IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)
Volume 12, Issue 1 (Jan. - Feb. 2022), pp.37-46
e-ISSN: 2319-4200, p-ISSN: 2319-4197
www.iosrjournals.org

Clustering Approaches in Device-to-Device Communications

Ezeh Ikechukwu H[1], Prof. V. E. Idigo[2], Dr. V. N. Okorogu[3]
[1]*Department of Cyber Security, Federal University of Technology, Owerri, Imo State, Nigeria.*
[2,3]*Department of Electronic and Computer Engineering, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria*

Abstract
For D2D communication to take advantage of various new services introduced by 3GPP project, D2D devices are required to form clusters. Clustering enables proximate devices to communicate and share common resources, thus saving scarce network resources such as bandwidth and energy. The implementation of cluster in cellular networks (especially 5G/6G) leads to better energy/spectral efficiency and traffic signaling reduction than the traditional non-clustered cellular networks. Though elaborate work on the survey and classification of clustering algorithms in cellular networks have been done by some literatures, but a review of details of clustering techniques in D2D communications is lacking. The several benefits of clustering demand that various aspects of clustering in D2D communications be reviewed. Thus the aim of this paper is to review various clustering approaches in D2D communications, including the techniques adopted and their performances.
Key Words: *D2D, Clustering, Cluster Head Selection, Cluster Algorithms, Social Aware*

I. Introduction

The previous cellular generations are network centric, meaning that they rely on base station for network setup and coordination. But 5G and future generations are expected to be device centric, which implies that the devices not the base station manage the set up of the network. Device-to-device (D2D) communication is an integral part of 5G network. It enables cellular devices to discover and communicate with one another without necessarily having the assistance of evolved NodeB (eNodeB) or base station [1][2].

D2D communication offers many benefits. Due to the fact that short links exist between the proximate communicating devices, power saving is enhanced in D2D system more than the traditional cellular networks. In addition the flexibility of D2D enables it to offload traffic form the core network. This further improves the performance of D2D significantly because the overload density of the cellular system is reduced, thus reducing the transmission delay, provides high data rate and enhances energy savings [3]. In event of disaster that disrupted core network infrastructure, D2D can provide alternate access to cellular services, thus playing a key role in national security and public safety services [4].

The authors in [5] pointed out that in addition to mobility management, resource allocation and connection management, cluster formation is one of the factors that determine the D2D effectiveness. According to [6], for D2D communication to take advantage of various new services introduced by 3GPP project, D2D devices are required to form clusters. Through clustering, proximate devices communicate and share common resources, thus saving scarce network resources such as bandwidth and energy. Clustering allows the network to be divided into groups of geographically proximate devices. This efficiently simplified and optimized network function [7].

Various literatures agreed that cluster implementation in cellular networks (especially 5G/6G) leads to better energy/spectral efficiency and traffic signaling reduction than the traditional cellular networks [8][6][9]. Clustering is an important scheme that helps in network resource utilization and scalability [10][11]. According to [12], future generation mobile networks will possess three major attributes, namely: ultra-densification (having larger number of device per a unit area), heterogeneity (having various and diverse range of devices) and variability (bursty or high peak rate arising from traffic generated by heterogeneous network elements). Thus, the authors in [12] pointed out that the future generation network can use clustering to cater for these attributes. This is possible because clustering can divide network elements that have common characteristics or behaviors into logical groups. The common characteristics include relative speed, social tie, or degree of centrality. Such grouping will enable similar nodes to optimally access resources, achieve network stability, minimize network congestion and enhance spectral efficiency.

Various literatures have described and investigated the application of clustering in various networks such as vehicle ad hoc networks (VANETs) [13], mobile ad hoc networks (MANETs) [14] and wireless sensor networks (WSNs) [15]. Also, the work by [12] did elaborate work on the survey and classification of clustering algorithms in 5G networks. But a review of details of clustering techniques in D2D communications is lacking in literatures. The several benefits of clustering demand that various aspects of clustering in D2D communications be reviewed. Thus the aim of this paper is to review various clustering approaches in D2D communications including techniques and performance.

II. Clustering Algorithms in D2D

Various algorithms exist for organizing data into clusters and each has its own strengths and weaknesses. There is no general consensus on the best algorithm to choose, and in practice achieving perfect separation of objects using a clustering algorithm is difficult. But it is expected from a good clustering algorithm to be able to generate cluster groups with distinct non-overlapping boundaries. Furthermore, clustering algorithms take numerous parameters and they have to cope with noisy, sometimes incomplete/sampled data, making them to have considerable varying performance for different applications and data types [16].

Various clustering algorithm have been applied in by authors to address clustering in D2D enabled networks. According to [17], clustering algorithms in D2D are categorized into hierarchical based, distance and similarity-based, graph theory-based, squared error-based and density based clustering algorithms. Some of the algorithms proposed for D2D enabled networks include [18], who applied K-means Algorithm to cluster D2D device based on the value of interference between them. In another work, Genetic Algorithm (GA) and K-means algorithm was adopted in [19]. LEACH and related enhancements have been proposed in many literatures such as [20][21][22]. Recent literatures such as [17] and [23] have adopted and applied machine learning to clustering in D2D and D2D communication in general.

It was pointed out in [24] that there are three key features of D2D clustering algorithms. First is the ability to select or identify CH, Second is the grouping or clustering of CMs (which involves the association/dissociation of CMs to or from CH). The third is communication, i.e. intra-and inter-cluster communication. In addition, clustering algorithms are required to target some parameters such as the cumulative throughput of the cluster, the QoS, spectrum utilization and the maximum number of CMs in a cluster.

