Khan, M. F., Yau, K.-L. A., Noor, R. M. and Imran, M. A. (2020) Survey and taxonomy of clustering algorithms in 5G. Journal of Network and Computer Applications, 154, 102539. (doi: 10.1016/j.jnca.2020.102539)

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

http://eprints.gla.ac.uk/207890/

Deposited on: 23 January 2020

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk

Survey and Taxonomy of Clustering Algorithms in 5G

Muhammad Fahad Khan[a,b], Kok-Lim Alvin Yau[a], Rafidah MD Noor[c] and Muhammad Ali Imran[d]

[a] *Department of Computing and Information Systems, Sunway University, Malaysia*
[b] *COMSATS University Islamabad (CUI), Pakistan*
[c] *Department of Computer System and Technology, University of Malaya, Malaysia*
[d] *School of Engineering, University of Glasgow, Glasgow, United Kingdom*

ARTICLE INFO

Keywords:
5G
Fifth generation
Cellular mobile communication
Clustering
Network topology.

ABSTRACT

The large-scale deployment of fifth generation (5G) is expected to produce a massive amount of data with high variability due to ultra-densification and the rapid increase in a heterogeneous range of applications and services (e.g., virtual reality, augmented reality, and driver-less vehicles), and network devices (e.g., smart gadgets and sensors). Clustering organizes network topology by segregating nodes with similar interests or behaviors in a network into logical groups in order to achieve network-level and cluster-level enhancements, particularly cluster stability, load balancing, social awareness, fairness, and quality of service. Clustering has been investigated to support mobile user equipment (UE) in access networks, whereby UEs form clusters themselves and may connect to BSs. In this paper, we present a comprehensive survey of the research work of clustering schemes proposed for various scenarios in 5G networks and highlight various aspects of clustering schemes, including objectives, challenges, metrics, characteristics, performance measures. Furthermore, we present open issues of clustering in 5G.

Nomenclature

5G	Fifth generation
AoA	Angle of arrival
BS	Base station
CDSA	Control-data separation architecture
CG	Cluster gateway
CH	Cluster head
CM	Cluster member
CoMP	Coordinated multi-point
D2D	Device-to-device
IoE	Internet of everything
IoH	Internet of humans
IoT	Internet of things
MIMO	Multiple-input and multiple-output
MSN	Mobile social network
NS	Neighbor set
QoS	Quality of service
RSS	Received signal strength
SC	Small cell
SINR	Signal-to-interference-plus-noise-ratio
UE	User equipment
V2X	Vehicle-to-everything
VANET	Vehicular ad-hoc network

1. Introduction

The tremendous growth of user equipment (UE) expecting to reach up to billions in number [1, 2], along with bandwidth-starving applications (e.g., video streaming, multimedia sharing, and online gaming), has contributed to 74% increment in data traffic over the years [3]. By 2020, data traffic is expected to increase by 8-fold [4] with the introduction of next-generation bandwidth-starving applications (e.g., augmented reality, virtual reality, and driver-less vehicle), and new services (e.g., smart home, smart healthcare, and smart city). Hence, there is a colossal demand for significantly higher network capacity and lower delay to support higher mobility UEs, leading to the introduction of the next-generation mobile wireless network, namely fifth generation (5G).

5G incorporates new technologies, including massive multiple-input and multiple-output (MIMO) [5], device-to-device (D2D) communication [6], coordinated multi-point (CoMP) [7], and beamforming [8], in order to explore and exploit mmWave [9] and underutilized spectrum for improved spectral efficiency [10], coordinate different kinds of network cells (e.g., macrocells, and small cells (SCs) including picocells and femtocells) for reduced interference [11, 12, 13], and achieve network virtualization for sharing of network-wide resources [14]. This caters for next-generation network scenarios characterized by ultra-densification, heterogeneous, and high variability, in order to achieve a better Quality of Service (QoS) of up to 10× higher data rate [15], up to 1000× lower delay [16], up to 99.999% higher reliabil-

*Corresponding author
18029090@sunway.edu.my,m.fahad@cuiatk.edu.pk (M.F. Khan); koklimy@sunway.edu.my (K.A. Yau); fidah@um.edu.my (R.M. Noor); muhammad.imran@glasgow.ac.uk (M.A. Imran)
ORCID(s):

ity and availability [17], up to 100× larger network coverage [18], and up to 10× longer battery lifetime [19]. As an example of the new technologies, D2D enables neighboring nodes to perform direct communication among themselves without passing through a base station (BS), which can offload traffic from the BS to reduce network congestion while reducing delay and energy consumption [6, 20].

D2D can use the to-be-opened mmWave bands from 3 GHz to 300 GHz [21], in addition to the currently used frequency bands from 300 MHz to 3 GHz. Despite its short wavelength causing high propagation loss and poor penetration through obstacles, it provides short-range D2D communication with data rate of up to multiple Gbps [22]. Hence, D2D can support the deployment of SCs through spatial reuse of frequency bands to cater for ultra-densification.

Clustering segregates nodes with similar interests (e.g., improving load balancing, social awareness, fairness of resource distribution, network lifetime, and spectral efficiency) or behaviors (e.g., similar geographical location, speed, and direction) in a network into logical groups in order to achieve various network objectives [23]. Of particular interest is clustering in the access networks, which is different from cell clustering performed in the core network [52]. Cell clustering segregates cells with similar interests or behaviors in a network into logical groups in order to improve network performance, particularly higher quality of service (QoS) (e.g., higher throughput and energy efficiency, as well as lower cost and delay) and higher intra-cluster signal-to-interference-plus-noise-ratio (SINR). The rest of this paper focuses on clustering in access networks, rather than cell clustering. Fig. 1 shows a cluster structure in a 5G network. In Fig. 1, there are three types of nodes. *Cluster head*, which serves as the leader of a cluster, manages and handles cluster-level operations (e.g., data aggregation, load distribution, resource allocation, local synchronization, and D2D transmission), as well as performs intra-cluster and inter-cluster communications. For data forwarding, *intra-cluster communication* involves interaction between a cluster head and a cluster member; while *inter-cluster communication* involves interaction between a cluster head and a neighboring cluster [24]. *Cluster member*, which is associated with the cluster head, performs intra-cluster communication. *Cluster gateway*, which is also associated with the cluster head, is a cluster member that can communicate with neighboring clusters, and so it performs inter-cluster communication. In Fig. 1, there are three clusters (i.e., C_1, C_2 and C_3). In cluster C_1, there are cluster head CH_1, cluster member $CM_{1,1}$, as well as cluster gateways $CG_{1,1}$ and $CG_{1,2}$ that provide inter-cluster communication with clusters C_2. Using D2D, in Fig. 1, cluster member $CM_{2,2}$ communicates with cluster gateway $CG_{2,1}$ in cluster C_2, and cluster member $CM_{3,1}$ communicates with cluster member $CM_{3,2}$ in cluster C_3 [25, 26, 27]. Each cluster can consist of a single network cell, or span across multiple network cells. For instance, SCs are deployed in ultra-dense networks that require higher data rate; each SC can form a disjoint cluster to reduce interference and contention, as well as improve the efficiency of resource utilization [28]. In each cluster, CH is elected and CGs are selected for inter-cell communication.

1.1. Our contributions

This paper provides a comprehensive survey of clustering algorithms in 5G networks. A taxonomy for clustering attributes, covering clustering objectives, challenges, metrics, characteristics, and performance measures, as well as a clustering framework, are presented. The clustering algorithms are classified, analyzed, and discussed based on the taxonomy and clustering framework. Open issues of this research topic are also outlined.

Traditionally, clustering has been proposed to achieve network stability and scalability in order to improve network performance, such as throughput, the fairness of resource distribution and load balancing, as well as the lifetime of a CH, a cluster, or a network, while providing support for routing. Various surveys has been conducted on clustering schemes in different *network types* (i.e., cognitive radio [47, 24], mobile ad-hoc [40, 42], wireless sensor [39, 41, 45, 48, 50, 51], vehicular ad-hoc [43, 44], Internet of things (IoT) [46, 49], and 5G [30]) in the literature. This paper focuses on clustering in access networks, rather than cell clustering in core networks. The clustering approaches presented in this paper should also be distinguished from the application of clustering to segregate data points into groups in machine learning. To the best of our knowledge, this paper is first of its kind to provide a survey on clustering algorithms in 5G. While a survey paper that focuses on clustering using coordinated multi-point (CoMP) [30], which is one of the main technologies in 5G, has been presented, this paper focuses on clustering using various technologies in 5G. Hence, this paper establishes a holistic foundation on clustering in 5G, and provides insights to guide research direction in this topic.

1.2. Organization of this paper

The rest of this paper is organized as shown in Fig. 2. Section 2 presents background and the motivation for clustering in 5G. Section 3 presents a taxonomy of clustering attributes in 5G. Section 4 presents a clustering framework. Section 5 presents various clustering algorithms in 5G based on the taxonomy and clustering framework. Section 6 presents open issues. Section 7 concludes the paper.

2. Background and Motivating the Need for Clustering in 5G Networks

With a massive amount of highly variable data generated by the tremendous growth of a diverse range of UEs, clustering has been proposed to organize network topology and summarize data in order to improve network performance (e.g., higher network scalability, spectral efficiency, data availability, and load balancing) [23, 31] based on new and traditional clustering metrics. In existing clustering schemes, various clustering metrics are used in cluster formation (e.g., CH election) and cluster maintenance (or re-clustering), includ-

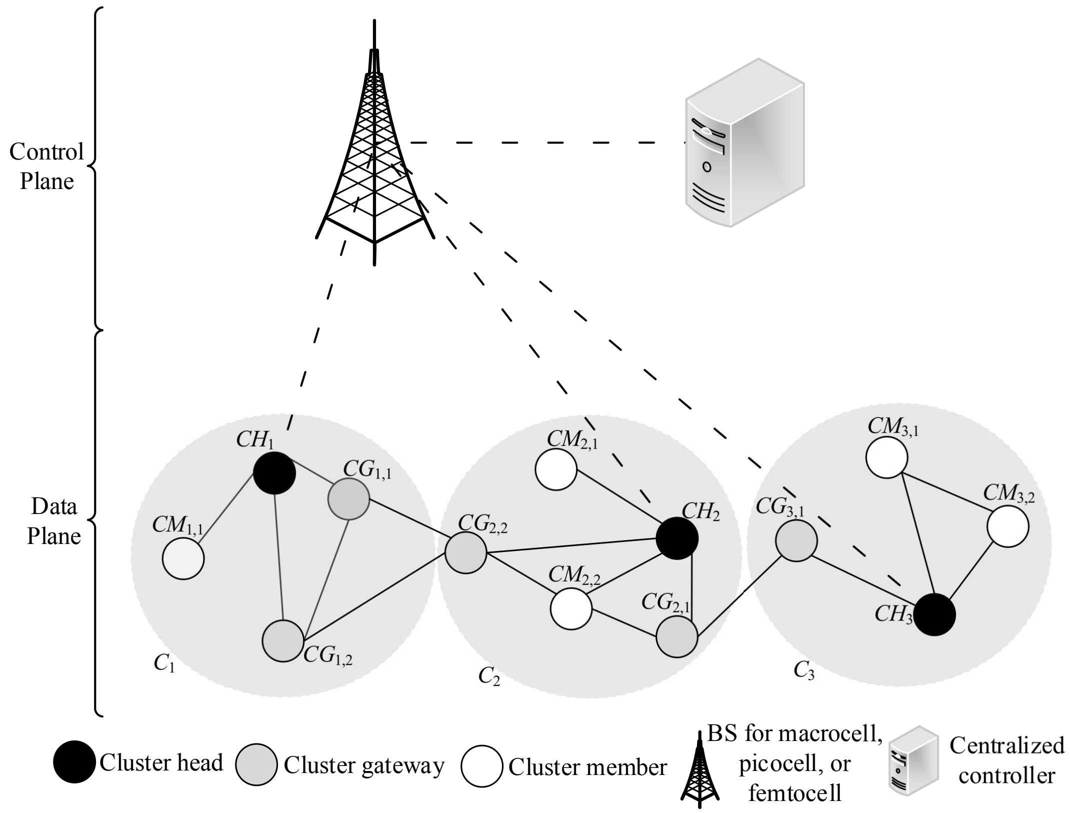

Figure 1: An example of a cluster structure in a 5G network. In 5G, the network is segregated into control plane (e.g., consists of BS and centralized controller) and data plane (e.g., consists of routers and switches) to handle network-wide and local tasks, respectively. More description of the 5G architecture is provided in Section 2.2. Gray shaded area represents a cluster boundary. Dashed line represents a connection between the BS of a network cell (e.g., macrocell, picocell, and femtocell) and a CH or a centralized controller. Solid line represents a connection between a pair of UEs.

ing node degree (or the number of neighboring nodes) [32], node ID [33, 34] (e.g., the highest and lowest node ID in max-min D-clustering [35]), mobility (e.g., the lower mobility in [36]), cluster size (e.g., the number of hops in a cluster in *k*-hop clustering [37]), and the capability to provide equal access to neighboring nodes [38]. Clustering schemes can use multiple clustering metrics, such as weighted clustering that uses four kinds of clustering metrics, namely the node degree, transmission power, mobility, and residual energy of a node [53]. While existing clustering schemes have shown network performance enhancement in various contexts, they are insufficient to cater to the needs of next-generation wireless mobile networks. The rest of this section presents background and the need for clustering in 5G. Section 2.1 presents three main characteristics of the next generation wireless mobile networks. Section 2.2 presents 5G architecture and new technologies. Section 2.3 presents how clustering can cater for the three main characteristics in the 5G context. Finally, Section 2.4 presents the cost of clustering.

2.1. What will the next generation wireless mobile networks be like?

The next generation wireless mobile networks are foreseen to possess three main characteristics.

Firstly, *ultra-densification* whereby there is a large number of active UEs per unit area (i.e., up to 100× [54]) generating a massive amount of data. In addition, a large number of active UEs with high mobility causes frequent link disconnections [71]. So, it is necessary to increase spectral efficiency [55], and hence network capacity and bandwidth availability [56]. As an example, D2D communication allows neighboring nodes to communicate with each other without passing through a BS, which can reduce control message exchange, and offload traffic from the BS, leading to increased bandwidth availability at BS [57]. As another example, SCs are deployed to cater for local traffic with reduced energy consumption [60]. Ultra-densification raises the challenge to provide up to 10× higher network capacity than that in 4G networks, while providing greater mobility support and efficiency in resource allocation.

