Discovering Energy Communities for Microgrids on the Power Grid

Yuan Hong
Department of Computer Science
Illinois Institute of Technology
10 W 31st St.,
Chicago, IL 60616
yuan.hong@iit.edu

Sanjay Goel
Department of Info Sec
& Digital Forensics
SUNY-Albany
Albany, NY 12222
goel@albany.edu

Haibing Lu
OMIS Department
Santa Clara University
500 El Camino Real
Santa Clara, CA 95053
hlu@scu.edu

Shengbin Wang
North Carolina
A & T State University
1601 E. Market Street
Greensboro, NC 27411
swang@ncat.edu

***Abstract*—Smart grid has integrated an increasing number of distributed energy resources to improve efficiency and flexibility of power generation and consumption as well as the resilience of the power grid. The energy consumers on the power grid (e.g., households) equipped with distributed energy resources can be considered as "microgrids" that both generate and consume electricity. To facilitate energy management, in this paper, we study the energy community discovery problems which identify multiple kinds of communities for the microgrids, such as homogeneous energy community (HEC), mixed energy community (MEC) and self-sufficient energy community (SEC). We present algorithms to discover such energy communities for microgrids, and finally experimentally validate the performance of the algorithms using real datasets.**

1. Introduction

Smart grid superposes a communication network on top of the electrical power network allowing massive sensor data collection from the grid as well as two-way metering of power for users [11]. While collecting and transmitting data across the grid, it allows for the integration of renewable energy resources at the individual consumer level [7]. It creates a paradigm where any individual consumer on the grid can also be a supplier of power: this facilitates the creation of microgrids. Microgrids are localized grids that can be separated from the larger power grid to operate autonomously and be self-sufficient in power. A microgrid typically consists of renewable (wind turbines, solar panels, etc.) and/or non-renewable (micro-turbines, fuel cells, etc.) energy resources, energy storage devices, and energy consuming devices/appliances, all of which are connected through a power and communication network [27]. A microgrid can be operated in a grid with the connected or islanded mode. In the islanded mode, it could be connected to other microgrids or operate independently. Therefore, microgrids can provide energy independence to individual communities or entities who intend to manage their own power generation and distribution [21]. Moreover, microgrids can provide resilience against large-scale failures across the grid: they can continue to operate if large-scale blackouts occur [21].

With autonomous energy, every microgrid may fully or partially feed their local demand. More importantly, numerous microgrids would have great flexibility to utilize their local energy (i.e., sharing) to collaboratively advance the energy management on the power grid, e.g., load balancing [19], energy exchange [28], and load shifting [17], [22]. Therefore, it is desirable to discover various microgrid communities that can efficiently implement their cooperation on the grid [20]. More specifically, based on every microgrid's local energy amount (supply) and its local consumption amount (demand load), we can derive its *Net Energy* as the amount of supply minus the demand load, which can be either *positive* or *negative* over a specified period for the energy supply and demand (from short-term to long-term). In this paper, given the net energy of microgrids, we study the problem of discovering different types of energy communities to facilitate different applications on the power grid (specific energy communities are defined in Section 2 and 3, respectively).

1.1. Literature Review

As the key building blocks on the smart grid, microgrids have attracted significant interests in both industry and academia in the past decade. In such context, many recent research were conducted to design microgrids and/or energy management schemes so as to improve the grid performance, such as load management [2], and demand response solutions [16], [29]. In addition, analysis of data collected from distributed microgrids (e.g., demand load, energy generation and storage) has advanced the energy management of the grid and individual microgrids [23], e.g., short term load forecasting for microgrids [3], and load shifting [22].

Furthermore, some cooperative models among distributed microgrids have been investigated in multiple applications, e.g., optimizing the power loss via a unified microgrid voltage profile [25], eliminating the central energy management unit and price coordinator via localized smart devices [5], and load management via exchanging and sharing local electricity [15], [28]. In this paper, we develop techniques to identify microgrid communities which can directly implement all these cooperations within each energy community to further advance grid performance.

Notice that the energy community discovery problems are significantly different from the prior community discovery problems studied in other contexts, such as geolocations in the spatial data [9] and social graphs [31]. The key difference is that the criteria of *grouping two microgrids into the same energy community should consider not only the spatial distances on the power grid but also their individual net energy amounts*. Moreover, some additional constraints may apply in real world, for example, (1) due to limited power supply to each of the discovered communities, the overall demand load of each community may have an upper bound, (2) in some communities, the difference between the overall demand and supply in each community might be required to be bounded to a small number (for optimizing the performance of the power grid) [19], or (3) some communities may need to possess a non-negative overall net energy to be self-sufficient. In this paper, we will investigate and tackle the above problems.