III. Clustering Techniques in D2D

In D2D communication, cluster formation can take various structures or shapes. Devices can form clusters and share information directly. Cluster members may or may not communicate directly with the base station or eNodeB. Figure 1 shows cluster formation structure that implements the utilization of CH by CMs for communication with eNodeb.

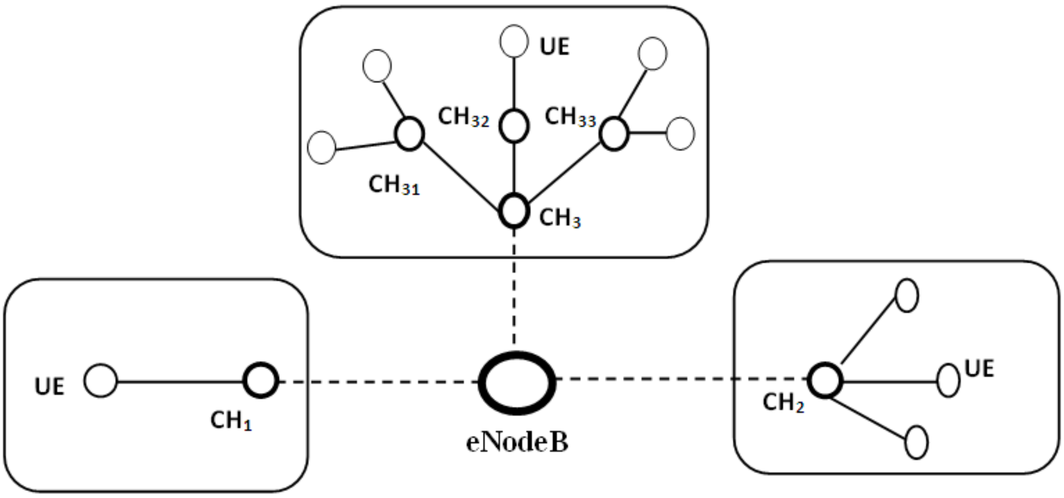

Figure 1: Various Cluster Structures with Cluster Heads [6]

In figure 1, it is assumed that the CMs (i.e. UEs) communicate with the eNodeB through the CHs (CH2, CH3) or through transit nodes (CH1, CH31, CH32, CH33). The shape of the cluster so formed could be tree-like (with CH and other transit nodes) or star-like (with single transit node). The structure so formed may or may not involve the functionality of the network [6].

There are various techniques, clustering metrics or criteria adopted in D2D clustering. But no matter the technique adopted, cluster formation seeks to enhance a set of objectives. According to [12], these objectives include cluster stability, QoS satisfaction, load balancing and social awareness. Failure of clustering technique to achieve QoS would lead to packet loss, latency and low throughput. The inability of a technique to ensure load balance will result in network congestion, and handover or clustering overheads. A low social aware clustering technique would likely increase the number of clusters and re-clustering occurrence.

Some literatures proposed clustering metric that rely on distance; others adopted network performance metrics such as throughput to form clusters, while some investigated social aware clustering methods. Other various metrics adopted for cluster formation and re-clustering include the number or size of cluster devices, device mobility, device geographic location, the transmit power and residual energy [12]. On the other, some literatures adopted the use of mixed metrics to ensure good satisfaction of clustering objectives.

It was pointed out in [12] that using geographic metric in cluster formation helps to achieve some objectives which include enhancement of cluster stability and creation of social awareness. These objectives are achieved because proximate devices would likely have common social interests, and can form cluster with reduced distance between the CH and the CMs thereby minimizing the energy dissipation of the CH which in turn prolongs its lifetime and that of the cluster. Similarly, using mobility and residual energy as metrics also enhance cluster stability by reducing the rate of occurrence of re-clustering. In addition, residual energy as a metric improves the QoS due to reduction in packet loss arising from disconnection that happens when device with less energy is chosen as a CH. Furthermore, social tie can be used as a base to identify the request or demand for cluster resources. Thus using social tie as metric helps to achieve the objective of social awareness and it indicates the strength and nature of relationships that exist among the nodes with regards to interests, attributes and behaviors.

In cluster formation that considers network performance, a number of literatures considered algorithms that assumed the spectrum utilization efficiency as the major clustering criterion, but the work in [6] suggested that there is need to also consider the channels' bandwidth between the CMs and the CH when implementing cluster algorithms. Contrary to some authors that used distance between the CH and the CMs as the metric for optimization of clustering, the authors in [6] used cluster throughput as the optimization metric and considered uniform and normal cluster members' distributions in their work. In addition, distribution of cluster members for large and small numbers of clusters was obtained using K-means and FOREL algorithms. Considering resource utilization and distribution, it was observed that K-means provides better clustering than FOREL algorithm. In addition, average throughput between the CH and the CM for normal distribution is more than two times larger than the uniform distribution.

Instead of a single metric, mixed metric was adopted by some authors to form cluster. For instance, an algorithm that uses distance and SINR as the parameters in cluster formation was proposed by [25]. The result from simulation indicates that energy efficiency and network capacity are improved. Similarly, a model that incorporated both social interactions and physical relationship among the D2D devices was proposed by the [8]. The physical relationship in this case is the distance between the devices. The authors opined that social interaction profiles of the user equipments (UEs) greatly affect the effectiveness and the efficiency of cluster formation in D2D. The results from the work showed that in terms of throughput, energy efficiency and power consumption, the proposed model has better performance than other models that depend only on the physical distance between D2D devices.