Secondly, *network heterogeneity* whereby there is a diverse range of UEs (e.g, laptops, smart gadgets, unmanned aerial vehicles (UAVs), and sensors), network cells (e.g., macrocells, picocells, and femtocells), and networks (e.g., Internet of everything (IoE), Internet of humans (IoH), and IoT) [21, 61]. So, it is necessary to cater for a diverse range of network entities with different capabilities and operating parameters (e.g., operating frequency band and transmission

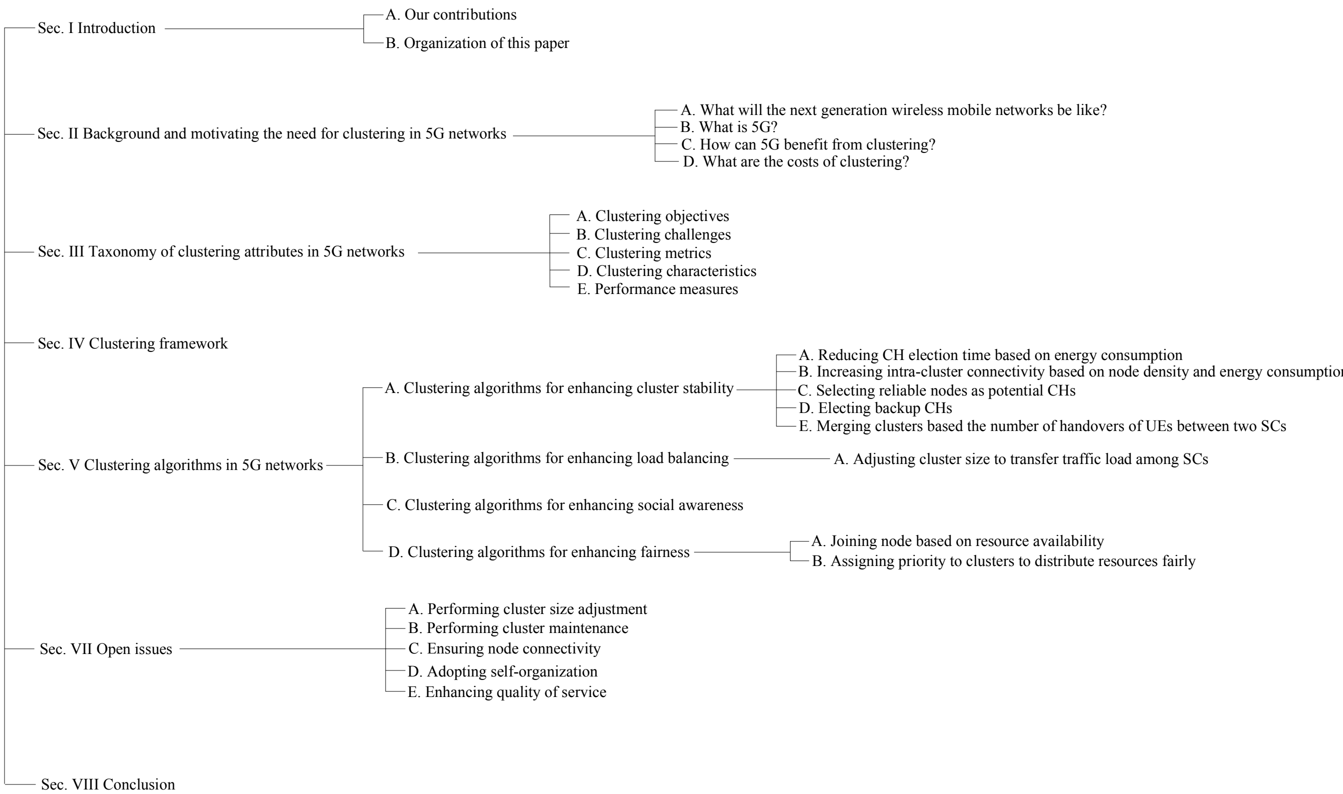

Figure 2: Organization of this paper.

power) [62]. As an example, the deployment of macrocells caters for long-range transmission, while the deployment of SCs caters for short-range transmission that enables efficient resource utilization via spatial reuse of frequency bands [63]. Network heterogeneity raises the challenge to provide up to 10× higher network capacity and 100× larger coverage than that in 4G networks [64].

Thirdly, *high variability* whereby the heterogeneous traffic generated by a diverse range of network entities and applications, and the spectrum availability [65], can be bursty with a high peak-to-mean ratio of up to 100 : 1 across space and time causing bottleneck and network congestion [69]. So, it is necessary to cater for the unpredictable and resource intensive network condition. As an example, mmWave bands provide data rate of up to multiple Gbps [22] to improve spectrum availability and peak-to-mean ratio [66]. High variability raises the challenge to achieve scalability, agility, and reconfigurability [67, 68] in order to cater for the scarcity and overabundance of network capacity caused by high variability in traffic and spectrum availability [69, 70].

2.2. What is 5G?

The expected advancement of 5G compared with its predecessor technology, namely 4G, in terms of some performance measures is given in Table 1. Generally speaking, mean data rate and peak data rate are expected to increase up to 10× and 20×, respectively; traffic capacity and energy efficiency are expected to increase up to 100×; and delay is expected to reduce to 1 ms, and so on, with the comparison of 4G. Clustering has been investigated to support mobile user equipment (UE) in access networks, whereby UEs form clusters themselves and may connect to BSs. However, clustering has not been investigated to support base stations (BSs) in the network core, whereby multiple BSs from different network cells form clusters themselves.

Architecturally, 5G network is based on the control-data separation architecture (CDSA) whereby the network is separated into two planes for better network management, as shown in Fig. 1. *Control plane* contains a centralized controller that observes and manages network-wide traffic and its fluctuations/ changes, and makes intelligent decisions for routing and clustering based on network-wide policies in an efficient manner. *Data plane* contains network infrastructure, such as routers and switches, that performs tasks following instructions, logics, or rules, given by the control plane [71]. For instance, the control plane performs routing to determine forwarding tables for routers, and the data plane performs data forwarding based on the forwarding tables. Hence, using a centralized controller from cloud that provides high computational power and shared resources, UEs can offload tasks to the centralized controller in the control plane. So, CDSA improves network performance (e.g., network stability, scalability, and flexibility) in the data plane [24].

5G incorporates six main new technologies:

Table 1
Comparison of 4G and 5G networks

Feature	Unit	4G Network	5G Network
Mean data rate	Megabits per second	10	100
Peak data rate	Gigabits per second	1	20
Area traffic capacity	Megabits per second per meter square	0.1	10
Energy efficiency	-	1×	100×
Connection density	Devices per kilometer square	10^5	10^6
Delay	Millisecond	10	1
Mobility	Kilometer per hour	350	500
Spectral efficiency	-	1×	3×

- *Massive MIMO* uses a large number of vertical and horizontal software-defined transmit-antennas (e.g., up to 16 directional antennas per sector) at BS to provide narrow beams over multiple beam-point angles in order to provide highly directional transmission. This improves network capacity up to 10% and extends network coverage up to 100×, as well as dramatically increases energy efficiency [72]. Hence, massive MIMO can support a higher number of active UEs [73, 74]. However, a higher number of transmit-antennas can increase the hardware and computational complexity/ cost. So, there is a trade-off between throughput enhancement and hardware and computational complexity/ cost [75], although a slight reduction of throughput can reduce the computational complexity to a larger extent using clustering [76]. Clustering has been proposed to optimize the trade-off via cooperative communication among the transmit-antennas [77] [74] in MIMO.
- *CoMP* enables overlapping network cells to coordinate among themselves in order to minimize inter-cell interference [78, 79] and improve the data rate and spectral efficiency of UEs' at the cell edge (or cell edge user) [11]. In CoMP, a large number of messages containing control information (e.g., channel state information) must be exchanged among network cells. However, coordination among network cells can increase computational complexity (e.g., for supporting beamforming, multiplexing, and synchronization among network cells) and overhead, resulting in increased bandwidth requirement. This is despite a larger number of network cells participating in a collaboration can reduce inter-cell interference. So, there is a trade-off between the level of inter-cell interference and the number of network cells participating in a collaboration (or the amount of overhead). Clustering at network cell-level has been proposed to localize (or to limit) collaboration among network cells within a cluster participate in a collaboration; in other words, the boundary of a cluster defines the collaboration area [80]. This enables BSs with directional antennas to coordinate resource usage among BSs and UEs, particularly at overlapping areas, in order to improve CoMP gain (e.g., which is based on spectral efficiency, energy efficiency, load balancing, and fairness of resource distribution) [81] [82, 83] and reduce overhead [30], otherwise a large-scale collaboration can congest the network. Nevertheless, other mechanisms can be applied to minimize inter-cell interference, such as using antenna down-tilt [80].
- *Exploration and exploitation of mmWave* enables network entities to explore and use the to-be-opened mmWave bands from 3 GHz to 300 GHz [21] in order to improve network performance (e.g., higher throughput, energy efficiency, and spectral efficiency). Since mmWave communication uses short wavelength, it provides a data rate of up to multiple Gbps [84]; however, it experiences high propagation loss and poor penetration through obstacles, so it requires line-of-sight (LOS) and is suitable for short-range communication only [85]. This means that node mobility can cause frequent handover and user association with BS in mobile networks, so it can cause rapid changes (or fluctuations) to the channel state of a node and the traffic load at the BS [86, 87, 88]. Clustering has been proposed to handle handover and user association with BS in order to optimize load balancing [89].
- *Exploration and exploitation of underutilized spectrum* enables network entities to explore and use underutilized high-quality and persistently available frequency bands, including the mmWave bands [73, 90], in the presence of dynamic channel availability. Messages containing information (e.g., the underutilized spectrum) must be exchanged among UEs to facilitate channel access [91]. Clustering has been proposed to localize collaboration among UEs within a cluster to reduce the effects of dynamic channel availability, and the flooding of messages, in order to achieve network stability and scalability, while supporting cooperative tasks (e.g., channel sensing and routing [92]).
- *Network virtualization* enables centralized entities (e.g., centralized controllers in the cloud) to pool together network resources and functionalities from heterogeneous network entities [93, 94] in order to provide shared network resources and advanced network functionalities [95]. While this increases network

scalability and flexibility [96], resource distribution to functionalities must be adjusted to achieve fairness in network virtualization [97]. This is because some nodes may consume majority of a resource pool, causing starvation among other nodes. Clustering has been proposed to impose a threshold on the amount of resources that CMs, which have different resource requirements, can request in order to achieve fairness of resource distribution among clusters [98][14] [99].

- *D2D communication* enables neighboring nodes to perform direct communication among themselves without passing through a BS. This helps to offload traffic from the BS [54] to reduce network congestion while reducing delay and energy consumption; however, D2D must be well managed to reduce interference among communications with and without D2D. There are *two* main issues: a) there is a tradeoff between communication with and without D2D [100], b) the dynamicity of channel conditions, network topology, and traffic amount can cause bottleneck and network congestion in D2D communication [101, 102]. Clustering has been proposed to: a) segregate nodes that initiate communications with and without D2D into different clusters in order to reduce network congestion [100, 103], b) ensure a CH has high channel quality in communication with BSs and CMs [104, 105].

2.3. How can 5G benefit from clustering?

Traditionally, clustering has been proposed to achieve network stability and scalability in order to improve network performance, such as throughput, the fairness of resource distribution, a better load balancing, as well as the lifetime of a CH, a cluster, or a network, while providing support for routing [29]. In 5G, clustering caters for the three main characteristics of next-generation wireless mobile networks (see Section 2.1). that warrant the need for clustering in access networks, which is the focus of this paper.

Firstly, in order to cater for *ultra-densification*, clustering segregates nodes in a network into logical groups based on common characteristics (e.g., angle of arrival (AoA), relative speed, degree centrality, and social relation) or metrics (e.g., fairness index), so that alike nodes can aggregate the massive amount of data [106] and access resources in an optimized manner [107, 108]. Since network disconnections can be caused by high mobility UEs [109], clustering segregates nodes in a network with similar behavior (i.e., similar speed) into logical groups, and each cluster may connect to different BSs in order to achieve network stability [110, 56].

Secondly, in order to cater for *network heterogeneity*, clustering segregates network entities into logical groups based on common characteristics (e.g., relative speed, degree centrality, and social relation) or metrics (e.g., fairness index), so that heterogeneous network entities can prolong their respective connections. As an example, in [112], there are heterogeneous UEs with different requirements for network resources [113], and clustering segregates SCs into logical groups in order to reduce interference. In other words, the UEs of the SCs are jointly served by the BSs of the SCs belonging to a single SC cluster in order to achieve fairness of resource allocation [114]. As another example, SCs are overlaid in macrocells to increase network capacity, and clustering segregates heterogeneous network cells into logical groups in order to offload computation tasks from macrocells. So, each logical group, being a *computation cluster* – much like a local cloud – consists of nearby SCs that contribute to the computation tasks [115] in order to improve network performance and user satisfaction in heterogeneous networks [116, 117].

Thirdly, in order to cater for *high variability*, clustering segregates network entities into logical groups whereby a CH and its CMs interact with each other via intra-cluster communication in D2D mode. As an example, in [118], clustering segregates nodes with similar behavior (i.e., traffic characteristics and network resource requirements) into logical groups to enable spatial reuse of frequency bands in order to improve spectral efficiency and reduce network congestion [21, 118].

2.4. What are the costs of clustering?

Cluster maintenance (or re-clustering [119]) requires message exchange [40], and so it incurs clustering cost. During re-clustering, a new CH is re-elected, and each non-clustered node associates itself with the new or an existing CH to join the cluster. Not only does frequent re-clustering reduce network scalability and stability, as well as the capability of clustering to cater for the three main characteristics (i.e., ultra-densification, network heterogeneity, and high variability) (see Section 2.3), but also increases clustering cost and reduces network performance. Nevertheless, re-clusterings are inevitable due to the underlying dynamic network topology; for instance, a CM moves out of the coverage of a cluster [120, 121]. There are three main types of clustering costs. Firstly, *bandwidth wastage* due to explicit clustering message exchange among nodes (or clustering overhead) in order to exchange information for clustering purpose. Secondly, *computational cost* (or time complexity) due to the time incurred for re-clustering, specifically from the dissolution of a cluster until all non-clustered nodes are clustered again. Thirdly, *energy consumption* due to clustering message exchange among nodes.

3. Taxonomy of Clustering Attributes in 5G Networks

This section presents a taxonomy of clustering attributes in 5G networks, as shown in Fig. 3. The rest of this subsection explains the taxonomy.

3.1. Clustering objectives

There are *four* main clustering objectives in 5G networks:

O.1 *Enhancing cluster stability*. In 5G networks, cluster stability increases with cluster lifetime and reduces with

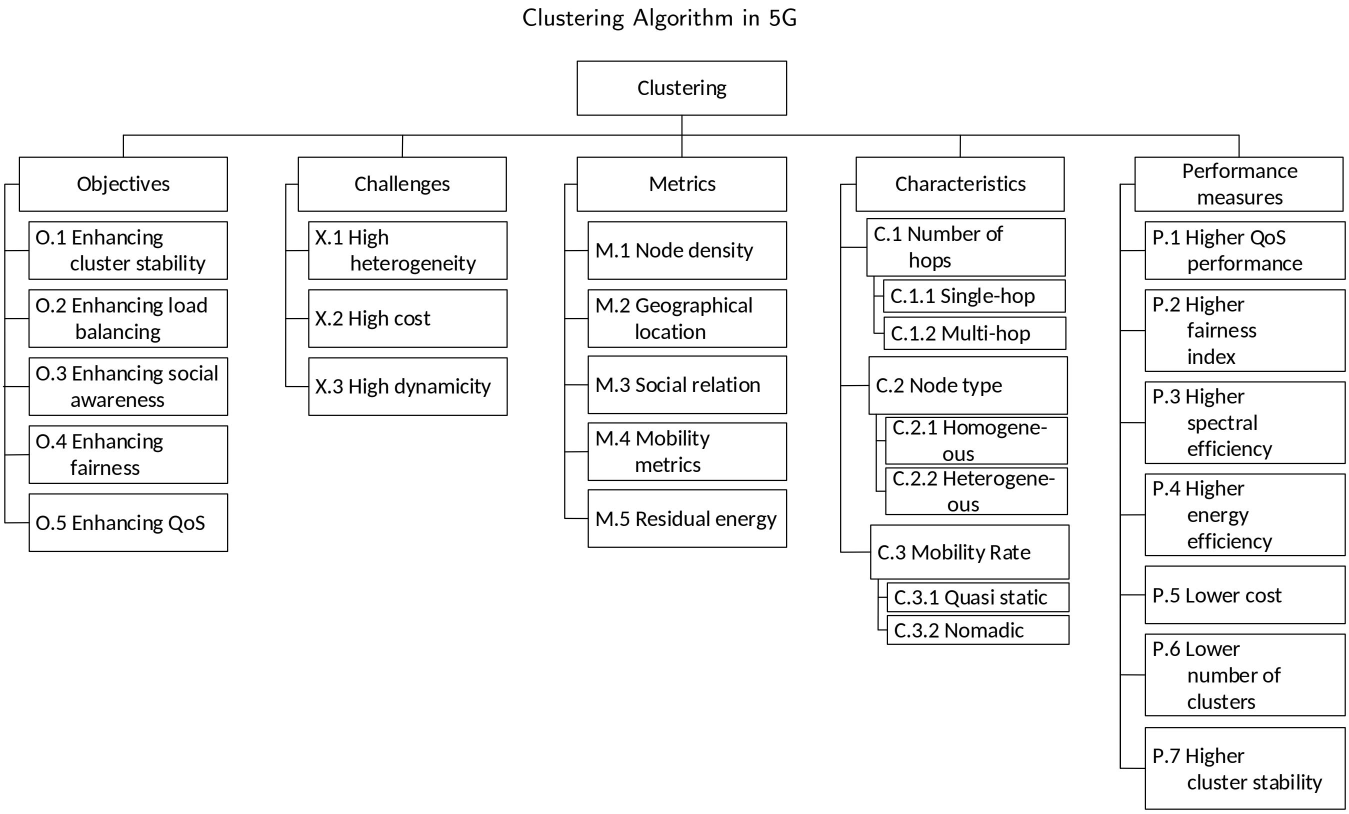

Figure 3: Taxonomy of clustering attributes in 5G.

the occurrence of re-clustering. Cluster stability is affected by: a) ultra-densification whereby higher node density increases the probability of a connected network remaining connected [122], and b) heterogeneity whereby a diverse range of network entities (i.e., UEs, network cells, and networks), which possess a diverse range of capabilities, overlap and increase interference [123]. The effects of a lower cluster stability are that it: a) increases packet loss rate, b) reduces cluster lifetime [124], c) increases clustering overhead or message exchange, and hence network congestion [103], and d) increases computational cost or time complexity [122]. Both (c) and (d) are incurred during re-clustering. 5G technologies improve cluster stability. As an example, D2D reduces access to network core leading to reduced network congestion and overhead, and so it increases cluster lifetime, and hence cluster stability. [125] [126].

