Machine Learning-Based Optimization of Cluster Formations in 6G Networks

Luis Antonio L. F. da Costa*, Rafael Kunst*, Rodrigo C. de Lamare †, Edison Pignaton de Freitas ‡
*University of Vale do Rio dos Sinos (UNISINOS)
{lalfcosta, rafaelkunst}@unisinos.br
†Pontifical Catholic University - Rio de Janeiro (PUC-RIO)
delamare@puc-rio.br
‡Federal University of Rio Grande do Sul, Brazil
edison.pignaton@inf.ufrgs.br

***Abstract*—The advent of 6G networks promises unprecedented advancements in communication technology, characterized by ultra-low latency, massive connectivity, and enhanced data rates. One of the critical challenges in realizing the full potential of 6G networks, especially in the context of FANETS, is optimizing cluster formations to ensure efficient resource allocation and network performance. This paper explores the application of machine learning techniques to address this challenge. We propose a novel framework that leverages machine learning techniques to dynamically optimize cluster formations in 6G networks. Our approach integrates network data, user mobility patterns, and traffic demands to predict optimal clustering configurations. Through extensive simulations, we demonstrate that our method exhibits promising performance results in terms of lower delay, jitter, and higher throughput. The results also highlight the potential of deep learning to enhance the adaptability and robustness of 6G networks, paving the way for future research and development in this emerging field.**

***Index Terms*—Flying Ad-Hoc Networks, 6G, Machine Learning, Network Clustering.**

I. Introduction

The sixth generation (6G) of wireless networks is ready to revolutionize communication by offering unparalleled data rates, ultra-low latency, and massive connectivity. Advancements in the fields of electronics, sensors, and communication systems has enabled the development of small unmanned aerial vehicles (UAVs). However, the capacity of a single UAV may be insufficient, hence the utilization of several UAVs is necessary to overcome performance constraints and create a more advanced system. A flying ad hoc network (FANET) is a network composed of a collection of small UAVs that are interconnected and are able to work together as a team to accomplish complex objectives [1].

Nevertheless, FANETs must ensure communication stability and have high scalability. Recently, researchers have used clustering techniques to address routing problems in FANETs [2]. In the clustering process, the network is categorized into small groups called clusters, with each cluster incorporating cluster head nodes (CHs) and cluster member nodes (CMs) [3]. The most important step in network clustering is the selection of the clusters heads, since the CH is responsible for managing the cluster and establishing inter-cluster and intra-cluster communication [4]. Therefore, key aspects such as choosing the best CHs, managing network topology, and the choice of the routing protocol for the CMs, are essential steps to improving network efficiency in the clustering process for FANETs.

In this work, we present a novel framework that leverages machine learning techniques to dynamically optimize cluster formations in 6G networks by analyzing user mobility patterns in order to predict optimal clustering configurations. Through detailed simulations, we demonstrate that our approach is able to reduce communication delay, jitter and offer high data throughput values in scenarios of both centralised and decentralised network topologies for 5G and 6G communication environments, respectively. The results also highlight the potential of deep learning to enhance the adaptability and robustness of 6G networks, paving the way for future research and development in this emerging field.

The remainder of this paper is organized as follows. Section II reviews essential background about network clustering in FANETs, as well as the relevant literature in this area. The proposed system model is presented in section III. The simulation scenario, performed experiments, and obtained results are presented and discussed in IV. Finally, concluding the paper, Section V presents final remarks and directions for future investigations.

II. Related Works

In order for FANET to meet the requirements of a wide variety of applications - such as rescue operations, traffic surveillance, agriculture, border supervision, forest fire observations etc [5] - the primary requirement is to maximise the coverage of a particular region of interest by a group of UAVs. To achieve this goal, path planning methods, which mainly involve routing mechanisms, are essential.

Nevertheless, routing mechanisms for FANETs must take important network constraints in consideration such as quick movement and limit energy in a highly dynamic topology environment. In order to construct and balance the network topology, the cluster head (CH) election, cluster formation, and cluster management must be performed efficiently.

