AI-Enabled Planning and Control for Aeronautical Ad-Hoc Networks

by

Mohsen Shahbazi Dastjerdi

Thesis Supervisor: Dr. Burak Kantarci

A thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
degree of Master of Applied Science
Electrical and Computer Engineering

School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In-Flight Entertainment and Connectivity (IFEC) is becoming a key trend and offering in-flight connectivity is one of the most essential demands of commercial airline passengers. A grand challenge is to provide in-flight connectivity in high altitudes and particularly in isolated locations, such as the oceans, where establishing an air-to-ground link is not possible. Moreover, the high speed and dynamic characteristics of such aircraft make this task difficult. Aeronautical Ad-Hoc Networking (AANET) intends to cope with this challenge by forming a network of airplanes having air-to-air (A2A) connections. However, the dynamic nature of such a network is likely to lead to unstable connections. The primary root cause of the majority of these stability issues is known to be the short life of A2A links which is the result of poor topology formation of aircraft.

Concentrating on aircraft clustering and making them more stable can improve connection lifetime and improve the stability and performance of the network. Therefore the main objective in making AANETs feasible should be to form the topology as clusters of aircraft. With this in mind, the thesis's proposition is twofold: First, unveil the benefits of density-based clustering to improve the AANET performance. To do so, a modified DBSCAN algorithm is employed for the clustering problem that exploits several features of real flight datasets. This method also includes a weighted scheme to reflect the relative importance of each feature of the final calculation. The proposed method improved the packet delivery ratio and end-to-end latency of the state-of-the-art clustering-based AANET solutions by 51 % and 30 %, respectively. In addition, the proposed approach reduces the number of cluster changes by 22%. Second, selecting a well-connected cluster head is the next stage in enhancing connection and stability. This thesis presents a new cluster head selection technique for AANETs that calculates the Neighbor Nodes within a given distance of each node and selects the node with the most connections as the new cluster head. In instances where a cluster head cannot interact directly with another cluster, a Gateway node is chosen to facilitate connection with other clusters. According to simulations, the suggested method increases packet delivery ratio by 3, end-to-end delay by 9 and throughput by up to 10% compared to the current state of the art. In addition, the proposed method reduces cluster head replacements by 17% and increases cluster head longevity by 8%.

Acknowledgements

I would like to express my utmost gratitude to my thesis advisor, Dr. Burak Kantarci, for his invaluable guidance, support, and encouragement throughout the course of my research. His knowledge and expertise in the field have been a constant source of inspiration, and his dedication to mentoring me has been instrumental in the successful completion of my thesis. I am truly grateful for his unhesitating belief in me and for the time and effort he has invested in my academic and professional growth. Thank you, Dr. Kantarci, for being an excellent mentor and for being an important part of my academic journey.

I am grateful to Dr. Murat Simsek for his invaluable contributions to the project. His insights and constructive feedback helped me to navigate the complexities of my research and to achieve a higher level of understanding of the subject. I would like to extend my heartfelt appreciation to my fellow SCVIC lab members, specially Ghazal and Jason, for their camaraderie and for providing a stimulating and supportive environment in which to conduct my research. I am deeply grateful for the friendships and collaborations that have developed during my time in the lab.

My last but deepest gratitude is to my parents for their unwavering love, support, and guidance throughout my academic journey. They are the shining stars that have guided me through life and have been my beacon of hope. Their belief in me and my abilities has been a continual source of motivation for me., and their encouragement and support have been crucial in helping me to pursue my goals and to complete my thesis. I appreciate the sacrifices they made and the time and work they put into my education. Their love and guidance have been the foundation that has helped me to become the person I am today. I would not have been able to achieve this accomplishment without their love and support. Thank you.

Table of Contents

List of Tables

List of Figures

Publications of the Candidate During MASc Studies

Outcomes of the thesis:

- **M. Shahbazi**, M. Simsek and B. Kantarci, “Density-Based Clustering and Performance Enhancement of Aeronautical Ad Hoc Networks” *2022 International Balkan Conference on Communications and Networking (BalkanCom), 2022, pp. 51–56*
- **M. Shahbazi**, M. Simsek and B. Kantarci, “AI-Enabled Cluster Head Selection Method through Modified Density Based Clustering in Aeronautical Ad Hoc Networks” *Elsevier Ad Hoc Networks 2023*

Chapter 1

Introduction

The global market for In-Flight Entertainment & Connectivity (IFEC) is anticipated to reach USD 6.1 billion by 2026, up from an estimated USD 4.7 billion in 2021 [1]. This market's expansion is primarily driven by an increase in demand for in-flight entertainment. Travelers are prepared to pay a premium for services, such as Wireless Fidelity (Wi-Fi) access, on board an aircraft. Modern passengers use more devices than ever before while in flight. According to SITA Passenger IT trends, 65% of passengers stream material on their own devices [2]. While according to a poll by Inmarsat, 55% of customers regard in-flight Wi-Fi to be a need, and 67% will rebook with an airline that offers high-quality in-flight Wi-Fi [3].

The development of in-flight Wi-Fi networks that leverage both satellite connection and Air-to-Ground (A2G) links is the result of the growing demand for Internet access during flight. However, these Existing in-flight connectivity solutions have intrinsic constraints, such as restricted coverage, expensive construction and operating costs, and significant delays, that diminish the quality of the end-user experience.

Aeronautical Ad-hoc Networks (AANETs) are a potential developing solution [4] for the future of in-flight connectivity. AANETs can provide low-cost broad coverage by joining current commercial aircraft to build a self-configured wireless network via multi-hop air-to-air (A2A) links that combine satellite and ground networks to create a large-scale, multi-hop wireless network capable of moving data as illustrated in Figure 1.1. Therefore, AANETs can provide Internet connectivity to airplanes that cannot directly connect to the terrestrial or satellite network [5]. In the case of satellite networks, some airplanes

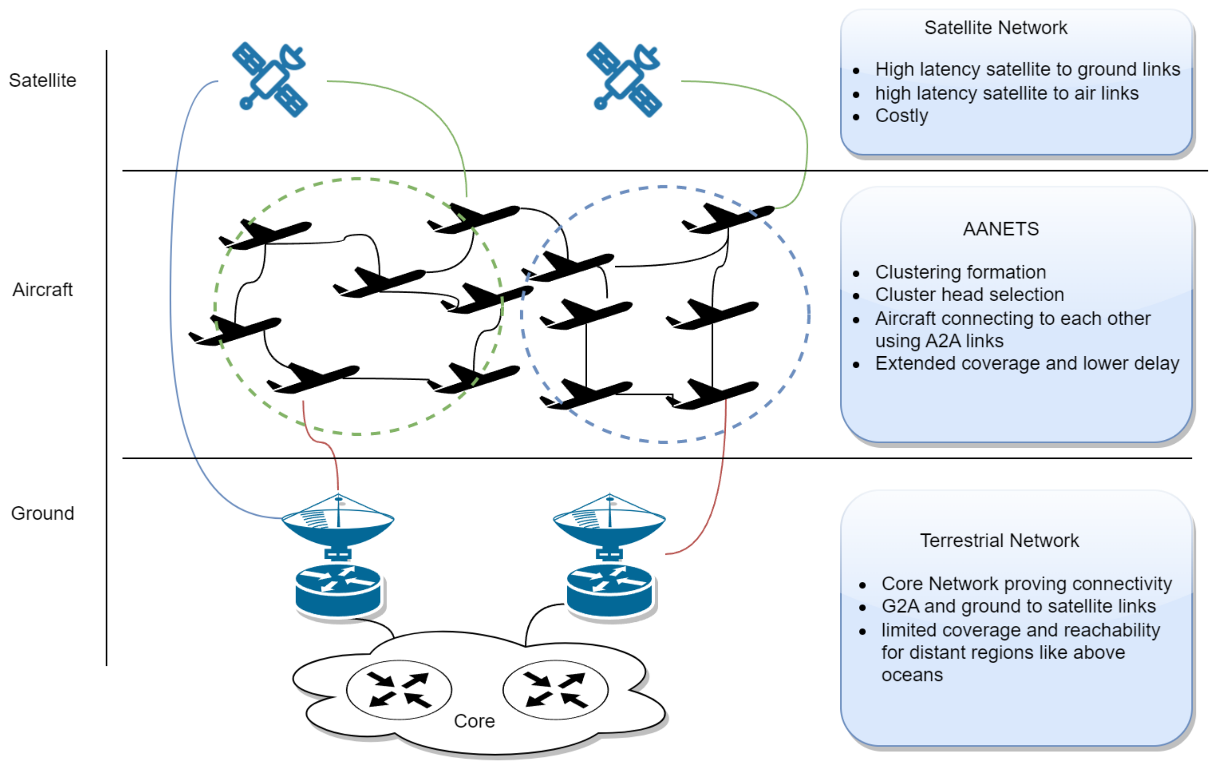


Figure 1.1: Topology of AANET and comparison to Satellite and Terrestrial Networks.

may not have coverage because of the curvature of the earth and also terrain and weather specifications.

This thesis investigates the strategies utilized in the construction of AANETs, especially in the form of clusters, to improve the performance of communication metrics and increase stability, hence making the formation of these networks more feasible.

