A Survey of Clustering Techniques in WSNs and Consideration of the Challenges of Applying Such to 5G IoT Scenarios

Lina Xu, Rem Collier, and Gregory M. P. O'Hare, *Member, IEEE*

***Abstract*—Wireless sensor network (WSN) systems are typically composed of thousands of sensors that are powered by limited energy resources. To extend the networks longevity, clustering techniques have been introduced to enhance energy efficiency. This paper presents a survey on clustering over the last two decades. Existing protocols are analyzed from a quality of service (QoS) perspective including three common objectives, those of energy efficiency, reliable communication and latency awareness. This review reveals that QoS aware clustering demands more attention. Furthermore, there is a need to clarify how to improve quality of user experience (QoE) through clustering. Understanding the users' requirements is critical in intelligent systems for the purpose of enabling the ability of supporting diverse scenarios. User awareness or user oriented design is one remaining challenging problem in clustering. In additional, this paper discusses the potential challenges of implementing clustering schemes to Internet of Things (IoT) systems in 5G networks. We indicate that clustering techniques enhanced with smart network selection solutions could highly benefit the QoS and QoE in IoT. As the current studies for WSNs are conducted either in homogeneous or low level heterogeneous networks, they are not ideal or even not able to function in highly dynamic IoT systems with a large range of user scenarios. Moreover, when 5G is finally realized, the problem will become more complex than that in traditional simplified WSNs. Several challenges related to applying clustering techniques to IoT in 5G environment are presented and discussed.**

***Index Terms*—5G, clustering, HetNets, Internet of Things (IoT), quality of service (QoS), quality of user experience (QoE), wireless sensor networks (WSNs).**

I. Introduction

Wireless sensor networks (WSNs) are networks composed of distributed micro-devices embedded with various sensing abilities (called sensors), which are used to monitor the environment and send the information back to the end users. WSN technologies were introduced more than 20 years ago and many projects have been proposed and undertaken that embrace this technology [1], [2]. Green computing [3] was introduced in 2008 with the purpose of reducing the use of limited resources and maximizing energy efficiency during the lifetime of a system. WSNs typically include a large number of sensors that are equipped with limited energy resources, but are required to operate without recharging or replacing batteries for extended periods of time. In order to prolong a networks longevity, clustering techniques have been introduced to achieve energy efficient communication between sensors.

Manuscript received March 31, 2017; revised June 8, 2017; accepted July 4, 2017. Date of publication July 12, 2017; date of current version October 9, 2017. *(Corresponding author: Lina Xu.)*

The authors are with the Department of Computer Science, University College Dublin, Dublin D4, Ireland (e-mail: lina.xu@ucd.ie; rem.collier@ucd.ie; gregory.ohare@ucd.ie).

Digital Object Identifier 10.1109/JIOT.2017.2726014

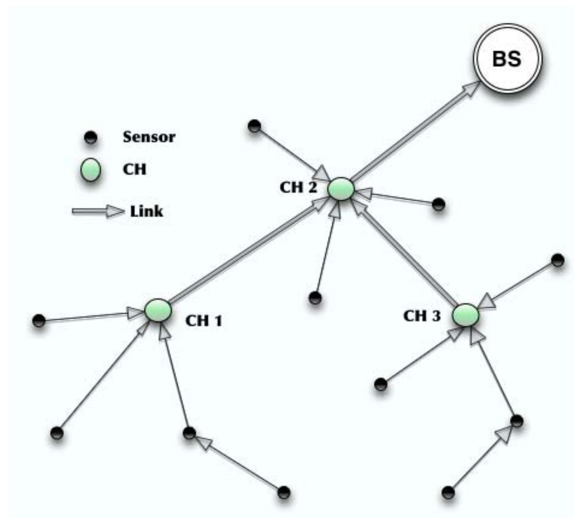

Fig. 1. Example of cluster-based WSN.

For large WSN systems, topology control ought to be applied to balance the network load, increase the network scalability, and prolong the network lifetime. Clustering techniques are also one of the approaches to topology control, which can organize a WSN into a cluster-based network. Task scheduling, data gathering and transmission power control (TCP) algorithms can be implemented in this structure in order to achieve specific objectives. A clustering algorithm can partition sensors into different clusters/groups, as shown in Fig. 1. In each cluster, a cluster header (CH) is elected to be in charge of generating a transmission schedule, gathering data from all the sensors in the cluster and transmitting the assembled data back to the base station (BS). Based on the clustered structure, the system can maintain a longer life by scheduling the duty cycle between the sensors within a cluster, without harming the functionality of the network. In addition to saving energy from scheduling, a sensor can also reduce energy consumption from communication since it only needs to communicate with a local CH rather than a far located BS. Clustering algorithms are proposed and have become essential in WSNs primarily for two reasons.

1) *Sleep Scheduling:* Scheduling the duty cycle between the sensors in a cluster cannot only prolong the network lifetime, but can also harmonize the power usage throughout the network.
2) *Density Control:* Activating partial sensors in the network can lower the density of the network and in turn avoid the conflict at the medium access control (MAC) layer. Achieving this can reduce the communication latency and energy consumption caused by retransmission [4].

Energy saving is the most original temptation motivating clustering. This type of heuristic—reducing long-distance communication ratio, inspires many other related studies. Afterwards, a wide body of clustering algorithms has become available. Different heuristics can be adopted depending on the clustering purposes. However, the ultimate goal for clustering techniques is to divide sensors in a WSN to different clusters. After joining a cluster, a sensor normally only needs to communicate with its own CH. A CH can communicate with the BS directly or through other CHs, as shown in Fig. 1. The routing between the sensors in the same cluster is called *intracluster routing*. The routing between the CHs and the BS is called *intercluster routing*. The routing scheme can be either single hop or multihop, and is dependent upon several factors, such as the objectives of a clustering algorithm or the communication capability of the sensors.

In general, a clustering algorithm has two main phases.

1) *CH Election:* The CH is the leader of a cluster and it is in charge of gathering data and transmitting the data to the BS. The CHs will consume more energy than normal sensors and therefore, run out of power sooner. CHs are normally rotated between different sensors to balance the power usage on each sensor. Which sensors ought to be elected as CHs requires careful investigation. Generally, one of three CH election schemes is adopted.
   a) *Deterministic:* In this scheme, the CHs are preset and placed at fixed locations in a network. When super nodes (with powerful processing ability and high energy storage) exist, the deterministic CH election can maintain the stability of the network and avoid the energy and time consumption associated with the frequent CH election. However, this case is not common since the sensors in fields are normally homogeneous and the super nodes can die for unexpected reasons.
   b) *Random:* CHs can be selected among sensors based on randomly generated values. If a network is homogeneous, random CH election scheme is a simple and beneficial strategy.
   c) *Adaptive:* Adaptive CH electing scheme provides an alternative approach from that of the random scheme. Instead of electing CHs based on random values, the adaptive CH election is based on some particular parameters, such as remaining residual energy or distance to the BS. With different system objectives, a clustering algorithm can use a specific combination of these parameters. An adaptive CH election approach is designed to adapt to the variations of the network and the environment.

   A CH election algorithm can be either *centralized* or *distributed*. For centralized algorithms, the BS normally has access to the information pertaining to the entire network. A centralized algorithm can provide a global optimal solution, but may consume more energy and time. For distributed algorithms, without the global information of the network, the sensors can only achieve a local optimal performance based on the local information. However, they can be more reliable, energy and time efficient.
2) *Cluster Formation/CH Selection:* After being elected, the CHs will advertise themselves by broadcasting their information to other sensors. Each sensor will gather the information from all the CHs within its communication range and then decide which CH to join based on some communication properties. Several metrics can be used to determine the communication properties between a sensor and a CH, such as communication cost, hop count or even physical distance. In some cases, the size of clusters is also considered when sensors join a cluster.

Algorithms with different CH election or cluster formation schemes will exhibit diverse performances. For example, compared to random CH election scheme, energy aware CH election, which can elect high residual energy sensors as CHs, can balance the power usage throughout a network better. However, the election process may need more energy and time. All clustering algorithms have their own specific CH election schemes and cluster formation schemes. Depending on the requirements of a system, different strategies may be applied in order to achieve specific objectives.

The intercluster communication and intracluster communication can be single hop or multihop. To realize multihop communication in clustering, several algorithms, such as max–min [5] and Khopca [6] are proposed. Max–min algorithm has introduced a heuristic to achieve k-hop cluster network structure with a time complexity of $O(k)$ rounds, reducing from $O(n)$ rounds in earlier work [7] (where n is the number of nodes). In this algorithm, a sensor can join a CH at most k wireless hops away.

Clustering is one of the major approach to green computing in WSNs which can be harnessed in many systems. Those techniques can extend the longevity of a WSN through partially activating the sensors in the network. Existing clustering algorithms, including review papers, focus on solving energy efficiency problems, ignoring other quality of services (QoSs) requirements, like transmission reliability or network latency. Reviewing existing studies from the QoS perspective is required. Besides, being able to detect the user's preference and being aware of the system context also become more attractive features in intelligent systems.

Since the concept of Internet of Things (IoT) has been proposed, the corresponding systems are widely deployed to assistant people's everyday life. IoT systems differ from WSN systems for its high diversity and usability. Instances for IoT

TABLE I
LIST OF PREVIOUS SURVEY PAPERS

Year	Survey papers
2006	Arboleda 2006 [8], Younis 2006 [9]
2007	Ameer 2007 [10]
2008	Kumarawadu 2008 [11], Deosarkar 2008 [12]
2009	Jiang 2009 [13]
2010	Lotf 2010 [14], Boyinbode 2010 [15]
2012	Naeimi 2012 [16], Aslam 2012 [17], Mundada 2012 [18] Ramesh 2012 [19], Liu 2012 [20]
2013	Sudhanshu 2013 [21], Jan 2013 [22], Jain 2013 [23] Kumari 2013 [24], Jindal 2013 [25], Subha 2013 [26]
2014	Dhawan 2014 [27], Nayyar 2014 [28], Afsar 2014 [29],
2015	Liu 2015 [30], Santar 2015 [31] Ouafaa 2015 [32], Zanjireh 2015 [33]
2016	Pradhan 2016 [34] [35]

include smart home/buildings, connected cars, etc. With the development of future networks/5G networks, migrating current IoT systems to those advanced communication platform is the trend toward wireless. In this paper, we have also discussed the differences between WSNs and IoT along with the challenges of applying clustering techniques into IoT systems based on 5G networks.

The remainder of this paper is organized as follows. Related work is delivered in Section II. The clustering techniques are analyzed in detail by classifying the existing approaches into two categories: 1) Voronoi structure-based and 2) non-Voronoi structure-based in Section III. Section IV presents QoS criteria in clustering algorithms from a network perspective, including energy efficiency, transmission reliability, and network latency. Section V indicates that one QoS criterion from a user's perspective—quality of user experience (QoE), is missing from current work. Since the clustering techniques are normally utilized in a large network, in order to examine the usability and scalability of an algorithm, a simulator is normally be used. The advantages and disadvantages of the different simulators are compared in Section VI. Finally, a comparison of 27 existing clustering algorithms is summarized and several design guidelines are drawn from the comparison in Section VII.

Section VIII has revealed the major differences between WSN systems and IoT systems. Section X discusses the possibility and challenges to apply clustering algorithms to IoT systems and the potential challenges, specially when deployed in 5G enabled environment. The conclusions of this paper is finally drawn in Section X.

II. Related Work

We have reviewed most of the existing survey papers for clustering in WSNs for the last decade, as shown in Table I. There are many overlapping studies and investigations. Many of them lack of deep analysis and comprehensive introduction. These four selected survey papers shown in Table II, [10], [20], [29], and [30], are good to start with when researchers are about to explore in this area. For the selected survey papers in Table II, as you can see, they include a large number of clustering algorithms. There are also some new studies in recently two years (2015 and 2016) on clustering that are not covered by the existing reviewing papers, such as RINtraR [36], SenCar [37], FL-low-energy adaptive clustering hierarchy (LEACH) [38], BEEM [39], and PathQuality [40]. Survey paper [35] has stated that the existing heuristics in the state of the art are diverse. However, most of them are following the same idea and many clustering algorithms have similar heuristic. The essential parts for the most well accepted heuristics are control information exchanging and CH detection and selection in one-hop range. Information exchanging allows decentralized clustering, which is the most adopted scheme in existing work. In order to exchange information between nodes, different systems may have their own requirements on the devices, such as synchronization (strong, loose, or none), location or energy level awareness.

LEACH and similar algorithms are sharing the same heuristic, which requires time synchronization and homogeneous development. The heuristics are mainly determined by the main objectives of the clustering algorithms, such as load balancing, energy efficiency, or mobility awareness. The common analysis topics in existing survey papers include: convergence time, node mobility, cluster overlapping, location awareness, energy efficient, failure recovery, balanced cluster/cluster size, cluster stability, cluster count, load balancing, deliver delay, intra/inter routing schemes, objectives and complexity, etc. However, besides energy efficiency, other discussed metrics are mostly about topology.

In this paper, we are mainly focusing on QoS oriented metrics to analysis clustering algorithm, including energy efficiency, transmission reliability and network latency. Besides, we indicate that QoE awareness also requires more studies when undergoing the transformation from WSN to IoT. We aim to provide guidelines for researchers who are new in this area to make their design decisions.

III. Clustering Techniques

Clustering algorithms normally structure networks into Voronoi diagrams. However, some algorithms apply non-Voronoi structures, like chain or spectrum structures. In this section, clustering techniques are classified into two categories based on the post network structure: 1) Voronoi-based approaches and 2) non-Voronoi-based approaches. The detailed implementations of several highly referenced clustering algorithms are also presented.