The rest of this paper is organized as follows. Section 2 and 3 illustrate how to discover different HECs, MECs and SECs. Section 4 presents the experimental results conducted on real world microgrid datasets. Finally, Section 5 gives the concluding remarks and discusses the future work.

2. Homogeneous Energy Communities

TABLE 1. FREQUENTLY USED NOTATIONS AND ACRONYMS

HEC	homogeneous energy community
MEC	mixed energy community
SEC	self-sufficient energy community
m_i, e_i	microgrid and its net energy
$Dis(m_i, m_j)$	distance between two microgrids on the grid
c_j	an energy community
$NE(m_i, m_j)$	net energy distance of two microgrids
M^+	set of microgrids with positive net energy
M^-	set of microgrids with negative net energy

We first look at the case that all the microgrids have excessive local energy (positive net energy), or all of them have to request external demand (negative net energy). Such community is defined as below.

Definition 1 (Homogeneous Energy Community (HEC)). A set of microgrids whose net energy are exclusively positive; or exclusively negative.

2.1. Discovering Fixed Number of HECs

In real world, the smart grid may plan to partition a set of homogeneous microgrids (exclusively positive or negative) into a fixed number of HECs. For instance, the grid intends to place K groups of new generators at K different substations respectively to provide power supply to some newly established microgrids (e.g., new constructions), then it is desirable to partition a set of microgrids with *negative net energy* into K different HECs such that the grid can increase the power supply to those HECs; or the grid plans to partition a set of microgrids with *positive net energy* into K different HECs such that the grid can establish K different energy banks to store the excessive energy at different locations respectively.

At this time, the grid will try to identify K different HECs on the power grid. Then, this energy community discovery problem can be considered as a clustering problem based on the distances of microgrids' geolocations on the electricity transmission network, where the number of clusters is given as K and the net energy of each HEC can be aggregated. The classic K-Means algorithm [24] can efficiently generate K HECs among N microgrids based on *their distances* on the power grid.

For all $j \in [1, K]$, any HEC c_j's net energy can be aggregated as $E_j = \sum_{\forall m_i \in c_j} e_i$. Then, such HECs could help the power grid better manage their energy. For instance, if $\forall i \in [1, N], e_i < 0$, K new energy resources with power supply amount $\forall j \in [1, K], |E_j|$ will be placed at the closest substations to the centroids of the HECs respectively.

2.2. Discovering HECs with Bounded Net Energy

Some real world constraints may require that each HEC's net energy (either positive or negative) should be bounded, e.g., the external supply to every HEC is limited due to capacity of generators. In these cases, the number of communities is unknown and some additional constraints should apply – i.e., in each HEC, if the HEC's net energy is positive, then it cannot exceed a positive upper bound L; otherwise (negative net energy), it cannot be less than $-L$ (viz. each HEC's external demand is no greater than L).

Without loss of generality, we consider the negative energy case – the external demand of each HEC should be bounded by L. To find such HECs, we extend the Density-based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [10] to "L-DBSCAN" by adding an upper bound L for the external demand of each HEC. Specifically, three different microgrids can be defined [10]:

- **Core microgrid**: a microgrid m_i has at least min microgrids within distance ϵ of it on the grid.
- **Reachable microgrid**: a microgrid m_j is reachable from microgrid m_i if there is a path $m_i, \ldots, m_j$, where the next microgrid is directly reachable from the previous microgrid on the path and all the microgrid except m_j are core microgrids.
- **Outlier**: not reachable from any other microgrids.

The basic idea of DBSCAN algorithm is to group together *reachable microgrids* by reaching them from the *core microgrids*: scanning neighbor microgrids from the core microgrids. However, different from the DBSCAN algorithm [10], discovering HECs should take into account each microgrid's external demand ($\forall i \in [1, N], |e_i|$) as well as the spatial distances on the grid. Our L-DBSCAN algorithm first groups microgrids based on the spatial distances between their geolocations (similar to DBSCAN). Then, with the bounded external demand L of each HEC, the L-DBSCAN algorithm will stop scanning microgrids for the current HEC once its aggregated external demand gets close to L, and then initialize a new HEC to continue scanning the microgrids based on their geolocations. Finally, all the outliers should be assigned to their nearest communities if the updated external load remains no greater than L. If no such communities found, L-DBSCAN groups the outliers to form new HECs.

