# Single Linkage Weighted Steepest Gradient Adaboost Cluster-Based D2D in 5G Networks

Sridhar Varadala and S. Emalda Roslin

*Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, India*

https://doi.org/10.26636/jtit.2023.167222

**Abstract — Efficiency of data transmissions with minimum latency levels and better resource utilization is a challenging issue in 5G device-to-device (D2D) environments. A novel technique referred to as single linkage steepest gradient gentle AdaBoost cluster-based device (SLSGAC) is introduced to improve device-to-device communications with minimum latency. The proposed technique uses the ensemble clustering approach to group mobile devices by constructing a set of weak clusters, based on the Minkowski single linkage clustering technique. In the weak clustering process, residual energy, bandwidth and SINR are estimated, and mobile devices are grouped based on the Minkowski distance measure. Results of the weak clustering process are combined to provide the final ensemble's clustering output by applying the steepest gradient function to minimize the error rate. For each cluster, a head is selected from among the group members to improve the data transmission rate and minimize latency. Simulations are conducted comparing the proposed technique with the existing methods based on such metrics as energy efficiency, data delivery ratio, packet loss rate, throughput and latency.**

***Keywords — 5G networks, D2D communication, Minkowski single linkage clustering.***

## 1. Introduction

Among the many modern solutions, device-to-device (D2D) communication is an upcoming technology for enhancing the performance of 5G cellular communication. D2D is currently the subject of many scientific works aimed at expanding the application range, improving parameters and performance. Below is a short review of the important works in this area.

The social-aware D2D video delivery model based on rapid sample-efficient measurement (DMSEM) was introduced in [1]. It relies on the fuzzy c-means clustering algorithm. However, the designed algorithm failed to consider the process of accounting for mobility to improve D2D data transmissions. A dynamic mode selection algorithm for improving the communication based on the fuzzy clustering process was designed in [2]. Unfortunately, good performance of efficient latency-aware data communication was not achieved. A novel cluster-based architecture was introduced in [3] for D2D mmWave communication to enhance the performance of 5G networks. The intended architecture succeeds in increasing throughput but not in increasing data delivery. A fuzzy clustering-based approach was introduced in [4] for a 5G ultra-dense network in order to improve the data transmission rate. But accurate clustering of mobile devices with multiple characteristics was not considered. A novel weighted clustering algorithm for the distributed architecture was proposed in [5] to improve throughput and minimize energy consumption. But the designed method failed to apply machine learning techniques to further improve data delivery. In [6], random-based clustering (RBC) and channel-gain-based clustering (CG-BC) approaches were introduced to realize D2D clustering. The clustering algorithm decreases the delay but no higher delivery rates and lower packet loss levels were achieved. In [7], a random linear network coding-based cooperative relaying (RNCC) system was developed for network coding supported by D2D transmissions based on resource distribution. The drawback of this method is that the multiple-hop model-based D2D transmission for network coding-assisted D2D communications was not provided.

A centralized non-uniform clustering routing protocol was introduced in [8], along with the power estimation as well as communication cost. Unfortunately, the performance of latency-aware data transmission was not verified. Various issues, solutions and challenges were reviewed in [9] to efficiently utilize the available resources and thus minimize latency and enhance data transmission rates.

A VNF resource allocation system was introduced in [10] to obtain the best number of clusters for reducing delay in data transmission. An artificial neural network-based channel model framework was introduced in [11] for 5G wireless communication. But cluster-based data delivery was not performed to minimize latency. A data-driven multi-objective optimization approach was introduced in [12] for the purpose of developing hyperdense 5G networks. Unfortunately, no analyses concerned with a wider span of network communications were performed in order to minimize latency and improve reliability. A MIMO-NOMA cellular system approach for D2D communications was introduced in [13]. But the efficiency of the algorithms was not improved.