To deal with the cluster head selection in an efficient way, the work of [6] proposed a clustered weighted scheme with a redundant cluster-head to ensure end-to-end communication in critical 6G infrastructures. The promising results obtained from MATLAB simulations showed that the proposed scheme is efficient in choosing a new cluster head when the network changes.

Besides machine learning techniques, nature-inspired algorithms can also provide suitable solutions to network clustering in FANETs. In the work of [7], the authors suggested the use of an intelligent clustering scheme based on the whale optimization algorithm (WOA) called ICW for FANETs, and having shown that their approach presents a low clustering time, as well as optimizing energy consumption, network longevity, packet delivery rate, routing overhead, and delay, when compared to similar approaches. In a similar fashion, the authors of [8]. also created a bio-inspired solution for network clustering in a FANET, with an inter-cluster routing protocol (ICRP), based on a hybrid ant colony algorithm that was able to optimize relay node selection and enhanced routing stability. The comparative simulations showed that this protocol surpasses ad-hoc on-demand distance vector(AODV), fuzzy-logic-assisted-AODV, and Enhanced-Ant-AODV routing protocols in packet delivery rate and end-to-end transmission delay.

III. Network Clustering

This section outlines the design details the proposed system model for network clustering, as well as the optimization problem envisioned for the cluster head selection, and the application scenario of both network topologies: a centralized topology for 5G networks, and then a distributed topology for 6G networks.

A. System Model

In order to perform network clustering [9], a system model was designed following a set of steps and tasks. Figure 1 summarizes the workflow of our proposed approach for network clustering in order to create a suitable communication environment for testing simulations in both 5G and 6G contexts.

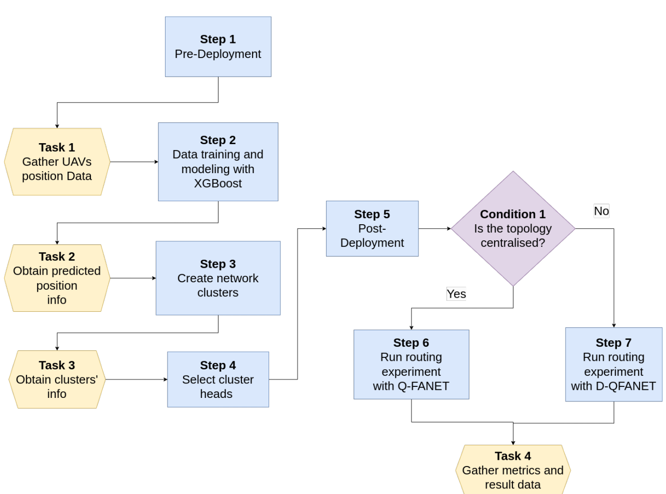


Fig. 1: System workflow for network clustering and routing simulation.

The network clusters are created using the Elbow method [10] and the Knee Point method [11] to automate the elbow method by mathematically identifying the point of maximum curvature, enhancing objectivity and reproducibility.

For training of the *XGBoost* [12] model to predict the final position of mobile stations, a set of parameters has been chosen:

- **Loss function:** Squared Error
- **Evaluation metric:** Root Mean Squared Error
- **Decision Tree maximum depth:** 6
- **Learning rate:** 0.1
- **Fraction of features to be randomly sampled for constructing each tree:** 1
- **Fraction of the training data to be randomly sampled for growing each tree:** 1

These parameters strike a balance between performance and preventing overfitting. They are commonly used in regression tasks such as predicting the final positions of elements in a grid based on input features [13].

B. Network Model

FANET's communication architecture is divided into two categories based on its connectivity, such as centralized and decentralized (distributed). In 5G networks, for example, a centralized network topology is necessary to handle complex interactions between various network components and efficiently manage network resources [14].