In addition, a methodology that uses LEACH algorithm and distance based resource allocation (RA) in D2D over LTE-A network was proposed by [21]. The authors proposed cluster formation technique that has the ability to form social aware clusters. This technique proposed by the authors allocates resources based on the QoS and the distances between the CMs and the CHs. The use of resource allocation minimizes signaling overhead and interference level. The inclusion of clustering algorithm to the methodology minimizes power consumption and improves the efficiency of the system. On the other hand the work by [24] investigated two clustering schemes. The schemes are dynamic formation of clusters by means of cluster head and a threshold-based cluster formation approach. The investigation of the two schemes showed that cluster head driven approach has high energy efficiency, but low coverage ratio. On the other hand, threshold approach handled tradeoff between energy efficiency and convergence ratio.

Another clustering scheme proposed in literature adopts energy harvesting D2D clustering model. This work described in [4] was proposed to provide energy savings during disaster or emergency situation for both UEs and CH by using the concept of power transfer/harvesting in D2D cluster formation. The idea is to have cluster formation among the participating UEs in a manner that would minimize the average energy consumption. The results form simulation showed that the proposed model enhances energy savings for both UEs and CH. In addition, the results of their work showed that there is a linear increase in energy consumption for clustered D2D, whereas there is an exponential increase in energy consumption for non-clustered D2D.

The summary of various metrics adopted in various literatures to cluster devices in D2D Communication is shown in Table 1.

Table 1: D2D Clustering Technique/metrics Adopted in Literatures

Literature	Clustering Technique/Metric	Key Conclusion
[6]	Cluster throughput	Increase in network resource utilization rate
[25]	Distance and SINR	Improvement in energy efficiency and network capacity
[8]	Distance and social interactions	The model outperforms algorithms that consider only distance as a metric
[24]	Cluster head and threshold-based cluster formation	Cluster head approach has high energy efficiency. Threshold approach handled tradeoff between energy efficiency and convergence ratio
[21]	LEACH algorithm, QoS and distance based RA	RA minimizes signaling overhead and interference level. Clustering algorithm minimizes power consumption
[4]	Power transfer and energy harvesting	Energy savings for both UEs and CH are enhanced.
[44]	Distance and channel conditions	Improvement in energy consumption, better area spectral efficiency (ASE)

IV. Physical versus Social Domain

In addition to physical factors such as mobility, the stability of links or clusters in D2D communication is affected by the degree of relationship among the users. When the degree of social relationship is not sufficient, it could result in frequent link failure or re-clustering, thereby reducing the users' quality of experience. Thus recent literatures have integrated social network and social domain information in the formation of D2D clusters.

The work by [26] pointed out that currently researchers are developing interest on the interplay between D2D communication and social-enabled networks/networking. Social-enabled D2D communications involves the discovery and utilizing the social interaction patterns that exist between the social network entities (people and objects) to improve the D2D communications' efficiency based on proximity information. For instance, the authors in [27] proposed the use of social ties as a motivation to enhance D2D communication. In this regard, the authors postulated that there is higher probability to select a node as a CH or relay from two peers if the tie between them is stronger.

A good number of literatures have considered physical attributes such as mobility, device location or link quality as major determining factors in D2D cluster formation. But recent studies have shown that D2D communication performance is strongly related to the social relationships that exist among the users. According to [28], the social relationship attributes can be reflected by degree of trust and the similarity of demands. In addition, social stability among the users affects the stability of clusters formed. The authors further pointed out that the stability of communications links among D2D users can only be achieved when both the physical and social links among them are strongly stable. Hence the physical and the social domains are related and the relationship that exists between them is shown in figure 2.

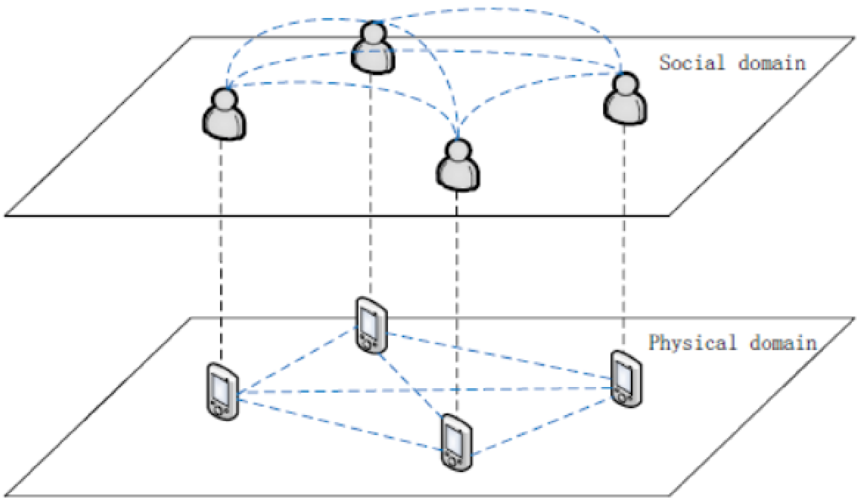

Figure 2: The Physical and social domains relationship of D2D communication users [28]

To cater for video file distribution, the authors in [28] proposed a user-demand/information aware D2D multicast clustering algorithm. The willingness of a user requesting for a video is evaluated based on the history of the user's viewing behavior and the popularity of the video. Both the social ties and users' channel quality information are taking into account. The result from their work showed that the system transmission rate is enhanced and hence the throughput, the energy and spectral efficiency are improved. Similarly, in a network, some users may have interest in same content. An algorithm that groups users which have interest in the same content into clusters was proposed by [9]. The result of their work showed a high reduction in energy when more than one CH is utilized.

To increase the system rate, admission-policy based D2D clustering scheme was proposed by [29]. In addition to two attractive functions presented which considered social interaction, location and energy balance, the authors also derived the probability of the arrival user joining a certain D2D cluster based on Chinese Restaurant Process (CRP) algorithm and designed a matching function to assign optimal D2D cluster for each arrival user. The results from simulation showed that the proposed scheme leads to good performance in terms of the system rate and the D2D clusters' stability.