A. Beyond the Structure

Before we discuss clustering algorithms based on the finalized network structure, a brief introduction to clustering heuristics is presented here. A heuristic is composed with the information and rules used to form clusters. Those information can be from the sensors themselves or data collected from the network. Many clustering algorithms are sharing similar heuristic with certain variations on how the information exchange, or the metrics applied. The most popular heuristic is single hop intercluster and intracluster communication, like LEACH. The heuristic of LEACH as one of the very first energy efficiency studies in clustering, has motivated many variations and extensions. Other heuristics may improve the

TABLE II
SUMMARIES FOR SELECTED EXISTING SURVEY PAPERS ON CLUSTERING FOR WSNs

Survey Paper	Major contributions	Covered Papers
Ameer [10] 2007	1. Earlier clustering work before LEACH mainly addressed node failure problem. Clustering algorithms after LEACH start to focus more on energy efficiency. 2. Multihop intra-cluster topology is rare. Multihop inter-cluster communication is well applied.	1997: Adaptive clustering [41] 1998: CLUBS [42] 2000: LEACH [43] 2001: HCC [44], TEEN [45], MOBIC [46] 2002: APTEEN[47], Deterministic [48], GS^3[49], LEACH-C [50], PEGASIS [51] 2003: EEHC [52] 2004: ACE [53],FLOC [54], HEED [55], SEP [56] 2005: BCDCP [57], CAWT, DWEHC [58], EDACH [59], EECS [60], EEUC [61], Ex-HEED [62], Gupta [63] TL-LEACH [64], TASC, TTDD [65], UCS [66] 2006: MOCA [67] 2007: CCS [68], CMEER [69], EACLE [70] EDC [71], PEACH [72], TCCA 2008: EEMC [73], EEDUC [74], MRPUC [75] PRODUCE [76], S-WEB [77] 2009: EEDCF [78], PEBECS [79] 2010: CCN [80], EAUCF [81], PANEL [82], Unequal-LEACH [83] 2011: AWARE [84], EADUC [85], EC [86], EECABN [87] LUCA [88], MBC [89], Spatial-clustering [90] 2012: EBCAG [91], EEBCDA [92], LEACH-DT [93] TCAC [94], UHEED [95] 2013: ACDA [96], DSBCA [97], LCM [98] 2014: EEDC [99]
Afsar [29] 2014	1. Energy efficiency/maxing lifetime is the the dominate objective for clustering. 2. No reviewed clustering algorithm can operate in a heterogeneous network. 3. Multihop intra-cluster communication is rarely supported. While multihop inter-cluster communication is well applied. 4. The CH is only in charge of data aggregation and transmission. 5. Device mobility is hardly concerned in solutions. 6. Distributed implementation is the mainstream in clustering.	
Liu [20] [30] 2012 2015	1. Algorithms have small delivery delays tend to have low energy efficiency. 2. Distributed implementation is the mainstream in clustering. 3. Tree and train based implementations have lower scalability than grid based ones.	

existing one from several perspectives, such as balance of cluster size, maximal hop numbers, or catering for specific scenarios. Regardless of the heuristic that an approach adopts, the network structure constructed at the end, can be classified into Voronoi-based ones or non-Voronoi-based ones.

B. Voronoi-Based Approaches

Twenty-seven clustering algorithms are reviewed in this paper and 88% of them are Voronoi-based. Chain-based and spectrum-based algorithms account for 8% and 4%, respectively, as show in Table III.

Voronoi diagrams (2-D or 3-D), as a very important data structure in computational geometry, are mainly used for solving clustering and scheduling problems in computer science. The formal definition for 3-D Voronoi diagrams amended from [101] and the definition for 2-D is as follows.

Let S denote a set of n points (called sites) in the Euclidean space R^3. For two distinct sites $p, q \in S$, the dominance of p over q is defined as the subset of the space being at least as close to p as to q. Formally

$$\mathrm{dom}(p, q) = \left(x \in R^3 \mid \delta(x, p) \leq \delta(x, q)\right)$$

for δ denoting the Euclidean distance function. The region of a site $p \in S$ is the portion of the space lying in all the dominances of p over the remaining sites in S. Formally

$$\mathrm{reg}(p) = \bigcap_{q \in S - \{p\}} \mathrm{dom}(p, q).$$

From the definition, for any point x that $x \in R^3 \cap x \notin S$ one can have $\delta(x, p) \leq \delta(x, q)$ where $p \in S \cap q \in S - \{p\}$ if $x \in \mathrm{reg}(p)$. Here S is called the seed-point set, which is the CH set in a WSN.

TABLE III
STRUCTURE REVIEW ON EXISTING CLUSTERING ALGORITHMS

Algorithm	Distributed	Structure	Intra cluster	Inter cluster
LEACH [43]	Yes	Voronoi	1-hop	1-hop
LEACH-C [50]	No	Voronoi	1-hop	1-hop
LEACH-TL [64]	Yes	Voronoi	1-hop	2-hop
HEED [55]	Yes	Voronoi	1-hop	multihop
EECS [60]	Yes	Voronoi	1-hop	1-hop
EEHC [52]	Yes	Voronoi	multihop	1-hop
DWEHC [58]	Yes	Voronoi	1-hop	multihop
PANEL [82]	Yes	Voronoi	multihop	1-hop
UCS [66]	Yes	Voronoi	1-hop	multihop
EEUC [61]	Yes	Voronoi	1-hop	multihop
ACE [53]	Yes	Voronoi	1-hop	1-hop
BCDCP [57]	No	Voronoi	1-hop	multihop
PEGASIS [51]	Yes	Chain one layer	multihop	1-hop
TEEN [45]	Yes	Voronoi	1-hop	multihop
APTEEN [47]	Yes	Voronoi	1-hop	multihop
TTDD [65]	Yes	Voronoi	1-hop	multihop
CCS [68]	Yes	Chain multi-layers	multihop	multihop
HGMR [100]	Yes	Voronoi	1-hop	multihop
S-WEB [77]	Hybrid	Spectrum	1-hop	multihop
PEACH [72]	Yes	Voronoi	1-hop	multihop
MRPUC [75]	Yes	Voronoi	1-hop	multihop
MOCA [67]	Yes	Voronoi	multihop	1-hop
FLOC [54]	Yes	Voronoi	1-hop	1-hop
Adaptive [41]	Yes	Voronoi	1-hop	multihop
AWARE [84]	Yes	Voronoi	multihop	1-hop
BEEM [39]	Yes	Voronoi	1-hop	multihop
RINtraR [36]	Yes	Voronoi	multihop	1-hop

Fig. 2 shows an example of a Voronoi diagram. The points on the surface will join the closest point from set S (the black dots). For example, M is the closest black dot for all the red dots in the cluster of M, compared to other black dots. Clustering algorithms can totally adopt this

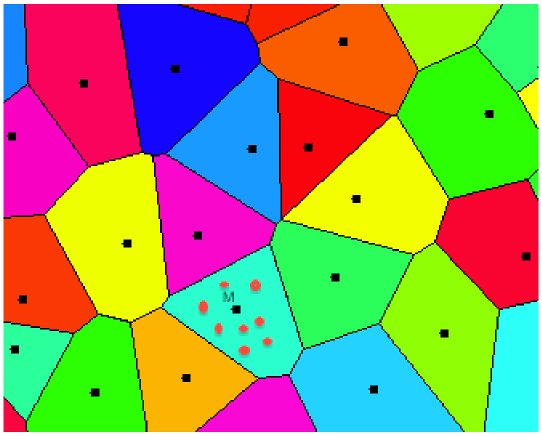

Fig. 2. Voronoi diagrams.

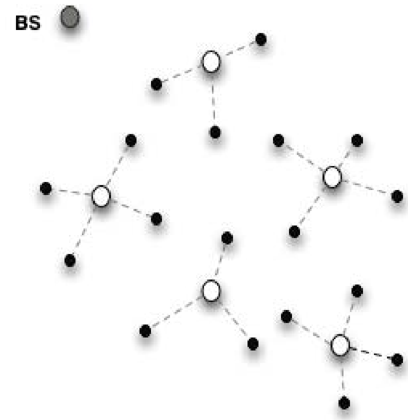

Fig. 3. Voronoi-based cluster network.

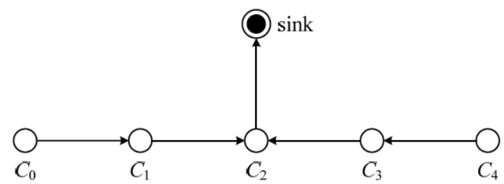

Fig. 4. PEGASIS single layer chain structure from [20].

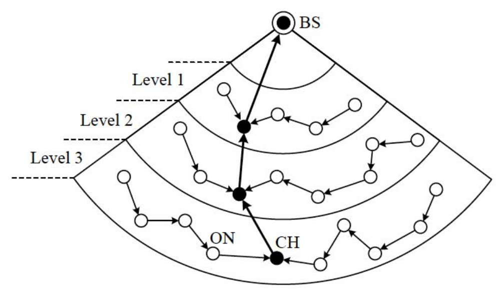

Fig. 5. CCS multilayer chain structure from [20].

idea, using Euclidean distance as a partition metric. However, other metrics, such as energy consumption for communication and hop count, can also be applied to replace the use of Euclidean distance. This type of clustering algorithm is defined as Voronoi structure-based approaches. This structure is convenient for relevant data fusion and energy saving for two reasons.

1) The sensors close to each other have a better chance of having relevant data.
2) Generally the power consumption on communication are low and the link quality is high.

Therefore, Voronoi diagrams are the dominating structure adopted in clustering algorithms.

LEACH [43] and hybrid energy-efficient distributed (HEED) [55] are two classic Voronoi structure-based clustering algorithms. Several common phases are included in Voronoi-based clustering algorithms.

1) The first step is to construct the CH set. Then the CHs will broadcast their information. After that, each sensor in the network will join one of the CHs based on the received information.
2) The sensors and their CH in the same cluster form a star shape network as shown in Fig. 3. In some scenarios, the sensors can commutation with the CHs through multihop connections.
3) A sensor normally can only belong to one cluster at one time. There is no overlapping between different clusters.

C. Non-Voronoi-Based Approaches

Chain and spectrum are two non-Voronoi structures that are used in clustering algorithms. In a chain structure-based network, no CHs are elected before the clusters are constructed. All the sensors are organized into chains as shown in Fig. 4 from PEGASIS and Fig. 5 from CCS. A sensor only needs to communicate with its left and right neighbors. The data flows in one direction. If a sensor receives data from its two adjacent neighbors on the chain, it will transmit the data to the BS, performing like a CH.

In PEGASIS [51], every sensor in the network transmits its data to one of its neighbors. Through this way, the gathered data is transferred from one node to another through a chain. A designated node (a node that receives data from both neighbors) will send the assembled data back to the BS. PEGASIS is claimed as a distributed algorithm. However, it is assumed that every single node has the global information of the network.

CCS [68] is a centralized clustering algorithm based on PEGASIS. Instead of using a single chain structure, CCS utilizes a multihop chain structure, as shown in Fig. 5. Regarding the BS as the center of the network, each sensor assigns itself a level number according to the signal strength received from the BS. Through this way, the sensors in the network are organized into a hierarchy structure. For each level, the sensors perform transmission and fusion in the same way as that in PAGASIS. The sensor that is elected as the CH will gather data from all the other sensors on the same level and then transmit the assembled data to the CH in the one-lower level. Once being assigned a level, a sensor will not change its level unless the location of the BS changes. This structure suffers

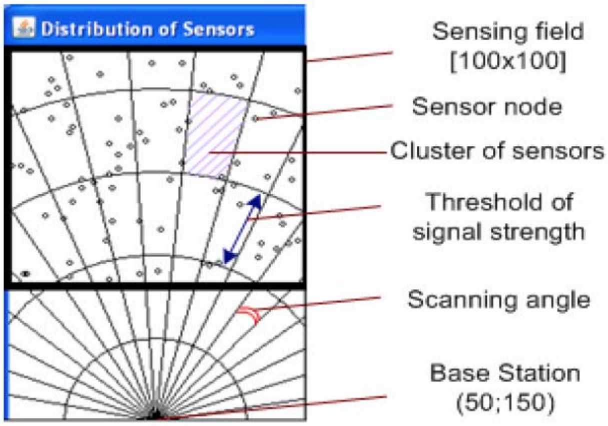

Fig. 6. S-WEB spectrum cluster structure [77].

from a problem that the sensors near the BS can die soon from forwarding packets for the sensors in the higher levels. Only total power consumption of the network is measured to evaluate this algorithm. There is no evidence showing a balanced power usage throughout the network.

In a spectrum structure-based network, the sensors are partitioned based on both distance and angle to the BS. The angle is captured from a scanning sweep from the BS at a specific time. Fig. 6 shows the method used in the S-WEB clustering algorithm. S-WEB [77] is a spectrum structure-based clustering algorithm. The first step in S-WEB is similar to that in CCS. All sensors are organized into layers based on their distance (measured by signal strength) to the BS. Then the BS does a 360° scanning sweep by sending out a signal at one angel at a specific time. The sensors are clustered into cells based on their layer and the scanning angle. In each cluster cell, the sensor with the highest residual energy will be elected as the CH. All the sensors have the responsibility of forwarding packets. The CHs may rotated between different sensors. However, the clusters are fixed during the entire operation of the network. The structure of the network is fixed after performing S-WEB, which cannot adapt to the dynamic changes in the network, like node failure or node death. The evaluation of S-WEB is only based on the comparison with a noncluster-based routing—direct routing.

D. *Influential Clustering Algorithms*

A wide body of research on clustering algorithms for WSNs is available; in particular, four noticeable survey papers have been published, shown in Table II. In Table IV, a list of existing clustering algorithms is analyzed from the perspectives that are not covered by the existing survey papers, including simulation environment, basic structure and QoS objectives (energy efficiency, transmission reliability, and network latency). For each algorithm, in order to evaluate their new approach, one or more existing algorithms are normally used as a benchmark to compare with. Table III shows the control manner (distributed, centralized, or hybrid), structure, and intracluster and intercluster routing scheme of each algorithm. The algorithms that are used as benchmarks are elaborated in the Table IV. As it can be seen from Table IV, both TTDD and HGMR support energy efficient, reliable and latency aware communication. However, TTDD is a solution based on a mobile BS that has a specific defined travel route. In such a case, the location of the sensors, the user scenario and the network condition have to be constant in order for TTDD to operate correctly. HGMR can only cluster the sensors that are involved in the current transmission. Furthermore, it uses multicast routing, which could significantly increase the traffic load in large networks. Therefore, LEACH and HEED are often selected as two benchmarks in existing studies for several reasons.

1) LEACH and HEED are referred to as benchmarks in 44% and 20% of the 27 algorithms reviewed in this paper, respectively.
2) They are two classic clustering algorithms that have inspired many other algorithms.
3) They are distributed algorithms and they have low requirements on the applied WSNs.
4) They are convenient to deploy in real networks and easy to extend in order to cater to different scenarios.
5) They are Voronoi structure-based, in which TPC can have the best performance [102].

These two algorithms will be discussed in detail. A detailed comparison between different algorithms will be drawn in Section VII.

1) LEACH: LEACH provides an elegant approach to clustering routing that has inspired many adaptations, like LEACH-C, LEACH-TL, and HEED. It is a decentralized Voronoi structure-based clustering algorithm. Instead using Euclidean distance as a metric, LEACH uses received signal strength. Most current radio chips provide a specific register to store the received signal strength value, called received signal strength indicator (RSSI). This value will be updated once a new packet is received. LEACH necessitates five steps to construct the cluster structure.

1) Each sensor elects itself to be a CH with a specific probability, which is set to be 5% in the experiments.
2) After the CHs have been elected, they will broadcast their information to the reset of the sensors in the network.
3) Based on the information received from all the CHs in its communication range, a sensor will decide which CH to join.
4) The CHs will create a transmission schedule for the sensors in their respective clusters. All the sensors in the same cluster communicate with their CH in a single-hop TDMA manner.
5) The CHs collect and fuse the data from the sensors and then send the assembled data to the BS using long-distance CDMA communication.

The first four steps together are called the set-up phase and the last one is called the steady-state phase. Since there is no sensing data transmission in the set-up phase, the related energy and time cost is considered as the overhead of LEACH. To minimize the impact of the overhead from CH election and cluster forming, the steady-state phase is set long enough compared to the set-up phase.