3. Mixed Energy Communities

Among thousands of microgrids on the power grid, some of them may have excessive energy while some others may request energy from external resources (e.g., main grid). Therefore, adjacent microgrids can share their locally generated electricity for reduced energy loss on transmission and better reliability and resilience of power supply [15], [28]. Such microgrids can form different energy communities to feed their local energy demands, which are beneficial to both the power grid and individual microgrids. Clearly, the net energy of the microgrids in the communities are mixed with negative and positive, thus denoted as "Mixed Energy Communities" (MECs).

Definition 2 (Mixed Energy Community (MEC)). A set of microgrids whose net energy are mixed with positive and negative.

The ideal case of the discovered MECs is that all the microgrids in the same MEC are geographically close to each other while balancing the local demand and supply of each MEC within a tight margin [19] (e.g., zero net energy [4]). In Section 3.1, we propose an algorithm to identify such MECs towards this goal. In Section 3.2, we present another algorithm to discover a special form of MECs on the grid – self-sufficient energy communities (SECs).

3.1. Discover MECs with Two Distance Thresholds

Similar to the HECs, each microgrid m_i's net energy is denoted as e_i, which can be either *positive* or *negative*. While grouping two microgrids (e.g., m_i and m_j) into an MEC, besides the spatial distance between them on the grid $Dis(m_i, m_j)$, we also have to consider their net energy e_i and e_j towards the load balancing of their community – the overall demand and supply should be balanced (ideally, equal to each other). For example, if one microgrid has a net energy e_i while the other microgrid has a net energy demand $-e_i$, such two microgrids can supply their demand using their local energy in the same community. Thus, we define a novel measure namely "Net Energy (NE)" distance of two microgrids m_i and m_j as:

$$NE(m_i, m_j) = |e_i + e_j| \tag{1}$$

If $e_i = -e_j$, we have $NE(m_i, m_j) = 0$. However, if $e_i = e_j$, we have $NE(m_i, m_j) = 2|e_i|$. The NE distance differs from other distance measures used in community discovery problems due to its unique feature: two opposite values (e.g., e_i and $-e_i$) are measured as "close".

For MECs discovery, we define *two maximum distance thresholds* for the normalized NE distances and the normalized spatial distances respectively[1]: $\epsilon, \epsilon' \in [0, 1]$. Then, we propose a novel agglomerative algorithm [30] to identify MECs by utilizing ϵ and ϵ' to specify the criteria for bounding the differences between the overall supply and demand of each community and the spatial distances between the microgrids in each community. Specifically, we let each microgrid find its closest microgrid (with an NE distance $\leq \epsilon$ and a spatial distance $\leq \epsilon'$) to form an MEC, update the MEC centroid's geolocation and net energy, and then hierarchically merge "small MECs" to form "large MECs" (for *pursuing better resilience*). The merging process terminates if the NE distance between two MECs' centroids exceeds ϵ or their spatial distance exceeds ϵ'. Algorithm 1 presents the details.

Algorithm 1 MECs Discovery

Input: maximum threshold of the NE distances ϵ
maximum threshold of the spatial distances ϵ'
Output: MECs

1: **while** any ungrouped microgrid m_i in $m_1, \ldots, m_N$ **do**
2: initialize a new MEC with m_i: $c_j = \{m_i\}$
3: **for** each ungrouped microgrid m_k **do**
4: compute MEC c_j's net energy: E_j and its centroid's geolocation μ_j
5: **if** $NE(E_j, m_k) \leq \epsilon$ and $Dis(\mu_j, m_k) \leq \epsilon'$ **then**
6: $c_j = c_j \cup m_k$ (add m_k to the MEC c_j)
7: update E_j and μ_j
8: considering each MEC c_j as a microgrid with net energy E_j and geolocation μ_j, repeat Line 1-7 to hierarchically merge the MECs based on ϵ and ϵ' until convergence

Therefore, the difference of the overall supply and demand of every MEC is bounded/balanced by ϵ, and the spatial distance between any microgrid and its MEC's centroid is bounded by ϵ'.

3.2. Discover Self-sufficient Energy Communities

Many real world cases require that the microgrids in each MEC can fully supply their demand with their local energy (e.g., large-scale blackouts). Therefore, it is also desirable to discover the "Self-sufficient Energy Communities (SECs)", defined as below.

Definition 3 (Self-sufficient Energy Community (SEC)). A set of microgrids with non-negative overall net energy.