In [14], an intensive benchmarking of the incorporation of D2D communication was discussed to improve energy efficiency, throughput and latency. The reinforcement learning-based latency-controlled D2D connectivity (RL-LCDC) method was introduced in [15]. Though the designed method minimizes the delay and energy consumption, the packet loss rate was not reduced. A distributed intelligent method was developed in [16] to enhance the data rate and reduce power consumption in D2D communications. But no latency-aware D2D communication was performed. Several machine learning techniques were introduced in [17] for 5G cellular networks to improve the data transmission rate. The sparse code multiple access (SCMA) method was introduced in [18] for enhancing the overall network performance of D2D-enabled cellular networks. The dynamic resource block (RB) sharing method was introduced in [19] to enhance the network throughput of the overall system. However, the designed method failed to minimize the latency of data communication.

To solve such issues, a technique called SLSGAC is introduced, offering the following contributions:

– it is based on the ensemble clustering approach to boost D2D communication,
– it uses weak clusters as well as the Minkowski single linkage clustering technique to divide the cellular network into different groups based on residual energy, bandwidth, and SINR. It uses the steepest gradient function into the ensemble clustering technique to minimize the error rate and obtain accurate clustering results,
– to improve data delivery and minimize latency, a cluster head is selected for each cluster,
– finally, a simulation is performed using several performance measures versus the traditional approaches.

The paper is organized into different sections. Section 2 describes the proposed SLSGAC technique. Section 3 presents the simulation of the proposed technique and conventional methods. Finally, Section 4 contains the summary of the paper.

## 2. Proposed Method

Figure 1 illustrates the data exchange between mobile devices and the base station. Two mobile devices communicate with each other in a direct manner and via the links provided from the base station. In such a way, communication is entirely handled by the base station.

In the course of direct communication between the devices, minimum utilization of resources is essential for increasing

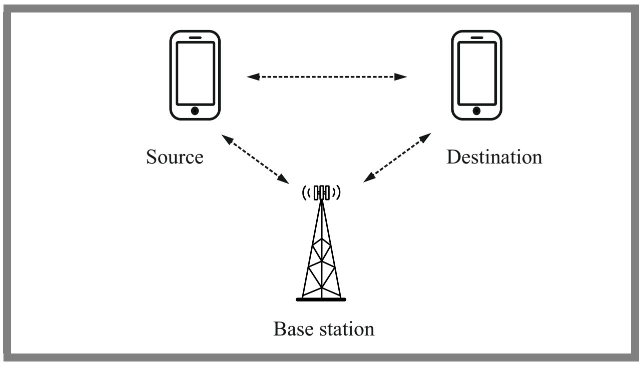


**Fig. 1.** Data flow in a 5G network.

the network's lifespan. In order to achieve resource-efficient data transmission rates, a novel cluster-based technique is introduced.

Let us consider a scenario in which 5G network comprises a number of mobile devices $d_1, d_2, d_3, \ldots, d_n$ distributed randomly within the network. Each mobile device has different computational resources available, such as energy $E$, bandwidth $B$, and signal to noise-plus-interference ratio (SINR).

Figure 2 illustrates the concept behind the SLSGAC technique to enhance data communication between mobile devices within a 5G network. The number of mobile devices is considered as input. Then, the ensemble clustering approach is applied to partition the entire hole network into different clusters depending on such resources as energy, bandwidth, and SINR. For each cluster, a head is selected among the individual members to additionally boost data transmission.