O.2 *Enhancing load balancing*. In 5G networks, load balancing improves with efficient distribution of network traffic and lower traffic variability. Load balancing is affected by: a) ultra-densification whereby higher node density increases the amount of data generated, b) heterogeneity whereby a diverse range of network cells increases the number of handovers, handover overhead, and computational cost, and c) high variability whereby higher variability causes a sudden demand for significantly higher network capacity. The effects of lower load balancing are that it: a) reduces cluster lifetime, b) increases clustering and handover overheads, c) increases delay, and d) increases network congestion [127]. 5G technologies improve load balancing. As an example, SCs can be deployed to handle traffic load, which otherwise would have been handled by macrocells. Each SC, being a cluster, can adjust its cluster size, which can be limited by an upper and a lower bound, based on node density and traffic load to prevent bottleneck at the cluster level [128]. As another example, CoMP transfers traffic load from: a) clusters with higher traffic load to those with lower traffic load, and b) clusters with lesser resources (e.g., femtocells) to those with more resources (e.g., picocells and macrocells) [30].

O.3 *Enhancing social awareness*. In 5G networks, social awareness, which defines the relationship between network entities [129, 130] (e.g., UEs and BSs [131]) increases with social relation and reduces with increasing dynamicity. Social awareness is affected by: a) ultra-densification whereby higher node density provides a higher value of social relation and more options in selecting influential nodes, and b) heterogeneity whereby a diverse range of interests (e.g., using the same set of services) and behaviors pose a great challenge to segregate alike network entities into logical groups. The effects of a lower social awareness are that it: a) increases the occurrence of re-clustering, b) increases clustering overhead, and c) increases the number of clusters in the network. 5G technologies improve social awareness. As an example, in [132], clustering segregates nodes with similar behavior (i.e., geographical closeness) into logical groups, and uses D2D for intra-cluster communication without passing through a BS in order to reduce network congestion. Nevertheless, the use of physical attributes (e.g., geographical closeness) is insufficient as some nodes may not be supportive or cooperative, or even share interests, and so social parameters (e.g., social relation) are used to predict the next course of actions and behaviors of nodes in highly dynamic networks [126, 133].

O.4 *Enhancing fairness*. In 5G networks, fairness increases with equal distribution of network resources without starving low-priority traffic [134]. Fairness is affected by: a) ultra-densification whereby higher node density reduces fairness due to the different resource requirements from the nodes can cause starvation among low-priority applications, b) heterogeneity whereby network entities with low-quality resources (e.g., poor channel quality) receive more network resources compared to those with high-quality resources[135], and c) high variability, whereby higher variability causes a sudden demand for significantly higher network capacity [136], compounding the resource scarcity issue and affecting the fairness of resource distribution. In addition, fairness is affected by the accuracy of the definition of network requirements (particularly social attributes such as user satisfaction), and the role of the network entities in a cluster (e.g., CH and CG incur higher energy consumption and so they have shorter node lifetime). The effects of a lower fairness are that it: a) reduce user satisfaction, b) reduce fairness index, and c) reduces cluster lifetime. 5G technology improve fairness. As an example, D2D increases fairness of resource distribution in which a nodes with good channel quality has more communication with BS on behalf of the entire cluster [104, 136].

O.5 *Enhancing QoS*. In 5G networks, QoS increases with cluster lifetime, fair distribution of network traffic and resources, as well as reduces with computational cost and the occurrence of re-clustering. QoS is affected by: a) ultra-densification whereby higher node density increases demand for network resources, b) heterogeneity whereby a diverse range of network cells increases interference and handover, and c) high variability whereby a sudden demand for significantly higher network capacity is required to maintain high QoS. The effects of a lower QoS are that it: a) reduces throughput, b) increases end-to-end delay, c) reduces cluster lifetime, and d) increases packet loss rate. 5G technologies improve QoS. D2D communication reduces interference. As an example, in [137, 138], a BS imposes a reuse price to each subchannel. Since a D2D transmission can cause interference to other transmissions, the price increases with more D2D transmission in a particular subchannel, and a higher price represents a higher resource reuse (i.e., the reuse of subchannels). The resource usage of high power nodes (i-e., UEs and BSs) and low power nodes (i-e., can be UEs and BSs) should be tightly coordinated to realize the maximum capacity and coverage benefits to achieve QoS. Hence, the reuse price is effective in controlling the interference level. D2D node pairs must compete with each other to adjust its transmission power in order to minimize its own price for a given reuse price vector broadcast by the BS, while achieving a better QoS [139].

3.2. Clustering challenges

There are *three* main clustering challenges that must be addressed during cluster formation and maintenance (or re-clustering) in 5G networks:

X.1 *High heterogeneity*, which is an intrinsic characteristic of 5G, poses difficulties in segregating nodes with different nature of heterogeneity, including interests or behaviors, in a network into logical groups [140, 141]. Resources (e.g., frequency band, power, and SCs) are distributed on the basis of resource requirements among heterogeneous nodes in order to fulfill a diverse range of resource requirements. High heterogeneity can be addressed by deploying SCs and promoting collaboration among them to reduce interference among SCs in order to maximize throughput [116].

X.2 *High cost*, which is incurred when re-clusterings become more frequent, poses difficulties in improving the efficiency of clustering and resource utilization. The costs are clustering overhead (i.e., due to clustering message exchange), computational cost (i.e., due to the time incurred for re-clustering), and energy consumption (i.e., due to increased clustering overhead or message exchange). Higher energy consumption can lead to shorter cluster lifetime causing partitioned network topology [142]. More descriptions about cost are presented in Section 2.4. High cost can be addressed by D2D communication as it reduces message exchange between UEs and BSs [116], and reduces unnecessary involvement from the core network, leading to better resource utilization at BSs [100].

X.3 *High dynamicity*, which is caused by nodal mobility, poses difficulties in maintaining a cluster structure, as well as maintaining operation during network failure (or being fault tolerance). A CM can maintain its status quo, join another cluster, or leave its existing cluster in a highly dynamic network with frequent network topological variations, resulting in shorter cluster lifetime and frequent re-clustering. High dynamicity can be addressed by exploiting various mobility patterns of nodes in clustering [143]. Fault tolerance improves with the capability in maintaining operation during re-clustering as a result of high dynamicity [144].

3.3. Clustering metrics

There are *five* main clustering metrics used for cluster formation and maintenance (or re-clustering) in 5G networks:

M.1 *Node density* represents the number of nodes per unit area, or per coverage area, and it reflects the number of neighboring nodes of a node and the traffic load of a cluster. In 5G networks, node density is high due to ultra-densification. In clustering, this metric affects resource allocation and the amount of intra-cluster communications. Using this clustering metric helps to achieve two main objectives in 5G. Firstly, it achieves the objective of enhancing cluster stability (O.1) because higher node density increases the probability of a connection exists between a CH and a CM. Secondly, it achieves the objective of enhancing load balancing (O.2) because traffic load can be transferred from clusters with higher traffic load to those with lower traffic load. This clustering metric has been used in [124, 132, 129, 30].

M.2 *Geographical location* represents the physical location of a node in a network. Using this clustering metric helps to achieve two main objectives in 5G. Firstly, it achieves the objective of enhancing cluster stability (O.1) because physically nearby nodes can form a cluster to reduce intra-cluster distance between a CH and its CMs, which helps to prolong the lifetime of the CH and the cluster, leading to reduced clustering costs, including clustering overhead, computational cost, and energy consumption. Secondly, it achieves the objective of enhancing social awareness (O.3) because physically nearby nodes tend to share similar interest. This clustering metric has been used in [126, 134, 132, 122, 129, 145, 131, 146, 124, 133] .

M.3 *Social relation* represents the strength of the relationship among nodes in terms of interests or behaviors. Examples of the clustering metrics are file centrality, degree centrality, closeness centrality, betweenness centrality, and bridging centrality [147]. In 5G networks, nodes with higher file centrality can quickly disseminate information [148, 149] even under ultra-dense and heterogeneous networks. Using this clustering metric helps to achieve two main objectives in 5G. Firstly, it achieves the objective of enhancing social awareness (O.3) because alike nodes form clusters; for instance, nodes with interests towards a specific multimedia content form a cluster and share the content. Secondly, it achieves the objective of enhancing fairness (O.4) since the actual demand for a resource of a cluster can be identified based on social relation. This clustering metric has been used in [132, 133, 131, 129].

M.4 *Mobility metrics* represents mobility characteristics, which can cause a CM to maintain its status quo, join another cluster, or leave its existing cluster in a highly dynamic network with frequent network topological variations. In 5G networks, mobility is higher due to higher dynamicity. Using this clustering metric helps to achieve two main objectives in 5G. Firstly, it achieves the objective of enhancing social awareness (O.3) because mobile nodes with similar mobility patterns (e.g., similar AoA) form a cluster to prolong the lifetime of a CH and a cluster. Secondly, it achieves the objective of enhancing cluster stability (O.1) because of the prolonged lifetime. This clustering metric has been used in [126, 150, 131].

M.5 *Residual energy* represents the remaining energy level of a node. In 5G networks, higher residual energy increases the network lifetime by increasing the probability of a connection between nodes remaining connected. Using this clustering metric helps to prolong the network lifetime and reduces the occurence of re-clustering, and so it achieves two main objectives in 5G. Firstly, it achieves the objective of enhancing cluster stability (O.1) because a node with higher residual energy is selected as a CH [129]. Secondly, it achieves the objective of enhancing QoS (O.5) because CH with higher residual energy reduces the need to transfer its responsibility to other nodes, leading to reduced packet loss rate. This metric has been used in [124, 103, 129].

3.4. Clustering characteristics

There are *three* main clustering characteristics in 5G networks:

C.1 *Number of hops* (or intra-cluster distance) represents the maximum number of hops between a CH and its CMs. There are two possible options as follows:

C.1.1 *Single-hop* clusters allow a CM to communicate with its CH only, which helps to increase cluster stability and reduce delay. In 5G, SCs can form single-hop clusters and use mmWave to provide short-range communication in order to provide higher data rate. However, single-hop clusters increases clustering overhead due to frequent handover in ultra-dense networks [151]. Single-hop clusters have been investigated in [132, 126, 131, 129, 136, 134, 133, 127, 128, 30].

C.1.2 *Multi-hop* clusters allow a CM to communicate with its CH in multiple hops, which helps to increase network scalability (i.e., reduce the number of clusters in the network). In 5G, massive MIMO at BS provides long-range communication with high data rate, and multi-hop clusters provide larger network coverage. However, multi-hop clusters increases clustering overhead due to explicit clustering message exchange between CHs, BSs, and UEs [124]. Multi-hop clusters have been investigated in [122, 124, 103].

C.2 *Node Type* represents the type of nodes in a cluster. There are two possible options as follows:

C.2.1 *Homogeneous* clusters contain network entities, such as network cells (e.g., macrocells and SCs) and UEs (e.g., smart gadgets and sensors), with similar capabilities (e.g., computational and storage capacities), which helps to increase cluster stability (i.e., connectivity due to alike nature). In 5G, D2D among homogeneous nodes can prolong the cluster lifetime because network congestion among BSs and UEs is reduced.

C.2.2 *Heterogeneous* cluster contain network entities with distinctive capabilities, which can introduces constraints (e.g., data rate and transmission range) on the clustering process. Subsequently, heterogeneous nodes can reduce network lifetime because of interference among network entities. In 5G, SCs can collaborate among themselves to reduce interference among heterogeneous nodes in order to maximize throughput [116]. More descriptions about heterogeneity are presented in Section 2.3.

C.3 *Mobility* represents the nature of movement of a cluster. There are two possible options as follows:

C.3.1 *Quasi static* clusters contain network entities (e.g., network cells) that are fixed or have low mobility (e.g., the movement of UEs within a specific region). These clusters are easy to maintain as compared to those with high mobility.

C.3.2 *Mobile/ nomadic* clusters contain network entities (e.g., vehicular nodes) that have high mobility. These clusters are more difficult to maintain as compared to those with low mobility. In 5G, high dynamicity is one of the main characteristics, and it can be addressed using mobility models in clustering.

3.5. Performance measures

There are *six* main performance measures for clustering in 5G networks. Table 2 shows the details of performance measures and performance metrics. Cluster lifetime has been investigated with respect to energy efficiency and cluster stability as shown in the table.

P.1 *Higher QoS performance* increases packet delivery rate (or reduces packet loss rate) from a source node to a destination node, and reduces end-to-end delay of delay-sensitive packets, as a result of more robust and reliable communication. So, it aims to achieve the objective of enhancing QoS (O.5) in resource distribution among network entities, while addressing the challenge of high cost (X.2) incurred during re-transmission and re-clustering, particularly clustering overhead, computational cost, and energy consumption.

P.2 *Higher fairness index* increases fairness of resource distribution. So, it aims to achieve the objective of enhancing fairness (O.4) in resource distribution among network entities, and enhancing QoS (O.5), while addressing the challenge of high heterogeneity (X.1) that causes unequal resource distribution among the clusters, and high cost (X.2), particularly clustering overhead.

P.3 *Higher spectral efficiency* optimizes resource utilization, particularly frequency bands and network capacity. So, it aims to achieve the objectives of enhancing fairness (O.4) in resource distribution among network entities, and enhancing QoS (O.5), while addressing the challenges of high cost (X.2), particularly clustering overhead, computational cost, and energy consumption.

P.4 *Higher energy efficiency* reduces node failure, and so it: a) increases the lifetime of a cluster and a CH, and reduces the occurrence of re-clustering (i.e., CH election and node joining), and b) minimizes network partition (or network topology disconnection) [152]. So, it aims to achieve the objectives of enhancing cluster stability (O.1), social awareness (O.2) among network entities, and load balancing (O.3) among clusters, while addressing the challenges of high cost (X.2), particularly clustering overhead, computational cost, and energy consumption, and high dynamicity (X.3) that creates network partition and causes re-clustering.

P.5 *Lower cost* reduces clustering overhead, computational cost, and energy consumption. This can be achieved by reducing re-clustering and unnecessary message exchange. So, it aims to achieve the objective of enhancing cluster stability (O.1), while addressing the challenge of high cost (X.2).

P.6 *Lower number of clusters* increases the coverage of a cluster to incorporate more CMs. In other words, cluster size increases, and this reduces inter- and intra-cluster communication [122, 128]. So, it aims to achieve the objectives of enhancing cluster stability (O.1), social awareness (O.2) among network entities, and load balancing (O.3) among clusters, while addressing the challenge of high cost (X.2) incurred during re-transmission, particularly clustering overhead, computational cost, and energy consumption.

P.7 *Higher cluster stability* increases cluster lifetime, and so it reduces the effects of: a) node failure, b) re-clustering, and c) network partitioning. So it aims to achieve the objective of enhancing cluster stability (O.1) among clusters, while addressing the challenge of high cost (X.2) incurred during re-clustering, particularly clustering overhead and computational cost.

4. Clustering Framework

A clustering framework consists of abstract-level steps from non-clustered to clustered networks with each node being a CH, a CM, or a CG. There are *four* main stages as summarized in Table 3.