Algorithm 1 identifies the MECs with balanced load and bounded spatial distance (by ϵ and ϵ' respectively). If the overall power supply of all N microgrids is greater than their demand, most of the MECs identified by Algorithm 1 can be SECs. However, if the overall power supply of

1. Microgrids m_i and m_j's spatial distance $Dis(m_i, m_j)$ and net energy distance $|e_i + e_j|$ can be normalized into $[0, 1]$, e.g., divided by $max(Dis(m_i, m_j))$ and $\sum_{i=1}^{N} |e_i|$ respectively.

all N microgrids is significantly less than their demand, many MECs identified by Algorithm 1 may not be able to feed themselves. To address this issue, we develop a novel algorithm to discover a subset of microgrids to form a number of SECs out of N microgrids. If the net energy turns larger, more microgrids (up to all N microgrids) will be involved in the SECs.

Specifically, among all the N microgrids, we denote the set of microgrids with positive net energy as M^+, and the set of microgrids with negative net energy as M^-. The algorithm first clusters all the microgrids in M^+ based on their geolocations, where each cluster can be considered as a "bigger microgrid" with aggregated positive net energy. In this stage, we extend the K-Means algorithm [24] to cluster such microgrids' geolocations by specifying different $K \in \{K_{min}, \dots, K_{max}\}$. Then, the algorithm repeats K-Means with different K values and chooses the best clustering result – the minimum sum of squared errors (SSE) of the spatial distances [30] in all the clustering results.

Denoting clustering result of M^+ as $c_1^*, \dots, c_K^*$, the net energy of any cluster $\forall j \in [1, K]$, c_j^* can be aggregated as $\sum_{\forall m_i \in c_j^*} e_i$. Then, $\forall j \in [1, K], c_j^*$ iteratively adds its centroid's closest ungrouped microgrid in M^- until its net energy drops close to 0. Finally, the updated $c_1^*, \dots, c_K^*$ are identified as K different SECs, as shown in Algorithm 2.

Algorithm 2 SECs Discovery

Input: M^+: set of microgrids with positive net energy
M^-: set of microgrids with negative net energy
$\{K_{min}, \dots, K_{max}\}$: possible values for K
Output: SECs
for $K = K_{min}, \dots, K_{max}$ **do**
  run K-Means for all microgrids in M^+ based on their geolocations to obtain $c_1, \dots, c_K$
choose the best clustering result with the minimum SSE for different K: $c_1^*, \dots, c_K^*$
for $j \in [1, K]$ **do**
  compute the centroid of c_j^* as μ_j^*
  while $\sum_{\forall m_i \in c_j^*} e_i \geq 0$ **do**
    find μ_j^*'s closest ungrouped microgrid in M^-, denoted as m_k
    $c_j^* = c_j^* \cup m_k$ (add m_k to the SEC c_j)
    update c_j^*'s net energy: $\sum_{\forall m_i \in c_j^*} e_i += e_k$ and μ_j^*
  if $\forall m_k \in M^-$ are grouped, then break
return the updated $c_1^*, \dots, c_K^*$ as SECs

Note that all the microgrids in M^+ are involved in the SECs, but not all the microgrids in M^- (depending on the net energy of the microgrids in M^+ and M^-). Furthermore, the net energy of most self-sufficient communities can be well balanced to form "Zero Net Energy" communities [8].

4. Experiments

4.1. Experimental Setup

Datasets. Our experimental simulations were conducted on three real world datasets: a spatial dataset and two power generation & consumption datasets. First, the spatial dataset of 115,475 cities/towns in the U.S. was collected by the US Geological Survey on July 7, 2012 and is available in National Imagery and Mapping Agency [1]. Second, two power generation & consumption datasets were collected by Richardson et al. [26] in East Midlands, UK and Barker et al. [6] in Massachusetts, US respectively. In our experiments, we integrate the spatial dataset with each of the power generation & consumption datasets. Table 2 shows the characteristics of the datasets.

TABLE 2. CHARACTERISTICS OF DATASETS

Datasets	Characteristics
Spatial Data	115,475 unique geolocations
UK Dataset Consumption (Watts)	16,060 microgrids average consumption rate: 953 max consumption rate: 2,891 min consumption rate: 140
UMass Dataset Generation/Consumption (Watts)	6,480 microgrids average generation rate: 776 average consumption rate: 1,045 max generation rate: 1,250 max consumption rate: 2,147 min generation rate: 355 min consumption rate: 192

Normalization. We use Euclidean distance to measure the spatial distance between any two microgrids on the grid. Both the Euclidean distances and the net energy (NE) distances are normalized into $[0, 1]$ in all the experiments.