TABLE IV
QoS REVIEW ON EXISTING CLUSTERING ALGORITHMS

Algorithm	Simulator	Energy	Reliability	Latency	QoE Awareness	Benchmarks
LEACH	MATLAB	Yes	No	No	No	Direct, MTE, Static
LEACH-C	NS-2	Yes	No	Yes	No	LEACH, Static
LEACH-TL	NS-2	Yes	No	No	No	LEACH
HEED	MATLAB*	Yes	No	No	No	LEACH
EECS	MATLAB	Yes	No	No	No	LEACH
EEHC	MATLAB*	Yes	No	No	No	MAX-MIN-D
DWEHC	NS-2	Yes	No	No	No	HEED
PANEL	TOSSIM	Yes	No	No	No	HEED
UCS	MATLAB	Yes	No	No	No	Equal cluster size
EEUC	MATLAB*	Yes	No	No	No	LEACH, HEED
ACE	MATLAB*	Yes	No	No	No	HCP
BCDCP	MATLAB	Yes	No	No	No	LEACH, LEACH-C, PEGASIS
PEGASIS	MATLAB*	Yes	No	Yes	No	Direct, LEACH
TEEN	NS-2	Yes	No	No	No	LEACH, LEACH-C
APTEEN	NS-2	Yes	No	No	No	LEACH, LEACH-C, TEEN
TTDD	NS-2	Yes	Yes	Yes	No	no comparison
CCS	MATLAB*	Yes	No	No	No	PEGASIS
HGMR	NS-2*	Yes	Yes	Yes	No	HRPM, GMR
S-WEB	MATLAB	Yes	No	No	No	Short, Direct
PEACH	MATLAB*	Yes	No	No	No	EEUC, HEED, LEACH, PEGASIS
MRPUC	MATLAB	Yes	No	No	No	HEED
MOCA	MATLAB	Yes	No	No	No	No comparison
FLOC	MATLAB	No	No	No	No	No comparison
Adaptive	MATLAB	Yes	Yes	No	No	PRNET, Cluster TDMA
AWARE	TOSSIM	Yes	Yes	No	No	Unaware LEACH
RINtraR	MATLAB	Yes	Yes	No	NO	LEACH
BEEM	MATLAB	Yes	No	Yes	NO	LEACH, HEED

* indicates that the paper did not specify a simulator, but the simulation can be performed in the given one

In an ideal environment without any obstacles, the value of RSSI between two sensors is determined by the distance and the transmission power of the transmitter. The relation between RSSI and distance can be represented by the log distance path loss model [103]

$$P_r(d)[\text{dBm}] = P_r(d_0)[\text{dBm}] - 10\eta \log\left(\frac{d}{d_0}\right) - X_\sigma \tag{1}$$

where $P_r(d)$ is the received transmission power measured in dBm at distance d. It is also the value for RSSI. $P_r(d_0)$ is RSSI at the reference distance d_0. η, also referred to as path loss exponent, is a constant parameter value determined by the environment. For example, in a free space its value is 2; while in a well constructed building, its value can be from 4 to 6. X_σ is a zero-mean Gaussian distributed random variable (in dBm) with a standard deviation value of σ. This variable is used only when there is a shadowing effect. Otherwise, this variable will be zero.

Fig. 7 depicts one such example whereby five CHs are elected in the network. The sensors are clustered into different groups based on their RSSI values to all the five CHs. As mentioned before, RSSI is determined by the distance in an ideal environment without any obstacles or noise, if transmitting data at a fixed transmission power. Therefore, in an ideal environment, the network structure after performing LEACH is a regular Voronoi diagram. In the real world, since RSSI is not totally determined by the distance due to the influence of the environment, the network structure will be a analogous Voronoi diagram instead of a strict one.

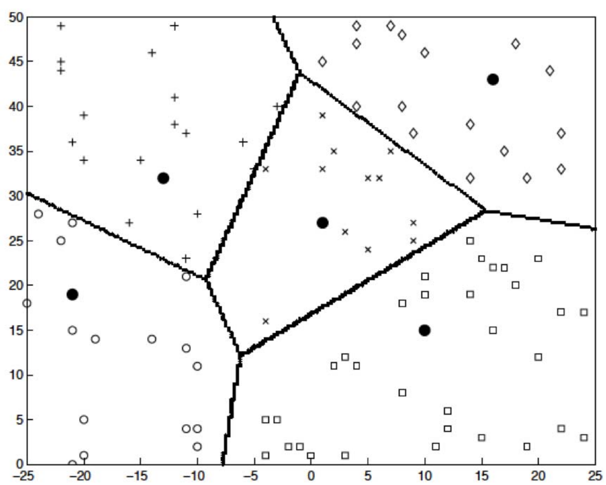

Fig. 7. LEACH clustering.

A node becomes a CH for the current round if the randomly generated number is less than the following threshold:

$$T(n) = \begin{cases} \frac{P}{1-P\left(r \bmod \frac{1}{P}\right)}, & \text{if } n \in G \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where G indicates the set of the nodes that have not been elected as CHs in the last $1/P$ rounds. P is the desired percentage of the CHs. Through experimental analysis, LEACH has

```
I. Initialize
S_nbr ← {v: v lies within my cluster range}
Compute and broadcast cost to ∈ S_nbr
CH_prob ← max(C_prob × E_residual/E_max, p_min)
is_final_CH ← FALSE

III. Finalize
If (is_final_CH = FALSE)
    If ((S_CH ← {v: v is a final cluster head}) ≠ φ)
        my_cluster_head ← least_cost(S_CH)
        join_cluster(cluster_head_ID, NodeID)
    Else Cluster_head_msg(NodeID, final_CH, cost)
Else Cluster_head_msg(NodeID, final_CH, cost)

II. Main Processing
Repeat
If ((S_CH ← {v: v is a cluster head}) ≠ φ)
    my_cluster_head ← least_cost(S_CH)
    If (my_cluster_head = NodeID)
        If (CH_prob = 1)
            Cluster_head_msg(NodeID,final_CH,cost)
            is_final_CH ← TRUE
        Else
            Cluster_head_msg(NodeID, tentative_CH,cost)
ElseIf (CH_prob = 1)
    Cluster_head_msg(NodeID,final_CH,cost)
    is_final_CH ← TRUE
ElseIf Random(0,1) ≤ CH_prob
    Cluster_head_msg(NodeID,tentative_CH,cost)
CH_previous ← CH_prob
CH_prob ← min(CH_prob × 2, 1)
Until CH_previous = 1
```

Fig. 8. HEED pseudocode [55].

proved that when $P = 5\%$, the system has the best normalized energy dissipation.

Compared to: 1) direct transmission to the BS with full transmission power; 2) minimum transmission energy (MTE) that uses minimum transmission power to complete the data transmission; and 3) static clustering with fixed CHs and network structures, LEACH can provide a better load balancing throughout the network and an improved networks longevity.

LEACH assumes that every sensor has two communication modes: 1) short distance and 2) long distance. If the RSSI between two sensors is higher than a threshold, the sensors can switch to short distance communication mode by lowering the transmission power level. Otherwise long-distance CDMA communication mode is required in order to guarantee the network connectivity. By long-distance communication, the CHs can connect to the BS with direct single hop. In a cluster, each sensor sends data to its CH in a TDMA manner to avoid the collisions with others. Outside the clusters, the CHs communicate with the BS by CDMA MAC protocol.

The basic idea of LEACH has inspired many researchers to propose their own solutions for clustering. It has become the most popular clustering algorithm because of its extensibility. However, LEACH, is not an environment adaptive or user-friendly protocol. In order to achieve specific expected performance, more work needs to be done to enhance the algorithm in order to be implemented in real WSNs.

2) HEED: HEED is a clustering algorithm based on LEACH that can support multihop intercluster communication. Instead of using random CH election scheme, HEED can select the sensors with high residual energy to be the CHs through an iteration CH election scheme. In this iteration CH election scheme, CHs are classified into two types: *tentative* or *final* CHs. Some initial CHs are selected randomly as *tentative* CHs. The probability to become a CH is

$$\mathrm{CH}_{\mathrm{prob}} = C_{\mathrm{prob}} \frac{E_{\mathrm{residual}}}{E_{\max}} \tag{3}$$

where E_{residual} is the current battery level and $E_{\max}$ is the initial battery level. C_{prob} is the optimal probability that a sensor will elect itself as a CH. However, C_{prob} is only used to limit the number of the initial CHs and it has no direct impact on the number of the *final* CHs. C_{prob} is set to be 5%, which is the same as that in LEACH.

The pseudocode of the CH iteration election process is shown in Fig. 8. After each iteration, every sensor doubles their probability of becoming a CH, which is $\mathrm{CH}_{\mathrm{prob}}$. The minimal value of $\mathrm{CH}_{\mathrm{prob}}$ is set to a specific threshold ($P_{\min} = 10^{-4}$) to bound the number of iterations. If a sensor is not covered by any CHs and its $\mathrm{CH}_{\mathrm{prob}}$ is higher than a randomly generated number, it will elect itself to be a *tentative* CH. If its C_{prob} has reached 1 and it is still not covered by any *tentative* or *final* CHs, it will claim itself as a *final* CH and then broadcast its information. A *tentative* CH becomes a *final* CH once its C_{prob} reaches 1. After its C_{prob} is increased to 1, a sensor will terminate the iteration process. A *tentative* CH will give up to be a CH if it discovers a *final* CH in its communication range. At the end, the sensors that are not covered by any *final* CHs will elect themselves as CHs.

The iterative CH election can elect those sensors with high residual energy to be CHs and therefore, balance the power usage throughout the network. To further reduce the transmission cost, a second parameter, that of, the energy consumption for intracluster communication, is considered when a sensor chooses its CH. This iteration CH election scheme guarantees that the probability of the phenomena that two sensors, within each other's communication range, both become CHs is rare.

Hence it can be deduced that the CHs are distributed evenly in the network. In that case, all the clusters will have similar size.

A single-hop TDMA intracluster communication scheme is applied, which is the same as that in LEACH. Different from LEACH, HEED adopts multihop intercluster communication scheme between the CHs and the BS. Significant overhead is involved in HEED due to the heavy broadcast in each iteration. Along with the energy decreasing while operating, the number of the iterations required in the initial phase consequently increases. Therefore, the clustering overhead and the network delay are also increased.

IV. QoS IN CLUSTERING

In general, over the top (OTT) applications have some specific QoS requirements for a given WSN. These requirements motivate the WSN algorithms and protocols to achieve particular QoS objectives. The common objectives in a system include network lifetime, transmission reliability, and network latency. Depending on the user requirements, a system may be designed to function in different ways. Nowadays, QoS awareness is essential in intelligent systems. Liu [20] emphasized that QoS supported clustering should be addressed in the future work. Table IV shows it is hard to consider all these three objectives and limited work has been done on that. The major reason for this is that clustering techniques are normally used to extend network lifetime. Other issues are often ignored. Transmission reliability and network latency problems have as yet failed to drawn sufficient research attention. More work need to be done to support QoS in clustering algorithms.

A. Network Lifetime

Clustering techniques are used as a means to prolong network lifetime. From when a system starts until the first/last sensor dies is normally referred to as the lifetime of a network. Network lifetime is crucial in some WSNs, such as wild life animal detection systems, as replacing sensors in these network is difficult and expensive. Such systems are required to operate without user maintenance for a long time. Therefore, longevity is always one of the most important considerations when designing related algorithms or protocols. How to extend network longevity is one of the most challenging topics in the WSN field. In general, it can be achieved from two perspectives.

1) *Power Usage Reduction:* If the power consumption on each individual sensor is reduced, consequently the lifetime of the entire network will be extended.
2) *Load Balancing:* The QoS of a system may be significantly undermined if some sensors that have vital responsibilities for the network connectivity or coverage die. Therefore, balancing the energy consumption on each sensor to balance the power usage throughout the network can maintain the QoS, which can improve longevity from another perspective.

As shown in Table IV, with the exception of FLOC, all other clustering algorithms have addressed the energy efficiency problem. FLOC primarily focuses on fast forming clusters locally instead of providing QoS clustering. Several techniques are normally utilized in existing clustering algorithms to prolong network lifetime.

1) *BS Location:* The BS is the final destination of all the data, so its location is critical to the overall performance. Considering the location of the BS can optimize the clustering structure and therefore, further reduce the power consumption from transmission.
2) *CH Rotation:* Since the CHs are in charge of data gathering, assembling, and transmission, they normally consume more energy than other normal sensors. To balance the energy consumption throughout the network, CH rotation between sensors is necessary.
3) *Energy Aware CH Election:* The sensors with high residual energy should be elected as CHs to further even the power usage.
4) *Cluster Size:* The CH of a big cluster consumes more energy than the one of a small cluster. To balance the cluster size, the CHs should be well distributed in the network. The chance that two sensors in each other's communication range both become CHs should be low.
5) *Data Fusion:* To reduce the energy consumption on transmission, the CHs should be able to fuse the data and send the assembled data to the BS.
6) *TPC:* Through TPC, a sensor can communicate with its CH with the minimal transmission power level while maintaining the transmission quality.
7) *Node Density:* Some work assumes that the sensors are deployed uniformly. The node density is a constant value in each single unit area. However, in most real-world WSN systems, the sensors are randomly distributed. Node density in some area may be higher than that in some other area. The sensors in a low density area will have a bigger impact on the network coverage than the ones in a high density area. Taking node density into account in clustering algorithms can balance the sensor distribution and maintain the network coverage [39].

It is difficult to cover all of the techniques in one single algorithm. Since the CHs normally consume more energy than the normal sensors, the CH election scheme has a big impact on the power consumption in a network. This explains why most existing work put a lot of effort into the CH election schemes. Several typical algorithms are summarized in Table V from the perspective of energy efficiency. LEACH and HEED are two classic Voronoi-based clustering approaches. DWEHC improved HEED by supporting multihop intercluster and intracluster communication. BEEM have proposed a new perspective of network longevity—coverage sensitive clustering. Regardless of the number of sensors, if losing network coverage, a system will lose part of the sensing data. That data may be critical in terms of the overall performance of the system. Therefore, coverage sensitive longevity is essential in WSN systems. In addition to energy efficiency, some systems also have high requirements on transmission reliability. For example, in a disaster monitoring system, once an event is detected, the information has to be transmitted successfully. Otherwise, serious damage may happen. In real-time systems, such as multimedia or road lighting, latency

TABLE V
ENERGY EFFICIENCY TECHNIQUES USED IN CLUSTERING ALGORITHMS

Algorithm	LEACH	HEED	DWEHC	BEEM	PEGASIS	CCS	S-WEB
Structure	Voronoi	Voronoi	Voronoi	Voronoi	Chain	Chain	Spectrum
Distributed	Yes	Yes	Yes	Yes	Yes	No	No
Node Density	No	No	No	Yes	No	No	No
CH Rotation	Yes	Yes	Yes	Yes	Yes	Yes	No
Data Fusion	Yes	Yes	Yes	Yes	Yes	Yes	No
TPC	Yes	Yes	Yes	Yes	No	No	No
Energy Aware	No	Yes	Yes	Yes	Yes	Yes	Yes
Cluster size	No	Yes	Yes	Yes	No	No	No
Sensing Abilities	None	Energy	Energy Location	Energy	Location	Engery Location	Energy Location
Simulation	MATLAB	MATLAB	NS-2	MATLAB	MATLAB	MATLAB	MATLAB
Benchmarks	Direct MTE, Static	LEACH	HEED	LEACH HEED	Direct LEACH	PEGASIS	Direct

awareness is essential for user experience. Coverage sensitive longevity, reliable transmission and latency awareness are important issues in terms of QoS. BEEM was introduced with an aim to increase WSN coverage sensitive longevity.