During the *first* stage, non-clustered nodes exchange messages among themselves, or receive broadcasts from BS, to gather clustering information about: a) neighboring nodes, and b) network topology. Neighboring nodes' information, such as social parameters (e.g., file centrality, degree central-

Table 2
Performance measures, performance metrics, and types of measurements

No.	Performance measure	Performance metric	Unit	Reference	Description of performance metric	Types of measurements	Objectives	Challenges
P1	Higher QoS performance	End-to-end delay	seconds	[122, 128, 132, 131, 129]	Time taken for a packet to be transmitted from source to destination across a network.	Mean	O.5	X.2
		Throughput	bits per second	[103, 136]	The amount of data being transferred from one location to another in a given amount of time.			
		Packet loss	number of packets	[103, 126]	Packet loss is the failure of transmitted packets to arrive at their destination.			
P2	Higher fairness index	Resource distribution	bits per second	[132, 136, 134]	Distribution of resources to maximize the network performance (e.g., bandwidth).	Probability distribution	O.4, O.5	X.1, X.2
P3	Higher spectral efficiency	Network capacity	bits per second	[124, 156, 127, 128, 132, 133, 134]	The amount of traffic that a network can process at any given time.	Maximum		
P4	Higher energy efficiency	Network lifetime	seconds	[124, 156, 127, 136, 134]	The time at which the first network node runs out of energy to send a packet.	Maximum & minimum	O.1, O.2, O.3	X.2, X.3
		Cluster lifetime	seconds	[129, 132, 133]	The time in which either dissolution or merging of cluster(s) occurs.			
P5	Lower cost	Computational cost	seconds	[103, 131]	The amount of resources (e.g., bandwidth) required for clustering.	Mean	O.1	X.2
		Energy consumption	joule	[122, 136]	The amount of energy consumption for clustering.			
P6	Lower number of clusters	Cluster size	number of entities	[124, 156, 122, 126, 128, 129]	The number of cluster entities (i.e., CM, CH, and CG) in a cluster.	Mean	O.1, O.2, O.3	X.2
P7	Higher cluster stability	Cluster lifetime	seconds	[124, 156, 122, 103, 126, 127, 132]	The time in which either dissolution or merging of cluster(s) occurs.	Maximum & minimum	O.1	X.2

Table 3
Stages of clustering algorithms

Stage	Details	Outcomes
First	Non-clustered nodes exchange messages among themselves or receive broadcasts from BSs	NSs and network topologies are formed
Second	Non-clustered nodes elect CHs	CHs are elected
Third	Non-clustered nodes join clusters	Clusters are formed
Fourth	Clusters form new clusters after the dissolution of existing clusters (or re-clustering)	New clusters are formed

ity, and closeness centrality) and physical parameters (e.g., geographical location, AoA, relative speed, and received signal strength (RSS)), can be gathered to generate a neighbor set (NS), which can be managed by BSs. A BS can advertise NS to nodes in its coverage. Using the neighboring nodes' information, each node forms a network topology. In [153], nodes exchange their geographical locations (i.e., x- and y-coordinates), and calculate Euclidean distance with their neighbor nodes. The Euclidean distance is used to determine cluster boundary so that physically nearby nodes can form a cluster. CMs of a cluster can elect a CH among themselves.

During the *second* stage, non-clustered nodes elect CHs to achieve clustering objectives (see Section 3.1) and address the clustering challenges (see Section 3.2) using clustering metrics (see Section 3.3) in order to form clusters with different characteristics (see Section 3.4). As an example, a clustering scheme is proposed to enhance social awareness (O.3) [132]. In [129], a CH is elected using a social relation metric (M.3) so that it has a high social influence among the non-clustered nodes, whereby data can be disseminated in the cluster quickly to reduce end-to-end delay. The proposed scheme has shown to increase cluster lifetime. As another example, a clustering scheme is proposed to increase cluster lifetime in order to enhance cluster stability (O.1). In [124], a CH is elected using a residual energy metric (M.5) so that it has a higher residual energy among non-clustered nodes, subsequently it broadcasts an energy threshold and non-clustered nodes that fulfill the threshold joins the cluster as CMs [154, 155]. The proposed scheme has shown to increase CH lifetime.

During the *third* stage, there are minor changes to the underlying cluster structure without affecting the entire cluster, such as a small number of non-clustered nodes joining a cluster, or a small number of clustered nodes leaving a cluster. The cluster size (or the number of CMs in a cluster) can be limited by thresholds. Two examples of clustering schemes are proposed to enhance load balancing (O.2) [128] [127]. In [127], the threshold is based on a social relation metric (M.3) whereby a maximum number of CMs with similar interests (i.e., multimedia content) can join a cluster. In [128], the thresholds are based on two factors, namely network capacity (i.e., signal-to-noise ratio) and traffic load (i.e., the number of message exchanges between CHs and CMs, and among CMs, via D2D communication). The node density metric (M.1) can also indicate the traffic load of a cluster. CMs can leave clusters with high load and join clusters with low load. Small clusters with similar interests can merge to form a single cluster [30]. The proposed scheme has shown to improve energy efficiency (P.4) and reduce the number of clusters in the network (P.6).

During the *fourth* stage, cluster maintenance (or re-clustering) occurs when there are major changes to the underlying cluster structure which are affecting the entire cluster, such as a large number of non-clustered nodes joining a cluster, or a large number of clustered nodes leaving a cluster. For instance, cluster maintenance enables a cluster to form a new cluster after the dissolution of an existing cluster. Re-clustering is essential: a) to manage a large number of nodes joining and leaving, which affects the entire cluster structure, in the dynamic 5G networks [126], b) to rotate the role of network entities in a cluster (e.g., CH and CG incur higher energy consumption, and so these roles are rotated among nodes in a cluster), and c) to adjust the cluster size in order to reduce traffic load of a cluster with high load for achieving load balancing [128]. Nevertheless, re-clustering in static networks, such as re-clustering of BSs, may not be necessary [132].

5. Clustering Algorithms in 5G Networks

This section presents a survey on clustering algorithms in 5G networks. The clustering algorithms are segregated into four categories based on the clustering objectives (see Section 3.1), namely enhancing cluster stability (O.1), enhancing load balancing (O.2), enhancing social awareness (O.3), and enhancing fairness (O.4). In addition to achieving the aforementioned clustering objectives, some of these clustering algorithms achieve the objective of enhancing QoS (O.5). The framework of each clustering scheme, which is based on the stages presented in Table 3, is summarized and presented in a table for ease of comparison among the clustering schemes. Table 4 shows the summary of the clustering schemes presented in this section.

5.1. Clustering Algorithms for Enhancing Cluster Stability

This section presents *five* clustering algorithms for achieving the objective of enhancing cluster stability.

Table 4
Summary of objectives, challenges, metrics, characteristics, and performance measures of clustering schemes for 5G networks

Reference	Year	Objectives					Challenges			Metrics					Characteristics						Performance measures						
		O.1 Enhancing cluster stability	O.2 Enhancing load balancing	O.3 Enhancing social awareness	O.4 Enhancing fairness	O.5 Enhancing QoS	X.1 High heterogeneity	X.2 High cost	X.3 High dynamicity	M.1 Node density	M.2 Geographical location	M.3 Social relation	M.4 Mobility metrics	M.5 Energy residual	C.1.1 Single-hop	C.1.2 Multi-hop	C.2.1 Homogeneous	C.2.2 Heterogeneous	C.3.1 Quasi static	C.3.2 Mobile	P.1 Higher QoS performance	P.2 Higher fairness index	P.3 Higher spectral efficiency	P.4 Higher energy efficiency	P.5 Lower cost	P.6 Lower number of clusters	P.7 Higher cluster stability
Lina *et al.* [124]	2017	×					×	×		×				×	×	×	×		×				×	×		×	×
Khan *et al.* [122]	2018	×					×		×	×	×		×	×		×	×			×	×					×	×
Sanaa *et al.* [103]	2018	×				×	×	×	×	×			×	×		×		×		×	×				×		×
Duan *et al.* [126]	2017	×				×		×	×		×		×		×		×			×	×					×	×
Ying *et al.* [127]	2017	×						×					×		×		×		×				×	×	×		×
Ali *et al.* [128]	2016		×					×		×				×	×			×		×	×		×			×	
Zhang *et al.* [133]	2016	×		×				×			×	×			×			×		×	×	×	×	×			×
Qi *et al.* [131]	2018			×					×		×	×			×			×		×	×				×		
Tulu *et al.* [129]	2018			×		×	×			×	×	×		×	×			×		×	×		×			×	
Zhao *et al.* [132]	2017			×				×				×			×			×		×			×	×			
Asadi *et al.* [136]	2016				×			×		×	×			×	×			×		×	×	×		×	×		
Huang *et al.* [134]	2017				×			×			×				×			×		×		×	×	×			

5.1.1. Reducing CH election time based on energy consumption

Lina *et al.* [124, 156] propose a clustering scheme that minimizes energy consumption during CH election to prolong network lifetime in order to increase cluster stability. Smart Balanced Energy Efficiency for Multi-hop clustering scheme, or Smart-BEE(M), uses MIMO, for homogeneous (C.2.1) and quasi static (C.3.1) nodes to form single- (C.1.1) and multi-hop (C.1.2) clusters. The main objective is to enhance cluster stability (O.1) among network entities. The clustering scheme addresses two challenges, namely high heterogeneity (X.1) whereby a diverse range of applications (e.g., low- and high-bandwidth transmission) and highly dynamic IoT devices communicate in a real-time manner, and high cost (X.2) whereby there is a high energy consumption. In general, Smart-BEE(M) increases network lifetime by reducing energy consumption and distance of inter-cluster communication. Smart-BEE(M) covers three of the four stages in the clustering framework as summarized in Table 5. During the *first stage*, non-clustered nodes generate NS using physical parameters. During the *second stage*, non-clustered nodes elect CHs using the residual energy metric (M.5) to extent network lifetime and coverage [157, 158] in order to enhance cluster stability and network scalability. Consider an out-of-range node that declares itself as a temporary CH. It is assigned a C_{prob} value that represents the probability of the node being selected as a CH. The C_{prob} value doubles every iteration until it reaches 1, so a C_{prob} value initialized with a value closer to 1 can reduce energy consumption during CH election. When $C_{prob} = 1$, the node either becomes a MN of a new cluster in its neighborhood, or it declares itself as a CH given that it has sufficient residual energy. During the *third stage*, a non-clustered node joins a cluster: a) if it has sufficient residual energy based on a residual energy metric (M.5), and b) if it is closest to its CH based on a node density metric (M.1). During data communication, a CH, which is equipped with multiple interfaces and uses MIMO, select the right interfaces for communication (e.g., video transmission uses high-bandwidth interface, and data transmission uses low-bandwidth interface, to reduce energy consumption), and can operate at several channels simultaneously. The CH can compress data and send it to BS. Smart-BEE(M) has shown to improve spectral efficiency (P.3), energy efficiency (P.4), reduces the number of clusters in the network (P.6), and improve cluster stability (P.7).

Table 5
Stages of clustering algorithm [124]

Stage	Details and Outcomes
First	Non-clustered nodes generate NSs and network topology by exchanging messages among themselves using physical parameters.
Second	Non-clustered nodes elect CHs using residual energy to extend the network lifetime.
Third	Non-clustered nodes, which have sufficient residual energy and are geographically closer to the respective CHs, join clusters.

5.1.2. Increasing intra-cluster connectivity based on node density and energy consumption

Khan *et al.* [122] propose a clustering scheme that elects reliable CH to prolong network connectivity in order to increase cluster stability. This scheme aims to achieve reliable communication in vehicular networks for homogeneous (C.2.1) and mobile (C.3.2) UEs to form multi-hop (C.1.2) clusters in V2X networks. The main objective is to enhance cluster stability (O.1) among network entities. The clustering scheme addresses two challenges, namely high heterogeneity (X.1) whereby there are UEs with different capabilities, and high dynamicity (X.3) whereby there are UEs with high mobility. In this scheme, UEs form clusters among themselves and connect to femtocell BSs. Long-range communication (e.g., between a CH and a femtocell BS, and between a femtocell BS and a macrocell BS) uses 5G, whereas short-range communication (e.g., CH-CH and CH-CM) uses wireless local area network (e.g., IEEE 802.11p). Each macrocell BS connects with a SC (i.e., a femtocell BS) within its coverage. The proposed clustering scheme covers three of the four stages in the clustering framework as summarized in Table 6. During the *first stage*, non-clustered UEs generate NS using physical parameters (e.g., UE ID, as well as *x*- and *y*-coordinates). During the *second stage*, non-clustered UEs elect CHs based on the mobility metric (M.4) whereby a prospective CH has the smallest average distance to neighboring UEs, and residual energy metric (M.5), in order to increase network connectivity for improving cluster stability and lifetime. During the *third stage*, a non-clustered UE joins a cluster: a) if it is closest to its neighboring UEs, which are clustered, based on a node density metric (M.1), and b) if it is closest to its CH based on a geographical location (M.2). A clustered UE can leave a cluster after: a) it broadcasts a beacon message to its entire cluster, and b) its counter expires after 180 seconds, causing other UEs in the cluster to remove its entry. The centralized controller coordinates with clusters to execute the first to third stages again whenever important UEs (e.g., CH) leave the cluster. This scheme has shown to improve QoS performance (P.1) (i.e., lower delay and cost during re-transmission and re-clustering, particularly clustering overhead, and energy consumption), reduce the number of clusters in the network (P.6) [159, 160, 161, 162], and improve cluster stability (P.7).

5.1.3. Selecting reliable nodes as potential CHs

Sanaa *et al.* [103] propose a clustering scheme that identifies potential CHs among UEs, and elects CHs that prolong long-term network connectivity in order to increase cluster stability. This scheme aims to form stable clusters with D2D communication for heterogeneous (C.2.2) and mobile (C.3.2) UEs to form multi-hop (C.1.2) clusters. The main objectives are to enhance cluster stability (O.1), and enhances QoS

Table 6
Stages of clustering algorithm [122]

Stage	Details and Outcomes
First	Non-clustered nodes generate NSs and network topology by exchanging messages among themselves using physical parameters (i.e., node ID and location information).
Second	Non-clustered nodes elect CHs using the smallest average distance to neighboring nodes in order to increase network connectivity.
Third	Non-clustered nodes, which are geographical closer to the respective CHs, join clusters.

Table 7
Stages of clustering algorithm [103]

Stage	Details and Outcomes
First	Non-clustered nodes generate NSs and network topology by exchanging messages among themselves using physical parameters.
Second	Non-clustered nodes elect potential CHs using historical mobility pattern and residual energy to increase cluster stability. In absence of potential CHs, the centralized controller act as the CH.
Third	Non-clustered nodes, which are closer to clustered neighboring nodes, join clusters.

(O.5) among network entities. The clustering scheme addresses three challenges, namely high heterogeneity (X.1) because different UEs have different transmission ranges, high cost (X.2) whereby there is a high energy consumption and overhead due to re-transmission [163, 164], and high dynamicity (X.3) due to high mobility of UEs. In this scheme, reliable UEs are identified based on mobility, and clusters are formed among the reliable UEs, contributing to lower data re-transmission, and hence lower energy consumption. The proposed scheme covers three of the four stages in the clustering framework as summarized in Table 7. During the *first stage*, non-clustered UEs generate NS using physical parameters to develop a network topology consisting the UEs in the network. During the *second stage*, non-clustered UEs elect CHs based on mobility metrics (M.4) and residual energy (M.5) whereby UEs with predictable mobility patterns (i.e., based on historical mobility pattern) and high residual energy are marked as potential CHs. The non-clustered UEs select centralized controller as the CH in the absence of potential CHs [165, 166]. During the *third stage*, a non-clustered UE joins a cluster if it is closest to its neighboring UEs, which are clustered, based on a node density metric (M.1) for longer-term connectivity. This scheme has shown to improve QoS performance (P.1) (i.e., higher throughput and lower packet loss), reduce cost (P.5), and improve cluster stability (P.7).