1.1 Motivation

In an AANET, data packets are carried from source to destination utilizing A2A connections, and each aircraft can forward the packet to the next hop without the requirement for a centralized network architecture [6]. This provides AANETS benefits over existing terrestrial networks, such as extending coverage due largely to highly scattered commercial

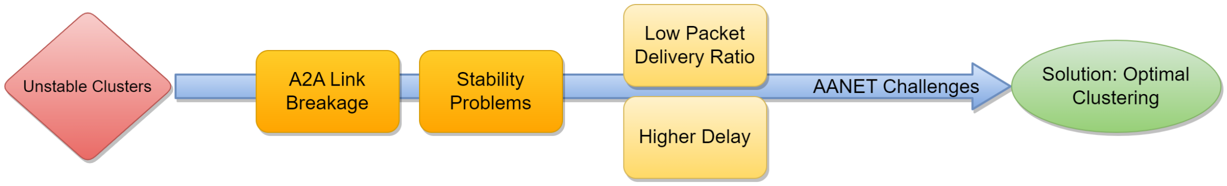


Figure 1.2: Main Challenges Regarding AANET Formation

aircraft, reducing latency and communication costs compared to satellites, and providing aircraft with seamless connectivity [7]. Delay is an essential factor in wireless communication, and although there are several factors that can be used to calculate the total delay, propagation delay is the most significant. Depending on factors such as distance, the propagation latency of a satellite connection for aircraft networks can range from several hundred milliseconds to one second or more when compared to a ground or air connection. A ground-to-air or A2A connection, in contrast, can have a latency of less than 50 milliseconds. [4, 8]

Despite having advantages, the deployment of AANETs faces several obstacles. Compared to other mobile ad hoc networks (MANET), AANETs have unique properties that pose unique challenges for network formation and administration. It is crucial to analyze these properties of AANETs before examining the challenges and determining the role of AI in resolving them.

First, AANETs are a dense network that can span hundreds of kilometers and can consist of hundreds of aircraft communication with Air-to-Ait links. Typical A2A links have ranges up to 400 Nautical Mile. Other forms of ad hoc networks, such as VANETs, are two-dimensional, but AANETs are three-dimensional since aircraft height is also a factor. This renders the data multi-dimensional and significantly more difficult to manipulate [9]. Therefore, we believe it could be efficient to employ AI algorithms to manage this data in order to find trends and create valuable clusters of aircraft with similar features.

Second, by definition, ad hoc networks are unstructured and do not require a centralized organization to operate the network [10]. This is advantageous since each aircraft may independently receive and forward traffic, allowing it to reach distant locations. This, along with the high speed and dynamic nature of AANETs, makes the formation of topologies an extremely difficult operation. Using unsupervised machine learning techniques, AI can assist in resolving this obstacle. As previously mentioned, AI may assist in the formation of meaningful clusters from unstructured data, and unsupervised learning is an appropriate technique for this purpose.

These two factors impair the stability of the AANET topology by lowering the lifetime of the A2A links between a pair of airplanes, which leads to a lower packet delivery ratio in the network and a longer end-to-end latency for packets that are routed. These issues with network stability may be traced back to sub-optimal clusters as shown in Figure 1.2. Therefore, the primary function of AANET should consist in establishing robust and manageable topologies in the form of clusters.

This thesis builds on the motivation that the root cause of AANET stability issues, which is the short life span of A2A links, need to be addressed by identifying suboptimal clusters and making use of unsupervised learning algorithms to form robust clusters. Further down, the performance of clusters can be enhanced by choosing a well-connected node as the cluster head to manage the intra-cluster and inter-cluster communications more efficiently.

1.2 Objectives

Having identified the root cause of problems in the stability of AANETs, in this thesis, we investigate the initial step in resolving the issue. With this in mind, we aim to obtain more structured clusters by utilizing machine learning algorithms, namely unsupervised learning to form clusters comprised of aircraft with comparable features.

The primary objective of the clustering-based AANET topology construction is to group airplanes with similar characteristics. The aircraft are clustered due to their commonalities in latitude, longitude, direction, and altitude. Therefore, airplanes in the same cluster are able to create durable A2A linkages, which boost network stability and communication performance. Various classes of machine learning techniques, including hierarchical clustering algorithms, density-based clustering algorithms, and innovative self-organizing maps, can be used to cluster the data. Each of these techniques has advantages and disadvantages that must be considered when selecting an algorithm for clustering, as detailed in chapter 3.

As will be discussed in chapter 4, in this thesis, to achieve the goal, we utilize density-based clustering, specifically a modified DBSCAN framework for clustering the aircraft that exploits numerous parameters of real flight data, including latitude, longitude, altitude, direction, and velocity. In lieu of a traditional distance metric such as Euclidean or Haversine, this approach generates a pre-computed distance matrix and feeds it to DBSCAN. This approach also incorporates a weighting mechanism to represent the proportional relevance of each distance computation component. Our objective is not only to enhance the clustering procedure but also to use the high-dimensional characteristics of a real flight dataset. Most clustering methods in the literature only employ a small number of these characteristics, which is suboptimal.

1.3 Contributions

After a complete literature analysis and assessment of the present state of the art, we conducted a detailed study of the underlying cause of stability issues in AANET formation. Utilizing a new clustering technique, we reevaluate and improve the efficiency of clusters utilized in the AANET architecture. We also proposed a cluster head selection algorithm to further boost the performance and stability of clusters. The following is a summary of our contributions to this thesis:

1. **Modified DBSCAN framework for clustering-based AANETs:** The first contribution of this thesis to cluster-based AANETs is the development of a density-based clustering method for grouping airplanes. We proposed a modified DBSCAN algorithm that groups aircraft with similar characteristics based on their position, velocity, direction, and altitude. Instead of utilizing a basic metric such as Haversine or Euclidean, a pre-computed distance matrix of the dataset's specified characteristics is created and provided to DBSCAN as the metric. A weighting system is also employed to highlight the significance of each feature to the system. The methods enhance the performance of AANET compared to K-means in terms of packet delivery ratio by 51% and end-to-end latency by up to 30%, as well as the clusters' stability.

2. **A novel cluster head selection method that builds on top of the modified DBSCAN:** The second contribution of the thesis is the proposal of a novel cluster head selection algorithm that selects a well-connected node as the new cluster head. Since the updated DBSCAN generates dense clusters and all nodes within a cluster are able to interact with one another, the suggested Cluster head selection is designed

to pick the most connected node as the cluster head, which should have a reliable connection to all nodes within the cluster. The method proposed utilizes the reverse haversine formula to choose the ideal candidate for the CH. In circumstances when a CH cannot interact directly with another cluster, the notion of a Gateway Node (GN) is also considered in order to simplify inter-cluster communication. Coupled with the suggested clustering method, we are able to detect a rise in the packet delivery ratio up to 9% and a 17% improvement in cluster stability compared to the current state of the art.

1.4 Thesis Outline

The dissertation consists of six separate chapters ordered as follows:

Chapter 2 contains the context and literature review. In Section 2.1, AANETs are compared to various forms of ad hoc networks. In Section 2.2, clustering methods in ad hoc networks and specifically AANETs are examined. The focus of Section 2.3 is on cluster head selection techniques currently available in the literature. Section 2.4 examines different sorts of research on the administration of AANETs, such as routing and security.

Chapter 3 delves deep into the role of AI algorithms in the development of AANETS and the design of this thesis. In Section 3.1, we examine several unsupervised clustering techniques utilized for data clustering in AANETs and ad hoc networks. This thesis's foundation is density-based clustering, which is discussed in depth in Section 3.2.

In Chapter 4, our proposed approach for clustering-based AANETs is presented. Following a brief introduction of density-based clustering approaches and a specification of the problem, Section 4.1 elaborates on the proposed architecture and gives a comprehensive explanation of its components. Comparing the performance of the proposed approach to that of existing state-of-the-art clustering algorithms is carried out in Section 4.2 through the use of performance metrics and the outcomes of simulated experiments.

Chapter 5 is the final contribution to the thesis. The chapter suggests implementing a cluster head selection mechanism on top of the suggested clustering approach. The chapter begins by discussing the significance of selecting an optimum cluster head and the associated difficulties. Section 5.1 presents an overview of the system, followed by a description of the procedure used to pick the CH. Section 5.2 consists of a comprehensive examination of the scenarios used to achieve the results.

Finally, Chapter 6 concludes the procedures outlined in the thesis and provides a conclusion. In addition, we examine potential future improvements that might be made to the AANETs and conclude the thesis.

Chapter 2

Background and Literature Survey

The urge to connect is more prevalent than ever in the current society. Internet connectivity is no longer restricted to computers and mobile devices; IoT devices in homes and all types of vehicles are now considered smart devices. Therefore, both academics and industry are working towards the realization of smart, connected cities. Automobiles, trains, and aircraft all make up a significant portion of the linked world. Thanks to developments in ad hoc networks and the creation of vehicular ad hoc networks (VANETs), various types of cars may now link not only to the internet but also to other vehicles, emergency systems, etc. [11]. Aeronautical ad hoc networks (AANETs) are a subset of VANETs that utilize commercial aircraft to create an ad hoc network to give access to airplanes in remote places where direct contact from the ground is not possible.

To comprehend the effects of clustering on the construction and maintenance of AANETs, it is necessary to first understand the distinctions and parallels between the AANET and other forms of ad hoc networks. In the sections that follow, we examine the current state of the art in terms of clustering method algorithms and cluster head selection strategies. While most of the current algorithms in the literature are focused on VANETs and the focus of the thesis is on AANETs; it is only recently a few appeared that are customized particularly for AANETs, it is still vital to examine them since AANET is a subset of VANET and fundamentals are the same.