In summary, social ties among the D2D users play important role in cluster formation. In addition, algorithms that take into account the social factors among users seem to have better performance than algorithms that considered only the physical factors.

V. Cluster Head/ Relay Selection

Devices in a cluster can operate in offline condition without being connected to the eNodeB. While some of the CHs only coordinate the CMs, some also act as a relay and communicate to other CHs or the base station. Traffic circulates directly between the cluster devices, but can be transmitted outside of the cluster to the eNodeB with the aid of CH serving as a relay.

(A) Cluster Head Selection

An important feature of a clustering algorithm is the ability to select appropriately a CH from the rest of the devices. The CH serves as a coordinator, coordinating other CMs [10][11][6]. Also, the authors in [8] added that in situation where there is transmission failure; CH can support the CMs to retransmit any stored information.

There are some rules or decisions that govern the selection of a particular member of a cluster as the CH. But it was pointed in [6] that CH is optimally and careful selected. This is because the choice of CH greatly affects parameters such as QoS, network efficiency, energy efficiency or dissipation offered to the CMs of the cluster. Thus as relates to energy dissipation and mobility, the devices selected as CH must be reliable according to [5]. The need for CH reliability is to ensure that connections are not terminated throughout the communication session.

There are various criteria assumed in literatures in the choice of CH selection. In literatures such as [30][31], CH selection was considered based on the distance between the CMs. But the work in [6] opined that that the selection of CH should be based majorly on the expected QoS parameters. The reason being that variant nature of the radio channel features exist as well as the channel resources between the CH and CM for each device within a cluster. On the other hand, authors in [30] stated that in scenario such as in disaster and public safety, CH should be optimally selected so that in event of damaged or dysfunctional base station, the selected CH can assume some of the roles of the failed base station.

While some literatures considered only one factor or parameter in CH selection, others considered weighted approach, in which the devices are assigned weights based on certain measures. The device with suitable weight is chosen as the CH. For instance [33] proposed a weighted CH selection algorithm. The factors the author proposed in selecting a suitable CH are the optimal number of CMs to support, the received signal strength, the mobility tendency of the potential CH, the cumulative time a device can serve as CH and the device capability. Each device calculates its own weight and broadcast the same to the neighbors. A comparison is made between the received weight and the device's own weight. The device with minimum weight is selected as the CH after all the devises have made their broadcast and comparison. The results from the work showed that the choice of selecting CH achieves high discovery and communication rate, though more energy is consumed during the discovery process.

Similarly, the work by [25] adopted the combination of distance and SINR as the choice of selection of CH and clustering of CMs to the CHs. A UE with a distance of R/2 or more from the base station is considered as the potential CH, where R is the radius of the cell. If the distance between a CH and other UEs is less than 0.025R, 0.1R or 0.25R, the UEs form clusters around the CH. Similar approach is adopted in the case of SINR, where a UE with SINR above average is a potential CH. The result of the work showed improvement in energy efficiency and system capacity. Also, the work by [29] optimized the CH selection based on the three factors namely: the communicating ability of the UE (represented as UE energy), the distance of influence and the social trust (or tie) between the proximate UEs. The authors compared the performance of the proposed scheme (which uses the three factors) with another proposed scheme that relied only on the average distance of the UE in selection of CH. From the comparison, it was shown that the inclusion of social trust enabled the average social trust of the proposed scheme to be larger than that of the scheme that relied on the distance. In addition, at higher rate or higher energy consumption, the transmission reliability of the distance based CHs selection scheme lags behind the three factor scheme.

A distributed dynamic CH selection and clustering scheme based on an improved K-means algorithm was proposed by [22]. The authors adopted a network assisted clustering by the base station, which group each device according to the quality of the channel and the location of the device. The proposed scheme was simulated and analyzed, and the result showed that the efficiency of the network can be improved by the proposed scheme. On the other hand the work in [34] made a comparison between the CHs selected based on distance from the base station and CHs selected based on the Received Signal Strength using Self-Organizing Map algorithm. It was shown from the comparison that the choice of threshold values or the parameters used in SOM algorithm affect the number of CH selected.

The summary of CH techniques adopted by various literatures is shown in table 2.

(B) Relay Selection

In addition to selecting CH that coordinates the operations of other members of a cluster, CH could also be chosen to act as a relay. Except the direct communication between proximate UEs, other D2D communications such as inter-cluster communication and communication between a cluster and a base station is made possible through a CH acting as a relay [1]. It was pointed out in [22] that the critical factors that affect energy efficiency of the system are the clustering algorithm used and the method adopted in CH selection. The use of relay increases energy efficiency and in addition, it extends communication range, and hence the total network coverage area. A relay must be optimally and efficiently chosen, especially in cooperative D2D systems. And the choice of relay selection could be centralized or distributed. Centralized relay selection is done by the base station, but this choice causes excessive load on the base station. On the other hand, the choice of distributed relay selection ensures that inappropriate relays are not selected [1].

The authors in [35] pointed out that some literatures that studied relay in LTE and other cellular systems adopted the strategy of selection instead of clustering in the choice of relays. Just as in selection of CH from a cluster, there are various approaches adopted in literatures in the selection of a relay from a cluster. Some literatures chose relay based on the minimum distance, while some others are based on the least path-loss. The authors in [35] stated that the choice of path-loss over distance is better option since it offers better link quality estimation. Other literatures make use of social aware or social relationship in the choice of relay selection.