5.1.4. Electing backup CHs

Duan *et al.* [126] propose a clustering scheme that elects backup CHs to prolong network connectivity in order to increase cluster stability and network lifetime. This scheme is designed for homogeneous (C.2.1) and mobile (C.3.2) nodes to form single-hop (C.1.1) clusters. The main objective is to enhance cluster stability (O.1) and QoS (O.5) among network entities. The clustering scheme addresses two challenges, namely high cost (X.2) and complexity whereby there are a high amount of overhead and a massive amount of data generated by a large number of nodes [167, 168], and high dynamicity (X.3) whereby there are UEs with high mobility [169]. In this scheme, UEs form clusters among themselves and connect to BSs. The CHs can aggregate and forward traffic pattern information to BSs in a single hop, and subsequently the CHs forward the information to a centralized controller. This enables the centralized controller to predict real-time traffic patterns based on AoA and RSS. Long-range communication (e.g., between a CH and a BS) uses 5G to exchange clustering messages, whereas short-range communication (e.g., among a CH and its CMs in intra-cluster communication) uses wireless local area network (e.g., IEEE 802.11p) [77] to exchange control messages for updates (e.g., geographical locations). This scheme covers all of the four stages in the clustering framework as summarized in Table 8. During the *first stage*, non-clustered UEs generate NS using physical parameters (e.g., geographical location) and form clusters. During the *second stage*, clustered UEs elect CHs, as well as backup CHs [170], based on the mobility metric (M.4) whereby the speed of the CH (backup CH) is closest (the second closest) to the average speed of the cluster in order to increase the lifetime of a CH and a cluster. Backup CHs are necessary to prevent the over dependency on a single node (i.e., the CH). For instance, the backup CH becomes the CH of a cluster when the existing CH leaves the cluster. During the *third stage*, a non-clustered UE joins a cluster if both are located at the same region and are moving in the same direction based on a geographical location (M.2) and a mobility (M.4) metrics, respectively, which increases the connection time of a link in a node pair. During the *fourth stage*, a CH performs cluster maintenance (or re-clustering) by removing some of the UEs from its cluster whenever its network capacity fails to cater for the traffic load at the CH [150]. During data communication, a UE, which is equipped with multiple interfaces and uses MIMO connects to CHs or CMs via directional links. This scheme has shown to improve QoS performance (P.1) (i.e., lower packet loss), reduce the number of clusters in the network (P.6), and improve cluster stability (P.7).

5.1.5. Merging clusters based the number of handovers of UEs between two SCs

Ying *et al.* [127] propose a clustering scheme that merges two clusters to prolong network connectivity. This helps to reduce the number of handovers, and hence energy consumption, contributing to a longer network lifetime, and hence improving cluster stability. This scheme is designed for homogeneous (C.2.1) and quasi static (C.3.1) nodes (i.e., a centralized controller and UEs) to form single-hop (C.1.1)

Table 8
Stages of clustering algorithm [126]

Stage	Details and Outcomes
First	Non-clustered nodes generate NSs and network topology by exchanging messages among themselves using geographical location.
Second	Non-clustered nodes elect CH and backup CH using the average speed of the cluster to increase the lifetime of CH and avoid dependency on a single node.
Third	Non-clustered nodes, which are located in same geographical location and moving in the same direction, join clusters.
Fourth	Clusters form new clusters after removing nodes that fail to handle traffic load.

Table 9
Stages of clustering algorithm [127]

Stage	Details and Outcomes
First	Non-clustered nodes generate NS, which is depicted in the form of graph, that contain nodes linked with values that increase with the number of handovers.
Second	Centralized controller act as the CH.
Third	Non-clustered nodes, which are connected with CH, join clusters.
Fourth	Femtocells are merged based on the number of handovers.

clusters. Each cluster consists of a centralized controller as the CH, and UEs as the CMs. Hence, a single CH manages multiple clusters. The main objective is to enhance cluster stability (O.1) among network entities. The clustering scheme addresses the challenge of high cost (X.2) caused by clustering overhead and energy consumption. This is because, as the number of SCs (i.e., femtocells) increases, the inter-cell interference level and the number of handovers of mobile UEs from one femtocell to another increases [171]. The proposed clustering scheme covers all the four stages in the clustering framework as summarized in Table 9. During the *first stage*, non-clustered nodes generate a NS or a network topology, which is represented by a graph $G = [F, H]$, where: a) F is a set of nodes that represents the physical parameters (e.g., geographical location) of UEs and femtocells, and b) H is a set of links that represents a mobility metric (M.4) that increases with the number of handovers of UEs. During the *second* and *third* stages, the centralized controller acts as a CH and the UEs become the CMs based on the network topology formed in the first stage. During the *fourth stage*, when the number of handovers of UEs from a particular femtocell to a neighboring femtocell is greater than a threshold, these two femtocells are merged to reduce interference and handover between the femtocells. This scheme has shown to improve spectral efficiency (P.3), energy efficiency (P.4), cost (P.5), and improve cluster stability (P.7).

5.2. Clustering Algorithms for Enhancing Load Balancing

This section presents *a* clustering algorithm for achieving the objective of enhancing load balancing.

5.2.1. *Adjusting cluster size to transfer traffic load among SCs*

Ali *et al.* [128] propose a clustering scheme that adjusts its cluster size (i.e., the number of nodes in a cluster) to self-organize traffic load in order to achieve load balancing in 5G networks. This scheme is designed for heterogeneous (C.2.2) and mobile (C.3.2) nodes to form single-hop (C.1.1) clusters. The main objective is to enhance load balancing (O.2). The clustering scheme addresses the challenge of high cost (X.2) caused by clustering overhead and energy consumption. In general, this scheme adjusts the cluster size dynamically based on the traffic load of a cluster, and can transfer traffic load from SCs with higher traffic load to those with lower traffic load. In CoMP, larger cluster size increases the CoMP gain (e.g., which is based on spectral efficiency, energy efficiency, load balancing, and fairness of resource distribution) which increases available resources to cater for high traffic load; however, too large a cluster can increase clustering overhead and energy consumption [82, 172]. Therefore, the cluster size must be adjusted accordingly to achieve load balancing. This scheme covers all of the four stages in the clustering framework as summarized in Table 10. During the *first stage*, non-clustered nodes generate NS using physical parameters (e.g., nodes' geographical location) to develop a network topology. During the *second stage*, non-clustered nodes elect CHs using the residual energy metric (M.5) whereby a prospective CH has a higher residual energy. During the *third stage*, a non-clustered node joins a cluster if it is closest to the CH based on a node density metric (M.1). The traffic load of a SC is monitored in each iteration. During normal operation, the traffic load is below a pre-defined threshold. When the traffic load exceeds the pre-defined threshold, the cluster size increases in order to increase the CoMP gain leading to more available resources. During the *fourth stage*, when a cluster becomes too large causing high clustering overhead and energy consumption, cluster maintenance (or re-clustering) is initiated to form new clusters and the traffic load of SCs is equally distributed among the newly formed clusters in order to increase cluster lifetime. The re-clustering process is repeated among clusters that have exceeded the pre-defined threshold until the traffic load of all newly formed clusters are less than the threshold, which helps to transfer traffic load from SCs with higher traffic load to those with lower traffic load. This scheme has shown to improve QoS performance (P.1) and spectral efficiency (P.3), and reduce the number of clusters in the network (P.6).

Table 10
Stages of clustering algorithm [128]

Stage	Details and Outcomes
First	Non-clustered nodes generate NSs and network topology by exchanging messages among themselves using geographical location.
Second	Non-clustered nodes elect CHs using residual energy to extend the network lifetime.
Third	Non-clustered nodes, which are closer to clustered neighboring nodes, join clusters.
Fourth	Clusters form new clusters when traffic load is high, which causes high clustering overhead and energy consumption, in order to extend the cluster lifetime.

Table 11
Stages of clustering algorithm [133, 132, 131, 129]

Stage	Details and Outcomes
First	Non-clustered nodes generate NSs and network topology using social attributes, such as social connectivity [129], social interest [131], the total number of messages exchanged among themselves [132], and similar content usage [133].
Second	Non-clustered nodes elect CHs using the number of social connections [129], similar mobility [131], and social interaction based on content [132], as well as node's cache with high hit rate [133].
Third	Non-clustered nodes, which are closer to clustered neighboring nodes [129] [133] and CHs [131] [132], join clusters.

5.3. Clustering Algorithms for Enhancing Social Awareness

This section presents clustering algorithms for achieving the objective of enhancing social awareness.

5.3.1. Joining node based on social interests

In general, there are four kinds of clustering schemes [133, 132, 131, 129] that increase social awareness (e.g., UEs belong to the same community, or have similar interests) by identifying strong social relation among network entities. The clustering schemes share a similar clustering algorithm, and the main differences among them are the social attributes and clustering metrics used in the clustering algorithm. These schemes are designed for heterogeneous (C.2.2) and mobile (C.3.2) UEs to form single-hop (C.1.1) clusters. The main objective is to enhance social awareness (O.3), in addition to cluster stability (O.1) [132], and QoS (O.5) [129] among network entities. These clustering schemes address different challenges, namely: a) high dynamicity (X.3) whereby there are UEs with high mobility [131], b) high cost (X.2) (i.e., clustering overhead due to ultra-densification [133, 132]), and c) high heterogeneity (X.1) whereby there are diverse range of communities or interests in the network [129]. In these schemes, UEs form clusters among themselves and connect to macrocell BSs. Long-range communication (e.g., between a CH and a macrocell BS) uses 5G to exchange information about clusters, whereas short-range communication (e.g., CH-CM) uses wireless local area network (e.g., IEEE 802.11p) [77] to identify neighboring UEs and exchange clustering messages [174, 173]. These schemes cover three of the four stages in the clustering framework as summarized in Table 11. During the *first stage*, non-clustered UEs generate NS using social attributes (e.g., the number of message exchanges [132], social interest [131], social connectivity [129], and same content [133]). During the *second stage*, non-clustered UEs elect CHs based on social relation metrics (M.3), which represent the similarity of the community and interests of a UE with neighboring UEs in terms of: a) UEs' contents [132], b) UEs' mobility metric (M.3) (i.e., relative velocity and relative distance) [131], c) UEs' entropy of betweenness centrality (or the number of social connections between a UE and other UEs in the network) so that information can be disseminated in a cluster quickly since CH and CMs have strong social relation[129], and d) UEs' cache with high hit rate [133]. During the *third stage*, a non-clustered UE joins a cluster: a) if it is closest to its neighboring UEs, which are clustered, based on a node density metric (M.1) [129], b) if it is closest to its CH based on a geographical location metric (M.2) [132, 131, 129], so it leads to a prolonged network lifetime, and c) if it has a strong social relation with its neighboring UEs based on a social relation metric (M.3), which must fulfill a pre-defined threshold based on trust value [175]. The trust value increases with the number of message exchanges (i.e., higher number of message exchanges, or communications, increases the effectiveness in improving privacy and security, so it leads to a higher trust value)[132, 131]. The proposed schemes have shown to improve: a) QoS performance (P.1) (e.g., higher throughput, and the number of satisfied users, as well as lower delay and packet loss rate) [132, 131, 129], b) fairness index (P.2) [132], c) spectral efficiency (P.3) [132, 133], d) energy efficiency (P.4) (e.g., longer cluster lifetime) [132, 129, 133], e) cost (P.5) (i.e., clustering overhead) [131], f) number of clusters in the network (P.6) [129], and g) improve cluster stability (P.7).

5.4. Clustering Algorithms for Enhancing Fairness

This section presents *two* clustering algorithms for achieving the objective of enhancing fairness.

5.4.1. Joining node based on resource availability

Asadi *et al.* [136] propose a clustering scheme that achieves a fair distribution of resources by joining a UE with less resource to a cluster with more resource for the benefit of the UE, and vice-versa. This scheme is designed for heterogeneous (C.2.2) and mobile (C.3.2) UEs to form single-hop (C.1.1) clusters. The main objective is to enhance fairness (O.4) among network entities. The clustering scheme addresses the challenge of high cost (X.2) due to

Table 12
Stages of clustering algorithm [136]

Stage	Details and Outcomes
First	Non-clustered nodes generate NSs and network topology by exchanging messages among themselves using geographical location.
Second	Non-clustered nodes elect CHs using high residual energy to extend the network lifetime.
Third	Non-clustered nodes with access to high-quality channels join clustered nodes with low-quality channel access in order to achieve fairness.

clustering overhead caused by ultra-densification in 5G. In this scheme, UEs form clusters among themselves and connect to macrocell BSs. Long-range communication (e.g., between a CH and a macrocell BS) uses 5G to exchange information about clusters, whereas short-range communication (e.g., CH-CM) uses wireless local area network (e.g., IEEE 802.11ad) [176, 177]. The macrocell BS monitors clustering activities (i.e., CH election and node joining) in order to achieve fairness [101, 178, 179, 180]. The proposed clustering scheme covers three of the four stages in the clustering framework as summarized in Table 12. During the *first stage*, non-clustered UEs generate NS using physical parameters (i.e., geographical location). During the *second stage*, non-clustered UEs elect CHs using the residual energy metric (M.5) whereby a prospective CH has a higher residual energy. During the *third stage*, a non-clustered UE joins a cluster: a) if it is closest to its neighboring UEs, which are clustered, based on a node density metric (M.1), and b) if it is closest to its CH based on a geographical location (M.2). In addition, a non-clustered UE with access to high-quality channel joins a cluster with low-quality channel (or with less resource) for the benefit of the cluster, while a non-clustered UE with access to low-quality channel joins a cluster with high-quality channel (or with more resource) for the benefit of the UE, in order to achieve fairness and improve the cluster throughput. This scheme has shown to improve QoS performance (P.1) (i.e., higher throughput), fairness index (P.2), energy efficiency (P.4), and cost (P.5) (i.e., lower energy consumption).

5.4.2. *Assigning priority to clusters to distribute resources fairly*

Huang *et al.* [134] propose a clustering scheme that allocate resources (i.e., channels) among clusters in a SC (i.e., femtocell) fairly to minimize inter-cluster interference, and hence inter-cell interference [181]. The channels are allocated based on the priority of a cluster. This scheme is designed for heterogeneous (C.2.2) and mobile (C.3.2) UEs to form single-hop (C.1.1) clusters. The main objective is to enhance fairness (O.4) among clusters in a femtocell. The clustering scheme addresses the challenge of high cost (X.2) due to clustering overhead caused by interference among clusters in a femtocell. In this scheme, UEs from a femtocell form clusters among themselves and connect to the femtocell BS. The proposed clustering scheme covers three of the four stages in the clustering framework as summarized in Table 13. During the *first stage*, non-clustered UEs generate NS using physical parameters (e.g., geographical location). During the *second stage*, non-clustered UEs elect CHs based on the geographical location metric (M.2), which is the distance between a femtocell BS and neighboring UEs. A UE with the shortest distance is elected as the CH [182]. During the *third stage*, a traditional k-means algorithm is applied to form clusters [183]. Resource is allocated based on the priority of a cluster determined using three criteria, whereby a cluster is given a higher priority if it has: a) higher interference level, b) higher traffic load, and c) packets with longer waiting time. For instance, high-quality channels are allocated to clusters with higher priority to achieve a fair distribution of resources among clusters in a femtocell. This scheme has shown to improve fairness index (P.2), spectral efficiency (P.3), and energy efficiency (P.4).

Table 13
Stages of clustering algorithm [134]

Stage	Details and Outcomes
First	Non-clustered nodes generate NSs and network topology by exchanging messages among themselves using geographical location.
Second	Non-clustered nodes elect CHs, which are closer to femtocell BSs.
Third	Non-clustered nodes join clusters using traditional k-means algorithm. Priority is assigned to each cluster based on interference level, traffic load, and packet waiting time. Higher priority clusters receive high amount of resources to achieve fairness.

6. Open Issues

This section presents open issues that can be pursued in this research topic.

6.1. Performing cluster size adjustment

Large cluster size: a) increases network scalability by increasing the number of CMs in a cluster (or reduces the number of clusters in a network) leading to reduced clustering overhead (or clustering message exchanges) which helps to address the challenge of high cost, and b) increases cluster stability by reducing the number of re-clusterings [184]. Hence, large cluster size consumes lower network resources; however, the congestion level of a cluster increases as the CH handles more traffic from its CMs [185, 186].

As an example, due to the ultra-densification of nodes in 5G networks, there are two main challenges in performing cluster size adjustment: a) high heterogeneity (X.1) whereby segregating nodes with similar interests and behaviors can be difficult, and so there are a higher number of clusters in the network, and b) high cost (X.2) whereby there is an exponential increase in the number of message exchanges in intra- and inter-cluster communications, and so more network resources are required. In other words, higher number of nodes increases the number of clusters and message exchanges in the network, and hence requiring more network resources and reducing network performance [187]. Cluster size adjustment is necessary to adjust the number of clusters and message exchanges in the network.

As another example, due to the high mobility of UEs in 5G networks, there are two main challenges in performing cluster size adjustment: a) high dynamicity (X.3) whereby UEs' with high mobility causes significant changes to the underlying network topology (e.g., nodes move out of the coverage of a cluster) [188], and b) high cost (X.2) whereby clustering overhead and energy consumption are incurred in node joining and re-clustering procedures. The challenges reduce the cluster lifetime. However, cost is low in quasi static networks as compare to mobile networks. Cluster size adjustment is necessary to adjust the number of message exchanges and energy consumption in the network. Mobility pattern can also be applied to predict new positions of nodes for the node joining procedure.