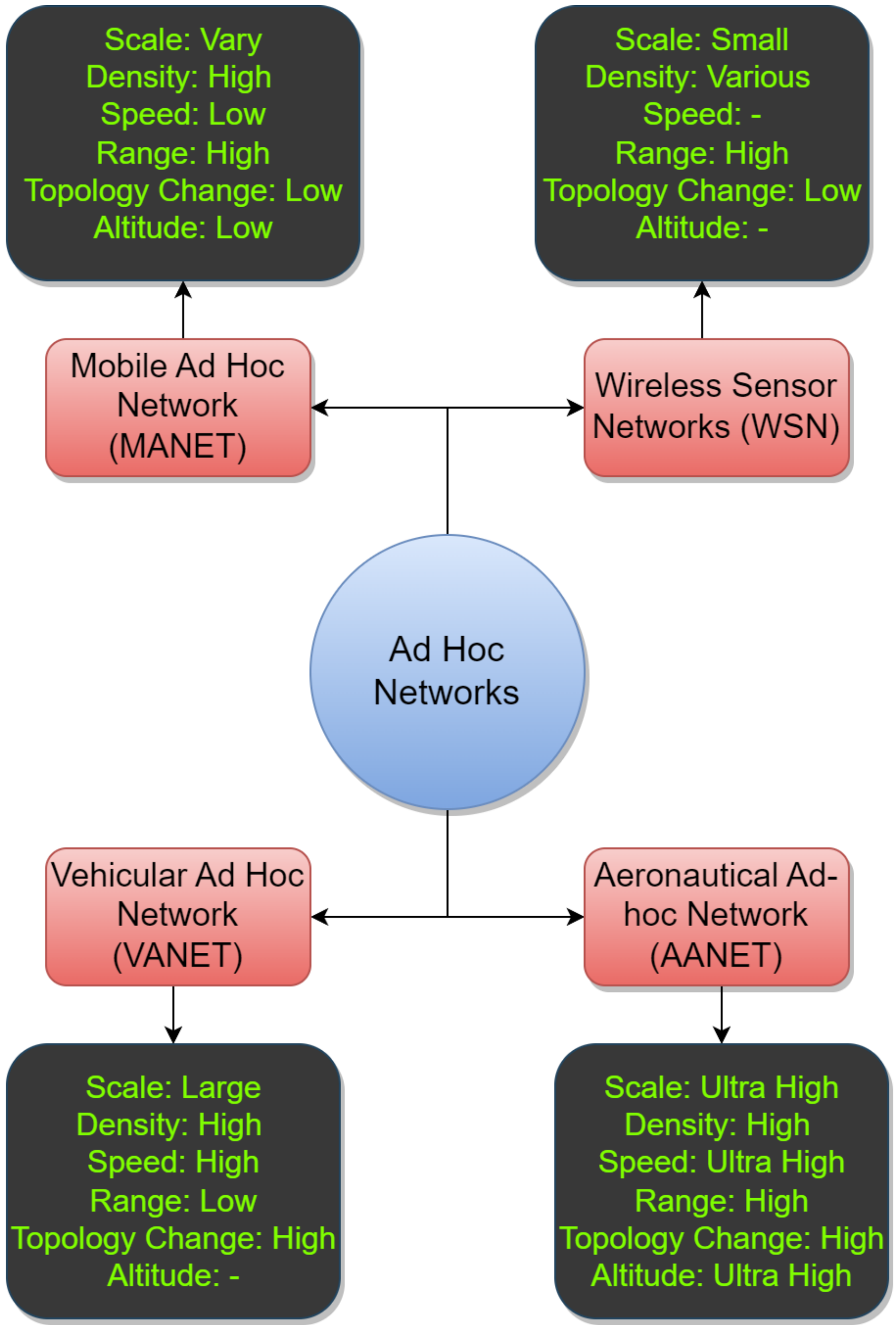


Figure 2.1: Comparison of Ad Hoc Networks

2.1 Types of Ad-hoc Networks

Ad-hoc networks are a form of the wireless network in which devices interact directly without requiring a central router or access point. Each device in an ad-hoc network functions as a router, passing data packets to other devices [12–14]. This provides for flexibility and simplicity of deployment, but it may also lead to a reduction in security and an increase in complexity. Ad-hoc networks are frequently utilized when a regular wired or wireless infrastructure is unavailable or difficult to establish.

While there are various ad hoc networks in existence today as described in Figure 3.2, they may be generally categorized as follows:

Mobile Ad Hoc Network (MANET)

The devices in a mobile ad-hoc network (MANET) are mobile and able to roam around [15–18]. Due to the mobility of the devices, network topology might often change, making routing more complicated. Each device in a MANET must function as both a host and a router, passing data packets to other devices. Routing in a MANET is generally determined by the distance between devices and their movement patterns and is performed using a number of different methods. MANETs are utilized for a variety of purposes, including disaster recovery, the military, and rescue operations in locations where a permanent infrastructure is unavailable [19, 20]. Their application may also be extended to vehicular networks, where automobiles can interact as nodes. The characteristics of a Mobile Ad-hoc Network (MANET) include the following:

- Dynamic topology: The network topology might vary regularly when devices move, which makes routing more difficult.
- Limited resources and communication range: Devices in a MANET have limited processing power, memory, and battery life, which can negatively impact the network's performance. Additionally, they often have a restricted communication range, which might impact network connectivity.
- Self-organization: MANETs are self-organizing, which means that devices can join or leave the network dynamically without requiring a central controller.

- Multihop: Communication between devices in MANETs is often multihop, meaning that data packets are sent from one device to the next until they reach their destination.

- No fixed infrastructure: MANETs do not rely on a fixed infrastructure, making them appropriate for scenarios in which a traditional wired or wireless infrastructure is unavailable or cannot be quickly installed.

In a clustered MANET, the devices are grouped into clusters, with one device serving as the cluster head. The cluster leader is accountable for coordinating the actions of the devices inside the cluster, including resource management and routing [21]. Typically, the clustering process in a MANET consists of the following steps:

- Cluster formation: Clusters of devices are formed depending on their location, mobility, and other characteristics. Typically, the nearest devices are clustered together within the same cluster.

- Cluster head selection: Each cluster has a single device designated as its cluster head. Typically, the option is determined by the device's battery life, its location, and other factors.

- Management and maintenance: The cluster leader oversees the cluster's shared resources, such as bandwidth and energy.r

Clustering may increase the scalability and efficiency of a network by decreasing the number of devices required for resource routing and management. It also increases network stability by isolating the impacts of device mobility and failures inside the cluster and by decreasing the number of needed route changes.

Wireless Sensor Networks (WSN)

Wireless Sensor Networks (WSNs) are networks of compact, low-power wireless sensor nodes used for data collection, processing, and transmission [22–24]. These nodes are often outfitted with sensors, processors, and wireless communication capabilities and may be used to monitor a variety of physical phenomena, including temperature, humidity, light, sound, and motion. By utilizing low-power electrical components and communication protocols, WSNs are meant to be energy-efficient in order to extend the network's lifetime over a vast region. WSNs are utilized in several applications, including environmental monitoring, industrial process control, healthcare, and smart cities [25, 26]. They are also utilized by the military and during search and rescue missions. Due to their capacity to collect and send data, WSNs can give vital information for decision-making and aid in the enhancement of the monitoring system's efficiency.

While WSNs and MANETs are both types of wireless networks, there are significant differences between them. WSNs are designed for monitoring and sensing and are comprised of low-power, low-cost devices with a star network topology, whereas MANETs are designed for communication and are comprised of mobile devices with a peer-to-peer network topology. In the case of clustering, the procedure is nearly the same; however, in the case of WSN, the cluster head has additional responsibilities, such as collecting data from all cluster nodes and sending it to the sink node or base station.

Vehicular Ad Hoc Network (VANET)

A VANET, also known as a Vehicular Ad-hoc Network, is a sort of Mobile Ad-hoc Network (MANET) comprised of vehicles equipped with wireless communication equipment, such as radios or cellular modems [27–29]. These cars transmit information such as traffic conditions, road dangers, and vehicle position by communicating with each other and with roadside infrastructure. VANETs are distinguished by the following characteristics:

- High mobility: Vehicles in a VANET have a high rate of movement, which can make routing and communication more difficult.
- Dynamic topology: As cars travel across the network, their topology might often change, making routing more complicated.
- Short communication range: Typically, vehicles in a VANET have a limited communication range, which might hinder network connectivity.

VANETs are utilized in several applications, including traffic control, navigation, and information and entertainment services [11, 30]. They may also be utilized for Intelligent

Transport System (ITS) applications [31–35], which aim to increase traffic efficiency, decrease congestion, and improve the driving experience overall. Another novel use case of autonomous vehicles is assisting in the detection of diseases such as COVID-19, as suggested in [36]. In contrast to MANET, vehicles on the road travel at high speeds, but the mobility of devices in a MANET might vary. In addition, VANETs have a very dynamic network topology due to the continual movement of vehicles, whereas the topology of a MANET can alter depending on the position and mobility of its devices. Due to the fast speeds of vehicles and the close proximity between them, VANETs also have a shorter communication range compared to MANETs.

Aeronautical Ad-hoc Network (AANET)

As described in Chapter 1, an Aeronautical Ad-hoc Network (AANET) is a type of ad hoc network built exclusively for communication between aircraft in flight. AANETs let aircraft connect with one another and with ground-based infrastructure, such as air traffic control systems and weather stations, to enhance safety, navigation, and flight efficiency [37]. AANETs can be used for applications such as air traffic control, navigation, and weather information sharing, in addition to providing in-flight connectivity for passengers. A typical implementation of aircraft can consist of hundreds of aircraft in dozens of clusters that utilize A2A links to communicate. these types of links are capable of transmitting data up to 400 Nautical Mile. Flying ad hoc network (FANET), a comparable sort of ad hoc network, is essentially an ad hoc network formed by drones to facilitate communication and collaboration among themselves. Drones serve as both network nodes and data relays in a FANET. They create wireless connections between themselves to form a self-organizing network. These connections can be made using a variety of communication technologies, including Wi-Fi, Bluetooth, and radio frequencies. The drones communicate with one another in order to share information, coordinate their movements, and efficiently disseminate duties. Compared to FANET, AANETs are different. FANET mostly includes UAVs that fly at lower altitudes and have energy restrictions whereas AANET is made of commercial aircraft. Also while AANET and VANET share the majority of the same properties, there are significant variances. First, the AANET architecture is extremely dynamic, and it may be regarded as more complicated than an automobile. Second, aircraft travel in three-dimensional space, and the addition of altitude makes data management and network administration more difficult than with VANET's two-dimensional structure.