Figure 2 details the centralized network topology proposed for the 5G network architecture, with our proposed approach creating the clusters and selecting the appropriate cluster heads. It comprises a centralized server connected to a group of several UAVs. The server is then connected to the UPF (User Plane Function) module that connects to the Data Network in the Data Plane. This module also connects directly into the 5G core network, which consists of several key components/modules, each serving a specific purpose in facilitating advanced telecommunications services. The Q-FANET routing algorithm [15] feeds the routing information through the 5G core architecture and passes it to the network. Finally, the Data Plane is composed of a network of several connected switches that are responsible for handling the routing of large amounts of data, if necessary.

Unfortunately, the current 5G architecture does not actually support network intelligence. Hence, one of the goals of the future 6G core networks is to provide network intelligence to manage its network service function and the element of the architecture. For this reason, a new module is present in the 6G core network, the Network Data Analytics Function (NWDAF) module. NWDAF is responsible for collecting, analyzing, and providing network data analytics. Therefore, with the generated routing information, the NWDAF module determines the routing policy for the controller, which can then establish the optimal routing scheme for the distributed topology.

Figure 3 details the distributed network topology proposed for the 6G network architecture, composed of three distributed servers, each one connected to different groups of several

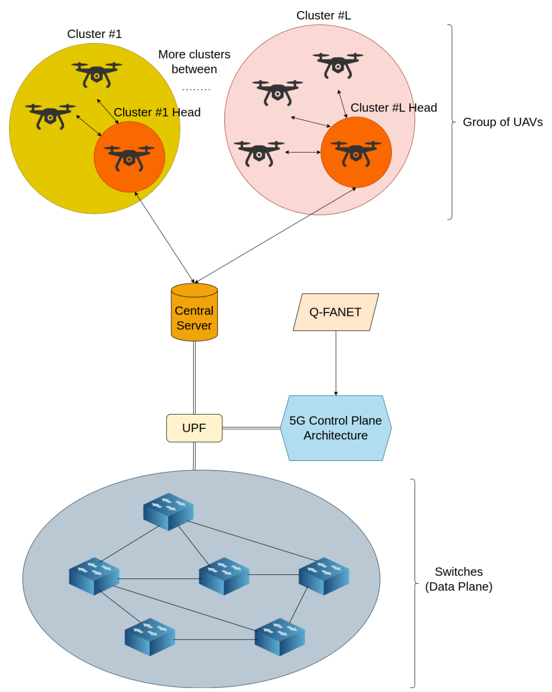


Fig. 2: A centralized network topology design in 5G Core Networks.

UAVs. The server is then connected to the UPF module that connects to the Data Network in the Data Plane. This module also connects directly into the 6G core network, with the D-QFANET routing algorithm [16] feeding routing information through the 6G control plane architecture and passing it to the network. Finally, the Data Plane is similar in design and function to the one present on the centralized topology for 5G networks.

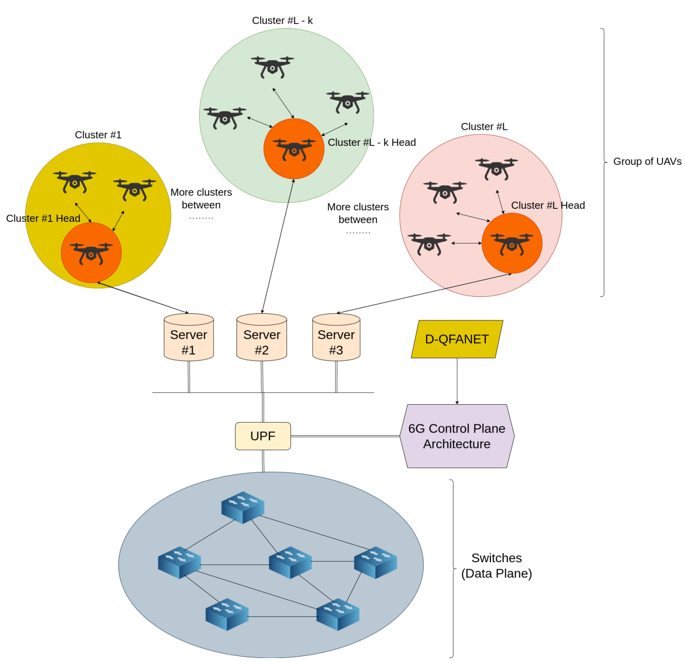


Fig. 3: A distributed network topology design in 6G Core Networks.