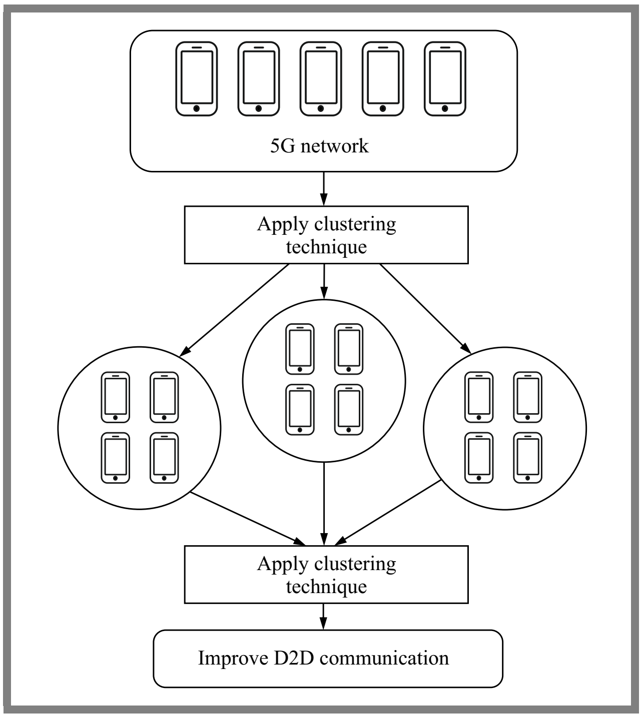


**Fig. 2.** Diagram of architecture of SLSGAC technique.

The single linkage AdaBoost clustering technique is an ensemble meta-algorithm that primarily converts weak learners' output into that of strong learners. A weak learner $z_1, z_2, z_3, \ldots, z_k$ is defined as a base clustering technique that is only slightly correlated with the true clustering results. In contrast, a strong learner is a clustering technique that offers true clustering outcomes. The weak clustering results are merged and the final, improved clustering results are obtained.

Figure 3 illustrates the basic process of ensembling clustering results. The mobile devices are presented, as input, to the clustering technique. Then $k$ weak learners are constructed using the Minkowski single linkage clustering technique to partition the entire network into different groups depending on energy, bandwidth, signal to interference, and noise ratio.

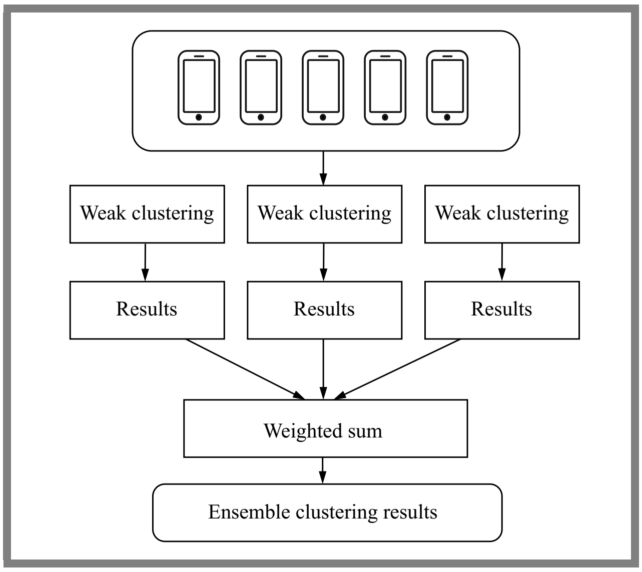


**Fig. 3.** Basic process of ensembling clustering results.

In the first step, energy is measured in terms of power and time as:

$$E_{device} = p_t t \ , \tag{1}$$

where $E_{device}$ denotes the energy of the mobile device, $p_t$ indicates the transmitting power of the mobile device, $t$ represents time in seconds. The remaining (i.e. residual) energy of each device is measured as:

$$RE_{device} = E_t - E_C \ , \tag{2}$$

where $RE_{device}$ indicates the remaining energy of the device in Joules, $E_t$ indicates the total energy, and $E_C$ denotes the consumed energy.

Bandwidth is estimated as the maximum amount of data transmitted in a given time slot and it expressed as:

$$BW = \frac{D_p}{t} \ , \tag{3}$$

where $BW$ denotes the bandwidth, $t$ is time in seconds, and $D_p$ denotes a data packet.

The signal to interference and noise ratio $R_{sin}$ is estimated as:

$$R_{sin} = \frac{s_p}{n_p + i_p} \ , \tag{4}$$

where $s_p$ denotes signal power, $n_p$ indicates noise power, and $i_p$ denotes interference power.