PEGASIS and CCS are chain-based approaches. S-WEB has a spectrum-based structure. PEGASIS has analyzed the impact of the BS's location on the performance of an algorithm, but it fails to account the location of the BS as a factor in clustering. Since CCS and S-WEB both have layer-based network structure, the location of the BS can affect the clustering results. HEED indicates that the sensors do not need any special component to measure the residual power. The battery level can be estimated from the energy consumption on sensing, processing and communication. However, this estimation, also called power management, is complicated to be implemented on real sensors. By using the location information, in DWEHC, the sensors are evenly distributed into the clusters. Therefore, the power consumption on the CHs is better balanced, compared to that in HEED.

From the above studies, it can be seen that sufficient work has been done to improve longevity from the traditional perspective of network lifetime. The coverage of the network is always neglected. BEEM has introduced a new concept, coverage sensitive longevity, which not only evaluate the network longevity using the number of alive sensors but also using the network coverage. We argue this concept should replace the transitional definition for network lifetime.

B. Network Transmission Reliability

In general, clustering algorithms are designed with specific objectives in mind, for example, lifetime extending, load balancing, or scalability increasing [10]; however, transmission reliability is often overlooked. Little existing work considers transmission reliability as a criterion when evaluating the performances of the approaches. For many WSN deployments, such as disaster or military scenarios, communication reliability is essential in terms of QoS. Besides, for some civil applications, like critical infrastructure monitoring, transmission reliability is a crucial metric. In such systems, transmission quality has a high priority from the user's perspective. However, enabling reliable transmission in cluster-based networks has received limited attention. Individual clustering algorithms use contrasting metrics to decompose the networks into interconnected clusters. Examples of such metrics include distance, hop count and cluster size [15]; however, link quality metrics are not considered.

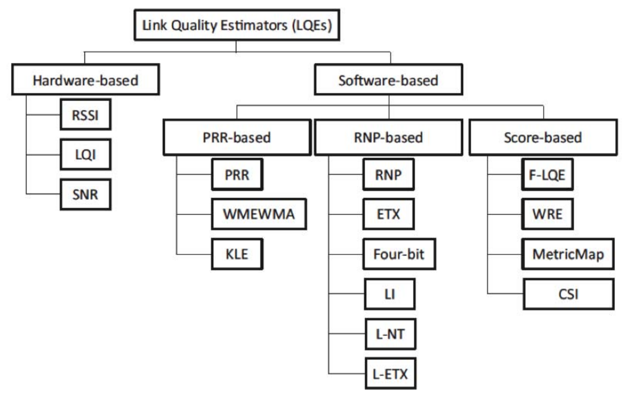

Fig. 9. Current available LQEs [104].

Baccour *et al.* [104] literally reviewed most of the available link quality estimators (LQEs) that can be used as link quality metrics. The current available LQEs are shown in Fig. 9. Packet receive rate (PRR), radio signal strength indicator (RSSI), and link quality indicator (LQI) are three common metrics used to estimate link quality. PRR is computed as the ratio of the number of successfully received packets to the number of transmitted packets. RSSI indicates the signal strength of the received packets. Its value is stored in the RSSI register, which is available in most popular radio chips. LQI meanwhile presents the correctness of the received packets. For the CC2420, LQI is measured based on the first eight symbols of a received packet. LQI is a good LQE when a large quality of data is available. In contrast, a good RSSI estimation can be obtained over a small number of measurements and can converge quicker than LQI [105]. Compared to RSSI and LQI, PRR has a higher correlation with transmission quality.

LEACH uses RSSI as a metric to cluster sensors. However, LEACH adopts TCP scheme, using RSSI to reduce energy consumption as an energy metric rather than an LQE metric. Moreover, LEACH assumes that all the sensors have long-distance single hop communication ability to avoid the complexity of multihop communication. TTDD presents an approach that can provide a reliable transmission with the help of the mobile BS. As mentioned before, HGMR is a routing protocol rather than a clustering algorithm. Even though HGMR claims that it can provide reliable transmission, it cannot be applied to the entire network. Although AWARE concludes that their solution can achieve a better PRR, it is not clear why this is the case. Additionally, the design of AWARE does not explicitly address the issue of reliable transmission. Adaptive clustering provides a code division access scheme for multimedia systems based on mobile radio networks, rather than based on WSNs. Using the standby routing, Adaptive clustering shows a low average packets loss rate but high end-to-end (E2E) network delay. RINtraR has proposed a solution for intraclustering routing to improve transmission quality by transmitting data through high quality multihop route rather than poor quality single hop route. Even though clustering techniques have been a hot research topic for more than 20 years, reliable transmission supported clustering has not drawn enough attention. Providing transmission reliability for intracluster or intercluster routing requires more work in the future.

C. Network Latency

Low network latency is an essential criterion in some WSN systems, such as road lighting systems. Once detecting a car on the road, the system needs to turn on the lights immediately. In such a latency sensitive system, maintaining network delay under an acceptable threshold is one of the major objectives. In a cluster-based network, the sensors need to transmit data to the CHs and then the CHs can transmit the data to the BS. Therefore, the network latency of the data transmission in a cluster-based network is longer than that in a direct transmission-based network. The cluster-based structure can extend network lifetime, but conversely increase network latency. For these reasons, for latency sensitive systems based on clustering structure, the objective normally is to support latency awareness, rather than necessarily to minimize it. QoE awareness in such a circumstance is to be able to customize the network latency to meet users' requirements.

Table IV shows that most existing work fails to consider latency when evaluating the performance of their approaches. Centralized clustering algorithm LEACH-C can balance the cluster size by distributing the CHs evenly over the network. From the experimental results of the throughput per unit energy and the throughput per unit time, it can be seen that LEACH-C is more energy and latency efficient, compared to LEACH. PEGASIS is a power efficient and delay sensitive algorithm. It states that minimizing energy or delay in isolation has drawbacks on the performance of the system. Therefore, PEGASIS uses energy×delay as a metric to select two out of its neighbors to form a chain structure. In the chain structure, each node can assemble the data that comes in from one neighbor on the chain with its own data and then transmit the assembled data to the other neighbor. PEGASIS indicates that the network delay for a packet is dominated by the number of transmission times (or called hop count) since: 1) there are no queuing delays and 2) the processing and propagation delays are negligible compared to the multihop transmission delay.

TTDD also provides for short network latency clustering by utilizing the mobility of the BS. It cannot be implemented in a network with a fixed BS. HGMR is a single-hop intercluster- and intracluster-based routing algorithm. By using multicast, it can support short latency transmission. HGMR is not specified as a clustering algorithm. With the exception of PEGASIS, no other algorithms account latency as a metric when performing clustering. Although the experimental performances may show good results on network latency, the reason and some further analysis are still missing.

Several approaches in clustering that are not covered by the existing surveys can also support short latency communication. Nikolidakis *et al.* [106] adopted the same idea as LEACH-C, through balancing cluster size to reduce network latency. Aioffi *et al.* [107] used several multiple mobile BSs to collect data from the network. By utilizing the border nodes, Tufail [108] proposed a solution that can provide alternative routes for some sensors in the network to reduce latency. As a result, the border nodes will consume more energy than other normal sensors. Reference [109] is focusing on fast cluster formation instead of fast data transmission. Reference [110] can provide low latency communication for the heterogeneous networks where supper nodes are deployed. The BS is assumed to locate in the center of the sensing field and the super nodes are deployed around the BS in a uniform manner. In a heterogeneous network with supper nodes deployed, no CH election is necessary since the supper nodes are the default CHs. Dousse *et al.* [111] and Li *et al.* [112] proposed solutions to provide short latency scheduling rather than short latency communication. Padmanabhan and Labeau [113] revealed that the geographical locations of the sensors and the node density both have an impact on the network latency in a cluster-based network.

V. Context Aware/User Configurable Features in Clustering

Besides the three QoS objectives discussed in Section IV, context awareness and user configurability are also important criteria from the user side. Those are often used to evaluate QoE. As indicated in Table IV, currently, none of the existing clustering algorithms can provide an interface for the users to interact with the systems. Xu *et al.* [114] presented a solution to achieve user configurability by integrating the improvements on network lifetime, transmission reliability and network latency. Since clustering is a major approach to energy efficiency in WSNs [115], most of the existing work has emphasis on energy saving, ignoring the diverse requirements from the users. The users normally expect that the WSN performs in a particular way, through which to benefit their applications furthest. For example, in a multimedia system, the users want immediate response in a cluster-based network rather than simply reducing power consumption. On the other hand, the users may highlight transmission reliability in a disaster detection system. Scenarios such as these involve a

TABLE VI
COMPARISON OF EXISTING SIMULATORS

Simulator	RSSI calculation	Specified for WSN	Power Model	Limitations
MATLAB	Log Distance Path Loss Model, calculated from distance	No	Numeric Model	No radio communication
NS-2	Friis equation, calculated from distance, antenna height	No	No	Slow for large network
TOSSIM	Not supported	Yes	No	Constrained to Berkeley nodes

tradeoff between competing demands. Therefore, power consumption is no longer the only consideration. An advanced WSN should be sufficiently intelligent to understand the users' preferences and adapt to these changes. Currently, there is no work showing any interest in adapting to user preference in clustering algorithms. User scenarios and system context are overlooked, without which QoE awareness is hard to be guaranteed.

VI. SIMULATION

Since clustering techniques are usually utilized in a large-scale network and it is not convenient to deploy a real one for testing, the evaluation of an algorithm is normally undertaken in a simulation. The performance of a proposed solution in the real network can be predicted from the simulation results. The selected simulator should not only be applicable for testing, but also convenient to use and modified.

A. Comparison of Simulators Utilized in Clustering

Within this paper, 27 different clustering algorithms are under discussed. All of the 27 clustering algorithms tested their approaches in simulations. 60% of the tests are based on MATLAB, followed by 28% based on NS-2. The remaining experiments were implemented in TOSSIM or other simulation tools constructed by the researchers themselves. MATLAB and NS-2 are commonly used simulation tools in clustering [116], [117]. Abuarqoub *et al.* [117] reviewed all the papers published in SENSORCOMM 2011 that were not limited in clustering area. It shows that MATLAB and NS-2 are still the most used simulators. Since MATLAB is convenient to use and easy to extend to add a power measurement component, the evaluations for most of the existing work were performed in MATLAB or can be accomplished in MATLAB (if not specified) when only energy consumption is being evaluated.

As a system grows, more QoS services will be required from the clustering algorithms. To evaluate the QoS performance of an algorithm, the applied simulator should either support the measurements for communication properties or can be extended to add the corresponding functionalities. Some metrics that are used in clustering algorithms can only be captured from the communication connection between sensors, such as RSSI. RSSI is often used as an LQE. Ideally, it is determined by transmission power, distance and other parameters, such as antenna height and environment impact factor. However, in the real world, it is highly influenced by the outside surroundings, like obstacles and noises.

The comparison of three popular simulators, which are MATLAB, NS-2, and TOSSIM, is presented in Table VI. Neither TOSSIM nor NS-2 have a power model that can simulate power consumption in WSNs. Since there is no radio communication is supported in MATLAB, the RSSI value between two sensors is computed from the distance by adopting log distance path loss model [103]. In this model, RSSI is mainly determined by the distance and transmission power. This model can only be applied to the networks that are deployed in an ideal environment, where no obstacles or noise exist. NS-2 is a general network simulator, which has a radio communication model. However, it is not specially designed for WSNs. The simulation processing becomes slow when the network scales over to 100 nodes [116]. In NS-2, the RSSI value is calculated from the Friis transmission equation, which is only accurate for long-distance communication scenarios. Now NS-3 is becoming more popular and it is should be used to replace NS-2. TOSSIM is specially designed to simulate WSNs that are composed with only TinyOS nodes. Currently, only sensors like Micaz, TelosB that are based on TinyOS. The simulator allows the users to test and verify the code that will run on real sensors. However, it is more focusing on examining the performance of an individual node rather than the whole network, and the sensor platform is constrained to a specific hardware type. TOSSIM is based on TinyOS, which has a deep learning curve. This is also the primary reason why it is not widely used. Besides, it can neither model RF cancellation nor support RSSI value generation.

As a result of the above analysis, MATLAB is the most popular one for simulating WSN clustering systems for the following reasons.

1) MATLAB can adopt log distance path loss model to calculate RSSI values in order to be used in algorithms.
2) It provides a numeric power consumption model, which can be easily modified to fit into different scenarios.
3) The processing time for large networks is quick.
4) It is not constrained to a specific type of sensors or networks. A more general evaluation can be performed.
5) MATLAB suits for agile programming. New models or components can be implemented and plugged in freely.

TABLE VII
ENERGY DISSIPATED IN MATLAB

Protocols	Transmission/Receiving (E_{elect})	Transmit Amplifier (ϵ_{amp})
LEACH	$50nJ/bit$	$100pJ/bit/m^2$
HEED	$50nJ/bit$	$10pJ/bit/m^2$ ($d < d_0$) $0.0013pJ/bit/m^4$ ($d > d_0$)

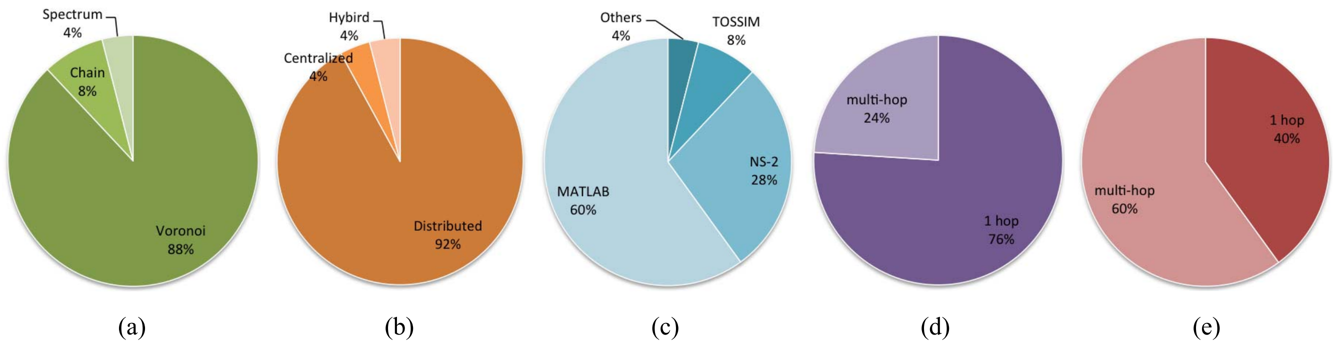

Fig. 10. Statistics analysis of 27 studied clustering algorithms. (a) Structure. (b) Control manner. (c) Simulator. (d) Intracluster routing. (e) Intercluster routing.