As can be observed for both the network topologies, the network clustering approach is able to eliminate the necessity of Access Points, more specifically, RAN (Radio Access Network) antennas, and NG-RAN (Next-gen Radio Access Network) antennas, for 5G and 6G, respectively. The cluster heads become the main communication receptors that will forward information packets from their respective UAV clusters to the servers, and backwards, creating a ad-hoc or infrastructureless network, where devices communicate directly with each other.

C. Cluster head selection

The main objective of the network clustering is to create a network environment that offers the lower communication delay possible between the mobile stations and host servers. Hence, the goal is to minimize an objective function, e.g, the chosen cluster head should minimize the weighted sum of distances and maximize the signal power to all other stations within the same cluster, while satisfying the constraint that there is exactly one cluster head per cluster. This problem is mathematically expressed by Equation 1:

$$\min_{w,x_i} \sum_{l=1}^{L} \Big(\sum_{i=1}^{M} \sum_{j=1, j\neq i}^{M} d_{ij} x_i - w \sum_{i=1}^{M} \sum_{j=1, j\neq i}^{M} p_{ij} x_i \Big) \quad \text{subject to} \sum_{i=1}^{M} x_i = 1, \tag{1}$$

where the variables are defined as:

- x_i: Binary variable indicating whether station i is chosen as the cluster head (1 if chosen, 0 otherwise)
- d_{ij}: Distance between station i and station j within the same cluster
- p_{ij}: Signal power of station i in relation to station j within the same cluster
- w: Weight factor that balances the importance of distance and signal power.

For model simplification purposes, the solution is obtained by a composite score computed for each station that considers the minimal distance and the stronger signal power from the other mobile stations inside the cluster. Moreover, this solution will identify the station within each cluster that optimizes the trade-off between minimizing the distance to other stations and maximizing the signal power to them, resulting in the lowest possible communication delay within the cluster.

First, the Received Signal Strength Indicator (RSSI) between the candidate cluster head and all other stations in the cluster is measured. Then, the Euclidean distance between the candidate cluster head and all other stations is computed. Next, the score for each station is computed as the difference between the average signal strength and the average distance to other stations in the cluster, giving priority to stations that are closer to others while also having strong signal strength. The station with the highest score within each cluster is selected as the cluster head, ensuring that the selected cluster head has both a strong signal and is centrally located relative to the other stations within the cluster.

IV. Experiments and Results

In order to evaluate the performance of the proposed system model for network clustering with different network architectures and topologies, a set of experiments is proposed, using the Mininet Wi-Fi simulator [17]. This group of simulations is not yet integrated with the core architectures of both 5G and 6G networks and deal only with the centralised and decentralised topologies.

A. Simulation setup

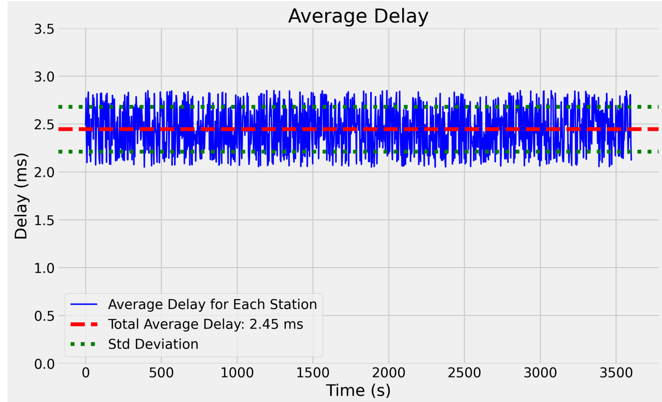
The simulation is based on a a 3GPP video traffic model [18], which involves generating packets based on a realistic video streaming pattern. The packet size and inter-arrival times are determined according to the video encoding (e.g., H.264/AVC). For this setup, is set an average packet size of 1024 bytes (with a variation of 256 bytes) and a average inter-arrival time of $30ms$ based on a Poisson process. Moreover, all the packets are sent through the network using UDP (User Datagram Protocol) sockets, in order to meet the stringent latency requirements of ultra-fast networks for 5G and 6G communication environments.