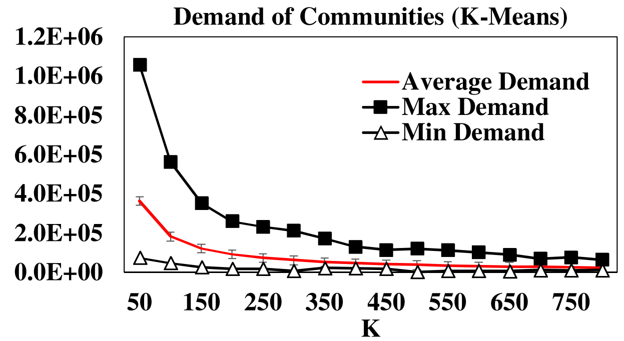
4.2. Discovering HECs

The algorithms have identical performance to discover HECs with negative and positive net energy. Without loss of generality, we evaluate the case of negative net energy (external demand) for discovering HECs.

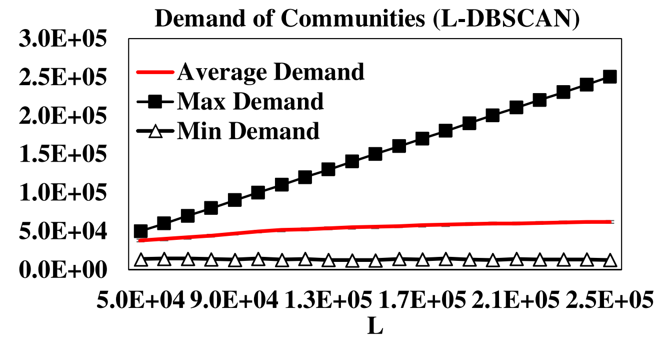
We first implement the K-Means algorithm [18] and L-DBSCAN to discover HECs from 16,060 microgrids, and then aggregate the external demand in each HEC. In literature, the performance of K-Means and DBSCAN algorithm on clustering has been well studied using measures such as sum of squared errors (SSE) and silhouette coefficient [30] to evaluate the cohesion and separation of the clusters. Therefore, we do not report the spatial cohesion and separation of the HECs on the grid here. Figure 1 shows the external demand of the discovered HECs in two different cases.

First, as shown in Figure 1(a), the average, maximum and minimum external demands of the HECs decline as K increases, where parameter K varies $\in [50, 800]$. Note that if $K = 50$, a large HEC ($\sim 1,250$ microgrids) can be identified to request external energy (with an amount $\sim 10^6$ Watts), then the external demands of the HECs drop significantly as K increases.

Second, while using L-DBSCAN to discover HECs with negative energy, we set a reasonable value for the normalized minimum distance (Euclidean) $\epsilon = 0.1$ and the core



(a) External Demand (Negative Net Energy) of HECs (in Watts) vs. Number of HECs K



(b) External Demand (Negative Net Energy) of HECs (in Watts) vs. Net Energy Bound in Each HEC L

Figure 1. External Demand of the Discovered HECs

microgrid's minimum number of neighbors $min = 10$. Table 3 presents the number of discovered HECs in the same experimental setting. We can observe that the number of HECs decreases as L increases, because each HEC can involve more reachable microgrids with a higher L. Furthermore, Figure 1(b) shows the average, maximum and minimum external demands of all the HECs. The maximum external demand of all the HECs always equals L since the net energy bound L is the major constraint besides the distances of microgrids' geolocations. However, the average and minimum external demand of all the HECs tend flat as L increases. In reality, the HEC with the minimum external demand only includes a small number of microgrids, where not many microgrids can be reachable from other (core) microgrids. Thus, the minimum external demand of such HEC is far less than L in general.

TABLE 3. NUMBER OF HECS FOR L

L Watts	# of HECs	L Watts	# of HECs
50,000	233	150,000	66
60,000	207	160,000	62
70,000	181	170,000	55
80,000	154	180,000	52
90,000	129	190,000	47
100,000	107	200,000	46
110,000	92	210,000	43
120,000	85	220,000	42
130,000	77	230,000	40
140,000	71	240,000	39

4.3. MECs Discovery

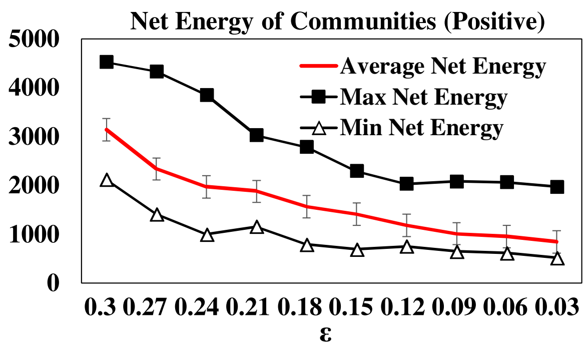
We conduct the experiments for the MECs discovery algorithm on the UMass Smart* Microgrid dataset [6].