B. Power Consumption Model in MATLAB

LEACH indicates that different assumptions about the radio characteristics will affect the performance of an algorithm. Two most referenced clustering algorithms, LEACH and HEED adopted different models to compute energy consumption for long-distance communication. A shown in Table VII, in LEACH, the calculation of transmission power consumption is simplified as

$$E_{Tx} = E_{\text{elec}} * k + \epsilon_{\text{amp}} * k * d^2 \quad (4)$$

where k is the message length measured in bits and d is the distance from the transmitter to the receiver. In the future work, LEACH research team provided an advanced study on radio power consumption [50]. The calculation is specified as

$$E_{Tx} = E_{\text{elec}} * k + \epsilon_{\text{amp}} * k * d^n. \quad (5)$$

When $d < d_0$, $\epsilon_{\text{amp}} = 10$ pJ/bit/m^2, and $n = 2$. When $d > d_0$, $\epsilon_{\text{amp}} = 0.0013$ pJ/bit/m^2, and $n = 4$. The value of d_0 is a constant distance, which is determined by the surrounding environment. HEED has also adopted this power consumption model.

VII. Comparison of Existing Clustering Algorithms

Many survey papers have been published during the last decade. The comparison between different algorithms is generally focusing on clustering objectives, CH election, and intracluster and intercluster routing. Table IV shows the comparison of cluster structure, QoS supported situation, simulation environment, and benchmark algorithms in parallel experiments for the 27 clustering algorithms reviewed in this paper. These criteria are not included in the existing survey papers. The statistics analysis of the 27 clustering algorithms for network structure, control manner, simulation environment, and intracluster and intercluster routing is shown in Fig. 10.

From the comparison of existing work, several conclusions, which will contribute to construct the solutions in later studies, are drawn.

1) Voronoi diagrams are easy to implement and convenient to maintain as a structure for clustering, comparing with other structures. In a Voronoi structure, the sensors in the same cluster are normally physically close to each other. Those sensors will have a high chance to have relevant data. The communication cost between close sensors can be low.
2) Distributed clustering manner is more tolerant and robust.
3) The most often used simulator is MATLAB for its shallow learning curve and high extensibility.
4) Multihop intercluster communication is essential to guarantee the connectivity of a network without enabling long-distance communication ability on the sensors.
5) LEACH and HEED are two popular algorithms that are referred to as benchmarks in evaluations.

Besides, some problems existing in current work are also unveiled.

1) Multihop interclustering routing is implemented in many existing work. However, multihop intracluster communication is not well supported.
2) As is evident, in QoS supported clustering algorithms, energy efficiency is well accomplished, while transmission reliability and network delay are less studied.
3) None of the 27 algorithms can provide interface for the users to interact with the algorithms in order to customize the performance on energy efficiency, network delay, and transmission reliability. QoE awareness is missing.

Through the above analysis and comparison, it is convincing that a Voronoi-based distributed clustering algorithm, which can support QoS and QoE awareness, enabling multihop intracluster and intercluster routing, is demanded. Specifically, we conclude that user oriented/QoE aware design and QoS supported services should attract more attention in the future study of clustering for WSNs.

In the following sections, the audience will amazingly find that those problems in clustering techniques for WSNs are exactly aligned with what we need to address applying such techniques to IoT systems in 5G environment. As WSNs evolve into IoT and communication technologies march to 5G, clustering techniques also need to grow new charms in order to adapt to those changes. In the next sections, first the transformation between WSNs and IoT is presented. The possibility and advantages to migrate IoT systems to 5G platform is also discussed. We further present the challenges that researchers are facing if applying clustering techniques in IoT in 5G.

VIII. From WSNs to IoT

As it is already presented in Section I, a WSN is a network composed of autonomous wireless micro-devices, which can monitor the surrounding environment, record the data and transmit the information throughout the network to a central BS, which is also sometimes referred to as a sink node [1], [2]. WSNs were originally proposed for military surveillance purposes [118]. Due to its first success, this technology was then envisioned for other purposes, such as habitat monitoring [119], weather monitoring [120], agriculture monitoring [121], and wildlife monitoring [122]. WSN technologies were introduced more than 20 years ago and many projects have been proposed and undertaken that embrace this technology. However, more effort was still required before WSNs became a truly mainstream technology. WSN technologies envisaged as enabling another digital revolution in order to facilitate people's daily life. Only a small number of users were showing interest in WSN area. Because of the complexities associated with it, application developers were required to be domain experts to build quality services based on WSNs. For this reason, only limited usages and few applications were available for normal users. Hence, researchers were facing the dilemma that WSN technologies could benefit and advance society, but bringing them from the laboratory to the marketplace.

A similar and successful paradigm to compare against is the case of the Internet. The Internet was invented in the late 1960s. However, it did not become universally popular until 1995 when the Internet access was more free and convenient. The number of Internet users has increased impressively for the last decade owing to numerous OTT applications developed for it, such as World Wide Web, electronic mail and social networking.[1] The Internet has brought great convenience to society and its importance is self-evident.

The revolution that is necessary to ensure WSN technologies flourish in a similar manner requires more effort in simplifying design, implementation, deployment, and usability. The societal impact of WSNs is defined as the number of the users, which itself is determined by the quality and quantity of the available applications. Currently, it is extremely urgent to inspire developers to build more useful OTT applications to improve people's life quality or experience. Therefore, the concepts of IoT [123] and Internet of Everything [124] were proposed and now they are extremely popular. The instances include smart buildings/home, smart cities, smart roads and connected cars. WSNs were the former existence and the technology foundation for IoT. The fundamental difference between WSNs and IoT is that the dynamics and diversities in IoT are much higher than that in WSNs, which includes the following aspects.

1) *Applications:* Traditional WSN applications are mainly focusing on monitoring the environment and then collecting the data. The sensors are deployed and fixed in the field. The volume of the data transmitted in the network is small. However, IoT applications can range from smart kitchen to smart city monitoring. All the applications in these systems have specific requirements on the network performance. For instance, smart road lighting applications in smart road system require short network delays. Smart monitoring applications in smart city system require high network bandwidth.
2) *IoT Devices:* WSN systems normally have homogeneous sensors, in terms of sensing ability, processing capability and power supply. This feature has simplified the design of the relevant protocols, such as routing protocols. However, for IoT systems, this feature is no longer applied. For example, in a smart home system, the fridge, the vacuuming robot and the light sensing devices all have their own characteristics in this system. Furthermore, the devices in IoT systems are no longer limited to those basic sensors.
3) *Communication Ability:* Since the features of the connected devices in IoT systems vary greatly, their communication ability, including communication range, communication band, and communication power consumption can be different.
4) *Number of Connected Devices:* By the end of 2020, according to Cisco, there will be over 50 million connected devices. Managing such a large number of devices along with the generated data is challenging in IoT systems.
5) *Objectives:* As mentioned, WSN systems are mainly used for monitoring environment and collecting data. User profile and system context are rarely under discussion. One of the most important objectives for IoT systems is to improve people's life and their personal experience. User oriented and context aware design is demanded in IoT systems.

IX. Clustering for IoT Systems in 5G

Among the current clustering strategies for WSNs, many of them assume the sensors in the networks are homogeneous. If a heterogeneous network is considered, the degree of the diversity is still rather low (for example, some heterogeneous network simply assumes that part of the sensors have double power supplies than others). This assumption has become unrealistic recently, especially in IoT systems. The challenges when migrating IoT systems to 5G platform have

[1] [Online]. Available: http://www.internetworldstats.com/stats.htm

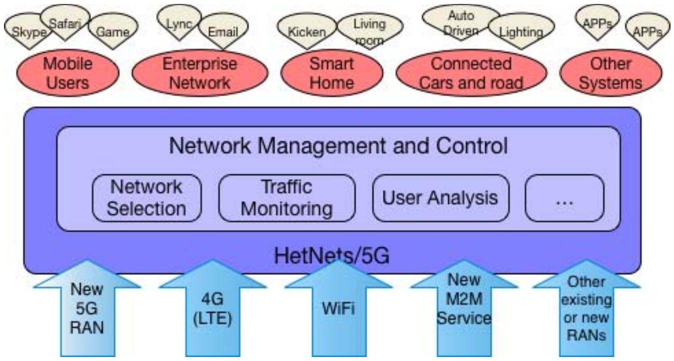

Fig. 11. Overview on 5G network.

been discussed a lot from a general perspective [125]. In this section, we will show the necessity of clustering techniques in IoT systems that are based on 5G networks.

A. *Overview on 5G*

For the first time in history LTE has brought the entire mobile industry to a single technology footprint resulting in unprecedented economies of scale. The converged footprint of LTE has made it an attractive technology baseline for several segments that had traditionally operated outside the commercial cellular domain. There is a growing demand for a more versatile machine-to-machine (M2M) platform. The challenge for industrial is the lack of convergence across the M2M architecture design that has not materialized yet. It is expected that LTE will remain as the baseline technology for wide area broadband coverage also in the 5G area.

Fig. 11 presents an overview of 5G network infrastructure. Mobile operators now aim to create a blend of pre-existing technologies covering 2G, 3G, 4G, WiFi, and others to allow higher coverage and availability, and higher network density in terms of cells and devices with the key differentiator being greater connectivity as an enabler for M2M services [126]. New machine type communication technologies are also invited, such as LTE-M and NB-IoT. An array of antennae supporting high-order multi-input, multioutput (MIMO) is installed in a device and multiple radio connections are established between the device and the cellular BS allowing paralleled data transmission. Meanwhile, operators, vendors and academia are combining efforts to explore technical solutions for 5G that could use frequencies above 6 GHz and reportedly as high as 300 GHz. This platform will need to provide a network management and control layer to coordinate the activities from the application layer and the services from underlying infrastructure. This layer should be implemented between network transport layer and application layer. It provides functionalities/components as such network selection, traffic monitoring, user analysis, etc. For example, smart network selection can be implement as the following way. The network selection component can match OTT usage to a suitable network interface based on the characteristics of the OTT application itself, the network conditions and the user profile. The traffic monitoring component can provide the network condition information. As you can see, those components will work together to achieve intelligent 5G.

On top of the 5G network, many systems can be supported, for example, smart homes and enterprise networks. In those systems, running multiple applications at the same time is also common. The network selection should know the applications and user scenarios in order to allocate suitable underlining network services for communication. This information can be accomplished by the user analysis component. Besides, the network context itself, such as congestion level, is also important when selecting the network interface. Therefore, traffic monitoring is necessary. 5G aims to enhance the degree of automatic adaptation and configuration. User oriented and QoE aware design is one of the main tasks. The management layer should decouple software functions from the hardware resource layer. Meanwhile, it also needs to provide network performance analysis and optimization. The 5G networks aim to advance IoT technologies further in the market [127].

3GPP standard/5G-based backhaul has become popular as a great solution for connectivity problem in IoT systems. Munoz *et al.* [128] indicated that the next generation of mobile networks (5G), will need not only to develop new radio interfaces or waveforms to cope with the expected traffic growth but also to integrate heterogeneous networks from E2E with distributed cloud resources to deliver E2E IoT and mobile services. Fantacci *et al.* [129] provided a backhaul solution through mobile network for smart building applications. The proposed network architecture will improve services for users and also will offer new opportunities for both service providers and network operators. Piri and Pinola [130] utilized LTE uplink resource as the communication backhaul for IoT systems. Hindia *et al.* [131] used LTE-Femtocell composed networks to collect data from health-caring IoT systems. They considered a heterogeneous network in which a macrocell tier is overlaid with a very dense tier of small cells. Kim [132] presented three scheduling algorithms for a two-tier HetNet to maximize network throughput of IoT devices. Jungnickel *et al.* [133] highlighted new applications for optical wireless communication as a mobile backhaul for WiFi, LTE, and 5G and as a new access technology in IoT systems where it enables secure and reliable communications at low latency.

The next generation networks/5G is definitely the ultimate means to connect everything together, specially for IoT systems [134], [135]. Small cells take a huge role in such infrastructure [136], [137]. Besides, software defined network and network function virtualization techniques are also widely used in such scenarios [138], [139] to improve the scalability. Smart backhaul solutions have been proposed to improve users' utility/QoE [140], [141]. Combining cloud technology with the backhaul is also becoming dominant [142]–[144]. Tehrani *et al.* [145] indicated that network selection/resource allocation in 5G network is a challenging research topic that needs to be addressed. In addition to the infrastructure, resource management in the backhaul is also critical [146]. References [147]–[149] proposed corresponding solutions for that. Ran *et al.* [150] aimed to balance the workload of different remote radio heads in the backhaul to alleviate the pressure

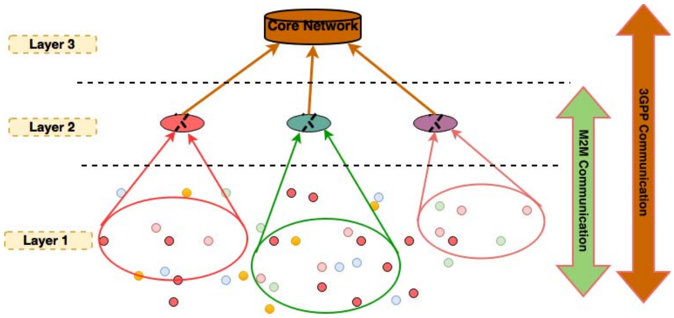

Fig. 12. 3GPP assisted IoT backhaul without clustering.

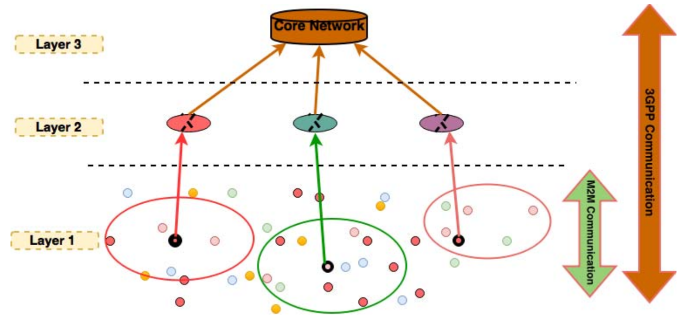

Fig. 13. 3GPP assisted IoT backhaul with clustering.

on the transmission links. Reliability can also be improved through cognitive radio [151].

As we can see from the above evidence, 5G can provide a feasible and reliable backhaul infrastructure for many IoT systems. It is sensible to design our future protocols for IoT systems based on that infrastructure, keeping QoS/QoE awareness as one design principle.

B. Need for Clustering for IoT in 5G

As cellular involved backhaul can provide full connectivity, it is becoming the main trend. As show in Fig. 12, there are normally three layers in the communication system. Layer 1 is composed with the sensors/devices in the field. Layer 2 is deployed by the mobile operators beside the field, normally in the form of small or micro cells, supporting 3GPP standard communication. Layer 3 is the evolved packet core network (core network), where all the information and data would be collected. The backhaul infrastructure is considered as the combination of layers 2 and 3. Layer 1 is referred to as the last hop in wireless communication. 3GPP communication can be applied to all these three layers. The M2M communication only exists in layers 1 and 2. Fig. 12 shows a standard infrastructure for 3GPP involved IoT backhaul. Nowadays, the 3GPP organization, mobile operators and academics are trying to realize 5G by the end of 2020. 5G with an enabler for M2M communication, has features such as 1–10-Gb/s speed, 1-ms latency, and 100% coverage and reliability, which aims to support and provide good QoE for a large range of applications and usages. Hence, this infrastructure shown in Fig. 12 has been well accepted and studied as mentioned in Section IX-A. Bassoy *et al.* [152] introduced a novel two-stage reclustering algorithm to reduce high load on cells in hotspot areas and improve user satisfaction. Smart coordinating multiple access points can improve the overall spectral efficiency.