Lastly, compared to a car, airplanes travel at a significantly faster speed and there is a greater speed and direction difference between the two aircraft, which in turn causes more connection failures and decreases the network's stability.

2.2 Clustering Methods

Now that we have reviewed and contrasted different forms of ad hoc networks, we may study the various clustering techniques utilized in the literature. The unstructured, ad hoc nature of the AANET is one of its primary features. This complicates the construction of the AANET topology, causes stability issues, and reduces the lifespan of the A2A between aircraft. AI can play a significant role in the formation of clusters [38–41]. The stability issues stated in chapter 1 can be resolved by implementing clustering algorithms, which will result in a more stable topology and a longer lifetime for A2A links. Regardless of the fact that AANET is a fairly new concept, various studies have been conducted to introduce the clustering methods tailored for mobile ad hoc networks (MANET) into AANET. Most of the research falls under the category of unsupervised learning, whose primary objective is to discover previously undiscovered data patterns. Consequently, it seeks to investigate the data in order to identify significant structures and determine classes.

Maglaras et al. [42] present two grouping algorithms for vehicle networks based on velocity, location, and physical attributes. Sociological pattern clustering (SPC) and route stability clustering (RSC) are the methodologies. The second strategy is a customized variant of the first that makes use of the social behavior of vehicles. Both techniques leverage historical data and the semi-Markov modeling approach to exploit the aforementioned social tendencies and generate stable clusters. RSC is then used to establish stable clusters based on long-term probability after SPC has classified the cars into social groups with comparable behavior. The results demonstrate a decrease in the number of cluster heads and an increase in cluster longevity, indicating greater stability.

CAPS (Cluster Algorithm for Publish / Subscribe) is proposed in [43] to address the issue of clustering in AANET and to establish the groundwork for future publish and subscription systems, as the performance of such a system is largely reliant on the cluster topology. CAPS is a 1-hop clustering technique designed to decrease controlling overhead and enhance the network's overall stability. Their recommended solution is comprised of three key steps: By using "hello" signals containing the position and velocity of each aircraft, the initial phase consists of all aircraft discovering their neighbors. This is followed by the selection of an ideal cluster head based on WCA or SBCA metrics for the newly created clusters. The final phase is the maintenance routine to ensure the stability and functionality of the clusters. The proposed approaches reduce the number of cluster heads and cluster head modifications in comparison to MOBIC. Moreover, the longevity of cluster heads rises dramatically.

To overcome the management challenges of hierarchical network architectures, the authors of [44] suggest a clustering strategy based on the Honeycomb Division. Hierarchical topologies can result in straightforward routing and reduced administrative burden, but they are largely dependent on the success of the clustering algorithm. The suggested technique has three goals: determining the size of clusters and addressing the blindness of existing algorithms utilizing the idea of zones. Secondly, to address the bottleneck issue of cluster heads, they employ numerous heads to distribute the load. They employ the well-known spanning tree method to limit the random overlap between clusters and the number of hops in order to prevent the packet from being transmitted numerous times and producing loops. In fact, the findings demonstrate that the proposed technique delivers a greater packet delivery ratio and lower latency.

Bilen et al. [45] suggest a modified variant of K-means for clustering AANET in order to extend the life of Air-to-Air links. The ultra-dynamic nature of AANETs affects the network's stability; thus, airplane clusters must be tuned to improve the internet connection efficiency for passengers. Included in the proposed K-means modification are distance and direction estimates. In order to compute the distance between two pairs of aircraft, the haversine formula is used, which may make use of latitude and longitude information. LVQ mapping [46] is used to map new aircraft to existing clusters in order to ensure that cluster maintenance is done effectively. The simulation with varied K values demonstrates a 75 percent increase in airplane stability over the standard K-means algorithm and a higher label rate.

Researchers in [7] offered a way for managing network partitioning and, more particularly, the multi-dimensional profile of AANETs. As opposed to a 2D environment like a circle, AANET's topology is 3D and spreads out like a sphere. In this instance, the network partitions are categorized as multidimensional clustering. The suggested technique uses HAPs to monitor the status of AANET before utilizing 3D Voronoi space tesselation to form stable clusters of aircraft. The proposed method takes isolated nodes into account and aids in resolving the network disconnectivity issue. The ultimate objective of this project is to achieve pervasive monitoring in the AANET.

Wang et al. [47] model the aircraft communication network as single-hop and double-hop. The first implies that the source aircraft can connect directly with the destination, whereas the second needs intermediary nodes to relay packets to their destinations. To improve performance, the suggested technique estimates the maximum number of simultaneous connections conceivable. Simulations were performed to evaluate the throughput and delay performance, and it was determined that the two-hop model is superior to the one-hop model.

[48] introduces a universal clustering approach that may be used to efficiently cluster any sort of vehicle, including automobiles and aircraft. The authors suggested a pair of stability-driven algorithms designed to decrease the number of cluster head transitions while extending the lifetime of cluster members. The first method is known as Dynamic Doppler Velocity Clustering (DDVC) and utilizes the Doppler value (DV) metric to build stable clusters. Additionally, cluster maintenance is performed with this algorithm. The Dynamic Link Duration Clustering (DLDC) technique utilized link duration predictions based on location and velocity to generate clusters. Both techniques are responsible for identifying cluster leaders and adding new nodes to existing clusters. Unique to this framework is the system's compatibility with vehicles that lack an onboard GPS device. In comparison to one-hop clustering approaches, the simulation shows that the suggested algorithms yield more consistent clusters.

The objective of the study conducted by Megas et al. [49] is to make internet connectivity for airplanes a reality, particularly over the oceans. They propose a heuristic method with two phases. In addition, the problem is stated as mixed integer linear programming (MILP) in order to increase the number of airplanes while maintaining a given data flow threshold. Phase 1 of the proposed methodology entails establishing the network topology in the form of clusters and then removing part of the linkages between them in order to satisfy the node degree requirement. In phase 2, bandwidth allocation calculations are performed to ensure the flow for all nodes. For a low-density network, the article claims to obtain just 8% below the ideal outcome; however, for a high-density network, the connection percentage drops from 70% to 40%.

Bharany et al. [50] take into account the energy limitation and low computational capability of flying ad hoc networks and provide an energy-efficient clustering technique.

The suggested technique is based on Low-Energy Adaptive Clustering Hierarchy (LEACH) and is intended to address the power consumption issues of LEACH-based algorithms. In the first step, referred to as the setup phase, clusters are constructed, and each node chooses whether or not to become the cluster head. The second phase is responsible for the transfer of actual data between nodes. As a result, the suggested technique achieves a substantial reduction in energy usage in comparison to conventional LEACH while simultaneously enhancing the delivery ratio.

A three-phased topology formation system is proposed in [51] for AANETs. The purpose of this research is to overcome the difficulties of short link lifetime in the air and hence increase the overall network performance in AANETs. The suggested technique consists of three phases: in the first phase, aircraft with similar characteristics are grouped together. This process is accomplished by utilizing a stability index and Moran Index. The second step involves the construction of air-to-air connections between aircraft. For this assignment, a delay-bounded form of the spanning tree is employed. The last step involves picking an optimum cluster head. Theta Index and the G/G/1 queueing model are utilized to create the centrality index for this job, hence selecting a well-connected node as the cluster leader. Extensive tests are conducted for aircraft takeoff, cruising, and landing maneuvers, and the findings indicate improvements in packet delivery ratio and end-to-end latency.

2.3 Cluster Head Selection Methods

Cluster heads (CH) are unique nodes responsible for intra-cluster and potentially inter-cluster communication. Choosing an optimal CH is a difficult challenge for AANETs. Maximizing the lifespan of CHs will result in a more efficient network with a more stable topology which will further enhance the packet delivery ratio. Establishing good clusters is a crucial prerequisite for choosing suitable cluster heads [52]. Various research has been conducted on cluster head selection for Vehicular Ad-hoc Networks (VANETs). Despite the fact that the majority of these researches are conducted on other types of ad hoc networks, they can serve as a baseline for AANETs as the fundamentals are the same.

Distributed Clustering Algorithm Based on Dominating Set (DCA-DS) is proposed in [53]. DCA-DS employs the concept of node span to build a dominant set for the internet of vehicles. The node span is determined by counting the number of unclustered neighbors each node has. The suggested technique, which is a simple heuristic algorithm, selects the ideal cluster head for each cluster using a greedy strategy on node spans. In addition, the gateway idea is studied in this study in order to link to other clusters. Furthermore, the suggested method includes a maintenance mechanism to guarantee that cluster heads stay optimum.

Beheshti et al. [54] propose Location-Aware Clustering in Vehicular Ad-hoc Networks (LAC-VANET) to attempt to improve the propagation and security of ITS systems. The suggested technique consists of three stages: in stage one, each vehicle should estimate its position based on the signal strength indicator received (RSSI). The second step consists of delivering information to surrounding nodes so that they can create the network's topology and clusters. In the third and final stage, fuzzy logic is used to choose the cluster head based on the least latency and link duration. In comparison to previous approaches, an improvement in throughput and delivery ratio can be noticed as a result of enhancing the cluster head selection procedure.

The grasshoppers' optimization-based node clustering algorithm (GOA) is introduced in [55] for optimum cluster head selection. This study tries to address the issues posed by ad hoc networks' dynamic topology changes and constantly mobile nature, particularly in environments with a high node density. GOA employs the swarming behavior of grasshoppers to optimize the cluster head selection procedure, resulting in fewer clusters and a minimal number of Cluster heads. The proposed method is compared against CACONET [56] and GWOCNET and, in all the scenarios, achieves a better performance.