4.3.1. Discovering MECs with ϵ and ϵ'. Recall that the net energy of all the 6,480 microgrids (overall power generation minus overall power consumption) is *negative*. To test the effectiveness of Algorithm 1 in two different cases (1) positive net energy and (2) negative net energy, we extract two subgroups of microgrids from the 6,480 microgrids, each of which includes 2,000 microgrids and has positive and negative net energy respectively. For simplicity of notations, these two subsets of microgrids are named as "Positive" and "Negative" respectively.

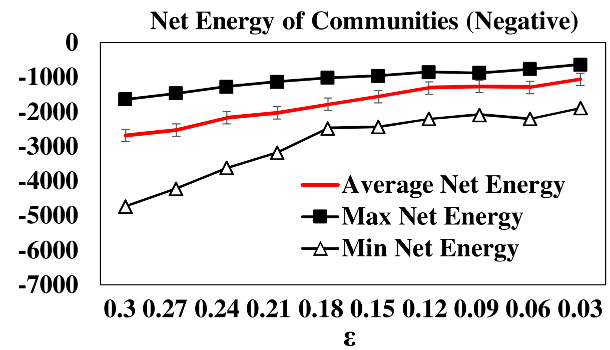
First, we implement Algorithm 1 with $\epsilon \in [0.03, 0.3]$ where the normalized spatial distance threshold ϵ' is fixed as a reasonable value 0.05. Then, Figure 2(a) shows the average, maximum and minimum net energy of all the communities generated from "Positive" where $\epsilon \in [0.03, 0.3]$. As ϵ drops from 0.3 to 0.03, the allowed maximum differences between the overall demand and overall supply in every MEC decline significantly. The average, maximum and minimum net energy then decrease close to 0 as ϵ decreases. Thus, the demand and supply of the MECs become better balanced with a net energy closer to 0. On the contrary, Figure 2(b) demonstrates the results for "Negative", which present a reverse trend as "Positive", but still tend to better balanced load (net energy lies closer to 0) as ϵ decreases.

Second, we fix $\epsilon = 1$ and $\epsilon' = 0.05$ in Algorithm 1, which then removes the constraint of NE distances and turns into a regular agglomerative clustering problem based on geolocations. Then, we compute the SSE in such case as the *benchmark SSE* (say SSE_0) and test how the spatial distances (viz. SSE) within each MEC vary for different levels of balanced load (different ϵ). Specifically, we fix $\epsilon' = 1$ (then Algorithm 1 only specifies the maximum NE distance threshold ϵ and ignores spatial distances), generate the MECs with $\epsilon \in [0.03, 0.3]$ for two inputs "Positive" and "Negative" respectively, and compute the corresponding SSE for each MEC. Then, we define a new measure SSE ratio as $\frac{SSE}{SSE_0}$ and plot all the results in Figure 3. Clearly, the SSE increases as ϵ declines – an MEC with better balanced load would include the furthest microgrids from each other if not bounding the spatial distances within each MEC.

4.3.2. SECs Discovery. In the experimental dataset, 34.4% (2,229) and 65.6% (4,251) of the 6,480 microgrids have positive and negative net energy respectively (leading to negative overall net energy). We implement Algorithm 2 on all 6,480 microgrids to discover the SECs, where $K = \{50, 60, \ldots, 200\}$. The characteristics of the discov-



(a) Net Energy in the MECs vs. Normalized Net Energy Distance Threshold ϵ (2,000 Micrograms with Positive Net Energy)



(b) Net Energy in the MECs vs. Normalized Net Energy Distance Threshold ϵ (2,000 Microgrids with Negative Net Energy)

Figure 2. MECs Discovery

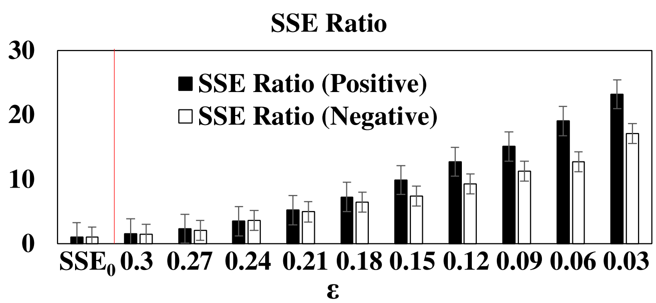


Figure 3. SSE Ratio vs. Normalized Net Energy Distance Threshold ϵ

ered SECs are depicted in Table 4.