Based on the above parameters, the single linkage clustering technique is applied to group the sensor nodes based on the Minkowski distance. The method identifies the minimum distance between mobile devices based on the energy, bandwidth, and SINR. The distance is formulated as:

$$D_{ij} = \left(|d_i - d_j|^r\right)^{\frac{1}{r}} \ , \tag{5}$$

where $D_{ij}$ denotes the distance between two mobile devices $d_i$ and $d_j$, $r$ denotes the Minkowski distance of order ($\geqslant 1$). The minimum distance between two mobile devices is grouped and clusters are formed:

$$Q = \arg\min(D_{ij}) \ , \tag{6}$$

where $Q$ denotes a base clustering result, $\arg\min$ denotes an argument of the minimum function. Consequently, weak learner results are obtained. In order to attain the ensemble clustering results, the weak learners are summarized:

$$Z = \sum_{i=1}^{k} Q_i \ , \tag{7}$$

where $Z$ denotes ensemble clustering results, $Q_i$ denotes weak classification results. The weighted sum of the weak learner results is obtained as:

$$Z = \sum_{i=1}^{k} Q_i \beta_t \ , \tag{8}$$

where $\beta_t$ indicates the weight of clustering results and the weight is an integer. For each weak clustering result obtained, the training error is estimated as:

$$E(f) = (R_A - Q_i)^2 \ , \tag{9}$$

where $E(f)$ denotes an error function, $R_A$ denotes the actual clustering result, $Q_i$ denotes the clustering results observed.

Next, the proposed technique updates the initial weight based on the error value. Finally, the steepest gradient descent is applied to find the weak clustering results with a minimum error among the multiple weak clustering results:

$$Y = \arg\min E(f) \ , \tag{10}$$

where $Y$ indicates the strong clustering results, $\min E(f)$ denotes the minimum argument of the error function of the clustering results. In this way, all sensor nodes are grouped into a particular cluster. For each cluster, the mobile device with the highest level of energy, bandwidth, and SINR is selected as a cluster head to ensure efficient data transmission from the source to the destination.

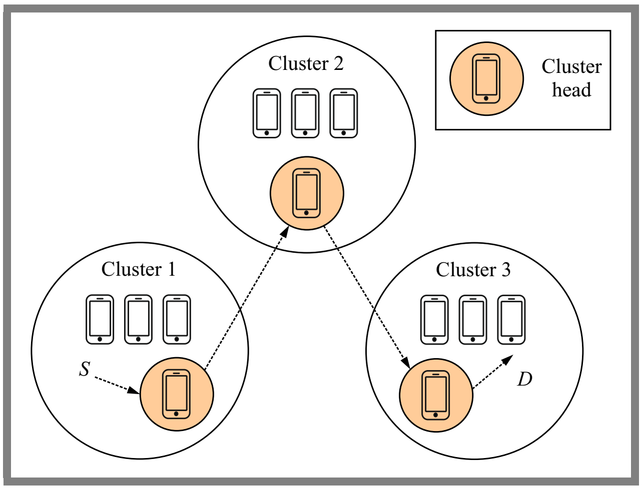


**Fig. 4.** Single linkage weighted steepest gradient AdaBoost cluster-based D2D communication.

Figure 4 illustrates the process of cluster-based data transmission from source mobile device $S$ to destination $D$ via a cluster head. As shown in Fig. 4, the source device in cluster 1 sends the data to the destination in cluster 3. This process is carried out via the cluster head, resulting in increases in the data delivery and in minimized latency. Algorithm 1 describes the technique's algorithmic phase.

**Algorithm 1.** Description of the proposed technique.

**Input**: number of mobile devices: $d_1, d_2, d_3, \ldots, d_n$
**Initialization**

**1: For** each mobile device $d_i$
**2:** Construct $k$ weak learners $z_1, z_2, z_3, \ldots, z_k$
**3:** **Measure** $RE_{device}$, $BW$, SINR
**4:** **Estimate** Minkowski distance $D_{ij}$
**5:** Find minimum distance $\arg\min[D_{ij}]$
**6:** Group sensor nodes into clusters
**7:** Obtain clustering results
**8: End for**
**9:** Sum all weak clustering results $\sum_{i=1}^{k} Q_i\beta_t$
**10:** Initialize weight $\sum_{i=1}^{k} Q_i\beta_t$
**11:** Compute error function $E(f)$
**12:** Update weight
**13:** Find the weak learner with minimum error $\arg\min E(f)$
**14:** Return (strong clustering results)

The ensemble boost clustering technique computes the number of weak learners. Next, residual energy, bandwidth, and signal to noise ratio of each device are estimated and the mobile devices are grouped into a particular cluster. The weak learner's results are summarized to create a strong one in order to achieve accurate clustering results. After combining the weak learners' results, the weight is initialized and measures the training error function. Then, the initial weight gets updated based on the error value. Finally, the steepest gradient function is applied to find the weak learner with the minimum error value. Once the devices are grouped, a cluster head is selected for efficient communication. Then, data transmission is performed via the cluster head to achieve a higher delivery ratio and minimum latency.