In the simulations, 100 UDP packets of various sizes are sent from each of the 25 mobile stations to their respective cluster heads, and to the central server (centralized topology) or to the distributed servers (distributed topology), for a period of $3600s$. In both scenarios, each UAV is assigned a random signal power between 60 and 80 dBm. Since both simulation scenarios will be run with the same clustering model, the network topologies will have the same number of clusters (3) and selected cluster heads. The complete set of parameters used for the experiments is the same used in [16].

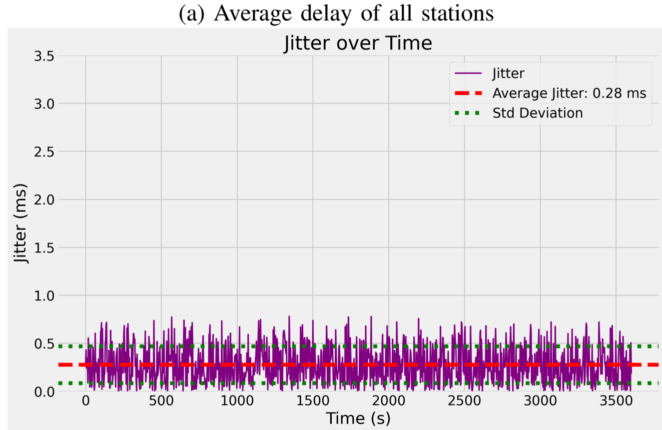
B. Results and Discussion

By comparing the two sets of results obtained in Figures 4 and 5, it can be observed that the proposed system model approach presents better performance and improvements in terms of low delay (16.3%) and jitter (51%) and high throughput (15.5%) when compared to the centralized topology scenario. This decentralized approach can lead to more efficient use of network resources, as each cluster manages its own traffic, reducing congestion on shared paths. The even distribution of traffic load among different cluster heads allows for faster packet processing, resulting in:

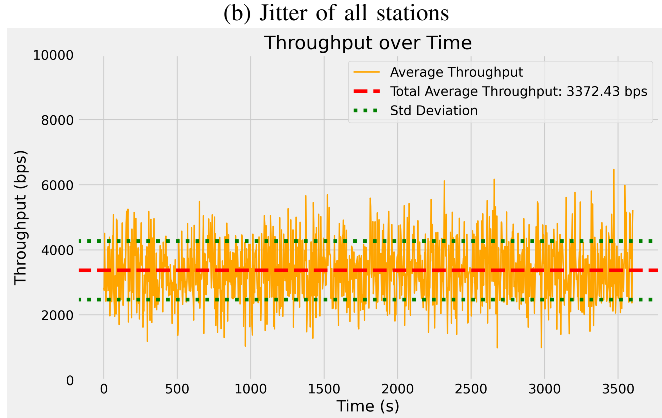
- **Lower delay:** Less congestion and shorter transmission paths between mobile stations and the cluster heads reduce the overall time for packet delivery.
- **Lower jitter:** The consistent and balanced distribution of traffic leads to more uniform packet transmission times, reducing variability in packet arrival times.
- **Higher throughput:** Because traffic is distributed more evenly, the network can handle more data at a faster rate without overwhelming specific nodes.



(a) Average delay of all stations



(b) Jitter of all stations



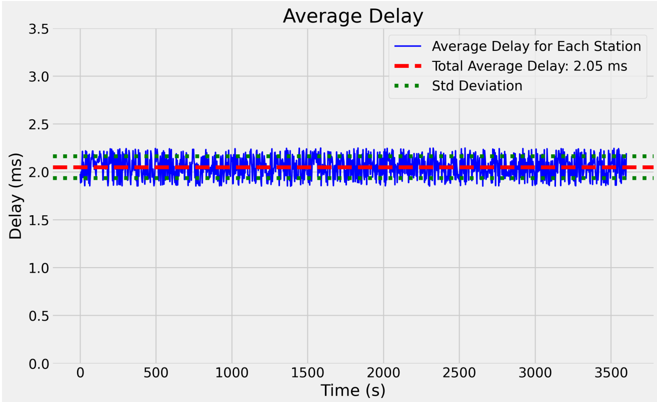
(c) Average Throughput of all stations

Fig. 4: Delay, jitter and throughput results for the centralized topology experiment.