In such a 3GPP involved IoT backhaul, we indicate that clustering techniques are still necessary and beneficial for the following reasons.

1) *Energy Efficiency:* In IoT systems, many sensors/devices are still deployed remotely, requiring lifetime for years. 5G aims to provide M2M communication, allowing devices to have up to ten years' battery life. Therefore, energy efficiency still is a challenging problem in IoT systems regarding of QoS. In addition, from green computing perspective, with 50 billion connected devices, if one device could reduce 1% energy consumption, it can save the world \$1 billion worth electricity. Xu *et al.* [153] introduced a clustering algorithm for IoT systems in MIMO scenarios to extend network longevity and maintain network coverage.
2) *Distributed Processing:* In recently ten years, big data has been a really hot topic and people generally expect treasures in the data. However, not all the data is useful. Treasure hunting in massive amounts of meaningless data can be costing. In addition, data transmission through the network and data maintaining in the server are also expensive. It is essential to filter out worthless or redundant data from the source, rather than transmitting it back to the core network.
3) *Management Hierarchy:* As we have discussed, the major difference between WSNs and IoT is the diversities from the devices themselves to the OTT applications. Besides, with such a large number of devices, an extensible and dynamic hierarchic structure can achieve effective and efficient management.

To relieve the problems caused by the above issues in IoT systems, clustering techniques can still be applied. As shown in Fig. 13, the three layers are still organized as before. However, at layer 1, the clusters are formed and one CH is elected for each cluster. The CH in each cluster will still be in charge of data gathering and confusion. With clustering in the field, first, the M2M communication is only happening in layer 1. The number of cross layer communication is largely reduced, which in turn saves a lot of energy on the devices. Second, clustering structure is beneficial for data gathering and data processing locally. The CH can dismiss redundant data and avoid overloading the 3GPP backhaul network. Third, the cluster structure in layer 1 can be further stratified, based on the size of the network, the device types, application types, communication abilities, etc. The hierarchy structure can be extended deeper depending on the requirements.

C. Challenges in Clustering Toward 5G

As we have discussed in the last section, clustering techniques are required and beneficial even for highly dynamic IoT systems. With 3GPP standard communication, the deployment becomes more flexible and the connectivity problem is solved straightway. When migrating IoT systems to 5G networks, if applied, several challenging problems in clustering are needed to be addressed.

The first challenge comes from the fundamental nature of IoT systems—the vastly diversity. The things in the field are highly heterogeneous. Some of the nodes can have low capabilities and some of them can be extremely advanced. In order to connect everything in smart cities, cheap and energy efficient transmit-only devices will be massively deployed [154] alone with other highly advanced sensors. Due to the hardware limitation, those devices cannot receive information from others, which makes decentralized clustering challenging. One opposite example to a transmit-only device is a super sensors. It is a more powerful node with the task of collecting sensing data and, in some cases, with the physical capability to act as a relay node [155]. Because of their high strength in memory and computing abilities, they normally are treated as CHs in networks. Therefore, comparing with traditional WSN, IoT systems are more complex and comprehensive. The OTT applications are no longer just about data collection and transmission. We need to consider more complicated scenarios and user cases. When clustering, from the perspective of reducing redundant data, sensors/devices with similar usage should be grouped together.

The second challenge is that cost for transmission. The energy cost is still a critical concern in IoT systems when deployed in 5G networks. Besides, since mobile network is also involved, the financial cost should be well controlled. For example, the use of LTE will be more expensive than that of WiFi. In practice, LTE can be used as the default communication means. If possible (the CH is in range and the user requirements can be satisfied), some of the devices can switch to Bluetooth or ZigBee, which are much power saving solutions. In such a condition, in order to utilize MIMO techniques, the devices in the same cluster should well distributed to use different network interfaces. This approach cannot only avoid interfering but also balance the load on those networks.

The third challenge is how to improve user utility. QoE has been emphasized greatly when formulating 5G related policies and regulations. Users are requiring specialized services according to their behaviors. In order to cater for characterized usage, user profile should be in consideration. The user profile can include information, such as user priority, user behaviors, user scenarios, etc. The first issue needs to be addressed is how to formularize user utility as a measurable metric. Then the clustering scheme should consider the user requirements for the networks. For example, if TDMA MAC protocol is applied in each cluster, it is not a good approach to group all the users that have high requirements on latency into the same cluster. Those users should be in a relevantly smaller cluster. In extreme case, they should be able to transmit with the layer 2 access points directly.

The fourth challenge is how to utilize the intelligent components in the core network. In order to make the 5G network smart and be aware of the context, extra components in the core network have been proposed to provide additional functionalities. The information from those components will be available for devices to query. For example, a network monitoring component can detect the congestion levels for all the network access interfaces. Be aware of this information should also help the designed clustering algorithms to be smart. Supposedly a device needs high network throughput for the application and uses LTE as the default communication interfaces. It should not switch to WiFi once it receives notifications from the core network indicating that currently WiFi network is congested [156].

The fifth challenge is how to manage and utilize mobility in the networks. Some work has been done specifically addressing mobility problem in WSNs [155], [157], [158]. With the development of IoT systems, supporting the connectivity of the fast moving things (such as vehicles in the connected cars system) and utilizing the mobility to improve the communication efficiency are challenging.

One of the general objectives for 5G is to improve QoS and QoE. Clustering techniques, in order to adapt to 5G and more complicated scenarios in IoT systems, the above challenges need to be addressed. The corresponding research in those directions should be further investigated. Cross layer design is also highly recommended to address multiple issues collaboratively. We sincerely indicate that more advanced studies should be undertaken in the current scenario with 5G rather than the traditional WSN usages.

X. Conclusion

In this paper, a general-to-specific review of clustering algorithms in WSNs was conducted. First, we presented a summarized survey for clustering techniques in WSNs. The services provided by existing algorithms were analyzed from four QoS angles: network lifetime, transmission reliability, network latency, and the QoE awareness perspective. Upon this analysis and the comparison, several findings have been revealed.

1) Limited work concerns network coverage when evaluating network lifetime.
2) Latency awareness and transmission reliability are not well supported in clustering.
3) User scenario/profile awareness has drawn little attention in current clustering research.
4) Limited work shows interest for clustering in heterogeneous networks with high degree of diversities.

These findings should motivate future work in clustering. As QoS has become a crucial evaluator in intelligent WSN systems. Unfortunately, existing work only proposes solutions to extend network lifetime from a narrow perspective. Moreover, limited work has been undertaken in clustering with a main objective to improve transmission reliability or reduce network latency. However, being able to optimize network performance from one single dimension among energy efficiency, transmission quality, and network latency is not sufficient to enable intelligence in clustering. It has been stated that an intelligent system should be able to understand the user's requirements and adapt to the changes within the system, especially when WSNs are evolving to IoT, with an aim to support a larger number of users and more complex user scenarios. It is extremely challenging if considering a higher degree of heterogeneous for the devices and the networks.

QoE supported features, such as user configurability, should be implemented in order to advance clustering techniques to again adaptivity to various conditions and user scenarios. Clustering algorithms now need to cater for multiple and differing scenarios with varying user preferences, which is essential when implementing a smart clustering strategy to understand the environment and the users. To the end, we discussed the need for clustering techniques in IoT systems and the challenges when migrating to 5G platforms. Future research directions are also outlined indicating where this technique should progress in order to profit IoT systems in the future network.

REFERENCES

[1] C.-Y. Chong and S. P. Kumar, "Sensor networks: Evolution, opportunities, and challenges," *Proc. IEEE*, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Netw.*, vol. 38, no. 4, pp. 393–422, 2002.

[3] S. Murugesan, "Harnessing green it: Principles and practices," *IT Prof.*, vol. 10, no. 1, pp. 24–33, Jan. 2008.

[4] L. Xu, D. T. Delaney, G. M. P. O'Hare, and R. Collier, "The impact of transmission power control in wireless sensor networks," in *Proc. 12th IEEE Int. Symp. Netw. Comput. Appl. (NCA)*, Cambridge, MA, USA, 2013, pp. 255–258.

[5] A. D. Amis, R. Prakash, T. H. P. Vuong, and D. T. Huynh, "Max-min D-cluster formation in wireless ad hoc networks," in *Proc. IEEE 19th Annu. Joint Conf. Comput. Commun. Soc. (INFOCOM)*, vol. 1. Tel Aviv, Israel, 2000, pp. 32–41.

[6] M. R. Brust, H. Frey, and S. Rothkugel, "Adaptive multi-hop clustering in mobile networks," in *Proc. 4th Int. Conf. Mobile Technol. Appl. Syst. 1st Int. Symp. Comput. Human Interact. Mobile Technol.*, Singapore, 2007, pp. 132–138.

[7] D. Baker and A. Ephremides, "The architectural organization of a mobile radio network via a distributed algorithm," *IEEE Trans. Commun.*, vol. COM-29, no. 11, pp. 1694–1701, Nov. 1981.

[8] L. M. C. Arboleda and N. Nasser, "Comparison of clustering algorithms and protocols for wireless sensor networks," in *Proc. Can. Conf. Elect. Comput. Eng.*, Ottawa, ON, Canada, May 2006, pp. 1787–1792.

[9] O. Younis, M. Krunz, and S. Ramasubramanian, "Node clustering in wireless sensor networks: Recent developments and deployment challenges," *IEEE Netw.*, vol. 20, no. 3, pp. 20–25, May/Jun. 2006.

[10] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Comput. Commun.*, vol. 30, nos. 14–15, pp. 2826–2841, 2007.

[11] P. Kumarawadu, D. J. Dechene, M. Luccini, and A. Sauer, "Algorithms for node clustering in wireless sensor networks: A survey," in *Proc. 4th Int. Conf. Inf. Autom. Sustainability*, Colombo, Sri Lanka, Dec. 2008, pp. 295–300.

[12] B. P. Deosarkar, N. S. Yadav, and R. P. Yadav, "Clusterhead selection in clustering algorithms for wireless sensor networks: A survey," in *Proc. Int. Conf. Comput. Commun. Netw.*, Dec. 2008, pp. 1–8.

[13] C. Jiang, D. Yuan, and Y. Zhao, "Towards clustering algorithms in wireless sensor networks-a survey," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Budapest, Hungary, Apr. 2009, pp. 1–6.

[14] J. J. Lotf, M. Hosseinzadeh, and R. M. Alguliev, "Hierarchical routing in wireless sensor networks: A survey," in *Proc. 2nd Int. Conf. Comput. Eng. Technol.*, vol. 3. Chengdu, China, Apr. 2010, pp. V3-650–V3-654.

[15] O. Boyinbode, H. Le, A. Mbogho, M. Takizawa, and R. Poliah, "A survey on clustering algorithms for wireless sensor networks," in *Proc. Netw. Based Inf. Syst. (NBiS)*, Takayama, Japan, 2010, pp. 358–364.

[16] S. Naeimi, H. Ghafghazi, C.-O. Chow, and H. Ishii, "A survey on the taxonomy of cluster-based routing protocols for homogeneous wireless sensor networks," *Sensors*, vol. 12, no. 6, pp. 7350–7409, 2012.

[17] M. Aslam *et al.*, "Survey of extended LEACH-based clustering routing protocols for wireless sensor networks," in *Proc. IEEE 14th Int. Conf. High Perform. Comput. Commun. 9th Int. Conf. Embedded Softw. Syst.*, Liverpool, U.K., Jun. 2012, pp. 1232–1238.

[18] M. R. Mundada, S. Kiran, S. Khobanna, R. N. Varsha, and S. A. George, "A study on energy efficient routing protocols in wireless sensor networks," *Int. J. Distrib. Parallel Syst.*, vol. 3, no. 3, pp. 311–330, 2012.

[19] K. Ramesh and D. K. Somasundaram, "A comparative study of clusterhead selection algorithms in wireless sensor networks," *CoRR*, vol. abs/1205.1673, 2012. [Online]. Available: http://arxiv.org/abs/1205.1673

[20] X. Liu, "A survey on clustering routing protocols in wireless sensor networks," *Sensors*, vol. 12, no. 8, pp. 11113–11153, 2012.

[21] S. Tyagi and N. Kumar, "A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 36, no. 2, pp. 623–645, 2013.

[22] M. A. Jan and M. Khan, "A survey of cluster-based hierarchical routing protocols," *Int. J. Comput. Netw. Wireless Commun.*, vol. 3, no. 2, pp. 138–143, 2013.

[23] S. K. Gupta, N. Jain, and P. Sinha, "Clustering protocols in wireless sensor networks: A survey," *Int. J. Appl. Inf. Syst.*, vol. 5, no. 2, pp. 41–50, 2013.

[24] P. Kumari, M. P. Singh, and P. Kumar, "Survey of clustering algorithms using fuzzy logic in wireless sensor network," in *Proc. Int. Conf. Energy Efficient Technol. Sustainability*, Apr. 2013, pp. 924–928.

[25] P. Jindal and V. Gupta, "Study of energy efficient routing protocols of wireless sensor networks and their further researches: A survey," *J. Comput. Sci. Commun. Eng.*, vol. 2. Nagercoil, India, 2013, pp. 924–928.

[26] C. P. Subha, S. Malarkan, and K. Vaithinathan, "A survey on energy efficient neural network based clustering models in wireless sensor networks," in *Proc. Int. Conf. Emerg. Trends VLSI Embedded Syst. Nano Electron. Telecommun. Syst. (ICEVENT)*, Tiruvannamalai, India, 2013, pp. 1–6.

[27] H. Dhawan and S. Waraich, "A comparative study on LEACH routing protocol and its variants in wireless sensor networks: A survey," *Int. J. Comput. Appl.*, vol. 95, no. 8, pp. 21–27, 2014.

[28] A. Nayyar and A. Gupta, "A comprehensive review of cluster-based energy efficient routing protocols in wireless sensor networks," *Int. J. Res. Comput. Commun. Technol.*, vol. 3, no. 1, pp. 104–110, 2014.

[29] M. M. Afsar and M.-H. Tayarani-N, "Clustering in sensor networks: A literature survey," *J. Netw. Comput. Appl.*, vol. 46, pp. 198–226, Nov. 2014.

[30] X. Liu, "Atypical hierarchical routing protocols for wireless sensor networks: A review," *IEEE Sensors J.*, vol. 15, no. 10, pp. 5372–5383, Oct. 2015.

[31] S. P. Singh and S. C. Sharma, "A survey on cluster based routing protocols in wireless sensor networks," in *Proc. Int. Conf. Adv. Comput. Technol. Appl. (ICACTA)*, vol. 45. 2015, pp. 687–695. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050915003695

[32] I. Ouafaa, L. Jalal, K. Salah-ddine, and E. H. Said, "The comparison study of hierarchical routing protocols for ad-hoc and wireless sensor networks: A literature survey," in *Proc. Int. Conf. Eng. MIS*, Istanbul, Turkey, 2015, p. 32.