The work by [35] adopted clustering to select relay using Basic Sequential Algorithmic Scheme (BSAS) together with power control scheme. The adoption of this approach showed that the capacity of the system increases and the energy consumption is enhanced compared to other relaying methods. On the other hand, the authors in [36] use Stackelber game approach to select the best relay that requires less energy and has the ability to provide enhanced spectral efficiency to UEs that in out of coverage areas. The authors did not adopt the popular individual or central relay selection, but use Stackelber game theory to select relay from competing UEs. The results from their work showed that in addition to complexity reduction, the spectral efficiency, the energy efficiency and the total capacity of the system are greatly enhanced. Similarly, the work by [37] used game theory to propose a relay selection based on distance and energy level of a device. The authors showed that when compared with other algorithms, the algorithm proposed helps improve effectively the throughput and the coverage, and it reduces the probability of connection interruption.

Other authors adopted the social aware approach and the social relationship between users to select suitable relay. It was noted in [26] that the efficiency of D2D network can be enhanced by considering the social attributes and the pattern of interaction between the users, which are not only man but machines as well. The drive for social D2D communication hinges on two factors – stability and reliability [26]. The stability concern emanates from the idea that the type of relationship between peers would determine the probability, nature and the length of connection especially when mobility is involved. In particular, a potential relay should be selected based on the type of relationship it shares with the end users. On the other hand, the reliability concern comes from the fact that there are several configurable smart devices in today's market, and to protect integrity of data and leakage of privacy, the choice of a relay should be optimally selected. One of the ways of achieving this is adopting trustworthiness selection procedure among peers.

Various literatures have investigated the choice of social tie in relay selection. For instance, the work by [38] proposed a scheme to enhance the stability of D2D communication by considering the mobility tendency of the relay. The authors proposed optimal stopping method by considering the centrality of the probable relay and the social relationships that exist between the relay and the end users. The results from their work showed that the data traffic relayed can be maximized and the system stability can be achieved both in short and long term. Similarly, the authors in [39] argued that to minimize negative effects of relay selection on the entire network, the suitability of the relay (in terms of power control and resource allocation) and the social tie among the users must be considered. In addition to power control and resource/channel allocation scheme adopted, the authors proposed the use of regional boundaries and the social relationship among users in the selection of the relays. It was observed that the algorithm proposed can improve to a great extent the system data rate and the performance of the D2D heterogeneous network.

In relation to social attributes and relationships, the reliability and the stability of relay selection in D2D depends on the social trust among the users. According to [26], in D2D communications, social relationships among peers can be exploited to build a high degree of trustworthiness to enhance stability and reliability. In this regard, the authors in [40] developed a cooperative D2D relay scheme based on social trust. The work considered both physical and social distances among the UEs. The results obtained showed that when compared with direct transmission scheme, the developed scheme can offer a significant throughput gain.

The idea of building social trust among peers is related to risk associated with security and privacy. A malicious device can pose as a friendly one, and thus there is need for peer devices to have a level of trust among themselves. The identification of trustworthy UEs can be enhanced through social awareness of devices according to [41]. Thus trust is a metric measured based on the social ties among the UEs. Some literatures have proposed a social trust-based scheme to limit a malicious node from corrupting or having access to private messages. The authors in [41] adopted this approach to model a reliable social trust-based scheme in 5G network. In the proposed scheme, a device that wants to upload content would select a relay based on the level of trust among its peers. The work showed that the proposed scheme can filter malicious nodes, enhances energy consumption, and achieved much gain in the content delivery or uploading time.

In addition, social tie could be in the form of some users having the same content interest. Such users or devices can be grouped into a cluster and a CH selected from the cluster. Using device to multi-device (D2MD) communication, the selected CH can serve to relay and multicast the content obtained from eNodeB to the CMs. In such multicasting, the channel condition among the CMs and the CH affects the data rate negatively. Though increasing the transmit power can enhance the data rate, the system energy efficiency can deteriorate. An option is for the eNodeB to send the content independently to each user, but the energy efficiency of such approach is less than the cooperative D2D option. But grouping adjacent users into cluster is a better strategy, in which the CM that has the best channel condition on average compared with other CMs is selected as the CH. This CH receives the content from eNodeB and serves as a relay to distribute or multicast the content to other CMs [9]. Though using a single CH in D2D content multicasting is a good strategy, the work by [9] pointed out that better performance (in terms of reduction of energy per bit) is obtained by using more than one CH in a cluster to simultaneously transmit the content.

The summary of relay techniques adopted by various literatures is shown in table 2.

Table 2: Summary of Cluster Head and Relay Selection Techniques in D2D Clustering

Literature	CH	Relay	Criteria	Key Conclusion
[30]	✓		Distance between the CMs	Increase in density of active cluster increases area spectral efficiency (ASE)
[6]	✓		Throughput between CH and CMs	Average throughput between CH and CM for normal distribution of CMs larger than the uniform distribution
[32]	✓	✓	Beacon broadcast & pre-defined metrics	Range of CH is extended, high discovery ratio & low latency
[33]	✓	✓	Weighted Factors: mobility, signal strength, etc	High discovery and communication rate
[44]	✓	✓	Weighted Factors: distance, channel condition, etc	Improvement in system performance
[25]	✓	✓	Distance and SINR	Improvement in energy efficiency and system capacity
[29]	✓	✓	Weighted Factors: energy, distance, social trust	Enhancement in average social trust, transmission reliability
[22]	✓	✓	Channel quality & device location	Enhancement in network efficiency
[35]	✓	✓	Basic Sequential Algorithmic Scheme (BSAS) & power control scheme	Improvement in system capacity & energy consumption
[36]	✓	✓	Stackelber game theory, based on less energy	Complexity reduction, improvement in spectral/energy efficiency, & system capacity
[37]	✓	✓	Game theory based on distance and energy	Improvement in throughput & coverage, reduction in the probability of connection interruption
[38]	✓	✓	Optimal stopping method based on relay centrality & social relationships	System stability enhancement & latency reduction
[39]	✓	✓	Power control, regional boundaries & social relationship	Improvement in system data rate & system performance
[40]	✓	✓	Social trust, physical & social	Significant throughput gain

			distances	
[41]	✓	✓	Level of trust among peers	Filtration of malicious nodes; enhancement in energy consumption, content delivery & uploading time
[9]	✓	✓	Selection of more than one CHs with different transmit power per cluster	Improvement in system performance, reduction in energy per bit