In short, cluster size adjustment ensures that there are sufficient network resources to cater for the right number of clusters in the network in order to enhance QoS (P.1), and energy efficiency (or increased cluster lifetime) (P.4), as well as to reduce cost (or overhead) (P.5). Cluster size adjustment can be performed using artificial intelligence techniques due to the dynamicity of the operating environment [189]. Due to the lack of investigation on cluster size adjustment in 5G networks, further studies can be pursued in this topic.

6.2. Performing cluster maintenance

Cluster maintenance or re-clustering, which is the fourth stage of the clustering framework (see Section 4), forms new cluster(s) after the dissolution of an existing cluster in order to maintain network performance as time goes by [190, 192]. This mechanism is essential to handle major changes in the underlying cluster structure that can affect an entire cluster, including: a) a large number of non-clustered nodes joining or leaving a cluster [191], b) CH migration (or the rotation of the CH role) that elects the most appropriate node to become the next CH, c) cluster merging that combines multiple neighboring clusters into a single cluster due to their similarities, and d) cluster splitting that segregates a single cluster into multiple clusters due to their differences [192]. 5G networks pose three main challenges to cluster maintenance: a) high heterogeneity (X.1) whereby there are a diverse range of nodes with different capabilities, b) high cost (X.2), particularly clustering overhead, computational power, and energy consumption, is incurred for selecting new CHs and node joining, and rotating the CH of a cluster, and c) high dynamicity (X.3) whereby nodes with high mobility increase the need for cluster maintenance. In addition, insufficient of resources can cause the dissolution of a cluster as a results of increased packet loss and reduced QoS of the entire cluster [193, 26].

In short, cluster maintenance ensures that network resources are efficiently utilized in order to enhance QoS (P.1), spectral efficiency (P.3), energy efficiency (P.4), as well as to reduce cost (P.5), and the number of clusters in the network (P.6). Due to the lack of investigation on cluster maintenance in 5G networks, further studies can be pursued in this topic.

6.3. Ensuring node connectivity

Intra- and inter-cluster connectivities among nodes reduce network partitions. Long-term connectivity among nodes in a cluster increases network lifetime leading to cluster stability [194]; while short-term connectivity reduces network lifetime causing cluster instability [195]. In ultra-dense networks, nodes are expected to be deployed nearer to each other, and so it increases the probability of a connected network remaining connected for long term [196]. However, connectivity among clustered nodes in 5G networks can reduce due to high node mobility. 5G networks pose three main challenges to achieving long-term connectivity, including: a) high heterogeneity (X.1) whereby there are a diverse range of nodes (i.e., with different interests and purposes) with different capabilities (i.e., mobility, transmission power, and social relations), b) high cost (X.2), particularly clustering overhead, is incurred due to ultra-densification, and c) high dynamicity (X.3) whereby nodes' or UEs' with high mobility increase the changes of the underlying network topology (e.g., nodes move out of the coverage of a cluster and cause dis-connectivity).

In short, ensuring node connectivity provides long-term connectivity in order to enhance QoS (P.1), spectral efficiency (P.3), energy efficiency (or prolong network lifetime) (P.4), as well as to reduce cost (P.5) and the number of clusters in the network (P.6). This can be performed using particle swarm optimization [197], artificial bees colony [198], and ant colony optimization, that extracts social behaviors of bird flocks, bees, and ant colonies, respectively [199]. In addition, the centralized controller can compute and predict future positions and directions of nodes based on the nodes' historical mobility pattern for achieving long-term connectivity [200]. Due to the lack of investigation on ensuring node-connectivity in 5G networks, further studies can be pursued in this topic.

6.4. Adopting self-organization

Due to ultra-densification in 5G networks, clustering a large number of nodes is difficult [201]. Traditionally, artificial intelligence approaches have been used to perform self-organization, whereby *macro-learning* is applied at the centralized controller and *micro-learning* is applied at the distributed entities. Macro-learning (e.g., K-means [202]) enables a centralized controller to learn about the preferences and priorities of a group of nodes (e.g., with similar geographical location or social relation) to achieve globally optimal solutions over time using network-wide information, while micro-learning enables each individual node to learn about the preferences and priorities of itself to achieve locally optimal solutions over time using neighborhood information. Both macro-learning and micro-learning approaches can be integrated to provide a hybrid approach so that the clustering algorithms can gain benefits from both [203]. Clustering messages, generated by a large number of network entities (e.g., sensors, meters, and tracking devices), must be minimized as local information, such as the congestion level, mobility pattern, and data rate of the node [26], must be gathered from the network entities by the centralized controller [204]. 5G networks pose three main challenges to self-organization: a) high heterogeneity (X.1) whereby there are a diverse range of nodes with different capabilities, b) high cost (X.2), particularly clustering overhead, is incurred for selecting new CHs and node joining under ultra-densification, and c) high dynamicity (X.3) whereby nodes with high mobility introduces unpredictability in self-organization.

In short, self-organization provides intelligence for achieving globally optimal solutions in order to enhance QoS (P.1), energy efficiency (or prolong network lifetime) (P.4), as well as reduce cost (P.5) and the number of clusters in the network (P.6). This can be performed using reinforcement learning that observes network statistics (e.g., traffic load, energy consumption, and mobility pattern), and subsequently selects and executes the right actions to form self-organized clusters [205]. In addition to performing continuous monitoring, unexpected events, such as those that occur during disaster and emergency, should be detected. Due to the lack of investigation on achieving self-organization in the context of clustering in 5G networks, further studies can be pursued in this topic.

6.5. Enhancing quality of service

Due to ultra-densification in 5G networks, clustering a large number of nodes, with some using real-time applications, demands a stringent level of user requirements and QoS [206]. As an example, a driver-less car requires safety and warning messages (i.e., information about vehicles and road infrastructure) sent within a tolerable delay of 1 ms [207], and this must be fulfilled by single-hop or multi-hop data packet transmission between a CM to a CH via D2D [208]. In order to provide more reliable communication and larger coverage, control message transmission uses licensed channel, while data transmission uses either licensed or unlicensed (e.g., IEEE 802.11p) channel. 5G networks pose three main challenges to QoS enhancement, namely high heterogeneity (X.1), high cost (X.2), and c) high dynamicity (X.3), all of which reduce QoS. Addressing these challenges can enhance QoS (P.1) (e.g., higher throughput and lower end-to-end delay), as well as reduce cost (P.5) and the number of clusters in the network (P.6). Further studies can be pursued in this topic.

7. Conclusion

In this paper, clustering schemes for the next-generation mobile wireless network, namely 5G, are presented. This article refreshes the topic of clustering, its motivation and background in 5G through a review of the limited research works in this topic. The clustering schemes discussed in this paper, which are mainly based on five types of objectives (i.e., enhancing cluster stability, load balancing, social awareness, fairness, and quality of service (QoS)) address three main challenges, namely high heterogeneity, high cost, and high dynamicity. The clustering schemes mainly use five types of clustering metrics, namely node density, geographical location, mobility metrics, social relation, and residual energy, for the clusterhead election and node joining procedures. These schemes are based on three main characteristics, namely the number of hops (i.e., single-hop and multi-hop), node type (i.e., homogeneous and heterogeneous), and mobility rate (i.e., quasi static and mobile/ nomadic). These clustering schemes have shown to provide five main performance measures, namely higher QoS performance, fairness index, spectral efficiency, energy efficiency, as well as lower cost and the number of clusters in a network. Moreover, the main open issues related to clustering schemes in 5G are outlined. The paper is expected to support and motivate researchers for further exploration and investigation in this research area.

Acknowledgment

This work was part of the project entitled "A Novel Clustering Algorithm based on Reinforcement Learning for the Optimization of Global and Local Network Performances in Mobile Networks" funded by the Malaysian Ministry of Education under Fundamental Research Grant Scheme FRGS/1/2019/ICT03/SYUC/01/1, and the Partnership Grant CR-UMSSTDCIS-2018-01 and RK004-2017 between Sunway University and University of Malaya.

References

[1] M. S. Corson, R. Laroia, J. Li, V. Park, T. Richardson, and G. Tsirtsis,“Toward proximity-aware internetworking,” *IEEE Wireless Communications*, vol. 17, no. 6, 2010.

[2] A. Maeder, P. Rost, and D. Staehle, “The challenge of M2M communications for the cellular radio access network,” in *Proceedings Wurzburg Workshop IP, Joint ITG Euro-NF Workshop "Visual Future Generation Network"*, 2011, pp. 1–2.

[3] M. Dighriri, G. M. Lee, and T. Baker, “Measurement and classification of smart systems data traffic over 5G mobile networks,” in *Technology for Smart Futures*. Springer, 2018, pp. 195–217.

[4] Huang, X.L., Ma, X. and Hu, F., *Machine learning and intelligent communications*, Mobile Networks and Application, vol. 23, no. 1, pp. 68-70, 2018

[5] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, “Position and orientation estimation through millimeter-wave MIMO in 5G systems,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1822–1835, 2018.

[6] F. S. Shaikh and R. Wismüller, “Routing in multi-hop cellular device-to-device (D2D) networks: A survey,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2622-2657, 2018.

[7] Q. Cui, H. Wang, P. Hu, X. Tao, P. Zhang, J. Hamalainen, and L. Xia, “Evolution of limited-feedback comp systems from 4G to 5G: Comp features and limited-feedback approaches,” *IEEE Vehicular Technology Magazine*, vol. 9, no. 3, pp. 94–103, 2014.

[8] I. Ahmed, H. Khammari, A. Shahid, A. Musa, K. S. Kim, E. De Poorter, and I. Moerman, “A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives,” *IEEE*

Communications Surveys & Tutorials, vol. 20, no. 4, pp. 3060-3097, 2018.

[9] V. Milosevic, B. Jokanovic, O. Boric-Lubecke, and V. M. Lubecke, "Key microwave and millimeter wave technologies for 5G radio," in *Powering the Internet of Things with 5G Networks* IGI Global, 2018, pp. 70–104.

[10] P. Pirinen, "A brief overview of 5G research activities," in *Proceedings IEEE 1st International Conference on 5G for Ubiquitous Connectivity (5GU)*, Akaslompolo, Finland, 2014, pp. 17–22.

[11] Q. C. Li, H. Niu, A. T. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *IEEE Vehicular Technology Magazine*, vol. 9, no. 1, pp. 71–78, 2014.

[12] A. Imran, A. Zoha, and A. Abu-Dayya, "Challenges in 5G: how to empower son with big data for enabling 5G," *IEEE Network*, vol. 28, no. 6, pp. 27–33, 2014.

[13] G. Wunder, M. Kasparick, S. ten Brink, F. Schaich, T. Wild, I. Gaspar, E. Ohlmer, S. Krone, N. Michailow, and A. Navarro, "5G now: Challenging the LTE design paradigms of orthogonality and synchronicity," in *Proceedings IEEE 77th Vehicular Technology Conference (VTC Spring)*, Dresden Germany, 2013, pp. 1–5.

[14] U. Habiba and E. Hossain, "Auction mechanisms for virtualization in 5G cellular networks: Basics, trends, and open challenges," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2264-2293, 2018.

[15] S. Chen, F. Qin, B. Hu, X. Li, Z. Chen, and J. Liu, "User-centric ultra-dense networks for 5G," in *User-Centric Ultra-Dense Networks for 5G*, Springer, 2018, pp. 1–3.

[16] I. Parvez, A. Rahmati, I. Guvenc, A. I. Sarwat, and H. Dai, "A survey on low latency towards 5G: RAN, core network and caching solutions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3098-3130, 2018.

[17] D. Soldani, Y. J. Guo, B. Barani, P. Mogensen, I. Chih-Lin, and S. K. Das, "5G for ultra-reliable low-latency communications," *IEEE Network*, vol. 32, no. 2, pp. 6–7, 2018.

[18] C. Ranaweera, P. Monti, B. Skubic, M. Furdek, L. Wosinska, A. Nirmalathas, C. Lim, and E. Wong, "Optical X-haul options for 5G fixed wireless access: Which one to choose" in *Proceedings IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Honolulu USA, 2018, pp. 1–2.

[19] M. Agiwal, N. Saxena, and A. Roy, "Towards connected living: 5G enabled internet of things (IoT)," *IETE Technical Review*, pp. 1–13, 2018.

[20] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: A survey," *IEEE Communications Magazine*, vol. 46, no. 9, 2008.

[21] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[22] J. Qiao, X. S. Shen, J. W. Mark, Q. Shen, Y. He, and L. Lei, "Enabling device-to-device communications in millimeter-wave 5G cellular networks," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 209–215, 2015.

[23] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1801–1819, 2014.

[24] M. Ozger, F. Alagoz, and O. B. Akan, "Clustering in multi-channel cognitive radio ad hoc and sensor networks," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 156–162, 2018.

[25] X. Zhang and J. Wang, "Heterogeneous statistical QoS-driven resource allocation for D2D cluster-caching based 5G multimedia mobile wireless networks," in *Proceedings IEEE International Conference on Communications (ICC)*, Kansas City USA, 2018, pp. 1–6.

[26] M. Gharbieh, A. Bader, H. El Sawy, H.-C. Yang, M.-S. Alouini, and A. Adinoyi, "Self-organized scheduling request for uplink 5G networks: A D2D clustering approach," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1197-1209, 2019.

[27] C. Kai, H. Li, L. Xu, Y. Li, and T. Jiang, "Energy-efficient device-to-device communications for green smart cities," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1542–1551, 2018.

[28] A. Hatoum, N. Aitsaadi, R. Langar, R. Boutaba, and G. Pujolle, "FCRA: Femtocell cluster-based resource allocation scheme for OFDMA networks," in *IEEE international Conference on Communications (ICC)*, Kyoto Japan, 2011, pp. 1–6.

[29] R. Rajaraman, "Topology control and routing in ad hoc networks: A survey," *ACM SIGACT News*, vol. 33, no. 2, pp. 60–73, 2002.

[30] S. Bassoy, H. Farooq, M. A. Imran, and A. Imran, "Coordinated multipoint clustering schemes: A survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 743–764, 2017.

[31] M. Zorzi, A. Gluhak, S. Lange, and A. Bassi, "From today's intranet of things to a future internet of things: A wireless-and mobility-related view," *IEEE Wireless Communications*, vol. 17, no. 6, 2010.

[32] M. Gerla and J. T.-C. Tsai, "Multicluster, mobile, multimedia radio network," *Wireless Networks*, vol. 1, no. 3, pp. 255–265, 1995.

[33] A. Ephremides, J. E. Wieselthier, and D. J. Baker, "A design concept for reliable mobile radio networks with frequency hopping signaling," *Proceedings of the IEEE*, vol. 75, no. 1, pp. 56–73, 1987.

[34] C.-C. Chiang, H.-K. Wu, W. Liu, and M. Gerla, "Routing in clustered multihop, mobile wireless networks with fading channel," in *Proceedings of IEEE SICON*, vol. 97, no. 1997, 1997, pp. 197–211.

[35] A. D. Amis, R. Prakash, T. H. Vuong, and D. T. Huynh, "Max-min d-cluster formation in wireless ad hoc networks," in *Proceedings IEEE 19th Annual Joint Conference of the Computer and Communications Societies (INFOCOM)*, Tel Aviv Israel, 2000, pp. 32–41.

[36] M. M. C. Morales, C. S. Hong, and Y.-C. Bang, "An adaptable mobility-aware clustering algorithm in vehicular networks," in *13th IEEE Asia-Pacific Network Operations and Management Symposium (APNOMS)*, Taipei Taiwan, 2011, pp. 1–6.

[37] F. G. Nocetti, J. S. Gonzalez, and I. Stojmenovic, "Connectivity based k-hop clustering in wireless networks," *Telecommunication Systems*, vol. 22, no. 1-4, pp. 205–220, 2003.

[38] C. R. Lin and M. Gerla, "Adaptive clustering for mobile wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 7, pp. 1265–1275, 1997.

[39] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad hoc Networks*, vol. 3, no. 3, pp. 325–349, 2005.

[40] J. Y. Yu and P. H. J. Chong, "A survey of clustering schemes for mobile ad hoc networks," *IEEE Communications Surveys & Tutorials*, vol. 7, no. 1, pp. 32–48, 2005.