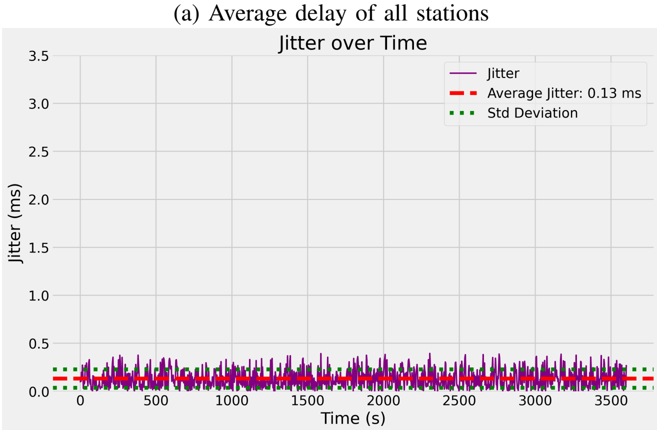
V. Conclusion and Future Work

This paper presented a novel intelligent system model for network clustering. It proposed the utilization of machine learning techniques to model and predict the mobile station's positions in FANETs for clustering, and the optimization of the cluster head selection to minimize the end-to-end delay. A set of experiments with two different topologies (centralized, for 5G environments, and decentralized, for 6G environments) showed that the proposed model can result in lower delay, jitter and increased throughput under a video traffic communication model in a decentralized network topology.

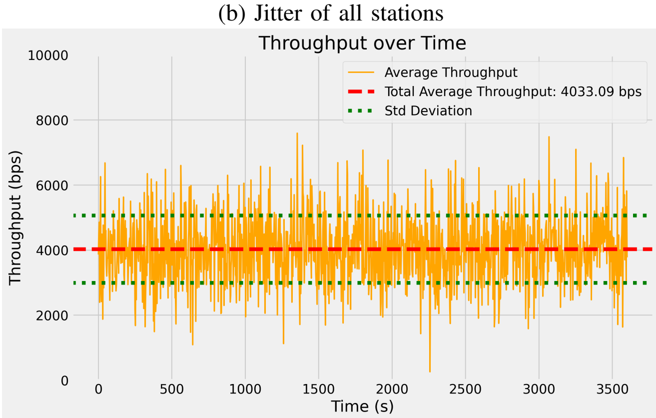
Future works include integration with simulated modules for 5G and 6G core networks in order to work with experiments closer to real-life scenarios. Furthermore, additional



(a) Average delay of all stations



(b) Jitter of all stations



(c) Throughput of all stations

Fig. 5: Delay, jitter and throughput values for the distributed topology experiment.

experiments with similar machine learning techniques to XGBoost such as AdaBoost (Adaptive Boosting), Light Gradient Boosting Machine (LightGBM), and another approach to clustering via K-Means with the Elbow method, such as DBSCAN (Density-Based Spatial Clustering of Applications with Noise).

References

[1] S. A. Hasan, M. A. Mohammed, and S. K. Sulaiman, "Flying ad-hoc networks (fanets): Review of communications, challenges, applications, future direction and open research topics," *ITM Web of Conferences*, 2024. [Online]. Available: https://api.semanticscholar.org/CorpusID:271022388

[2] O. T. Abdulhae, J. S. Mandeep, and M. Islam, "Cluster-based routing protocols for flying ad hoc networks (fanets)," *IEEE Access*, vol. 10, pp. 32 981–33 004, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:247642500