## 3. Simulation and Analysis

In this section, a Matlab simulation of the proposed SLSGAC technique is performed and the results are compared with existing DMSEM [1] and the dynamic mode selection method based on a fuzzy clustering algorithm [2], with performance measured in terms of energy efficiency, data delivery ratio, packet loss rate, throughput, and latency.

Energy efficiency is measured as the fraction of output energy required to perform communication:

$$EE = \frac{E_o}{E_I} \cdot 100\ [\%]\ , \tag{11}$$

where $EE$ denotes energy efficiency, $E_o$ indicates output energy, and $E_I$ denotes input energy.

The simulation of energy efficiency of various methods based on 100, 200, 300... 1000 mobile devices is presented in Fig. 5. The obtained values indicate that the energy efficiency of all three methods increases with the growing number of mobile devices. For comparison, the proposed SLSGAC technique provides better performance than the conventional methods.

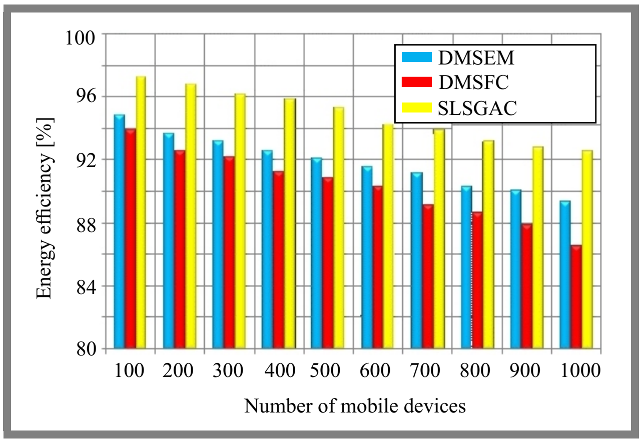


**Fig. 5.** Performance results – energy efficiency.

Data delivery ratio is a performance metric based on data packets successfully transmitted from the base station to a mobile device. This parameter is also used to estimate the confidential data transmission between devices:

$$DR = \frac{Data\ received}{Data\ sent} \cdot 100\ [\%]\ . \tag{12}$$

Figure 6 presents experimental data delivery outcomes for varying numbers of mobile devices. For each method, ten iterations are performed relating to different counts of mobile devices. While increasing the input count, the data delivery ratio for all the methods becomes lower. a comparison of all methods shows that the SLSGAC technique outperforms other approach in achieving a higher delivery ratio.

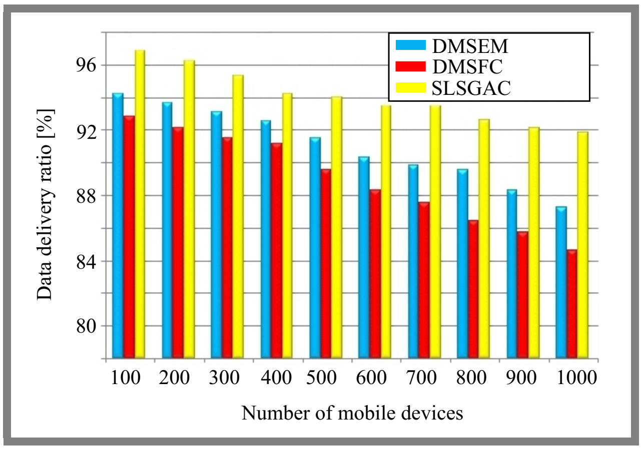


**Fig. 6.** Performance results – data delivery ratio.

Figure 6 illustrates the performance of the data delivery ratio versus the number of mobile devices in the simulation scenario. The SLSGAC technique achieves a delivery ratio of 96.8% in the first iteration, with 100 devices present.