[33] M. M. Zanjireh and H. Larijani, "A survey on centralised and distributed clustering routing algorithms for WSNs," in *Proc. IEEE 81st Veh. Technol. Conf. (VTC Spring)*, Glasgow, U.K., May 2015, pp. 1–6.

[34] N. Pradhan, K. Sharma, and V. K. Singh, "A survey on hierarchical clustering algorithm for wireless sensor networks," *ENERGY*, vol. 134, no. 4, pp. 30–35, 2016.

[35] V. Sucasas *et al.*, "A survey on clustering techniques for cooperative wireless networks," *Ad Hoc Netw.*, vol. 47, pp. 53–81, Sep. 2016.

[36] L. Xu, M. J. O'Grady, G. M. P. O'Hare, and R. Collier, "Reliable multihop intra-cluster communication for wireless sensor networks," in *Proc. Int. Conf. Comput. Netw. Commun. (ICNC)*, Feb. 2014, pp. 858–863.

[37] M. Zhao, Y. Yang, and C. Wang, "Mobile data gathering with load balanced clustering and dual data uploading in wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 14, no. 4, pp. 770–785, Apr. 2015.

[38] P. Nayak and A. Devulapalli, "A fuzzy logic-based clustering algorithm for WSN to extend the network lifetime," *IEEE Sensors J.*, vol. 16, no. 1, pp. 137–144, Jan. 2016.

[39] L. Xu, G. M. P. O'Hare, and R. Collier, "A balanced energy-efficient multihop clustering scheme for wireless sensor networks," in *Proc. 7th IFIP Wireless Mobile Netw. Conf. (WMNC)*, 2014, pp. 1–8.

[40] G. Sirisha, R. B. Babu, and K. R. Rao, "Establishing path quality management in wireless sensor networks through cluster head determination," *Indian J. Sci. Technol.*, vol. 9, no. 5, pp. 1–9, 2016.

[41] C. R. Lin and M. Gerla, "Adaptive clustering for mobile wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 15, no. 7, pp. 1265–1275, Sep. 1997.

[42] R. Nagpal and D. Coore, "An algorithm for group formation in an amorphous computer," in *Proceedings of the Tenth International Conference on Parallel and Distributed Systems (PDCS)*. Cambridge, MA, USA: MIT, 1998.

[43] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, vol. 2. Jan. 2000, p. 10.

[44] S. Banerjee and S. Khuller, "A clustering scheme for hierarchical control in multi-hop wireless networks," in *Proc. IEEE 20th Annu. Joint Conf. Comput. Commun. Soc.*, vol. 2. Anchorage, AK, USA, 2001, pp. 1028–1037.

[45] A. Manjeshwar and D. P. Agrawal, "TEEN: A routing protocol for enhanced efficiency in wireless sensor networks," in *Proc. 15th Int. Parallel Distrib. Process. Symp.*, San Francisco, CA, USA, 2001, pp. 2009–2015.

[46] P. Basu, N. Khan, and T. D. C. Little, "A mobility based metric for clustering in mobile ad hoc networks," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst. Workshop*, Mesa, AZ, USA, 2001, pp. 413–418.

[47] A. Manjeshwar and D. P. Agrawal, "APTEEN: A hybrid protocol for efficient routing and comprehensive information retrieval in wireless," in *Proc. Int. Parallel Distrib. Process. Symp. (IPDPS) Abstracts CD-ROM*, Fort Lauderdale, FL, USA, 2002, p. 8.

[48] M. J. Handy, M. Haase, and D. Timmermann, "Low energy adaptive clustering hierarchy with deterministic cluster-head selection," in *Proc. 4th Int. Workshop Mobile Wireless Commun. Netw.*, 2002, pp. 368–372.

[49] Z. Hongwei and A. Arora, "GS^3: Scalable self-configuration and self-healing in wireless networks," in *Proc. Annu. Symp. Principles Distrib. Comput. Mobile Wireless Commun. Netw.*, Monterey, CA, USA, 2002, pp. 58–67.

[50] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, Oct. 2002.

[51] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics," *IEEE Trans. Parallel Distrib. Syst.*, vol. 13, no. 9, pp. 924–935, Sep. 2002.

[52] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *Proc. IEEE 22nd Annu. Joint Conf. Comput. Commun. Soc. (INFOCOM)*, vol. 3. San Francisco, CA, USA, 2003, pp. 1713–1723.

[53] H. Chan and A. Perrig, *ACE: An Emergent Algorithm for Highly Uniform Cluster Formation*. Berlin, Germany: Springer, 2004, pp. 154–171. [Online]. Available: https://doi.org/10.1007/978-3-540-24606-0_11

[54] M. Demirbas, A. Arora, and V. Mittal, "Floc: A fast local clustering service for wireless sensor networks," in *Proc. Workshop Depend. Issues Wireless Ad Hoc Netw. Sensor Netw. (DIWANS/DSN)*, 2004, pp. 1–6.

[55] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Trans. Mobile Comput.*, vol. 3, no. 4, pp. 366–379, Oct./Dec. 2004.

[56] G. Smaragdakis, I. Matta, and A. Bestavros, "SEP: A stable election protocol for clustered heterogeneous wireless sensor networks," Dept. Comput. Sci., Boston Univ., Boston, MA, USA, Tech. Rep. BUCS-TR-2004-022, 2004.

[57] S. D. Muruganathan, D. C. F. Ma, R. I. Bhasin, and A. O. Fapojuwo, "A centralized energy-efficient routing protocol for wireless sensor networks," *IEEE Commun. Mag.*, vol. 43, no. 3, pp. S8–S13, Mar. 2005.

[58] P. Ding, J. Holliday, and A. Celik, "Distributed energy-efficient hierarchical clustering for wireless sensor networks," in *Proc. 1st IEEE Int. Conf. Distrib. Comput. Sensor Syst. (DCOSS)*, Marina Del Rey, CA, USA, 2005, pp. 322–339.

[59] K. T. Kim and H. Y. Youn, "Energy-driven adaptive clustering hierarchy (EDACH) for wireless sensor networks," in *Proc. Int. Conf. Embedded Ubiquitous Comput.*, Nagasaki, Japan, 2005, pp. 1098–1107.

[60] M. Ye, C. Li, G. Chen, and J. Wu, "EECS: An energy efficient clustering scheme in wireless sensor networks," in *Proc. 24th IEEE Int. Perform. Comput. Commun. Conf. (IPCCC)*, Phoenix, AZ, USA, 2005, pp. 535–540.

[61] C. Li, M. Ye, G. Chen, and J. Wu, "An energy-efficient unequal clustering mechanism for wireless sensor networks," in *Proc. IEEE Int. Conf. Mobile Adhoc Sensor Syst. Conf.*, Washington, DC, USA, 2005, p. 8.

[62] H. Huang and J. Wu, "A probabilistic clustering algorithm in wireless sensor networks," in *Proc. IEEE Veh. Technol. Conf.*, vol. 62. Dallas, TX, USA, 2005, pp. 1796–1798.

[63] I. Gupta, D. Riordan, and S. Sampalli, "Cluster-head election using fuzzy logic for wireless sensor networks," in *Proc. 3rd Annu. Commun. Netw. Services Res. Conf.*, 2005, pp. 255–260.

[64] V. Loscri, G. Morabito, and S. Marano, "A two-levels hierarchy for low-energy adaptive clustering hierarchy (TL-LEACH)," in *Proc. IEEE 62nd Veh. Technol. Conf. (VTC Fall)*, vol. 3. Dallas, TX, USA, 2005, pp. 1809–1813.

[65] H. Luo, F. Ye, J. Cheng, S. Lu, and L. Zhang, "TTDD: Two-tier data dissemination in large-scale wireless sensor networks," *Wireless Netw.*, vol. 11, nos. 1–2, pp. 161–175, Jan. 2005.

[66] S. Soro and W. B. Heinzelman, "Prolonging the lifetime of wireless sensor networks via unequal clustering," in *Proc. 19th IEEE Int. Parallel Distrib. Process. Symp.*, Denver, CO, USA, 2005, p. 8.

[67] A. Youssef, M. Younis, M. Youssef, and A. Agrawala, "WSN16-5: Distributed formation of overlapping multi-hop clusters in wireless sensor networks," in *Proc. Glob. Telecommun. Conf. (GLOBECOM)*, San Francisco, CA, USA, 2006, pp. 1–6.

[68] S.-M. Jung, Y.-J. Han, and T.-M. Chung, "The concentric clustering scheme for efficient energy consumption in the PEGASIS," in *Proc. 9th Int. Conf. Adv. Commun. Technol.*, vol. 1. 2007, pp. 260–265.

[69] T. Kang *et al.*, "A clustering method for energy efficient routing in wireless sensor networks," in *Proc. 6th WSEAS Int. Conf. Electron. Hardw. Wireless Opt. Commun. (WSEAS)*, 2007, pp. 133–138.

[70] K. Yanagihara *et al.*, "EACLE: Energy-aware clustering scheme with transmission power control for sensor networks," *Wireless Pers. Commun.*, vol. 40, no. 3, pp. 401–415, 2007.

[71] Q. Chen, J. Ma, Y. Zhu, D. Zhang, and L. M. Ni, "An energy-efficient K-hop clustering framework for wireless sensor networks," in *Proc. Eur. Conf. Wireless Sensor Netw.*, Delft, The Netherlands, 2007, pp. 17–33.

[72] S. Yi, J. Heo, Y. Cho, and J. Hong, "PEACH: Power-efficient and adaptive clustering hierarchy protocol for wireless sensor networks," *Comput. Commun.*, vol. 30, nos. 14–15, pp. 2842–2852, 2007.

[73] Y. Jin, L. Wang, Y. Kim, and X. Yang, "EEMC: An energy-efficient multi-level clustering algorithm for large-scale wireless sensor networks," *Comput. Netw.*, vol. 52, no. 3, pp. 542–562, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128607002927

[74] S. Lee *et al.*, "An energy-efficient distributed unequal clustering protocol for wireless sensor networks," *World Acad. Sci. Eng. Technol.*, vol. 48, pp. 443–447, 2008.

[75] B. Gong, L. Li, S. Wang, and X. Zhou, "Multihop routing protocol with unequal clustering for wireless sensor networks," in *Proc. Int. Colloquium Comput. Commun. Control Manag. (CCCM ISECS)*, vol. 2. Guangzhou, China, 2008, pp. 552–556.

[76] J.-H. Kim, S. H. Chauhdary, W.-C. Yang, D.-S. Kim, and M.-S. Park, "PRODUCE: A probability-driven unequal clustering mechanism for wireless sensor networks," in *Proc. 22nd Int. Conf. Adv. Inf. Netw. Appl. Workshops (AINAW)*, 2008, pp. 928–933.

[77] H. Le, D. Hoang, and R. Poliah, "S-Web: An efficient and self-organizing wireless sensor network model," in *Network-Based Information Systems* (LNCS 5186), M. Takizawa, L. Barolli, and T. Enokido, Eds. Berlin, Germany: Springer, 2008, pp. 179–188.

[78] T. Kaur and J. Baek, "A strategic deployment and cluster-header selection for wireless sensor networks," *IEEE Trans. Consum. Electron.*, vol. 55, no. 4, pp. 1890–1897, Nov. 2009.

[79] Y. Wang, T. Yang, and D. Zhang, "An energy efficient and balance hierarchical unequal clustering algorithm for large scale sensor network," *Inf. Technol. J.*, vol. 8, no. 1, pp. 28–38, 2009.

[80] N. Shigei, H. Morishita, and H. Miyajima, "Energy efficient clustering communication based on number of neighbors for wireless sensor networks," in *Proc. Int. Multi-Conf. Eng. Comput. Sci. (IMECS)*, Hong Kong, 2010.

[81] H. Bagci and A. Yazici, "An energy aware fuzzy unequal clustering algorithm for wireless sensor networks," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ)*, Barcelona, Spain, 2010, pp. 1–8.

[82] L. Buttyan and P. Schaffer, "PANEL: Position-based aggregator node election in wireless sensor networks," in *Proc. 4th IEEE Int. Conf. Mobile Ad Hoc Sensor Syst.*, Pisa, Italy, 2007, pp. 1–9.

[83] P. Ren, J. Qian, L. Li, Z. Zhao, and X. Li, “Unequal clustering scheme based LEACH for wireless sensor networks,” in *Proc. 4th Int. Conf. Genet. Evol. Comput. (ICGEC)*, Shenzhen, China, 2010, pp. 90–93.

[84] I. Urteaga, N. Yu, N. Hubbell, and Q. Han, “AWARE: Activity AWARE network clustering for wireless sensor networks,” in *Proc. IEEE 36th Conf. Local Comput. Netw. (LCN)*, 2011, pp. 589–596.

[85] J. Yu, Y. Qi, G. Wang, Q. Guo, and X. Gu, “An energy-aware distributed unequal clustering protocol for wireless sensor networks,” *Int. J. Distrib. Sensor Netw.*, vol. 7, no. 1, 2011, Art. no. 202145.

[86] D. Wei, Y. Jin, S. Vural, K. Moessner, and R. Tafazolli, “An energy-efficient clustering solution for wireless sensor networks,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3973–3983, Nov. 2011.

[87] W. Zhou, “Energy efficient clustering algorithm based on neighbors for wireless sensor networks,” *J. Shanghai Univ.*, vol. 15, no. 2, pp. 150–153, 2011.

[88] S. Lee, H. Choe, B. Park, Y. Song, and C.-K. Kim, “LUCA: An energy-efficient unequal clustering algorithm using location information for wireless sensor networks,” *Wireless Pers. Commun.*, vol. 56, no. 4, pp. 715–731, 2011.

[89] S. Deng, J. Li, and L. Shen, “Mobility-based clustering protocol for wireless sensor networks with mobile nodes,” *IET Wireless Sensor Syst.*, vol. 1, no. 1, pp. 39–47, Mar. 2011.

[90] Y. Ma, Y. Guo, X. Tian, and M. Ghanem, “Distributed clustering-based aggregation algorithm for spatial correlated sensor networks,” *IEEE Sensors J.*, vol. 11, no. 3, pp. 641–648, Mar. 2011.

[91] T. Liu, Q. Li, and P. Liang, “An energy-balancing clustering approach for gradient-based routing in wireless sensor networks,” *Comput. Commun.*, vol. 35, no. 17, pp. 2150–2161, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140366412002162

[92] J. Yuea, W. Zhang, W. Xiao, D. Tang, and J. Tang, “Energy efficient and balanced cluster-based data aggregation algorithm for wireless sensor networks,” *Procedia Eng.*, vol. 29, pp. 2009–2015, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877705812002639

[93] S. H. Kang and T. Nguyen, “Distance based thresholds for cluster head selection in wireless sensor networks,” *IEEE Commun. Lett.*, vol. 16, no. 9, pp. 1396–1399, Sep. 2012.

[94] D. P. Dahnil, Y. P. Singh, and C. K. Ho, “Topology-controlled adaptive clustering for uniformity and increased lifetime in wireless sensor networks,” *IET Wireless Sensor Syst.*, vol. 2, no. 4, pp. 318–327, Dec. 2012.