For flying ad hoc networks, such as UAVs, network fragility and collision risk are crucial issues. This notion is discussed in [57], in which the authors provide a CH selection procedure that applies a weighted model that takes energy, mobility, and distance between nodes into account. The proposed algorithm utilizes a weighted system to pick the best cluster heads that leverage the UAV characteristic. This study also takes energy concerns into account, which is crucial because the cluster head uses more energy than conventional nodes. The outcomes are a decrease in the number of cluster head replacements and an increase in the delivery ratio.

Hussain and Bingcai [58] propose a new cluster head selection method for VANETs.

Combining K-means with the Floyd-Warshall algorithm is utilized to determine cluster heads. K-means is responsible for clustering and grouping the vehicles using division points. After calculating the distances between each pair of nodes using the Floyd-Warshall method, the node with the lowest average of distances is selected as the new cluster head. The results indicate that this strategy generates superior connectivity compared to other comparable methods.

Duan et al. [59] consider 5G infrastructure and propose a cluster head selection approach suited to the high mobility and dynamic topology of ad hoc networks. Using Software-Defined Networking (SDN), the study proposes a dual cluster head technique. SDN is implemented to enable coordination between the base station or any form of operator and the ad hoc network, and because it provides a global view of the network, the clusters only need to be modified when necessary. In a dual cluster head configuration, the backup cluster head is always prepared to assume the functions of the primary cluster head; moreover, the backup cluster head manages the transition between the two cluster heads.

VMaSC [60] is proposed to address the dilemma of frequent ad hoc topology changes. The Vehicular Multi-hop algorithm for Stable Clustering (VMaSC) is a unique algorithm designed to ensure the stability of clusters in a VANET environment. The election of the cluster head is determined by measuring the relative mobility of each node relative to its neighbors. The objective is to pick the node with the lowest mobility as the cluster leader, hence extending the lifespan of both the cluster and the cluster leader. The suggested technique has been shown to extend the lifetime of both CHs and cluster members.

A center-Based Secure and Stable Clustering Algorithm is introduced in [61] to resolve the issues with existing MANET algorithms that are used for VANETs. The cluster head selection strategy described in this study is based on the assumption that VANETs are less stable than other forms of ad hoc networks due to high-speed movements. In order to select an appropriate CH, the relative mobility measure is introduced and compared to all cluster nodes. The node with the lowest mobility will therefore become the new CH.

2.4 Management And Routing Methods

Aside from clustering, topology construction, and cluster head selection, a few more studies on the administration and sustainability of AANETs have been conducted. These include subjects such as routing and packet handling, which is the emphasis of this section. Routing is crucial to the formation of a working ad hoc network, as an inadequate routing system would delay the transmission of data.

Yang et al. [62] present a framework geared to AANET's massive scale to combat the standard methods' high synchronization overhead. Current routing algorithms struggle with shifting flight paths and will produce loops which is why this study presents a low-cost routing algorithm. The virtual Topology-based Networking (VTN) technique is responsible for abstracting the physical topology into a virtual one in order to simplify the network. This is accomplished by assigning each aircraft a unique identity referred to as a VID, as well as defining a Path attribute (PA). The Trunk-Branch Cooperation assisted Routing (TBCR) method is then suggested to manage packet routing. The TBCR routing algorithm combines the loop-free features of VTN with the adaptability of geo-greedy routing. This research also takes intra-cluster and inter-cluster routing into consideration. As a result, both the E2E latency and the number of hops required for the packet to reach its destination are lowered.

AODV-LD [63] is offered as AANET's solution to the problem of air-to-ground communication. ADOV-LD, a link duration variation of traditional ADOV, is subdivided into two variants. AODV-LD-D is a deterministic method that utilizes aircraft motion to estimate the connection duration based on velocity and location. AODV-LD-S, meantime, is a stochastic method that employs link distributions derived from experience above the ocean. After replicating actual aircraft routes for seventy days, the research shows that the suggested strategy increased the routes' longevity and decreased their latency.

Researchers in [64] devised a solution to the problem of inter-intermittent linkages in a highly mobile and expansive AANET environment. This study models three types of data transmission: direct, linked relay, and opportunistic transmission. In addition, a dynamic graph model is utilized to examine the frequent topology changes and build a workable model of AANET. Then, an optimization problem is studied to reduce delay while taking bandwidth and throughput into account. The research demonstrates that the existence of an airplane buffer will have a favorable effect on delay, and the trade-off between latency and data quantity is studied. In [65], a continuation of this work is offered. To decrease the complexity of the problem, a basic dynamic graph is utilized in this version to record changes in the topology and disregard small modifications. AANET then formulates a coalitional game to maximize the transmission plan in light of the delay. The result reveals that a combination of aircraft buffer and ground station buffer will yield the optimum results.

A Q-Learning approach for routing in AANETs is proposed in [66]. This research seeks to address the high latency and packet loss rate in ad hoc networks, which are caused by connection failure and dynamic mobility. A QLR-based routing method is described that adapts the Bellman equation to Q-Learning and also updates the Markov-based hidden layer. The reward function is determined by air-to-air link capability and duration. Yen's shortest path is performed to compute the shortest distance utilized for packet routing. In comparison to existing approaches, the outcomes are offered in the form of decreased delay and enhanced delivery ratio.

Table 2.1: Summary of The Literature.

	AANET Specific		**Solution Offered**			**Improvements**		
	No	Yes	Clustering	CH Selection	Other	PDR	Delay	Stability
[42]	✓		✓					✓
[43]		✓	✓	✓				✓
[44]	✓		✓	✓		✓	✓	
[45]		✓	✓					✓
[7]		✓	✓				✓	
[47]		✓	✓				✓	
[48]	✓		✓					✓
[49]		✓	✓			✓		
[50]	✓		✓	✓				✓
[51]		✓	✓	✓		✓	✓	
[52]	✓			✓			✓	
[53]	✓			✓		✓		✓
[54]	✓			✓		✓	✓	
[55]	✓			✓				✓
[57]	✓			✓		✓		✓
[58]	✓		✓	✓		✓	✓	
[59]	✓			✓				✓
[60]	✓		✓	✓				✓
[61]	✓			✓				✓
[62]		✓			✓		✓	
[63]		✓			✓		✓	✓
[64, 65]		✓			✓	✓		✓
[66]		✓			✓	✓	✓	
This Thesis		✓	✓	✓		✓	✓	✓

Chapter 3

AI-Driven Clustering And Management of AANETs

Before exploring the clustering and administration of AANETS, it is crucial to explore the significance of AI in the formation of not only AANETS but also the majority of contemporary vehicular Ad-hoc networks.

There are several ways in which artificial intelligence (AI) may be utilized to improve the efficiency and performance of ad hoc networks. AI may be used, for instance, to optimize the routing of data packets in a network, increase communication reliability, and enable the network to adapt to changing conditions [67]. In particular, machine learning algorithms may be used to evaluate network traffic patterns and anticipate the optimal paths for data packets [68]. They may also be used to optimize the distribution of network resources, such as bandwidth and power so that all devices have access to the necessary resources. Overall, the employment of AI and machine learning algorithms in ad hoc networks can increase their efficacy, dependability, and flexibility, making them more usable and effective for a wide range of applications.

Unique properties of AANET justify the application of AI algorithms. For instance, AANETs have a high density and exist in 3D space; therefore, there are several data-related factors that must be addressed while constructing and operating the network [69]. In addition, the AANETs have no structure by default, do not rely on a central organization, and any aircraft can forward traffic. In order to make this communication as efficient and steady as possible, it is necessary to employ an AI algorithm for aircraft clustering.

3.1 Unsupervised Learning Methods For Clustering

Unsupervised learning is a sort of machine learning in which the training data for the model is not labeled. Instead, the model is supplied with a dataset and tasked with discovering patterns and correlations within it [70]. The tasks of clustering and dimension reduction are instances of unsupervised learning. Clustering is performed via unsupervised learning with the objective of grouping related data elements. Based on some measure of similarity, the method assigns each data point repeatedly to a cluster until the assignment of points to clusters stabilizes. Image segmentation, anomaly detection, customer segmentation, and natural language processing are a few examples of the many applications for clustering.

In the context of AANETs, unsupervised learning is particularly significant for forming topology and maintaining the airplane-based ad hoc network. As previously indicated, AANETs are naturally unstructured, and unsupervised learning can assist in determining patterns and classifications in aviation data, hence facilitating the formation of meaningful clusters. Consequently, aircraft with comparable positions, directions, speeds, or altitudes will be grouped, resulting in a longer lifespan for air-to-air connections and greater stability, as described in chapter 1. Important types of unsupervised clustering are discussed as follows.

3.1.1 Prototype-based Clustering

Prototype-based clustering, also known as exemplar-based clustering, is a sort of clustering algorithm that defines the groups using a set of sample instances, or prototypes. Each prototype represents a cluster and is often picked as the most representative or typical data point in the cluster [71]. The data points are thereafter allocated to the cluster corresponding to the nearest prototype. The k-means method is an example of a prototype-based clustering algorithm. In k-means, the prototypes are chosen as the centroids of the clusters, which are determined as the mean of all data points in the cluster [72]. The data points are then allocated to the cluster indicated by the nearest centroid. K-means can also be used for the analysis and classification of delay-sensitive data as suggested in [73]. Other examples of prototype-based clustering algorithms include the k-medoids and the Fuzzy c-means (FCM) method.