VI. Selection/Effects of Number of Cluster Members

The first feature of a clustering algorithm is the ability to select suitable CH from the rest UEs. But a good algorithm should be able to cluster or group other UEs or CMs around the selected CH. In most literature, the selection or clustering of CMs is based on distance between the CMs and the CH. Instead of physical distance, some literatures considered social distance or social relationship in clustering. For instance, the authors in [8] proposed clustering that incorporates both social and physical characteristics. Similarly, the work in [21] proposed clustering scheme that has ability to form socially aware clusters.

But the performance of the cluster is affected by the number of CMs per a cluster. Although having a larger number of CMs ensures savings in network resources as pointed out by [6], it has been noted that the number of CMs per a cluster depends on some factors, which include the CH channel throughput, the range of coverage, the cluster traffic intensity, the location and distribution of UEs [6][42].

The effects of variation of number of UE per cell, data transfer rate and the number of CMs per cluster on energy consumption for both clustering and cooperative clustering scenarios were investigated by [43]. The simulation results showed that cells that have greater number of UEs recorded higher energy consumption. For clustering scenario, energy consumption was not affected by the variation in the number of UE per cluster, but in cooperative clustering scenario, energy consumption increases by 25% for each additional UE per cluster. The work by [9] proposed that more than one CH can be used to simultaneously transmit content. In addition, the work considered the reception of power by the CMs from all the CHs within the cluster. From the simulation and analysis, the authors concluded that the proposed scheme performance improves for higher number of CMs.

VII. Conclusion

The major attributes of future cellular generations namely: ultra-densification, heterogeneity and variability can be taken care by efficient clustering technique or algorithm. Thus in this paper, we have reviewed various cluster formation techniques in D2D communications and their performances. In addition, a review of various criteria for cluster head/relay selection, the consideration of physical and social aware clustering and the effects of number of CMs per a cluster were done. Through this review work, it was discovered that clustering algorithm in D2D should not consider only the physical attributes (e.g. distance) but performance metrics (such as throughput) as well as social relationships that exist among the D2D users.

References

[1]. Gandotra, P. & Jha, R. K. (2016). Device-to-device communication in cellular networks: A survey. *Journal of Network Computer Applications*, *71*(C), 99–117. Retrieved from https://www.sciencedirect.com/science/article/pii/S1084804516301229?via%3Dihub

[2]. Condoluci, M., Militano, L., Orsino, A., Alonso-Zarate, J. & Araniti, G. (2015). LTE-Direct vs. WiFi-Direct for machine-type communications over LTE-A systems. *2015 IEEE 26th International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC): Workshop on M2M Communications: Challenges, Solutions and Applications*, 2298-2302. Retrieved from https://www.researchgate.net/publication/287644192_LTE-direct_vs_WiFi-direct_for_machine-type_communications_over_LTE-A_systems

[3]. Adnan, M. H., & Ahmad Zukarnain, Z. (2020). Device-To-Device Communication in 5G Environment: Issues, Solutions, and Challenges. *Symmetry*, *12*(11), 1762. MDPI AG. Retrieved from http://dx.doi.org/10.3390/sym12111762

[4]. Ali, K., Nguyen, H. X., Vien, Q., Shah, P., & Chu, Z. (2018). Disaster management using D2D communication with power transfer and clustering techniques. *IEEE Access: Special Section on Advances in Interference Mitigation Techniques for Device-To-Device Communications*, *6*, 14643-14654. Retrieved from https://core.ac.uk/download/pdf/160274938.pdf

[5]. Sharafeddine, S., & Farhat, O. (2018). A proactive scalable approach for reliable cluster formation in wireless networks with D2D offloading. *Ad Hoc Networks*, 77, 42-53. https://www.sciencedirect.com/science/article/abs/pii/S1570870518301483?via%3Dihub

[6]. Paramonov, A., Hussain, O., Samouylov, K., Koucheryavy, A., Kirichek, R., & Koucheryavy, Y. (2017). Clustering optimization for out-of-band D2D communications. *Hindawi: Wireless Communications and Mobile Computing, 2017*(Article ID 6747052), 1-11. Retrieved from https://doi.org/10.1155/2017/6747052

[7]. Bentaleb, A., Boubetra, A., & Harous, S. (2013). Survey of clustering schemes in Mobile Ad hoc networks. *Communications and Network*, *5*, 8-14. Retrieved from https://www.scirp.org/pdf/CN_2013071009593748.pdf

[8]. Wang, L., Araniti, G., Cao, C., Wang, W., & Liu, Y. (2015). Device-to-Device users clustering based on physical and social characteristics. *Hindawi Publishing Corporation: International Journal of Distributed Sensor Networks*, *2015*(Article ID 165608), 1-14. Retrieved from https://journals.sagepub.com/doi/pdf/10.1155/2015/165608

[9]. Ashraf, M., Yeo, W.-Y., Woo, M., & Lee, K.-G (2016). Smart energy efficient device-to-multidevice cooperative clustering for multicasting content. *Hindawi Publishing Corporation: International Journal of Distributed Sensor Networks*, *2016*(Article ID 3727918), 1-9. https://journals.sagepub.com/doi/10.1155/2016/3727918