[3] M. Hosseinzadeh, A. H. Mohammed, F. A. Alenizi, M. H. Malik, E. Yousefpoor, M. S. Yousefpoor, O. H. Ahmed, A. M. Rahmani, and L. Tightiz, "A novel fuzzy trust-based secure routing scheme in flying ad hoc networks," *Veh. Commun.*, vol. 44, p. 100665, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:261600695

[4] J. Lánský, S. Ali, A. M. Rahmani, M. S. Yousefpoor, E. Yousefpoor, F. A. Khan, and M. Hosseinzadeh, "Reinforcement learning-based routing protocols in flying ad hoc networks (fanet): A review," *Mathematics*, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:251747082

[5] H. Shakhatreh, A. H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N. S. Othman, A. Khreishah, and M. Guizani, "Unmanned aerial vehicles (uavs): A survey on civil applications and key research challenges," *Ieee Access*, vol. 7, pp. 48 572–48 634, 2019.

[6] G. Khayat, C. X. Mavromoustakis, A. Pitsillides, J. M. Batalla, E. K. Markakis, and A. Andreou, "Enhanced redundant scheme based on weighted cluster-head selection for critical 6g infrastructures," in *2022 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2022, pp. 1448–1453.

[7] M. Hosseinzadeh, J. Tanveer, F. Alanazi, K. Aurangzeb, M. S. Yousefpoor, E. Yousefpoor, A. M. Darwesh, S.-W. Lee, and A. M. Rahmani, "An intelligent clustering scheme based on whale optimization algorithm in flying ad hoc networks," *Veh. Commun.*, vol. 49, p. 100805, 2024. [Online]. Available: https://api.semanticscholar.org/CorpusID:270354260

[8] S. Yang, S. Wang, T. Li, T. Hu, Z. Xu, R. He, and B. Zhang, "Hybrid ant colony-based inter-cluster routing protocol for fanet," *Scientific Reports*, vol. 14, 2024. [Online]. Available: https://api.semanticscholar.org/CorpusID:271047601

[9] S. Mashdour, S. Salehi, R. C. de Lamare, A. Schmeink, and J. P. S. H. Lima, "Clustering and scheduling with fairness based on information rates for cell-free mimo networks," *IEEE Wireless Communications Letters*, vol. 13, no. 7, pp. 1798–1802, 2024.

[10] M. A. Syakur, B. K. Khotimah, E. Rochman, and B. D. Satoto, "Integration k-means clustering method and elbow method for identification of the best customer profile cluster," in *IOP conference series: materials science and engineering*, vol. 336. IOP Publishing, 2018, p. 012017.

[11] Q. Zhao, V. Hautamaki, and P. Fränti, "Knee point detection in bic for detecting the number of clusters," in *International conference on advanced concepts for intelligent vision systems*. Springer, 2008, pp. 664–673.

[12] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:4650265

[13] T. Chai and R. R. Draxler, "Root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature," *Geoscientific Model Development*, vol. 7, pp. 1247–1250, 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:123118384

[14] A. M. Alba, S. Janardhanan, and W. Kellerer, "Enabling dynamically centralized ran architectures in 5g and beyond," *IEEE Transactions on Network and Service Management*, vol. 18, no. 3, pp. 3509–3526, 2021.

[15] L. A. L. F. da Costa, R. Kunst, and E. P. de Freitas, "Q-fanet: Improved q-learning based routing protocol for fanets," *Comput. Networks*, vol. 198, p. 108379, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:238479274

[16] L. A. L. F. da Costa, R. Kunst, R. C. de Lamare, and E. P. de Freitas, "Routing performance evaluation in centralized and distributed fanets for 6g networks," *2024 19th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 1–6, 2024. [Online]. Available: https://api.semanticscholar.org/CorpusID:271940840

[17] R. dos Reis Fontes, S. Afzal, S. H. B. Brito, M. A. S. Santos, and C. E. Rothenberg, "Mininet-wifi: Emulating software-defined wireless networks," *2015 11th International Conference on Network and Service Management (CNSM)*, pp. 384–389, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:15898165

[18] "3gpp xr-01-xr traffic model and kpi," https://shorturl.at/vIzZm, (Accessed on 09/02/2024).