December 2021. He is currently an Innovation and Project Manager with StromDAO GmbH, Germany, to consult, coordinate, and lead business-to-business and research and development projects in energy communities, local energy and flexibility markets, peer-to-peer (P2P) energy markets, smart grids, smart homes, and electric vehicle (EV) charging strategies. His research interests include power systems and electricity market modeling, power flow and contingency analysis, local energy and flexibility markets design, peer-to-peer energy trading in smart grids, market-based coordination mechanisms, decentralized energy management systems, bidding strategies for autonomous home energy management systems, planning and operation of integrated energy systems, and application of machine learning algorithms on price and demand forecasting.

ESTHER MENGELKAMP received the Ph.D. degree (summa cum laude) from the Karlsruhe Institute of Technology. She is one of the renowned local energy market and P2P trading experts globally. She started researching local energy markets and P2P trading during her Ph.D. Afterward, she served in various positions in business and IT in two of the largest utilities in Germany, the EnBW and Vattenfall. She is currently responsible for the IT of the decentralized heat systems with Vattenfall Energy Solutions, Germany. Since 2019, she is the Founder and current CEO of MK Consulting, a consulting company focusing on bringing local energy markets and energy communities from the pilot and proof-of-concept stage to a scalable, sustainable product. She is consulting several companies—from start-ups to large companies with a focus on the DACH region. In addition, she is offering training on P2P and local energy markets in the Asia–Pacific region.

SARAH HAMBRIDGE received the B.S. degree in applied sciences–biomedical engineering from the University of North Carolina, Chapel Hill, in 2011, and the M.S. and Ph.D. degrees in electrical engineering from North Carolina State University, in 2013 and 2017, respectively. Her research interests include decentralized energy systems, peer-to-peer energy trading, autonomous control systems, transactive control, and energy system modeling and markets. She played a key role in the development of the Grid Singularity Energy Exchange as a Product Owner and a Principal Scientist, until 2022. She is currently employed in the digital grid area with Siemens Energy.

PETER TZSCHEUTSCHLER received the Diploma degree in electrical engineering and information technology from the Technical University of Munich (TUM), in 1998, and the Dr.-Ing. degree in 2005 (his thesis was titled "The Global Potential of Solar-Thermal Electricity Generation"). He joined the Chair of the Energy Economy and Application Technology, TUM, as a Research Associate. From 2009 to 2011, he was a Team Leader of the Group of the TUM International Graduate School for Science and Engineering on the topic "Building—User—Climate." From 2011 to 2014, he was an Operating Agent of Annex 54 "Integration of Microgeneration and Related Technologies in Buildings," an international research group within the framework of the Energy in Buildings and Communities Programme of the International Energy Agency. He is currently with the TUM Chair for Energy Economy and Application Technology. His research interest includes decentralized and renewable energy systems, with a special focus on building energy supply, cogeneration, energy management, and local energy markets.

THOMAS HAMACHER received the Ph.D. degree from the University of Hamburg, Hamburg, Germany, in 1994, for his work on baryonic beta decay after studying physics at the University of Bonn, Bonn, Germany, RWTH Aachen University, Aachen, Germany, and Columbia University, New York, NY, USA. He has been with the Max Planck Institute for Plasma Physics, Garching bei München, Germany, since 1996, most recently as the Head of the Energy and System Studies Group. From 2010 to 2013, he was the Acting Head of the Chair of the Energy Management and Application Technology, Technical University of Munich (TUM), Munich, Germany. In 2013, he was appointed as a Full Professor of renewable and sustainable energy systems at TUM. He is the Director of the Munich School of Engineering, Garching bei München. He is a member of the Environmental Science Centre (WZU), University of Augsburg, Augsburg, Germany. His research interests include energy and systems analysis, urban energy systems, integration of renewable energy into power grids, and innovative nuclear systems (including fusion). His additional research interest includes methods and fundamentals of energy models. He is also a member of the Energy Working Group of the European Physical Society (EPS).

● ● ●