[10]. Khatoon, N. & Amritanjali (2017). Mobility aware energy efficient clustering for MANET: A bio-inspired approach with Particle Swarm Optimization. *Hindawi: Wireless Communications and Mobile Computing, 2017*(Article ID 1903190), 1-12. Retrieved from http://downloads.hindawi.com/journals/wcmc/2017/1903190.pdf

[11]. Ishtiaq, A., Ahmed, S., Khan, M. F., Aadil , F., Maqsood, M., & Khan, S. (2019). Intelligent clustering using moth flame optimizer for vehicular ad hoc networks. *International Journal of Distributed Sensor Networks*, *15*(1), 1-13. Retrieved from https://journals.sagepub.com/doi/pdf/10.1177/1550147718824460

[12]. Khan, M. F., Yau, K., Noor, R.M., & Imran, M. (2020). Survey and taxonomy of clustering algorithms in 5G. *Journal of Network and Computer Applications*, *154*(102539). Retrieved from https://doi.org/10.1016/j.jnca.2020.102539

[13]. Cooper, C.S., Franklin, D., Ros, M., Safaei, F., & Abolhasan, M. (2017). A Comparative Survey of VANET Clustering Techniques. *IEEE Communications Surveys & Tutorials*, *19*(1), 657-681. Retrieved from https://ieeexplore.ieee.org/document/7572026

[14]. Hussain, K., Abdullah, A. H., Iqbal, S., Awan, K. M. & Ahsan, F. (2013). Efficient Cluster Head Selection Algorithm for MANET. *Journal of Computer Networks and Communications*, *2013*(Article ID 723913), 1-7. Retrieved from https://doi.org/10.1155/2013/723913

[15]. Wali, S., Ozuomba, S., & Kalu, C. (2019). Self-Organizing Map (SOM) Clustering of 868 MHz Wireless Sensor Network Nodes Based on Egli Pathloss Model Computed Received Signal Strength. *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, *6*(12), 12388-12397. Retrieved from http://www.jmest.org/wp-content/uploads/JMESTN42353511.pdf

[16]. Rodriguez, M. Z., Comin, C. H., Casanova, D., Bruno, O. M., Amancio, D. R., Costa, L. dF., & Rodrigues, F. A. (2019). Clustering algorithms: A comparative approach. *PLoS ONE*, *14*(1), 1-34. https://doi.org/10.1371/journal.pone.0210236

[17]. Aslam, S., Alam, F., Hasan, S. F., & Rashid, M. A. (2021). A machine learning approach to enhance the performance of D2D-enabled clustered networks. *IEEE Access*, 9, 16114-16132. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9328769

[18]. Huang, X., Zeng, M., Fan, J., Fan, X., & Tang, X. (2018a). A Full Duplex D2D Clustering Resource Allocation Scheme Based on a *K*-Means Algorithm. *Hindawi: Wireless Communications and Mobile Computing*, *2018*(Article ID 1843083), 1-8. Retrieved from https://www.hindawi.com/journals/wcmc/2018/1843083/

[19]. Lee, Y. H., Tseng, H. W., Lo, C. Y., Jan, Y. G., Chin, L. P., Song, T. C., Hsu, H. I. (2012). Using Genetic Algorithm with Frequency Hopping in Device to Device Communication (D2DC) interference mitigation. *2012 IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2012)*, November 4-7, 201-206. Retrieved from http://tkuir.lib.tku.edu.tw/dspace/retrieve/52760/B2.6.pdf

[20]. Midasala, V., Bhavanam, S.N., & Siddaiah, P. (2016). Performance analysis of LEACH protocol for D2D communication in LTE-advanced network. *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 1-3. Retrieved from https://www.researchgate.net/publication/316902746_Performance_analysis_of_LEACH_protocol_for_D2D_communication_in_LTE-advanced_network

[21]. Singh, M., Gupta, V. & Sona, D. R. (2018). Cluster formation using Leach Protocol and distance based resource allocation in D2D Communication". *International Journal of Electrical, Electronics and Data Communication (IJEEDC)*, *6*(6), 22-26. Retrieved from http://www.iraj.in/journal/journal_file/journal_pdf/1-478-153414118220-24.pdf

[22]. Zhao, Y., Liu, K., Xu, X., Yang, H., & Huang, L. (2019). Distributed Dynamic Cluster-Head Selection and Clustering for Massive IoT Access in 5G Networks. *Applied Sciences*, *9*(1), 132, MDPI AG, 1-15. Retrieved from http://dx.doi.org/10.3390/app9010132

[23]. Hashima, S.; ElHalawany, B.M.; Hatano, K.; Wu, K.; & Mohamed, E.M. (2021). Leveraging Machine-Learning for D2D Communications in 5G/Beyond 5G Networks. *Electronics*, *10*(169), 1-16. Retrieved from https://doi.org/10.3390/electronics10020169

[24]. Fodor, G., Parkvall, S., Sorrentino, S., Wallentin, P., Lu, Q, & Brahmi, N. (2014). Device-to-Device Communications for national security and public safety. *IEEE Access: Special Section on 5G Wireless Technologies: Perspectives of The Next Generation Mobile Communications And Networking*, *2*, 1510-1520. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6985517

[25]. Chaowen, L., Chun-long, H., Li, X., Huang, L., & Ying, L. (2018). Energy-efficient Device-to-Device communication overlaying cellular networks with cluster relays. *IOP Conf. Series: Journal of Physics:* Conf. Series 1069 012060, 1-9. Retrieved from https://iopscience.iop.org/article/10.1088/1742-6596/1069/1/012060/pdf

[26]. Nitti, M., Stelea, G. A. Popescu, V., & Fadda, M. (2019). When social networks meet D2D communications: A survey. *Sensors*, *19*(396), 2-16. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6359220/pdf/sensors-19-00396.pdf