TABLE 4. SELF-SUFFICIENT ENERGY COMMUNITIES

average net energy of all the SECs	47.6
max net energy of all the SECs	127
min net energy of all the SECs	0
number of SECs (all positive net energy): best K	90
number of micorgrids in all the SECs	3,923
average number of microgrids in the SECs	42.5
microgrids (with positive net energy) in the SECs	2,229
microgrids (with negative net energy) in the SECs	1,694

While discovering the SECs, all the 2,229 microgrids with positive net energy are clustered to form the communities of energy resources, and their electricity can supply additional 1,694 microgrids with negative net energy. The average and maximum net energy of all the communities are quite close to 0. This proves that the demand and supply of most SECs are well balanced (with net energy close to 0).

4.4. Efficiency

Finally, we evaluated the computational performance of all the algorithms based on different input sizes (number of microgrids), and plotted the results in Figure 4. Note that KM, L-DBSCAN, MEC, and SEC denote K-Means, L-DBSCAN, MECs Discovery, and SECs Discovery respectively. Specifically, two HECs discovery algorithms (K-Means and L-DBSCAN) are extremely efficient with fixed parameters (e.g., HECs number in the K-Means, three parameters in the L-DBSCAN). For discovering MECs and SECs, Algorithm 1 hierarchically groups close microgrids and merge the communities, and Algorithm 2 identifies the optimal communities with positive net energy for different K. Thus, they need relatively more runtime than K-Means and L-DBSCAN. Indeed, in Algorithm 2, clustering is only applied to the microgrids with positive net energy. Also, the step of finding microgrids with negative net energy in Algorithm 2 is highly efficient with a complexity of $O(N)$. Thus, both Algorithm 1 and 2 are still efficient, as shown in Figure 4.

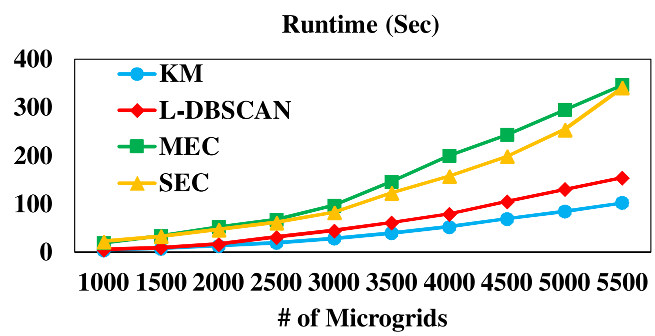


Figure 4. Computational Performance

5. Conclusion and Future Work

Energy communities formed by distributed energy resources (viz. microgrids) could facilitate the power grid to advance energy management and enable microgrids to find cooperative peer microgrids (e.g., sharing energy). In this paper, we have proposed a series of approaches to identify different energy communities for the microgrids, including homogeneous energy communities, mixed energy communities and self-sufficient energy communities. We have also validated the effectiveness and efficiency of the approaches using real world datasets.

From the economic perspective, not all the microgrids would be willing to share their excessive energy without benefits. For this reason, we plan to investigate the energy community discovery problems from an economic perspective, e.g., microgrids can sell their energy to each other at different times. Moreover, such local energy trade may also affect the global electric prices provided by the utility companies. In the future, we will explore effective models for the above research problems. Furthermore, beyond discovering communities based on the static energy generation and consumption at specific times, we will explore stochastic optimization models for energy community discovery based on time series power generation and consumption. Finally, the process of discovering energy communities requests all the microgrids to fully disclose their local data (i.e., demand and supply) to a trusted-third party, and thus results in privacy concerns in centralized [12] or distributed environment [13], [14], we plan to explore solutions to tackle such concerns in the future.

References

[1] http://www.math.uwaterloo.ca/tsp/.

[2] M. Ahmadi. Optimizing load control for a residential microgrid in a collaborative environment. *ProQuest Dissertations and Theses*, 1541222(3):66, 2013.

[3] N. Amjady, F. Keynia, and H. Zareipour. Short-term load forecast of microgrids by a new bilevel prediction strategy. *IEEE Trans. Smart Grid*, 1(3):286–294, 2010.

[4] P. Arboleya, C. Gonzalez-Moran, M. Coto, M. C. Falvo, L. Martirano, D. Sbordone, I. Bertini, and B. D. Pietra. Efficient energy management in smart micro-grids: ZERO grid impact buildings. *IEEE Trans. Smart Grid*, 6(2):1055–1063, 2015.

[5] N. R. Asr, U. Ojha, Z. Zhang, and M. Chow. Incremental welfare consensus algorithm for cooperative distributed generation/demand response in smart grid. *IEEE Trans. Smart Grid*, 5(6):2836–2845, 2014.

[6] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht. Smart*: An open data set and tools for enabling research in sustainable homes. In *the 2012 Workshop on Data Mining Applications in Sustainability*, 2012.