[41] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Communications*, vol. 30, no. 14-15, pp. 2826–2841, 2007.

[42] S. Chinara and S. K. Rath, "A survey on one-hop clustering algorithms in mobile ad hoc networks," *Journal of Network and Systems Management*, vol. 17, no. 1-2, pp. 183–207, 2009.

[43] X. Liu, "A survey on clustering routing protocols in wireless sensor networks," *Sensors*, vol. 12, no. 8, pp. 11 113–11 153, 2012.

[44] S. Vodopivec, J. Bešter, and A. Kos, "A survey on clustering algorithms for vehicular ad-hoc networks," in *IEEE 35th International Conference on Telecommunications and Signal Processing (TSP)*, Prague Czech Republic, 2012, pp. 52–56.

[45] M. M. Afsar and M.-H. Tayarani-N, "Clustering in sensor networks: A literature survey," *Journal of Network and Computer Applications*, vol. 46, pp. 198–226, 2014.

[46] A. M. Ortiz, D. Hussein, S. Park, S. N. Han, and N. Crespi, "The cluster between internet of things and social networks: Review and research challenges," *IEEE Internet of Things Journal*, vol. 1, no. 3, pp. 206–215, 2014.

[47] M. M. A. Osman, S. K. Syed-Yusof, N. N. N. A. Malik, and S. Zubair, "A survey of clustering algorithms for cognitive radio ad hoc networks," *Wireless Networks*, vol. 24, no. 5, pp. 1451–1475, 2018.

[48] S. Arjunan and S. Pothula, "A survey on unequal clustering protocols in wireless sensor networks," *Journal of King Saud University-Computer and Information Sciences*, 2017.

[49] L. Xu, R. Collier, and G. M. O Hare, "A survey of clustering techniques in WSNs and consideration of the challenges of applying such to 5G IoT scenarios," *IEEE Internet of Things Journal*, vol. 4, no. 5,

pp. 1229–1249, 2017.
[50] A. Chugh and S. Panda, "Strengthening clustering through relay nodes in sensor networks," *Procedia Computer Science*, vol. 132, pp. 689–695, 2018.
[51] A. S. Rostami, M. Badkoobe, F. Mohanna, A. A. R. Hosseinabadi, A. K. Sangaiah *et al.*, "Survey on clustering in heterogeneous and homogeneous wireless sensor networks," *The Journal of Supercomputing*, vol. 74, no. 1, pp. 277–323, 2018.
[52] Sun, Guolin, *et al.* "Delay-aware content distribution via cell clustering and content placement for multiple tenants." *Journal of Network and Computer Applications*, vol. 137, pp. 112–126, 2019.
[53] M. Chatterjee, S. K. Das, and D. Turgut, "Wca: A weighted clustering algorithm for mobile ad hoc networks," *Cluster Computing*, vol. 5, no. 2, pp. 193–204, 2002.
[54] X. Shen, "Device-to-device communication in 5G cellular networks," *IEEE Network*, vol. 29, no. 2, pp. 2–3, 2015.
[55] S. F. Yunas, M. Valkama, and J. Niemelä, "Spectral and energy efficiency of ultra-dense networks under different deployment strategies," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 90–100, 2015.
[56] J. Zhu, M. Zhao, and S. Zhou, "An optimization design of ultra dense networks balancing mobility and densification," *IEEE Access*, vol. 6, pp. 32 339–32 348, 2018.
[57] Kamel, Mahmoud, Walaa Hamouda, and Amr Youssef. "Ultra-dense networks: A survey." *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2522-2545, 2016.
[58] A. Mohamed, O. Onireti, M. A. Imran, A. Imran, and R. Tafazolli, "Control-data separation architecture for cellular radio access networks: A survey and outlook," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 446–465, 2016.
[59] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 336–361, 2013.
[60] S. M. A. El-atty and Z. Gharsseldien, "On performance of hetnet with coexisting small cell technology," in *IEEE 6th Joint IFIP Wireless and Mobile Networking Conference (WMNC)*, Dubai United Arab Emirates, 2013, pp. 1–8.
[61] L. Sanguinetti, A. L. Moustakas, and M. Debbah, "Interference management in 5G reverse TDD hetnets with wireless backhaul: A large system analysis," *IEEE journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1187–1200, 2015.
[62] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, "5G on the horizon: Key challenges for the radio-access network," *IEEE Vehicular Technology Magazine*, vol. 8, no. 3, pp. 47–53, 2013.
[63] N. Zhang, N. Cheng, A. T. Gamage, K. Zhang, J. W. Mark, and X. Shen, "Cloud assisted hetnets toward 5G wireless networks," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 59–65, 2015.
[64] Z. Wang, H. Li, H. Wang, and S. Ci, "Probability weighted based spectral resources allocation algorithm in hetnet under cloud-ran architecture," in *Proceedings IEEE International Conference on Communications in China-Workshops*, Xi'an China, 2013, pp. 88–92.
[65] M. Zhang, M. Mezzavilla, R. Ford, S. Rangan, S. Panwar, E. Mellios, D. Kong, A. Nix, and M. Zorzi, "Transport layer performance in 5G mmWave cellular," in *Proceedings IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, San Francisco USA, 2016, pp. 730–735.
[66] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The next generation wireless access technology*. Academic Press, 2018.
[67] P. Pirinen, "A brief overview of 5G research activities," in *Proceedings IEEE 1st International Conference on 5G for Ubiquitous Connectivity (5GU)*, Akaslompolo Finland, 2014, pp. 17–22.
[68] R. Ferrus, O. Sallent, J. Perez-Romero, and R. Agusti, "On 5G radio access network slicing: Radio interface protocol features and configuration," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 184–192, 2018.
[69] K.-L. A. Yau, J. Qadir, C. Wu, M. A. Imran, and M. H. Ling, "Cognition-inspired 5G cellular networks: a review and the road ahead," *IEEE Access*, vol. 6, pp. 35 072–35 090, 2018.
[70] C. Pan, M. Elkashlan, J. Wang, J. Yuan, and L. Hanzo, "User-centric C-RAN architecture for ultra-dense 5G networks: Challenges and methodologies," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 14–20, 2018.
[71] D. Kreutz, F. M. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, 2015.
[72] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE Journal of Selected topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, 2014.
[73] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 106–113, 2014.
[74] Y. Liu, M. Elkashlan, Z. Ding, and G. K. Karagiannidis, "Fairness of user clustering in MIMO non-orthogonal multiple access systems," *IEEE Communications Letters*, vol. 20, no. 7, pp. 1465–1468, 2016.
[75] M. K. Samimi, S. Sun, and T. S. Rappaport, "MIMO channel modeling and capacity analysis for 5G millimeter-wave wireless systems," in *Proceedings IEEE 10th European Conference on Antennas and Propagation (EuCAP)*, Davos Switzerland, 2016, pp. 1–5.
[76] M. Juntti, A. Padmanabhan, and A. Tölli, "Interference management via user clustering in two-stage precoder design," in *Proceedings IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Kalamata Greece, 2018, pp. 1-5.
[77] X. Duan, X. Wang, Y. Liu, and K. Zheng, "SDN enabled dual cluster head selection and adaptive clustering in 5G-VANET," in *Proceedings IEEE 84th Vehicular Technology Conference (VTC-Fall)*, Montreal QC Canada, 2016, pp. 1–5.
[78] A. Lozano Solsona, R. Heath Jr, and J. G. Andrews, "Fundamental limits of cooperation," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5213-26, 2013.
[79] C. D. T. Thai, P. Popovski, M. Kaneko, and E. De Carvalho, "Multiflow scheduling for coordinated direct and relayed users in cellular systems," *IEEE Transactions on Communications*, vol. 61, no. 2, pp. 669–678, 2013.
[80] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfrojd, and T. Svensson, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 44–51, 2014.
[81] B. Soret, A. De Domenico, S. Bazzi, N. H. Mahmood, and K. I. Pedersen, "Interference coordination for 5G new radio," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 131–137, 2018.
[82] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Communications Magazine*, vol. 49, no. 2, pp. 102–111, 2011.
[83] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges," *IEEE Communications Magazine*, vol. 50, no. 2, 2012.
[84] Z. Gao, L. Dai, D. Mi, Z. Wang, M. A. Imran, and M. Z. Shakir, "Mmwave massive-MIMO-based wireless backhaul for the 5G ultra-dense network," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 13–21, 2015.
[85] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4391–4403, 2013.
[86] Saha, Swetank Kumar, *et al.* "X60: A programmable testbed for wideband 60 ghz wlans with phased arrays." *Computer Communications*, vol. 133, pp. 77-88, 2019.
[87] L. Lu, X. Zhang, R. Funada, C. S. Sum, and H. Harada, "Selection of modulation and coding schemes of single carrier phy for 802.11

ad multi-gigabit mmWave WLAN systems," in *IEEE Symposium on Computers and Communications (ISCC)*,Kerkyra Greece, 2011, pp. 348–352.

[88] A. Gudipati, D. Perry, L. E. Li, and S. Katti, "Softran: Software defined radio access network," in *Proceedings ACM 2nd SIGCOMM workshop on Hot topics in software defined networking*, Hong Kong China, 2013, pp. 25–30.

[89] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges," *Wireless Networks*, vol. 21, no. 8, pp. 2657–2676, 2015.

[90] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez Jr, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, no. 1, pp. 335–349, 2013.

[91] J. Lunden, V. Koivunen, and H. V. Poor, "Spectrum exploration and exploitation for cognitive radio: Recent advances," *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 123–140, 2015.

[92] Y. Saleem, K.-L. A. Yau, H. Mohamad, N. Ramli, and M. H. Rehmani, "Smart: A spectrum-aware cluster-based routing scheme for distributed cognitive radio networks," *Computer Networks*, vol. 91, pp. 196–224, 2015.

[93] N. M. K. Chowdhury and R. Boutaba, "Network virtualization: state of the art and research challenges," *IEEE Communications magazine*, vol. 47, no. 7, 2009.

[94] Wen, Tao, and Wen Jiang, "Measuring the complexity of complex network by Tsallis entropy," *Physica A: Statistical Mechanics and its Applications,* vol. 526, pp. 121054, 2019.

[95] R. Jain and S. Paul, "Network virtualization and software defined networking for cloud computing: a survey," *IEEE Communications Magazine*, vol. 51, no. 11, pp. 24–31, 2013.

[96] I. F. Akyildiz, P. Wang, and S.-C. Lin, "Softair: A software defined networking architecture for 5G wireless systems," *Computer Networks*, vol. 85, pp. 1–18, 2015.

[97] B. Yi, X. Wang, K. Li, M. Huang *et al.*, "A comprehensive survey of network function virtualization," *Computer Networks*, 2018.

[98] A. Tootoonchian, A. Panda, C. Lan, M. Walls, K. Argyraki, S. Ratnasamy, and S. Shenker, "Resq: Enabling slos in network function virtualization," in *15th {USENIX} Symposium on Networked Systems Design and Implementation*, 2018.

[99] S. Abdelwahab, B. Hamdaoui, M. Guizani, and T. Znati, "Network function virtualization in 5G," *IEEE Communications Magazine*, vol. 54, no. 4, pp. 84–91, 2016.

[100] S. Mumtaz, K. M. S. Huq, and J. Rodriguez, "Direct mobile-to-mobile communication: Paradigm for 5G," *IEEE Wireless Communications*, vol. 21, no. 5, pp. 14–23, 2014.

[101] B. Zhou, H. Hu, S.-Q. Huang, and H.-H. Chen, "Intra cluster device-to-device relay algorithm with optimal resource utilization," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 5, pp. 2315–2326, 2013.

[102] A. Asadi, V. Sciancalepore, and V. Mancuso, "On the efficient utilization of radio resources in extremely dense wireless networks," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 126–132, 2015.

[103] S. Sharafeddine and O. Farhat, "A proactive scalable approach for reliable cluster formation in wireless networks with D2D offloading," *Ad Hoc Networks*, vol. 77, pp. 42–53, 2018.

[104] A. Asadi and V. Mancuso, "Dronee: Dual-radio opportunistic networking for energy efficiency," *Computer Communications*, vol. 50, pp. 41–52, 2014.

[105] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng, and S. Li, "Device-to-device communications underlaying cellular networks," *IEEE Transactions on Communications*, vol. 61, no. 8, pp. 3541–3551, 2013.

[106] M. Joud, M. Garcıa-Lozano, and S. Ruiz, "User specific cell clustering to improve mobility robustness in 5G ultra-dense cellular networks," in *Proceeding IEEE 14th Wireless On-Demand Network Systems and Services Conference (WONS)*, Isol France, February 6-8, 2018, pp. 45–50.

[107] V. Kumar, S. Yadav, D. Sandeep, S. Dhok, R. K. Barik, and H. Dubey, "5G cellular: Concept, research work and enabling technologies," in *Advances in Data and Information Sciences* Springer, 2019, pp. 327–338.

[108] L. C. Gimenez, P.-H. Michaelsen, and K. I. Pedersen, "UE autonomous cell management in a high-speed scenario with dual connectivity" in *Proceedings IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Valencia Spain, 2016, pp. 1–6.

[109] X. Ge, H. Cheng, G. Mao, Y. Yang, and S. Tu, "Vehicular communications for 5G cooperative small-cell networks." *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 7882–7894, 2016.

[110] H. Ibrahim, H. ElSawy, U. T. Nguyen, and M.-S. Alouini, "Mobility-aware modeling and analysis of dense cellular networks with *c*-plane/*u*-plane split architecture," *IEEE Transactions on Communications*, vol. 64, no. 11, pp. 4879–4894, 2016.

[111] D. López-Pérez, M. Ding, H. Claussen, and A. H. Jafari, "Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2078–2101, 2015.

[112] W. Hao and S. Yang, "Small cell cluster-based resource allocation for wireless backhaul in two-tier heterogeneous networks with massive MIMO," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 1, pp. 509–523, 2018.

[113] L. Liu, Y. Zhou, V. Garcia, L. Tian, and J. Shi, "Load aware joint comp clustering and inter-cell resource scheduling in heterogeneous ultra dense cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 2741–2755, 2018.

[114] S. K. Das and S. Tripathi, "Adaptive and intelligent energy efficient routing for transparent heterogeneous ad-hoc network by fusion of game theory and linear programming," *Applied Intelligence*, vol. 48, no. 7, pp. 1825–1845, 2018.

[115] J. Oueis, E. C. Strinati, and S. Barbarossa, "The fog balancing: Load distribution for small cell cloud computing," in *Proceedings IEEE 81st Vehicular Technology Conference (VTC Spring)*, Glasgow UK, 2015, pp. 1–6.

[116] E. Balevi and R. D. Gitlin, "A clustering algorithm that maximizes throughput in 5G heterogeneous F-RAN networks," in *Proceedings IEEE International Conference on Communications (ICC)*, Kansas City MO USA, 2018, pp. 1–6.

[117] J. Oueis, E. C. Strinati, S. Sardellitti, and S. Barbarossa, "Small cell clustering for efficient distributed fog computing: A multi-user case," in *Proceedings IEEE 82nd Vehicular Technology Conference (VTC Fall)*, Boston MA USA, 2015, pp. 1–5.

[118] A. A. Ateya, A. Muthanna, and A. Koucheryavy, "5G framework based on multi-level edge computing with D2D enabled communication," in *Proceedings IEEE 20th international conference on advanced communication technology (ICACT)*, Chuncheon-si Gangwon-do Korea (South), 2018, pp. 507–512.

[119] J. P. Singh, P. Dutta, and A. Chakrabarti, "Clustering modelling," in *Ad Hoc Networks* Springer, 2018, pp. 73–87.

[120] S. Basagni, "Distributed and mobility-adaptive clustering for multimedia support in multi-hop wireless networks," in *Proceedings IEEE 50th Vehicular Technology Conference (VTC)*, Amsterdam The Netherlands,, vol. 2. 1999, pp. 889–893.

[121] T. J. Kwon, M. Gerla, V. K. Varma, M. Barton, and T. R. Hsing, "Efficient flooding with passive clustering-an overhead-free selective forward mechanism for ad hoc/sensor networks," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1210–1220, 2003.

[122] Z. Khan and P. Fan, "A multi-hop moving zone (mmz) clustering scheme based on cellular-v2x," *China Communications*, vol. 15, no. 7, pp. 55–66, 2018.