Similarly, DMSEM [1] and dynamic mode selection based on a fuzzy clustering algorithm [2] generate delivery ratios of 94.2% and 92.8%, respectively. The results indicate that the SLSGAC technique provides better performance due to using the ensemble technique which groups the sensor nodes based on residual energy, bandwidth, and signal to interference ratio. For each cluster, a head is chosen for delivering the data packets from the source to the destination.

Packet loss rate is estimated based on the ratio between data packets lost and the number of data packets transmitted:

$$PLR = \frac{Data\ lost}{Data\ sent} \cdot 100\ [\%]\ . \tag{13}$$

Figure 7 shows a comparison of the packet loss rates. 100 mobile devices were used for conducting the simulation in the first iteration. For SLSGAC, the packet loss rate is 3.2%, whereas for DMSEM and dynamic mode selection based on fuzzy clustering algorithm, the results equal 5.8% and 7.2%, respectively. The SLSGAC technique minimizes the packet loss rate more effectively than the existing methods, due to the selection of cluster heads among the group members. The source mobile device sends the data to its destination via a neighboring cluster head. This process helps minimize data transmission losses.

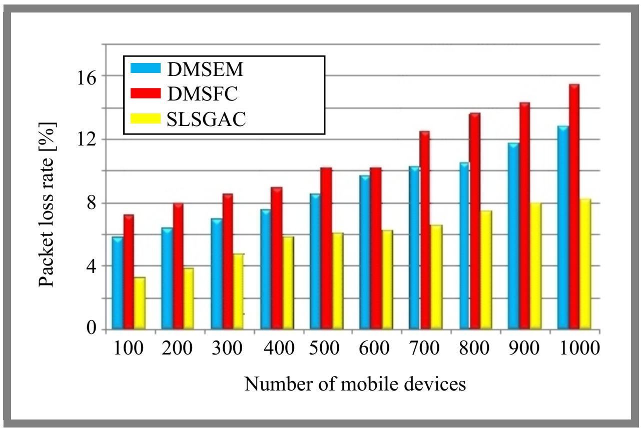


**Fig. 7.** Performance results – packet loss rate.

Throughput is estimated as the amount of data sent from one base station to mobile devices in a specified time frame. Throughput is computed in bits per second, using the following formula:

$$Throughput = \frac{Size\ of\ data\ received}{Time}\ [bps]\ . \tag{14}$$

Figure 8 shows that the proposed SLSGAC technique provides a higher network throughput due to the fact that higher residual energy requirements are applied when more bandwidth-efficient devices are selected as cluster heads, thus resulting in an improvement in data transmission speed.

Figure 9 illustrates the simulation results concerning latency measured for the number of mobile nodes varying from 100 to 1000. Latency of the SLSGAC technique is reduced when compared to other existing methods in 5G D2D communication. For 100 mobile devices, latency of the SLSGAC technique is 11.5 ms in the first iteration. Similarly,