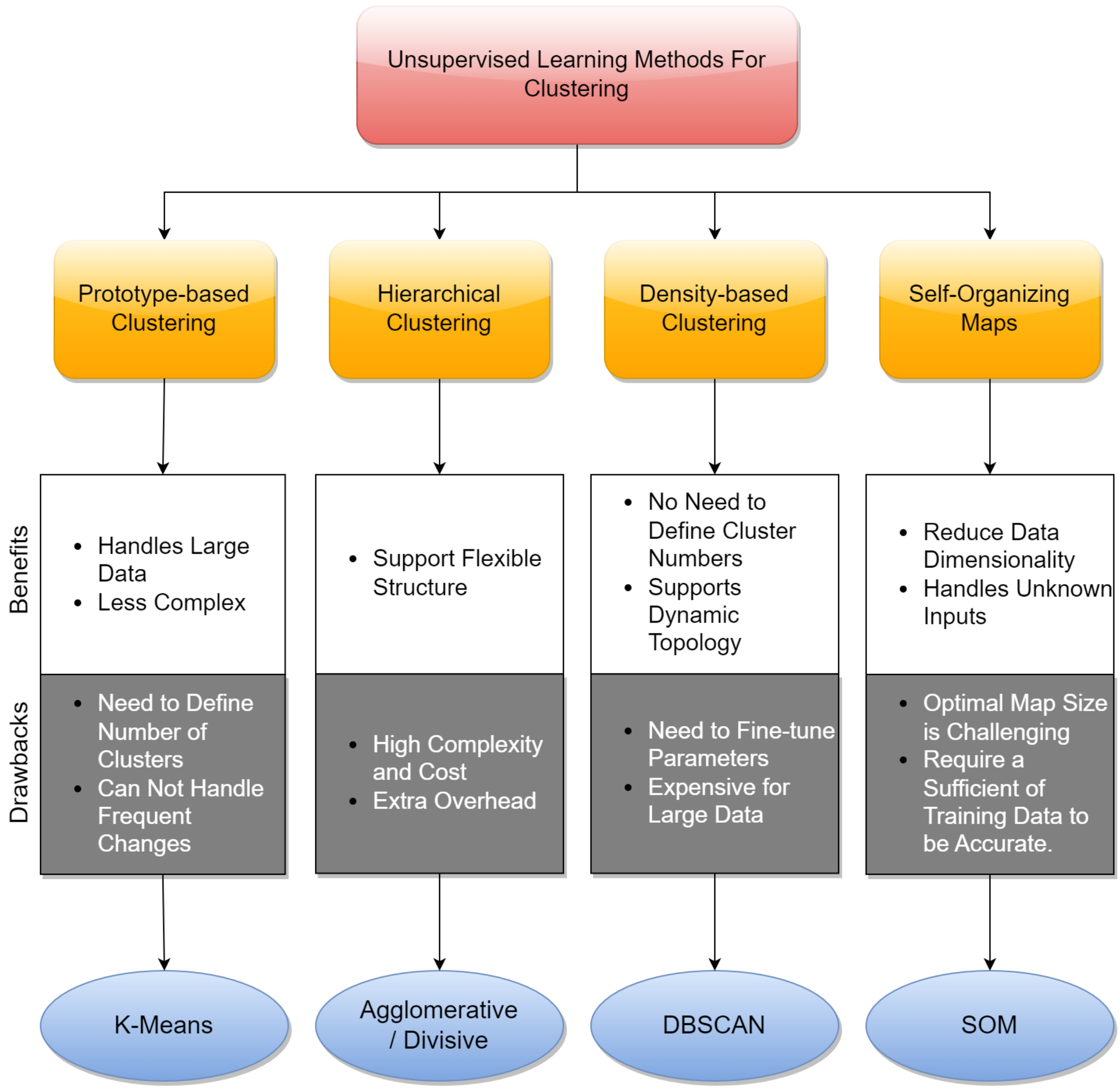


Figure 3.1: Unsupervised Approaches Used for Clustering in Ad Hoc Networks

This sort of clustering is simple to implement for AANETs and is effective with big data sets, resulting in enhanced performance in a dense and dynamic topology. The number of clusters must be determined before the algorithm's execution, which is counterintuitive and inappropriate for the dynamic nature of ad hoc networks.

3.1.2 Hierarchical Clustering

Hierarchical clustering is a sort of clustering method that creates subsets of the preceding cluster. Hierarchical clustering can be beneficial for exploratory data analysis when the number of clusters is unknown in advance [74], which is the case for AANETs or when the clusters have non-convex geometries. However, massive datasets might make it computationally costly. There are two primary hierarchical clustering strategies:

- Agglomerative: This method begins with individual data points as distinct clusters and merges the closest clusters together repeatedly until all data points belong to a single cluster. This technique generates a dendrogram, a tree-like structure used to display the hierarchy of groups.
- Divisive: This approach begins by grouping all data points into a single cluster, then divides the cluster into smaller clusters until each data point belongs to its own cluster.

It is difficult to maintain an accurate perspective of an ad-hoc network since nodes can roam freely, and the network architecture might change fast. Nodes may be grouped into clusters using hierarchical clustering, where each cluster is represented by a cluster head. The cluster leader is responsible for keeping member information and interacting with other cluster leaders to establish a hierarchical structure.

A distance-based technique is one method for implementing hierarchical clustering in AANETS, where the distance between aircraft is determined by the connectivity and quality of Air-to-Air links between them. The aircraft are then divided into clusters depending on their distances with the aim of minimizing communication overhead and maximizing network connection overall. A distributed method is an alternative in which each aircraft contributes to the construction of clusters by making judgments based on local data and interacting with surrounding nodes. This method may be more resistant to alterations in network topology and cluster head failures.

Hierarchical clustering can be advantageous for AANETs because it provides for a flexible and dynamic network structure and improves the scalability and efficiency of the network. However, this approach is computationally complex and costly, and it adds a great deal of overhead to network operations.

3.1.3 Self-Organizing Maps

A Self-Organizing Map (SOM) is an artificial neural network used for unsupervised learning of data pattern recognition. SOMs are taught to organize their weights in a manner that mirrors the data input structure. This is accomplished by iteratively modifying the network's weights depending on the input data and a learning algorithm until the network reaches a stable, well-organized state [75]. Frequently, SOMs are employed for dimensionality reduction, data visualization, and anomaly identification. A variant of SOM called Self Organizing Feature Map (SOFM) can also be used for fake task prevention in Mobile Crowdsensing [76–83]. The main advantages of using SOM include the following:

- Managing the high-dimensional data prevalent in ad hoc networks.
- SOMs may be used to uncover patterns and structures in data that are not immediately apparent.
- Being able to adapt to network changes because they can restructure themselves depending on fresh data.

Considering input data of (m, n), where m indicates the number of training steps and n represents the number of data features. SOM will utilize the weight size of (n, C), where C is the number of clusters. After iterating through each training step's input data, the weight will be modified. The fundamental algorithm used to update the weight is denoted by 3.1, where α is the learning rate, j is the winning vector, i is the n^{th} training feature, k is the n^{th} training from input data and, pre is the old weight.

$$W_{ij} = W_{ij}(pre) + \alpha(t)(x_{i^k} - W_{ij}(pre)) \tag{3.1}$$

Ad-hoc networks can utilize SOMs for clustering. Each node in the ad-hoc network is represented by a neuron in the SOM in a clustering technique utilizing SOMs. Using a learning algorithm, the SOM is taught to group neurons according to similarities in the properties of the nodes they represent. The generated neuronal clusters can then be utilized to cluster the nodes in the ad hoc network. This technique has the advantage of being able to react to changes in the network, which is one of the issues with AANETs since the SOM is able to restructure itself depending on new data and suitably manage newly added aircraft. Moreover, SOMs can manage high-dimensional data, which is one of the primary characteristics of AANET. In the case of AANETs, each neuron input to SOM represents an airplane, and SOM may translate these inputs to the reduced output and minimize dimensions. Due to the changeable topology of AANET, setting the map size is the most difficult aspect of employing SOM. SOMs also have the disadvantage of being computationally costly to train and run. In the AANET system, determining the learning rate and the size of the map might also be difficult. While SOM can function effectively in networks with low density, it could not yield ideal clusters.

3.1.4 Density-based Methods

Density-based clustering is the final category to examine for ad hoc network clustering. Density-based clustering is a form of clustering technique that unites data points that are densely placed (i.e., have a high density) and separates data points that are dispersed (low density or noise points). This clustering technique is beneficial for detecting clusters with varying forms or when the number of clusters is unknown beforehand. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) and OPTICS (Ordering Points To Identify the Clustering Structure) [84] are some examples of density-based clustering methods. Density-based clustering and, more specifically, DBSCAN, which is the basis of this thesis, will be discussed in detail in the next section.