[123] X. Ge, S. Tu, G. Mao, C.-X. Wang, and T. Han, "5G ultra-dense cellular networks," *IEEE Wireless Communications*, vol. 23, no. 1, pp. 72–79, 2016.

[124] L. Xu, G. O Hare, and R. Collier, "A smart and balanced energy-efficient multihop clustering algorithm (smart-beem) for MIMO IoT systems in future networks," *Sensors*, vol. 17, no. 7, p. 1574, 2017.

[125] B. Wang, Y. Sun, Q. Cao, S. Li, and Y. Wang, "Cooperative game-

based cheating in full-duplex relaying-based D2D communication underlaying heterogeneous cellular networks," *Mobile Networks and Applications*, vol. 23, no. 3, pp. 599–610, 2018.

[126] X. Duan, Y. Liu, and X. Wang, "SDN enabled 5G-VANET: Adaptive vehicle clustering and beamformed transmission for aggregated traffic," *IEEE Communications Magazine*, vol. 55, no. 7, pp. 120–127, 2017.

[127] Y.-y. Li, Z.-g. Ma, W.-b. Yan, and L.-y. Liu, "Weighted graph based clustering and local mobility management for dense small cell network with X2 interface," *Wireless Personal Communications*, vol. 95, no. 4, pp. 3771–3783, 2017.

[128] 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.

[129] M. M. Tulu, R. Hou, C. Li, and M. D. Amentie, "Cluster head selection method for content-centric mobile social network in 5G," *IET Communications*, vol. 12, no. 4, pp. 402–408, 2017.

[130] Z. Ning, X. Hu, Z. Chen, M. Zhou, B. Hu, J. Cheng, and M. S. Obaidat, "A cooperative quality-aware service access system for social internet of vehicles," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2506–2517, 2018.

[131] W. Qi, Q. Song, X. Wang, L. Guo, and Z. Ning, "SDN-enabled social-aware clustering in 5G-VANET systems," *IEEE Access*, vol. 6, pp. 28 213–28 224, 2018.

[132] P. Zhao, L. Feng, P. Yu, W. Li, and X. Qiu, "A social-aware resource allocation for 5G device-to-device multicast communication," *IEEE Access*, vol. 5, pp. 15 717–15 730, 2017.

[133] G. Zhang, K. Yang, and H.-H. Chen, "Socially aware cluster formation and radio resource allocation in D2D networks," *IEEE Wireless Communications*, vol. 23, no. 4, pp. 68–73, 2016.

[134] J. Huang, P. Zhou, K. Luo, Z. Yang, and G. He, "Two-stage resource allocation scheme for three-tier ultra-dense network," *China Communications*, vol. 14, no. 10, pp. 118–129, 2017.

[135] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober, and L. Hanzo, "User association in 5G networks: A survey and an outlook," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1018–1044, 2016.

[136] A. Asadi and V. Mancuso, "Network-assisted outband D2D-clustering in 5G cellular networks: theory and practice," *IEEE Transactions on Mobile Computing*, vol. 16, no. 8, pp. 2246–2259, 2017.

[137] H.-H. Nguyen, M. Hasegawa, and W.-J. Hwang, "Distributed resource allocation for D2D communications underlay cellular networks," *IEEE Communications Letters*, vol. 20, no. 5, pp. 942–945, 2016.

[138] O. Bazan and M. Jaseemuddin, "A conflict analysis framework for QoS-aware routing in contention-based wireless mesh networks with beamforming antennas," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3267–3277, 2011.

[139] S. Ehsan and B. Hamdaoui, "A survey on energy-efficient routing techniques with qos assurances for wireless multimedia sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 2, pp. 265–278, 2012.

[140] J. Zhao, C. Qiao, R. S. Sudhaakar, and S. Yoon, "Improve efficiency and reliability in single-hop wsns with transmit-only nodes," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 3, pp. 520–534, 2013.

[141] M. F. Munir and F. Filali, "Increasing connectivity in wireless sensor-actuator networks using dynamic actuator cooperation," in *Proceedings IEEE Vehicular Technology Conference (VTC)*, Singapore, 2008, pp. 203–207.

[142] L. Xu, J. Xie, X. Xu, and S. Wang, "Enterprise LTE and WIFI interworking system and a proposed network selection solution," in *ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS)*, Santa Clara CA USA, 2016, pp. 137–138.

[143] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122–130, 2014.

[144] 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 *Proceedings IEEE 11th International Symposium on Wireless Communications Systems (ISWCS)*, Barcelona Spain, 2014, pp. 858–862.

[145] Z. Su and Q. Xu, "Content distribution over content centric mobile social networks in 5G," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 66–72, 2015.

[146] W. Qi, Q. Song, X. Wang, L. Guo, and Z. Ning, "SDN-enabled social-aware clustering in 5G-VANET systems," *IEEE Access*, vol. 6, pp. 28 213–28 224, 2018.

[147] A. M. Vegni and V. Loscri, "A survey on vehicular social networks," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2397–2419, 2015.

[148] Z. Gao, Y. Shi, and S. Chen, "Measures of node centrality in mobile social networks," *International Journal of Modern Physics C*, vol. 26, no. 09, p. 1550107, 2015.

[149] D. Wei, X. Deng, X. Zhang, Y. Deng, and S. Mahadevan, "Identifying influential nodes in weighted networks based on evidence theory," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 10, pp. 2564–2575, 2013.

[150] X. Duan, X. Wang, Y. Liu, and K. Zheng, "SDN enabled dual cluster head selection and adaptive clustering in 5G-VANET," in *Proceedings IEEE 84th Vehicular Technology Conference (VTC-Fall)*, Montreal QC Canada, 2016, pp. 1–5.

[151] J. Deng, O. Tirkkonen, R. Freij-Hollanti, T. Chen, and N. Nikaein, "Resource allocation and interference management for opportunistic relaying in integrated mmWave/sub-6 ghz 5G networks," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 94–101, 2017.

[152] C.-K. Toh *et al.*, "Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks," *IEEE communications Magazine*, vol. 39, no. 6, pp. 138–147, 2001.

[153] M. Fahad, F. Aadil, S. Khan, P. A. Shah, K. Muhammad, J. Lloret, H. Wang, J. W. Lee, I. Mehmood *et al.*, "Grey wolf optimization based clustering algorithm for vehicular ad-hoc networks," *Computers & Electrical Engineering*, vol. 70, pp. 853-870, 2018.

[154] Y. Li, T. Wu, P. Hui, D. Jin, and S. Chen, "Social-aware D2D communications: Qualitative insights and quantitative analysis," *IEEE Communications Magazine*, vol. 52, no. 6, pp. 150–158, 2014.

[155] O. Semiari, W. Saad, S. Valentin, M. Bennis, and H. V. Poor, "Context-aware small cell networks: How social metrics improve wireless resource allocation," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 5927–5940, 2015.

[156] L. Xu, G. M. O'Hare, and R. Collier, "A balanced energy-efficient multihop clustering scheme for wireless sensor networks," in *Proceedings IEEE 7th Wireless and Mobile Networking Conference (WMNC)*, Vilamoura Portugal, 2014, pp. 1–8.

[157] O. Younis and S. Fahmy, "Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.

[158] S. Kassan, P. Lorenz, and J. Gaber, "Low energy and location based clustering protocol for wireless sensor network," in *Proceedings IEEE International Conference on Communications (ICC)*, Kansas City MO USA, 2018, pp. 1–6.

[159] K. Abboud, H. A. Omar, and W. Zhuang, "Interworking of DSRC and cellular network technologies for v2x communications: A survey," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 9457–9470, 2016.

[160] M. Ren, J. Zhang, L. Khoukhi, H. Labiod, and V. Vèque, "A unified framework of clustering approach in vehicular ad hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1401–1414, 2018.

[161] L. Song, D. Niyato, Z. Han, and E. Hossain, "Game-theoretic resource allocation methods for device-to-device communication," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 136–144, 2014.

[162] W. Saad, Z. Han, T. Basar, M. Debbah, and A. Hjorungnes, "Coalition formation games for collaborative spectrum sensing," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 276–297,

2011.

[163] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi, "Design aspects of network assisted device-to-device communications," *IEEE Communications Magazine*, vol. 50, no. 3, 2012.

[164] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 86–92, 2014.

[165] M. X. Goemans, "Minimum bounded degree spanning trees," in *47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, Berkeley CA USA, 2006, pp. 273–282.

[166] K. Jahed, M. Younes, and S. Sharafeddine, "Energy measurements for mobile cooperative video streaming." in *Wireless Days*, 2012, pp. 1–3.

[167] Athanasopoulou, Alexia, *et al.* "What technology enabled services impact business models in the automotive industry? An exploratory study." *Futures*,vol. 109, pp. 73-83, 2019.

[168] E. Feukeu, S. Ngwira, and T. Zuva, "Doppler shift signature for BPSK in a vehicular network: IEEE 802.11 p," in *Proceedings IEEE International Conference on Mechatronics and Automation (ICMA)*, Beijing China, 2015, pp. 1744–1749.

[169] A. Benslimane, T. Taleb, and R. Sivaraj, "Dynamic clustering-based adaptive mobile gateway management in integrated VANET 3G heterogeneous wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 3, pp. 559–570, 2011.

[170] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[171] I. Viering, M. Dottling, and A. Lobinger, "A mathematical perspective of self-optimizing wireless networks," in *Proceedings IEEE International Conference on Communications (ICC'09)*, Dresden Germany, 2009, pp. 1–6.

[172] V. Garcia, Y. Zhou, and J. Shi, "Coordinated multipoint transmission in dense cellular networks with user-centric adaptive clustering," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4297–4308, 2014.

[173] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Communications Magazine*, vol. 47, no. 12, 2009.

[174] L. A. Maglaras and D. Katsaros, "Social clustering of vehicles based on semi-markov processes," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 1, pp. 318–332, 2016.

[175] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wen, and M. Tao, "Resource allocation in spectrum-sharing ofdma femtocells with heterogeneous services." *IEEE Transactions on Communications*, vol. 62, no. 7, pp. 2366–2377, 2014.

[176] A. Asadi and V. Mancuso, "On the compound impact of opportunistic scheduling and D2D communications in cellular networks," in *Proceedings ACM 16th international conference on Modeling, analysis & simulation of wireless and mobile systems*, Barcelona Spain, 2013, pp. 279–288.

[177] "Wifi direct and LTE D2D in action," in *IEEE IFIP Wireless Days (WD)*, 2013, pp. 1–8.

[178] Hilal, Amr E., and Allen B. MacKenzie, "A distributed coalition game model for cooperation in MANETs", *Ad Hoc Networks*, vol. 85, pp. 46–59, 2019.

[179] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2706–2716, 2013.

[180] I. Chih-Lin, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: a 5G perspective," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 66–73, 2014.

[181] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2522–2545, 2016.

[182] Z. Donghai, Y. Jiang, G. Fei *et al.*, "K-means text clustering algorithm based on initial cluster centers selection according to maximum distance," *Application Research of Computers*, vol. 31, no. 3, pp. 713–719, 2014.

[183] D. K. Sharma, S. K. Dhurandher, D. Agarwal, and K. Arora, "kROP: K-means clustering based routing protocol for opportunistic networks," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–18, 2018.

[184] Z. Javed, K.-L. A. Yau, H. Mohamad, N. Ramli, J. Qadir, and Q. Ni, "Rl-budget: A learning-based cluster size adjustment scheme for cognitive radio networks," *IEEE Access*, vol. 6, pp. 1055–1072, 2018.

[185] G. Singh and G. Xiao, "Algorithms for energy-efficient clustering in wireless sensor networks," in *10th IEEE Singapore International Conference on Communication Systems (ICCS)*, Singapore, 2006, pp. 1–5.

[186] D. Nguyen, P. Minet, T. Kunz, and L. Lamont, "New findings on the complexity of cluster head selection algorithms," in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Lucca Italy, 2011, pp. 1–10.

[187] F. Aadil, W. Ahsan, Z. U. Rehman, P. A. Shah, S. Rho, and I. Mehmood, "Clustering algorithm for internet of vehicles (IoV) based on dragonfly optimizer (CAVDO)," *The Journal of Supercomputing*, pp. 1–26, 2018.

[188] Tabassum, Hina, Mohammad Salehi, and Ekram Hossain. "Fundamentals of Mobility-Aware Performance Characterization of Cellular Networks: A Tutorial," *IEEE Communications Surveys & Tutorials*, 2019.

[189] Y. He, F. R. Yu, N. Zhao, and H. Yin, "Secure social networks in 5G systems with mobile edge computing, caching, and device-to-device communications," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 103–109, 2018.

[190] M. Ren, J. Zhang, L. Khoukhi, H. Labiod, and V. Vèque, "A unified framework of clustering approach in vehicular ad hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1401–1414, 2018.

[191] A. K. Gupta and S. Prakash, "Secure communication in cluster-based ad hoc networks: a review," *Next-Generation Networks* Springer, 2018, pp. 537–545.

[192] A. Hajisami and D. Pompili, "Dynamic joint processing: Achieving high spectral efficiency in uplink 5G cellular networks," *Computer Networks*, vol. 126, pp. 44–56, 2017.

[193] W. Bai, Y. Li, T. Yao, and H. Zhang, "Performance analysis of dynamic re-clustering and resource allocation in ultra dense network," *IEEE Access*, vol. 6, pp. 60 891–60 899, 2018.

[194] A. Behnad, X. Wang, L. Hanzo, and T. J. Willink, "Connectivity-based centroid localization using distributed dense reference nodes," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6685–6689, 2018.

[195] M. H. Ling, K.-L. A. Yau, J. Qadir, and Q. Ni, "A reinforcement learning-based trust model for cluster size adjustment scheme in distributed cognitive radio networks," *IEEE Transactions on Cognitive Communications and Networking*, 2018.

[196] P.-H. Kuo and A. Mourad, "User-centric multi-rats coordination for 5G heterogeneous ultra-dense networks," *IEEE Wireless Communications*, vol. 25, no. 1, pp. 6–8, 2018.

[197] K. Vijayalakshmi and P. Anandan, "A multi objective tabu particle swarm optimization for effective cluster head selection in WSN," *Cluster Computing*, pp. 1–8, 2018.

[198] A. Rajagopal, S. Somasundaram, and B. Sowmya, "Performance analysis for efficient cluster head selection in wireless sensor network using RBFO and hybrid BFO-BSO," *International Journal of Wireless Communications and Mobile Computing*, vol. 6, no. 1, p. 1, 2018.

[199] M. Parsapoor and U. Bilstrup, "A simple ant colony optimization algorithm to select cluster heads", *Ad hoc Networks*, 2018.

[200] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.

[201] F. Ahmed, "Self-organization: A perspective on applications in the

internet of things," in *Natural Computing for Unsupervised Learning* Springer, 2019, pp. 51–64.

[202] Y. Zhao, K. Liu, X. Xu, H. Yang, and L. Huang, "Distributed dynamic cluster-head selection and clustering for massive IoT access in 5G networks," *Applied Sciences*, vol. 9, no. 1, p. 132, 2019.

[203] N. Kourtellis, G. D. F. Morales, and A. Bifet, "Large-scale learning from data streams with apache samoa," in *Learning from Data Streams in Evolving Environments* Springer, 2019, pp. 177–207.

[204] Hughes, B., Bothe, S., Farooq, H., & Imran, A, "Generative Adversarial Learning for Machine Learning empowered Self Organizing 5G Networks", in *IEEE International Conference on Computing, Networking and Communications (ICNC)*, pp. 282–286, 2019.

[205] M. J. Kaur, "A comprehensive survey on architecture for big data processing in mobile edge computing environments," in *Edge Computing* Springer, 2019, pp. 33–49.

[206] F. Schröder, "5G: New opportunities" in *Future Telco* Springer, 2019, pp. 63–74.

[207] U. Z. A. Hamid, H. Zamzuri, and D. K. Limbu, "Internet of vehicle (IoV) applications in expediting the implementation of smart highway of autonomous vehicle: A survey," in *Performability in Internet of Things* Springer, 2019, pp. 137–157.

[208] S. Zhong, H. Zhong, X. Huang, P. Yang, J. Shi, L. Xie, and K. Wang, "Connecting things to things in physical-world: Security and privacy issues in vehicular ad-hoc networks," in *Security and Privacy for Next-Generation Wireless Networks* Springer, 2019, pp. 101–134.