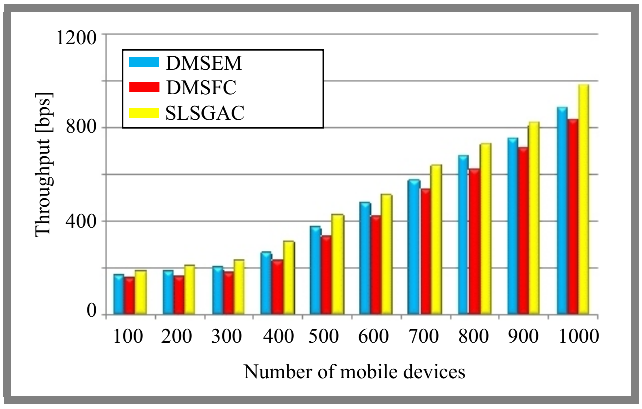


**Fig. 8.** Performance results – throughput.

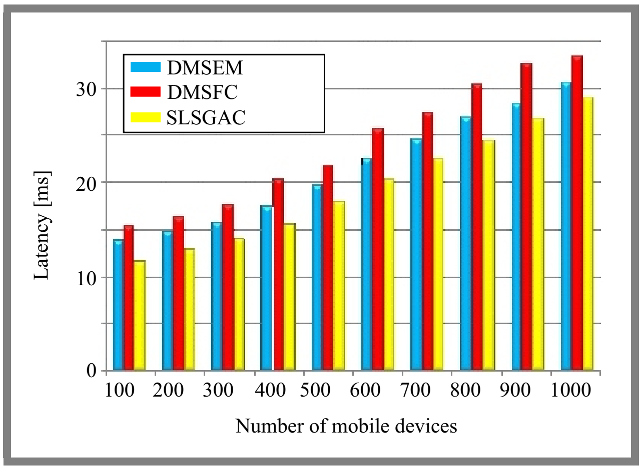


**Fig. 9.** Performance results – latency.

values of 13.7 ms and 15.3 ms were observed in the case of DMSEM and dynamic mode selection based on a fuzzy clustering algorithm, respectively. This is due to the application of the ensemble clustering technique, where residual energy, bandwidth, and signal to interference ratio are estimated and the clustering the mobile devices is based on the Minkowski distance measure. The weak clustering results are combined to provide the final ensemble clustering output by applying the steepest gradient function in order to find efficient mobile devices and achieve more efficient data communication between devices while maintaining minimum latency levels.

## 4. Conclusion

The simulation results prove that the proposed SLSGAC technique provides better performance than the traditional approaches in terms of achieving higher throughput, data delivery, and minimum latency. When compared to conventional cellular networks, D2D communication is anticipated to offer a number of advantages. Future networks are most likely to benefit from and be most successful with D2D technology. In the study, we focused on a thorough analysis of the D2D technologies now in use, along with their features, including device detection, interference management, security, mode selection, and power control. We examine a number of ideas that have been proposed in an effort to ensure secure device-to-device (D2D) connectivity in 5G. By highlighting the drawbacks or faults, as well as the suggested fixes, we emphasize the current solutions.

## References


[1] R. Zhang, S. Jia, Y. Ma, and Ch. Xu, "Social-aware D2D video delivery method based on mobility similarity measurement in 5G ultra-dense network", *IEEE Access*, vol. 8, pp. 52413–52427, 2020 (https://doi.org/10.1109/ACCESS.2020.2980865).

[2] A.A. Algedir and H.H. Refai, "Energy efficiency optimization and dynamic mode selection algorithms for D2D communication under HetNet in downlink reuse", *IEEE Access*, vol. 8, pp. 95251–95265, 2020 (https://doi.org/10.1109/ACCESS.2020.2995833).

[3] G. Mesbahi and A.G. Rahbar, "Cluster-based architecture capable for device-to-device millimeter-wave communications in 5G cellular networks", *Arabian Journal for Science and Engineering*, vol. 44, pp. 9719–9733, 2019 (https://doi.org/10.1007/s13369-019-03830-w).

[4] Q. Liu, Ch.F. Kwong, S. Zhang, L. Li, and J. Wang, "A fuzzy-clustering based approach for MADM handover in 5G ultra-dense networks", *Wireless Network*, vol. 28, pp. 965–978, 2022 (https://doi.org/10.1007/s11276-019-02130-3).

[5] S. Aslam, F. Alam, S.F. Hasan, and M. Rashid, "A novel weighted clustering algorithm supported through a distributed architecture for D2D enabled content-centric networks", *Sensors*, vol. 20, no. 19, Article number 5509, 2020 (https://doi.org/10.3390/s20195509).

[6] M. Gharbieh, A. Bader, H. ElSawy, H.-Ch. 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 (https://doi.org/10.1109/TCOMM.2018.2876008).

[7] X. Xiao, M. Ahmed, X. Chen, Y. Zhao, Y. Li, and Z. Han, "Accelerating content delivery via efficient resource allocation for network coding aided D2D communications", *IEEE Access*, vol. 7, pp. 115783–115796, 2019 (https://doi.org/10.1109/ACCESS.2019.2930728).