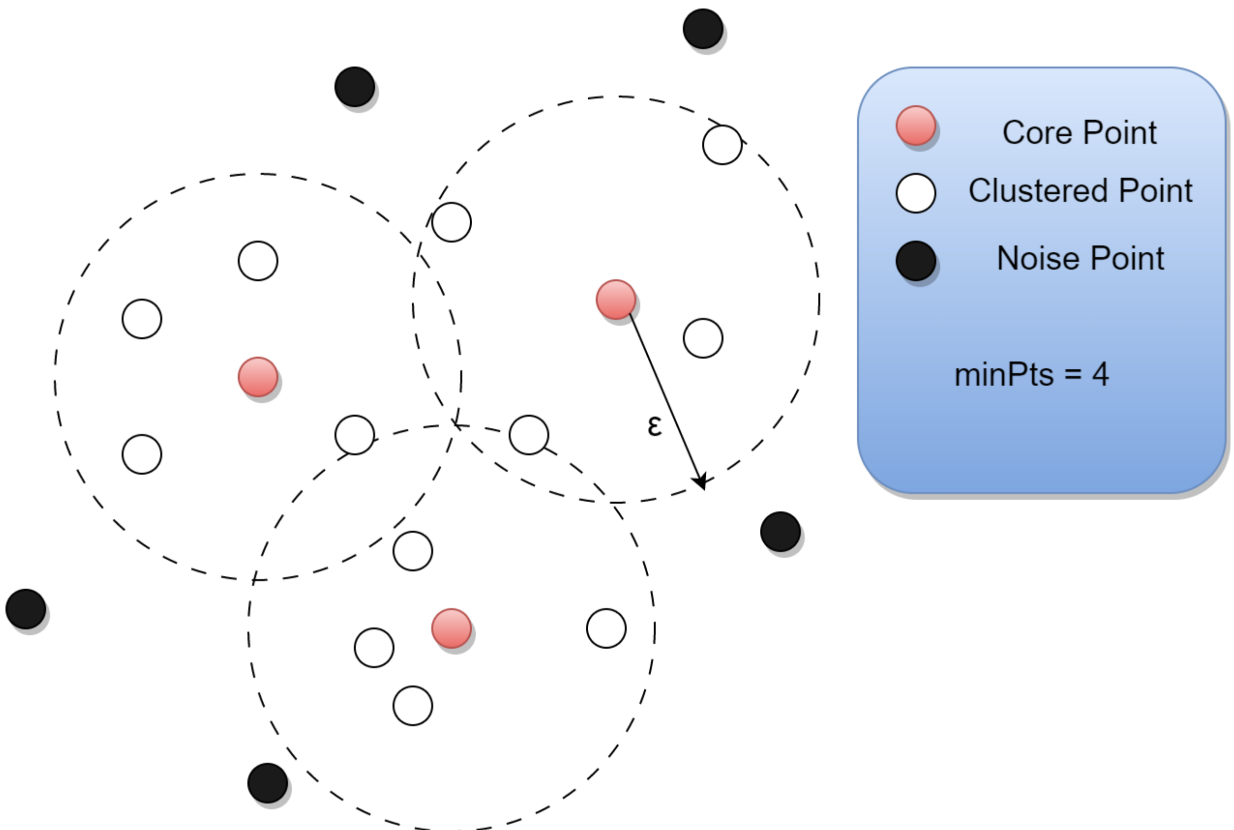


Figure 3.2: Example of DBSCAN Clustering

3.2 Density-based Clustering and DBSCAN

Density-based clustering is a non-parametric technique in which the groups in the data are regarded as high-density areas [85]. This approach does not require the number of clusters as an input parameter. Density-based clusters are not always collections of points with a high degree of within-cluster similarity as assessed by the distance function, but they can have any shape in the feature space. This feature improves the clustering performance of spatial data utilized in the formation of AANET clusters. The main advantage of this approach is:

- Density-based clustering algorithms, such as DBSCAN, do not rely on fixed cluster centers and are capable of adapting to the shifting placements of network nodes. This makes them suitable for ad-hoc networks, such as AANETs, in which nodes are extremely mobile, move quickly, and can regularly change places.
- Density-based clustering may detect clusters with variable forms and fluctuating densities, which is important in AANET, where the shape and density of clusters may fluctuate over time.
- By finding dense groups of nodes, density-based clustering may be used to discover and isolate nodes that are acting erratically or may be hacked, hence enhancing the security of ad hoc networks.

Density-based clustering algorithms function by detecting regions of the data space with a high density of data points and then classifying such regions as clusters. Typically, these algorithms have two essential parameters: a density threshold, which defines what constitutes a dense zone, and a distance metric, which measures the distance between data points.

In ad-hoc networks, density-based clustering can be used to cluster nodes that are located in close proximity. Nodes in ad-hoc networks are typically mobile and subject to rapid positional changes, making standard clustering algorithms that rely on established cluster centers and fixed bounds inappropriate. The construction of clusters for routing is one use of this class of algorithms. Nodes may be grouped into clusters using density-based clustering so that each cluster has a unique cluster head. The cluster head is responsible for transmitting data on behalf of the cluster's nodes, hence decreasing the number of hops necessary for data to reach its destination. In ad hoc networks, density-based clustering may also be utilized for security and intrusion detection. By finding dense clusters of nodes, it may be utilized to discover and isolate nodes exhibiting anomalous behavior or that are possibly compromised.

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a density-based clustering technique that separates densely located data points from sparsely distributed data points. It is especially helpful when finding clusters with non-convex geometries or when the number of clusters is unknown beforehand. The fundamental concept of DBSCAN is to combine points that are near-situated and separate those that are dispersed. It consists of two important parameters: the radius (ϵ) of the neighborhood around each data point and the minimum number of points (minPts) necessary to produce a dense zone. Within the radius of a dense zone, points are regarded to belong to the same cluster. These two parameters will result in nodes being classified into three groups:

- Core Point: A point is considered a core point if it is surrounded by at least minPts points in a radius of epsilon.
- Noise Point: A point is considered an outlier or noise point if it is neither a core point nor accessible from any core points.
- Member Point: The surrounding points that are deemed to be in the same cluster as the core point.

The algorithm begins by choosing a random point from the dataset and determining whether or not it is part of a dense zone. If so, the algorithm expands the cluster by adding any points within the radius of the dense zone that has a minimum of minPts points inside their radius. Repeat the procedure until no further points can be added to the cluster. In general DBSCAN algorithm follows these steps:

1. minPts and epsilon are determined before clustering begins.
2. The beginning point for clustering is picked at random, and the algorithm then utilizes the defined epsilon neighborhood region to look for members surrounding the initial point. If there are at least minPts points surrounding it, the node will be picked as a core point, and its neighbors will all belong to the same cluster.

3. Iterate throughout the remaining unvisited points that are not grouped in the previous step and repeat the process. The points that do not belong to a cluster are known as noise points.

4. The algorithm terminates after every point has been visited and categorized.

As described, using density-based clustering and DBSCAN for clustering in AANETs offers several advantages. In the next chapter, we present a modified version of DBSCAN that is customized to the requirements of AANETs and can enhance the performance of clustering and, therefore, communication quality.

Chapter 4

Modified DBSCAN Clustering for Aeronautical Ad-Hoc Networks

In light of the analysis of the current state of the art in Chapter 2, this chapter fills in the gaps in the literature by including additional features of the dataset when doing the clustering task. That being said, the state of the art calls for a new method that offers density-based clustering and utilizes key flight features so as to generate clusters that are more stable. Also as discussed in Chapter 3, density-based clustering has several advantages that makes it a good candidate for cluster-based AANET formation and is not investigated in the literature.

To maximize the effectiveness of AANETs, the network architecture must be designed as clusters to increase the stability and lifetime of A2A links [7]. However, unstable clusters will shorten the lifespan of A2A links, which is the primary cause of stability issues. Therefore, in an AANET with a shorter link lifespan, lower network stability is inevitable as a result of frequent connection and termination, as well as cluster changes [86]. This chapter builds on the motivation that the root cause of AANET stability issues needs to be addressed by identifying suboptimal clusters. With this in mind, we aim to obtain more structured clusters by utilizing machine learning algorithms to form clusters comprised of aircraft with similar characteristics. A modified DBSCAN that can exploit and use characteristics of the real flight data, including longitude, latitude, altitude, direction, and velocity, is utilized to build meaningful clusters. In the following sections, we present the proposed solution and then provide the simulation results that compare the performance

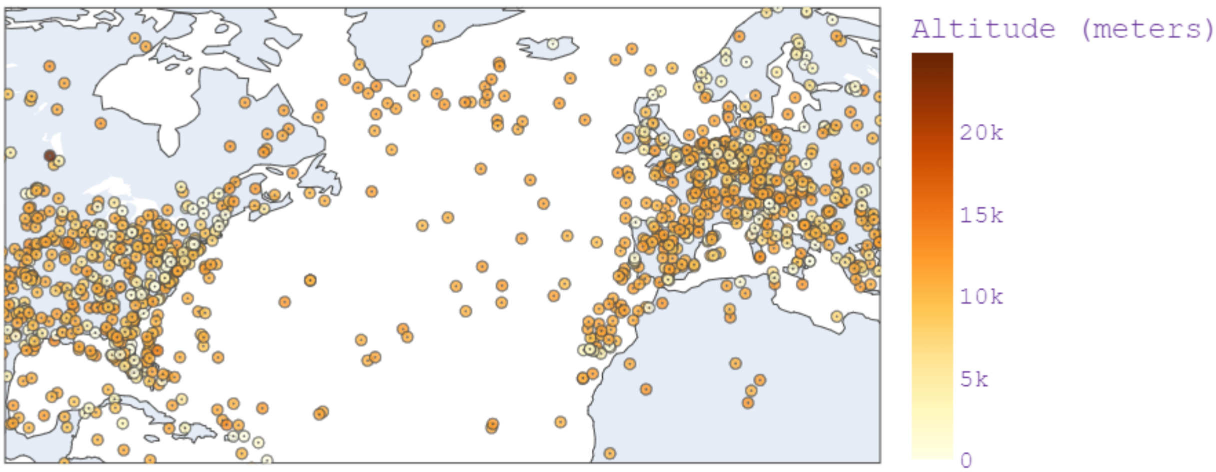


Figure 4.1: Visualization of flight data

of modified DBSCAN with the current state of the art.

4.1 Proposed Clustering Solution

4.1.1 Data model

To obtain relevant findings, we must experiment with an actual dataset; for this, the Aviationstack API is utilized [87]. Using the API, the necessary flight data to generate a real-world data set can simply be retrieved. The dataset includes various parameters, including timestamp, latitude, longitude, altitude (in meters), velocity, and direction (in degrees), as well as aircraft model, airline, departure, and arrival information. The data were extracted between 6:00 a.m. and 11:59 p.m. UTC Time on January 4, 2022. A sample slice of this data is visualized in Figure 4.1. The generated dataset comprised over 10000 rows; thus, the flights are limited to only include active flights over 10000 meters in elevation, removed irrelevant information, and converted the dataset to panda data frames prior to clustering.