[27]. Xu, W., Li, S., Xu, Y., & Lin, X. (2016) Underlaid-D2D-assisted cooperative multicast based on social networks. *Peer-to-Peer Networking & Applications*., *9*(5), 923–935. Retrieved from https://doi.org/10.1007/s12083-015-0348-9

[28]. Huang, H., Liu, B., Hu, M., Tao, Y., & Xiang, W. (2018b). User-Information-Aware D2D Multicast File Distribution Mechanism. *Sensors*, *18*(10), 1-20, 3389 MDPI AG. Retrieved from http://dx.doi.org/10.3390/s18103389

[29]. Cao, C. Wang, L., Song, M., & Zhang, Y. (2014). Admission policy based clustering scheme for D2D underlay communications. *IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*, 1-6. DOI: 10.1109/PIMRC.2014.7136488

[30]. Afshang, M., Dhillon, H. S., Han, P., & Chong, H. H. J. (2015). Coverage and area spectral efficiency of clustered device-to-device networks. *2015 IEEE Global Communications Conference (GLOBECOM 2015)*, San Diego, Calif, USA, pp. 1–6. DOI:10.1109/glocom.2014.7417450

[31]. Hassan, K. S., & Maher, E. M. (2013). Device-to-device communication distance analysis in interference limited cellular networks. *The Tenth International Symposium on Wireless Communication Systems (ISWCS 2013)*, Ilmenau, Germany. https://ieeexplore.ieee.org/document/6629748

[32]. Lu, Q., Miao, Q., Fodor, G., & Brahmi, N. (2014). Clustering Schemes for D2D Communications under Partial/No Network Coverage. *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*, 1-5. DOI: 10.1109/VTCSpring.2014.7022860

[33]. Zhou, Y. (2013). *Performance Evaluation of a Weighted Clustering Algorithm in NSPS Scenarios* (Thesis, Royal Institute of Technology (KTH), Stockholm, Sweden). Retrieved from http://www.diva-portal.org/smash/get/diva2:690435/FULLTEXT01.pdf

[34]. Udofia, K. M. (2020). Comparison of Signal Strength-Based and Distance-Based Cluster Head Selection for Cellular Network in the 2100Mhz Frequency Band. *Journal of Multidisciplinary Engineering Science Studies (JMESS)*, *6*(5), 3405-3415. Retrieved from http://www.jmess.org/wp-content/uploads/2020/09/JMESSP13420673.pdf

[35]. Maryam, H., Ghadah, A., & Nikos, D. & Win, M. (2019). Relay Selection Based Clustering Techniques for High Density LTE Network. *Wireless Networks*, *25*, 2305–2314. Retrieved from https://doi.org/10.1007/s11276-018-1658-7

[36]. Selmi, S., & Ridha, B. (2021). Energy efficient two-hop D2D communications underlay 5G networks: A Stackelber game approach. *Journal of Communications Software And Systems*, *17*(1), 50-58. Retrieved from https://hrcak.srce.hr/file/370278

[37]. Zhang, S. & Peng, Y. (2020). D2D communication relay selection algorithm based on Game Theory. *Procedia Computer Science: 3rd International Conference on Mechatronics and Intelligent Robotics (ICMIR-2019)*, *166*, 563-569. Retrieved from https://www.sciencedirect.com/science/article/pii/S1877050920301423/pdf?md5=5a8b9a50ea44383ba38e458e5494da09&pid=1-s2.0-S1877050920301423-main.pdf

[38]. Zhang, H., Wang, Z., & Du, Q. (2018). Social-aware D2D relay networks for stability enhancement: An optimal stopping approach. *IEEE Transactions on Vehicular Technology*, *67*, 8860-8874. DOI:10.1109/TVT.2018.2851779

[39]. Tian, C., Qian, Z., Wang, X., & Hu, L. (2019). Analysis of joint relay selection and resource allocation scheme for relay-aided D2D communication networks. *IEEE Access*, *7*, 142715-142725. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8852696

[40]. Zhang, M., Chen, X., & Zhang, J. (2014). Social-aware relay selection for cooperative networking: An optimal stopping approach. *2014 IEEE International Conference on Communications (ICC)*, 2257-2262. DOI:10.1109/ICC.2014.6883659

[41]. Militano, L.; Orsino, A.; Araniti, G.; Iera, A. (2017). NB-IoT for D2D-enhanced content uploading with social trustworthiness in 5G systems. *Future Internet*, *9*(31), 1-14. Retrieved from https://res.mdpi.com/d_attachment/futureinternet/futureinternet-09-00031/article_deploy/futureinternet-09-00031.pdf

[42]. Alnabhan, M. M., Al-jaafreh, S., Abadleh, A., Atoum, M., Hammouri, A., & Al-Dalahmeh, M. (2021). Enhanced D2D Communication Model in 5G Networks. *International Journal of Computing and Digital Systems*, *10*(1), 217-223. DOI: http://dx.doi.org/10.12785/ijcds/10012

[43]. Narottama, B., Fahmi, A., & Syihabuddin, B. (2015). Impact of number of devices and data rate variation in clustering method on Device-to-Device communication. *2015 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, 27-29 August. Bandung, Indonesia, 233-238. Retrieved from https://arfiantof.staff.telkomuniversity.ac.id/files/2017/04/12.-Impact-of-Number-of-Devices-and-Data-Rate-Variation-in-Clustering-Method-on-Deviceto-Device-Communication.pdf

[44]. Aslam, S., Alam, F., Hasan, S. F., & Rashid, M. (2020). A novel weighted clustering algorithm supported by a distributed architecture for D2D enabled content-centric networks. *Sensors*, *20*(5509), 1-25. Doi:10.3390/s20195509