[8] W. Chen, B. Zhang, X. Yang, W. Fang, W. Zhang, and X. Jiang, "C-EEUC: A cluster routing protocol for coal mine wireless sensor network based on fog computing and 5G", *Mobile Networks and Applications*, vol. 27, no. 5, pp. 1853–1866, 2022 (https://doi.org/10.1007/s11036-019-01401-9).

[9] M.H. Adnan and Z.A. Zuckarnain, "Device-to-device communication in 5G environment: Issues, solutions, and challenges", *Symmetry*, vol. 12, no. 11, 1762, 2020 (https://doi.org/10.3390/sym12111762).

[10] S. Song, Ch. Lee, H. Cho, G. Lim, and J.-M. Chung, "Clustered virtualized network functions resource allocation based on context-aware grouping in 5G edge networks", *IEEE Transactions on Mobile Computing*, vol. 19, no. 5, pp. 1072–1083, 2020 (https://doi.org/10.1109/TMC.2019.2907593).

[11] J. Huang, Ch.-X. Wang, L. Bai, J. Sun, Y. Yang, J. Li, O. Tirkkonen, and M.-T. Zhou, "A big data enabled channel model for 5G wireless communication systems", *IEEE Transactions on Big Data*, vol. 6, no. 2, pp. 211–222, 2020 (https://doi.org/10.1109/TBDATA.2018.2884489).

[12] B.B. Haile, E. Mutafungwa, and J. Hämäläinen, "A data-driven multi-objective optimization framework for hyperdense 5G network planning", *IEEE Access*, vol. 8, pp. 169423–169443, 2020 (https://doi.org/10.1109/ACCESS.2020.3023452).

[13] S. Solaiman, A. Nassef, and E. Fadel, "User clustering and optimized power allocation for D2D communications at mm wave underlaying MIMO-NOMA cellular networks", *IEEE Access*, vol. 9, pp. 57726–57742, 2021 (https://doi.org/10.1109/ACCESS.2021.3071992).

[14] H.H. Hussein, H.A. Elsayed, and S.M. Abd El-kade, "Intensive benchmarking of D2D communication over 5G cellular networks: Prototype, integrated features, challenges, and main applications", *Wireless Networks*, vol. 26, pp. 3183–3202, 2020 (https://doi.org/10.1007/s11276-019-02131-2).

[15] A.G. Sreedevi and T.R. Rao, "Reinforcement learning algorithm for 5G indoor device-to-device communications", *Transaction on Emerging Telecommunication Technologies*, vol. 30, no. 9, pp. 1–10, 2019 (https://doi.org/10.1002/ett.3670).

[16] I. Ioannou, V. Vassiliou, Ch. Christophorou, and A. Pitsillides, "Distributed artificial intelligence solution for D2D communication in 5G networks", *IEEE Systems Journal*, vol. 14, no. 3, pp. 4232–4241, 2020 (https://doi.org/10.1109/JSYST.2020.2979044).

[17] S. Sevgican, M. Turan, K. Gökarslan, H.B. Yilmaz, and T. Tugcu, "Intelligent network data analytics function in 5G cellular networks using machine learning", *Journal of Communications and Networks*, vol. 22, no. 3, pp. 269–280, 2020 (https://doi.org/10.1109/JCN.2020.000019).

[18] A. Sultana, I. Woungang, A. Anpalagan, L. Zhao, and L. Ferdouse, "Efficient resource allocation in SCMA-enabled device-to-device communication for 5G networks", *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5343–5354, 2020 (https://doi.org/10.1109/TVT.2020.2983569).

[19] P.K. Barik, A. Shukla, R. Datta, and Ch. Singhal, "A resource sharing scheme for intercell D2D communication in cellular networks: A repeated game theoretic approach", *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7806–7820, 2020 (https://doi.org/10.1109/TVT.2020.2991476).



**Sridhar Varadala**
Research Scholar
http://orcid.org/0000-0003-3715-8197
E-mail: varadalasri@gmail.com
Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, India

**S. Emalda Roslin, Ph.D.**
http://orcid.org/0000-0001-8699-9510
Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, India