[7] S. Bush, S. Goel, and G. Simard. Ieee vision for smart grid communications: 2030 and beyond roadmap. *IEEE Vision for Smart Grid Communications: 2030 and Beyond Roadmap*, pages 1–19, Sept 2013.

[8] N. Carlisle, O. V. Geet, and S. Pless. Definition of a ”zero net energy” community. *National Renewable Energy Laboratory Technical Report*, November, 2009.

[9] M. Coscia, F. Giannotti, and D. Pedreschi. A classification for community discovery methods in complex networks. *Statistical Analysis and Data Mining*, 4(5):512–546, 2011.

[10] M. Ester, H. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of KDD ’96*, pages 226–231, 1996.

[11] X. Fang, S. Misra, G. Xue, and D. Yang. Smart grid - the new and improved power grid: A survey. *IEEE Communications Surveys and Tutorials*, 14(4):944–980, 2012.

[12] Y. Hong, J. Vaidya, H. Lu, and M. Wu. Differentially private search log sanitization with optimal output utility. In *Proceedings of Extending Database Technology*, pages 50–61, 2012.

[13] Y. Hong, J. Vaidya, and H. Lu. Secure and efficient distributed linear programming. *Journal of Computer Security*, 20(5):583–634, 2012.

[14] Y. Hong, J. Vaidya, H. Lu, P. Karras, and S. Goel. Collaborative search log sanitization: Toward differential privacy and boosted utility. *IEEE Trans. Dependable Sec. Comput.*, 12(5):504–518, 2015.

[15] Y. Hong, S. Goel, and W. Liu. An Efficient and Privacy Preserving Scheme for Energy Exchange among Smart Microgrids. *International Journal of Energy Research*, 40(3):313–331, 2016.

[16] Y. Hong, W. Liu, and L. Wang. Privacy Preserving Smart Meter Streaming against Information Leakage of Appliance Status. *IEEE Trans. Information Forensics and Security*, 12(9):2227-2241, 2017.

[17] Y. Hong, S. Wang, and Z. Huang. Efficient Energy Consumption Scheduling: Towards Effective Load Leveling. *Energies*, 10(1):105, 2017.

[18] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):881–892, 2002.

[19] J. Kennedy, P. Ciufo, and A. Agalgaonkar. Intelligent load management in microgrids. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–8, July 2012.

[20] M. King. Community energy: Planning, development and delivery. *International District Energy Association*, 2012.

[21] R. Lasseter. Microgrids. In *Power Engineering Society Winter Meeting, 2002. IEEE*, volume 1, pages 305–308, 2002.

[22] X. Liang, X. Li, R. Lu, X. Lin, and X. Shen. Udp: Usage-based dynamic pricing with privacy preservation for smart grid. *IEEE Trans. Smart Grid*, 4(1):141–150, 2013.

[23] T. Logenthiran, D. Srinivasan, and T. Z. Shun. Demand side management in smart grid using heuristic optimization. *IEEE Trans. Smart Grid*, 3(3):1244–1252, 2012.

[24] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, pages 281–297, 1967.

[25] A. Maknouninejad and Z. Qu. Realizing unified microgrid voltage profile and loss minimization: A cooperative distributed optimization and control approach. *IEEE Trans. Smart Grid*, 5(4):1621–1630, 2014.

[26] I. Richardson, M. Thomson, D. Infield, and C. Clifford. Domestic electricity use: A high-resolution energy demand model. *Energy and Buildings*, 42(10):1878 – 1887, 2010.

[27] T. Roughan. Workshop on microgrid technologies and applications. In *National Grid*, 2013.

[28] W. Saad, Z. Han, H. V. Poor, and T. Basar. Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications. *IEEE Signal Process. Mag.*, 29(5):86–105, 2012.

[29] M. R. Sarker, M. A. Ortega-Vazquez, and D. S. Kirschen. Optimal Coordination and Scheduling of Demand Response via Monetary Incentives. *IEEE Trans. Smart Grid*, 6(3):1341–1352, 2015.

[30] P. Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Addison-Wesley, 2005.

[31] J. Yang, J. J. McAuley, and J. Leskovec. Community detection in networks with node attributes. In *2013 IEEE 13th International Conference on Data Mining*, pages 1151–1156, 2013.

Acknowledgments

This work is partially supported by the National Science Foundation under Grants No. CNS-1618221 and CNS-1745894. We acknowledge the real world data support from the Trace Repository at UMass and Center for Renewable Energy Systems Technology at Loughborough University. Meanwhile, we sincerely thank the anonymous reviewers for their very constructive comments.