[95] E. Ever, R. Luchmun, L. Mostarda, A. Navarra, and P. Shah, “Uheed—An unequal clustering algorithm for wireless sensor networks,” in *Proc. SENSORNETS*, Rome, Italy, 2012, pp. 185–193.

[96] Y.-C. Chen and C.-Y. Wen, “Distributed clustering with directional antennas for wireless sensor networks,” *IEEE Sensors J.*, vol. 13, no. 6, pp. 2166–2180, Jun. 2013.

[97] Y. Liao, H. Qi, and W. Li, “Load-balanced clustering algorithm with distributed self-organization for wireless sensor networks,” *IEEE Sensors J.*, vol. 13, no. 5, pp. 1498–1506, May 2013.

[98] S.-S. Wang and Z.-P. Chen, “LCM: A link-aware clustering mechanism for energy-efficient routing in wireless sensor networks,” *IEEE Sensors J.*, vol. 13, no. 2, pp. 728–736, Feb. 2013.

[99] M. Yu, K. K. Leung, and A. Malvankar, “A dynamic clustering and energy efficient routing technique for sensor networks,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3069–3079, Aug. 2007.

[100] D. Koutsonikolas, S. Das, Y. C. Hu, and I. Stojmenovic, “Hierarchical geographic multicast routing for wireless sensor networks,” in *Proc. Int. Conf. Sensor Technol. Appl. (SensorComm)*, Valencia, Spain, 2007, pp. 347–354.

[101] F. Aurenhammer, “Voronoi diagrams—A survey of a fundamental geometric data structure,” *ACM Comput. Surveys*, vol. 23, no. 3, pp. 345–405, 1991.

[102] M. Meghji and D. Habibi, “Transmission power control in multihop wireless sensor networks,” in *Proc. 3rd Int. Conf. Ubiquitous Future Netw.*, 2011, pp. 25–30.

[103] T. S. Rappapor, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.

[104] N. Baccour *et al.*, “Radio link quality estimation in wireless sensor networks: A survey,” *Trans. Sensor Netw.*, 2012.

[105] K. Levis, “RSSI is under appreciated,” in *Proc. 3rd Workshop Embedded Netw. Sensors*, vol. 3031. Cambridge, MA, USA, May 2006, pp. 239–242.

[106] S. A. Nikolidakis, D. Kandris, D. D. Vergados, and C. Douligeris, “Energy efficient routing in wireless sensor networks through balanced clustering,” *Algorithms*, vol. 6, no. 1, pp. 29–42, 2013.

[107] W. M. Aioffi, C. A. Valle, G. R. Mateus, and A. S. da Cunha, “Balancing message delivery latency and network lifetime through an integrated model for clustering and routing in wireless sensor networks,” *Comput. Netw.*, vol. 55, no. 13, pp. 2803–2820, 2011.

[108] A. Tufail, “Reliable latency-aware routing for clustered WSNs,” *Int. J. Distrib. Sensor Netw.*, vol. 8, no. 3, 2012, Art. no. 681273.

[109] T. Zhong, S. Wang, S. Xu, H. Yu, and D. Xu, “Time delay based clustering in wireless sensor networks,” in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2007, pp. 3956–3960.

[110] M. Yao, C. Lin, Y. Tian, L. Wu, and Y. Chen, “Energy and delay minimization in cluster-based wireless sensor networks,” in *Proc. IEEE Int. Conf. Green Comput. Commun. (GreenCom)*, Besançon, France, Nov. 2012, pp. 588–594.

[111] O. Dousse, P. Mannersalo, and P. Thiran, “Latency of wireless sensor networks with uncoordinated power saving mechanisms,” in *Proc. Mobihoc*, 2004, pp. 109–120.

[112] Y. Li, L. Guo, and S. K. Prasad, “An energy-efficient distributed algorithm for minimum-latency aggregation scheduling in wireless sensor networks,” in *Proc. IEEE 30th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Genoa, Italy, Jun. 2010, pp. 827–836.

[113] D. Padmanabhan and F. Labeau, “Effect of location on the latency in cluster-based WSNs,” in *Proc. IEEE Sensors*, Baltimore, MD, USA, Nov. 2013, pp. 1–4.

[114] L. Xu, D. Lillis, G. M. O’Hare, and R. W. Collier, “A user configurable metric for clustering in wireless sensor networks,” in *Proc. SENSORNETS*, 2014, pp. 221–226.

[115] N. Heo and P. K. Varshney, “Energy-efficient deployment of intelligent mobile sensor networks,” *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 35, no. 1, pp. 78–92, Jan. 2005.

[116] M. Korkalainen, M. Sallinen, N. Karkkainen, and P. Tukeva, “Survey of wireless sensor networks simulation tools for demanding applications,” in *Proc. 5th Int. Conf. Netw. Services (ICNS)*, 2009, pp. 102–106.

[117] A. Abuarqoub, F. Alfayez, M. Hammoudeh, T. Alsboui, and A. Nisbet, “Simulation issues in wireless sensor networks: A survey,” in *Proc. 6th Int. Conf. Sensor Technol. Appl. (SENSORCOMM)*, 2012, pp. 222–228.

[118] K. Romer and F. Mattern, “The design space of wireless sensor networks,” *IEEE Wireless Commun.*, vol. 11, no. 6, pp. 54–61, Dec. 2004.

[119] J. Polastre, R. Szewczyk, A. Mainwaring, D. Culler, and J. Anderson, “Analysis of wireless sensor networks for habitat monitoring,” in *Wireless Sensor Networks*, C. S. Raghavendra, K. M. Sivalingam, and T. Znati, Eds. Norwell, MA, USA: Kluwer Acad., 2004, pp. 399–423. [Online]. Available: http://dl.acm.org/citation.cfm?id=1013825.1013844

[120] G. Booch *et al.*, *Object-Oriented Analysis and Design With Applications*, 3rd ed. Upper Saddle River, NJ, USA: Addison-Wesley, 2007.

[121] L. Ruiz-Garcia, L. Lunadei, P. Barreiro, and I. Robla, “A review of wireless sensor technologies and applications in agriculture and food industry: State of the art and current trends,” *Sensors*, vol. 9, no. 6, pp. 4728–4750, 2009.

[122] K. H. Pollock *et al.*, “Large scale wildlife monitoring studies: Statistical methods for design and analysis,” *Environmetrics*, vol. 13, no. 2, pp. 105–119, 2002.

[123] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, “Internet of Things for smart cities,” *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.

[124] D. Evans, *The Internet of Everything: How More Relevant and Valuable Connections Will Change the World*, Cisco IBSG, San Jose, CA, USA, 2012, pp. 1–9.

[125] M. Chiang and T. Zhang, “Fog and IoT: An overview of research opportunities,” *IEEE Internet Things J.*, vol. 3, no. 6, pp. 854–864, Dec. 2016.

[126] D. Warren and C. Dewar, “Understanding 5G: Perspectives on future technological advancements in mobile,” GSMA, London, U.K., Tech. Rep., Dec. 2014.

[127] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[128] R. Munoz *et al.*, “The CTTC 5G end-to-end experimental platform: Integrating heterogeneous wireless/optical networks, distributed cloud, and IoT devices,” *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 50–63, Mar. 2016.

[129] R. Fantacci, T. Pecorella, R. Viti, and C. Carlini, “A network architecture solution for efficient IoT WSN backhauling: Challenges and opportunities,” *IEEE Wireless Commun.*, vol. 21, no. 4, pp. 113–119, Aug. 2014.

[130] E. Piri and J. Pinola, "Performance of LTE uplink for IoT backhaul," in *Proc. 13th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Las Vegas, NV, USA, Jan. 2016, pp. 6–11.
[131] M. Hindia, T. A. Rahman, H. Ojukwu, E. B. Hanafi, and A. Fattouh, "Enabling remote health-caring utilizing IoT concept over LTE-femtocell networks," *PLoS One*, vol. 11, no. 5, 2016, Art. no. e0155077.
[132] W. Kim, "Adaptive resource scheduling for dual connectivity in heterogeneous IoT cellular networks," *Int. J. Distrib. Sensor Netw.*, vol. 12, no. 4, 2016, Art. no. 6036952.
[133] V. Jungnickel *et al.*, "Optical wireless communication for backhaul and access," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Valencia, Spain, Sep./Oct. 2015, pp. 1–3.
[134] M. Grob, "The road to 5G: Providing the connectivity fabric for everything," in *Proc. IEEE Hot Chips Symp. (HCS)*, Aug. 2015, pp. 1–26.
[135] M. R. Palattella *et al.*, "Internet of Things in the 5G era: Enablers, architecture, and business models," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 510–527, Mar. 2016.
[136] A. Cimmino *et al.*, "The role of small cell technology in future smart city applications," *Trans. Emerg. Telecommun. Technol.*, vol. 25, no. 1, pp. 11–20, 2014.
[137] R. Fantacci, T. Pecorella, R. Viti, and C. Carlini, "Short paper: Overcoming IoT fragmentation through standard gateway architecture," in *Proc. IEEE World Forum Internet Things (WF IoT)*, Seoul, South Korea, Mar. 2014, pp. 181–182.
[138] A. Hakiri, P. Berthou, A. Gokhale, and S. Abdellatif, "Publish/subscribe-enabled software defined networking for efficient and scalable IoT communications," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 48–54, Sep. 2015.
[139] K. Wang, Y. Wang, D. Zeng, and S. Guo, "An SDN-based architecture for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 25–31, Feb. 2017.
[140] P. Huskov, T. Maksymyuk, I. Kahalo, and M. Klymash, "Smart backhauling subsystem for 5G heterogeneous network," in *Proc. Exp. Designing Appl. CAD Syst. Microelectron.*, Feb. 2015, pp. 481–483.
[141] A. Lea, K. Negus, H. Tapse, B. Varadarajan, and R. Vaughan, "Spectrum options for wireless backhaul of small cells," in *Proc. IEEE 8th Eur. Conf. Antennas Propag. (EuCAP)*, The Hague, The Netherlands, 2014, pp. 3310–3311.
[142] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5G heterogeneous cloud radio access networks," *IEEE Network*, vol. 29, no. 2, pp. 6–14, Mar./Apr. 2015.
[143] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 126–135, Dec. 2014.
[144] M. Hassanalieragh *et al.*, "Health monitoring and management using Internet-of-Things (IoT) sensing with cloud-based processing: Opportunities and challenges," in *Proc. IEEE Int. Conf. Services Comput.*, Jun. 2015, pp. 285–292.
[145] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: Challenges, solutions, and future directions," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 86–92, May 2014.
[146] I. Loumiotis, E. Adamopoulou, K. Demestichas, and M. Theologou, *Optimal Backhaul Resource Management in Wireless-Optical Converged Networks*. Cham, Switzerland: Springer Int., 2015, pp. 254–261. [Online]. Available: https://doi.org/10.1007/978-3-319-19743-2_36
[147] S. Andreev *et al.*, "Intelligent access network selection in converged multi-radio heterogeneous networks," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 86–96, Dec. 2014.
[148] G. Yu, Y. Jiang, L. Xu, and G. Y. Li, "Multi-objective energy-efficient resource allocation for multi-RAT heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2118–2127, Oct. 2015.
[149] W. Xia *et al.*, "Bandwidth allocation in heterogeneous networks with wireless backhaul," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.
[150] C. Ran, S. Wang, and C. Wang, "Balancing backhaul load in heterogeneous cloud radio access networks," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 42–48, Jun. 2015.
[151] R. Bonnefoi, C. Moy, and J. Palicot, "Advanced metering infrastructure backhaul reliability improvement with cognitive radio," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, 2016, pp. 230–236.
[152] S. Bassoy, M. Jaber, M. A. Imran, and P. Xiao, "Load aware self-organising user-centric dynamic comp clustering for 5G networks," *IEEE Access*, vol. 4, pp. 2895–2906, 2016.
[153] L. Xu, G. M. O'Hare, and R. W. Collier, "A smart and balanced energy-efficient multihop clustering algorithm (Smart-BEEM) for MIMO IoT systems in future networks," *MDPI Sensors*, vol. 17, no. 7, 2017, Art. no. E1574.
[154] J. Zhao, C. Qiao, R. S. Sudhaakar, and S. Yoon, "Improve efficiency and reliability in single-hop WSNs with transmit-only nodes," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 3, pp. 520–534, Mar. 2013.
[155] M. F. Munir and F. Filali, "Increasing connectivity in wireless sensor-actuator networks using dynamic actuator cooperation," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, 2008, pp. 203–207.
[156] L. Xu, J. Xie, X. Xu, and S. Wang, "Enterprise LTE and WiFi interworking system and a proposed network selection solution," in *Proc. ACM Symp. Archit. Netw. Commun. Syst. (ANCS)*, Santa Clara, CA, USA, 2016, pp. 137–138.
[157] R. Lent and E. C.-H. Ngai, "Lightweight clustering in wireless sensor-actuator networks on obstructed environments," in *Proc. IEEE 3rd Int. Symp. Wireless Pervasive Comput. (ISWPC)*, 2008, pp. 26–32.
[158] O. Banimelhem, M. Mowafi, E. Taqieddin, F. Awad, and M. Al Rawabdeh, "An efficient clustering approach using genetic algorithm and node mobility in wireless sensor networks," in *Proc. IEEE 11th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Barcelona, Spain, 2014, pp. 858–862.

Lina Xu received the B.E. degree from the Software Engineering School, Fudan University, Shanghai, China, and B.Sc degree in computer science from University College Dublin (UCD), Dublin, Ireland, though a Fudan and UCD Joint Program, and the Ph.D. degree in computer science from UCD in 2014.

From 2014 to 2016, she was with HP Labs, as a Researcher. She is currently an Assistant Professor with UCD. Her research interests include the Internet of Things (IoT), 5G networks, smart networking, machine learning, etc.

Rem Collier received the primary degree in pure and applied mathematics from the University of Bristol, Bristol, U.K., in 1994, the M.Sc. degree in computation (examination and dissertation) and M.Phil (research) from the University of Manchester Institute for Science and Technology (UMIST), Manchester, U.K., in 1995 and 1996, respectively, and the Ph.D. degree in computer science from University College Dublin (UCD), Dublin, Ireland, in 2001.

After completing his Ph.D. degree, he spent a year working in industry in the area of web-based e-learning content management platforms for course authoring and localization. In 2002, he returned to UCD, where he was a Post-Doctoral Researcher, and in 2003, became an Assistant Lecturer. In 2005, he became a College Lecturer.

Gregory M. P. O'Hare (M'08) received the B.Sc, M.Sc, and Ph.D. degrees from the University of Ulster, Coleraine, Northern Ireland.

He was the Head of the Department of Computer Science, University College Dublin (UCD), from 2001 to 2004. Prior to joining UCD, he was on the faculty of the University of Central Lancashire (1984–1986) and the University of Manchester (1986–1996). He is currently a Professor with the School of Computer Science and Informatics, UCD. His research interests are distributed artificial intelligence and multi-agent systems (MAS), and mobile and ubiquitous computing, autonomic systems, and wireless sensor networks.

Dr. O'Hare is a Fellow of the British Computer Society. He is a member of the ACM and AAAI. He is a Chartered Engineer.