4.1.2 Clustering model

In AANETs, K-means is the most prevalent clustering algorithm because of its simplicity [4]. K-means is an unsupervised machine learning algorithm that can cluster N observations (rows in an array of coordinates) into k groups. Nevertheless, K-means is not an optimal approach for latitude-longitude spatial data since it aims to minimize variance rather than geodetic distance, and at latitudes on the spherical surface, distant from the equator, there is severe distortion. Considering density-based characteristics of geographical data, Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is utilized as a more viable candidate in this study. Although The algorithm would still function, its outcomes are subpar, and there is not much that can be done to enhance them as the only variable that we can tune is the number of clusters.

DBSCAN is a robust algorithm for clustering utilized in a variety of machine-learning applications. It groups together dense data regions inside the feature space. Two parameters are required for the model: ϵ and $MinPts$. A point that has surrounding $MinPts$ within an ϵ-radius will be grouped with them. Moreover, if a point is not inside the range of a core point, it will be grouped on its own and referred to as a noise point. DBSCAN outperforms K-means for spatial latitude-longitude data. There is no need to define the number of clusters, and DBSCAN is resilient against outliers due to the noisy data points.

It is feasible to employ the Haversine distance rather than the default euclidean distance. In this case, the coordinates need to be converted to radians and fed to DBSCAN to enable using the Haversine formula for determining the distance between two points. In this instance, the ϵ parameter is relevant since it reflects the distance between nodes. Datasets such as the real flight dataset that is used in this work consist of additional features beyond latitude and longitude. If they are integrated and transformed into radians, the epsilon parameter no longer reflects a physical distance; tuning it also becomes more difficult and incoherent. However, it is feasible to include other features of the dataset in the clustering process with a custom distance matrix that exploits different types of dataset features. This approach also employs a weighting mechanism to indicate the significance of each computation factor, as seen in Eq. 4.1-4.2.

For determining the distance between points in the standard implementation of DBSCAN, we can just use a metric such as Euclidean or Haversine, but it is also possible to utilize a precomputed distance matrix instead of simply feeding the algorithm the name of a metric to determine the distance. It is feasible to use a custom distance matrix that

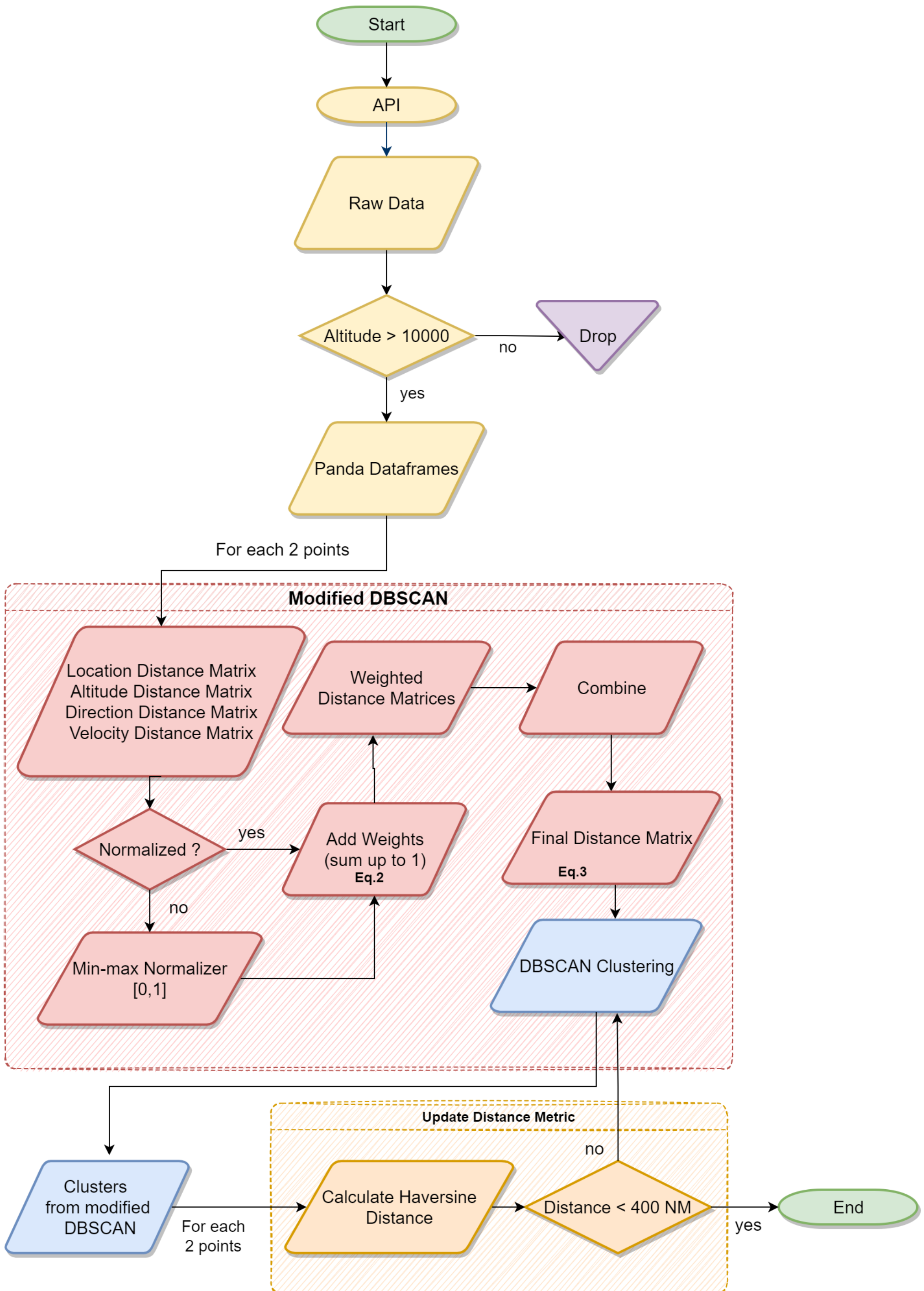


Figure 4.2: Proposed Model with Modified DBSCAN-based Clustering in AANET

uses different types of dataset features. This approach also employs a weighting system to indicate the significance of each computation factor as seen in equations 4.1 and 4.2.

$$D_{ij} = w_1 \cdot D_{ij}^{X_1} + w_2 \cdot D_{ij}^{X_2} + ... + w_n \cdot D_{ij}^{X_n} \tag{4.1}$$

$$\sum_{k=1}^{n} w_k = 1 \text{ and } 0 \leq D_{ij} \leq 1 \tag{4.2}$$

For each attribute X, the model computes a weighted sum of the precomputed distance matrices. Each pair of indices, i and j, represents two data points. Each of the n feature matrices should be re-scaled to the interval [0,1] if they were not initially in the same value scale. The n weights add up to 1, and each weight ($w \in \{\alpha, \beta, \gamma, \delta\}$) indicates the contribution of a particular feature to the final distance matrix D. The final distance matrix comprises distances in the interval [0,1] from which ϵ must also be chosen. In particular, the distance matrices are separately calculated and re-scaled, and their contribution to the final distance can then be computed.

$$D_{ij} = \alpha \cdot D_{ij}^{location} + \beta \cdot D_{ij}^{altitude} + \gamma \cdot D_{ij}^{direction} + \delta \cdot D_{ij}^{velocity} \tag{4.3}$$

In this work, the combination of 4 major features helps us obtain more stable clusters. The mathematical model used for the flight data is given in equation 4.3. $D1 = D_{ij}^{location}$ is the geographic distance between two aircraft (i and j) as calculated by the Haversine formula. $D_{ij}^{altitude}$, $D_{ij}^{Direction}$, and $D_{ij}^{velocity}$ are the difference between the altitude, speed, and direction of two aircraft, respectively. Min-max normalization is used to re-scale each of these four matrices to the interval [0,1]. The sum of the variables $\alpha, \beta, \delta,$ and γ is equal to one and represents the weights. The resulting distance matrix is fed as the pre-computed metric to the DBSCAN.

The ideal weight distribution is depicted in Table 4.1. A grid-search [88] is conducted on the values of $\alpha, \beta, \delta, \gamma$ (weights of different features) and DBSCAN parameter values (ϵ and $minPts$). For each combination, the Silhouette coefficient metric is evaluated. Given the constraints of AANETs, the combination of all features (first row) outperforms the configurations with a single feature. It is evident that even using the location as a single feature performs well but, the combination of features is superior. These values are specific

Table 4.1: The Optimal Balance of the Weights

α	β	γ	δ	**Score**
0.5	0.3	0.1	0.1	0.847
1.0	0.0	0.0	0.0	0.751
0.0	1.0	0.0	0.0	0.623
0.0	0.0	1.0	0.0	0.412
0.0	0.0	0.0	1.0	0.468

to the data we utilized also the α has a higher value than β because the location is made of the combination of both the latitude and longitude.

4.2 Performance Study

This section describes the simulation environment and how the real data is fed into the simulations. Following the description of the environment, simulation results regarding clustering factors and network performance are presented.

4.2.1 Environment

To test the AANET performance as the outcome of density-based clustering, OMNET++ and INET framework are utilized to simulate air traffic using real data obtained from the Aviationstack API. Due to the fact that AANETs are generally feasible in remote areas where there is no ground infrastructure, we evaluate a 2000 x 2000 NM (nautical mile) testing area over the North Atlantic Ocean [89]. The layer two protocol is emulated by Using the INET framework [90] and AODV to route the packets, the dynamic nature of A2A networks is considered, and the simulated link capacity is set to 1 Mbit/s and the packet size to 64 kb. At an altitude of 10,000 meters, the maximum transmission range of A2G and A2A connections is 200 and 400 NM, respectively. However, the real communication range is likely to be shorter, and it can presume that if the distance between two aircraft is less than 400NM, there is a chance of establishing an A2A link. Python is utilized for data processing and clustering, while Plotly and Seaborn libraries are utilized for graph plotting.

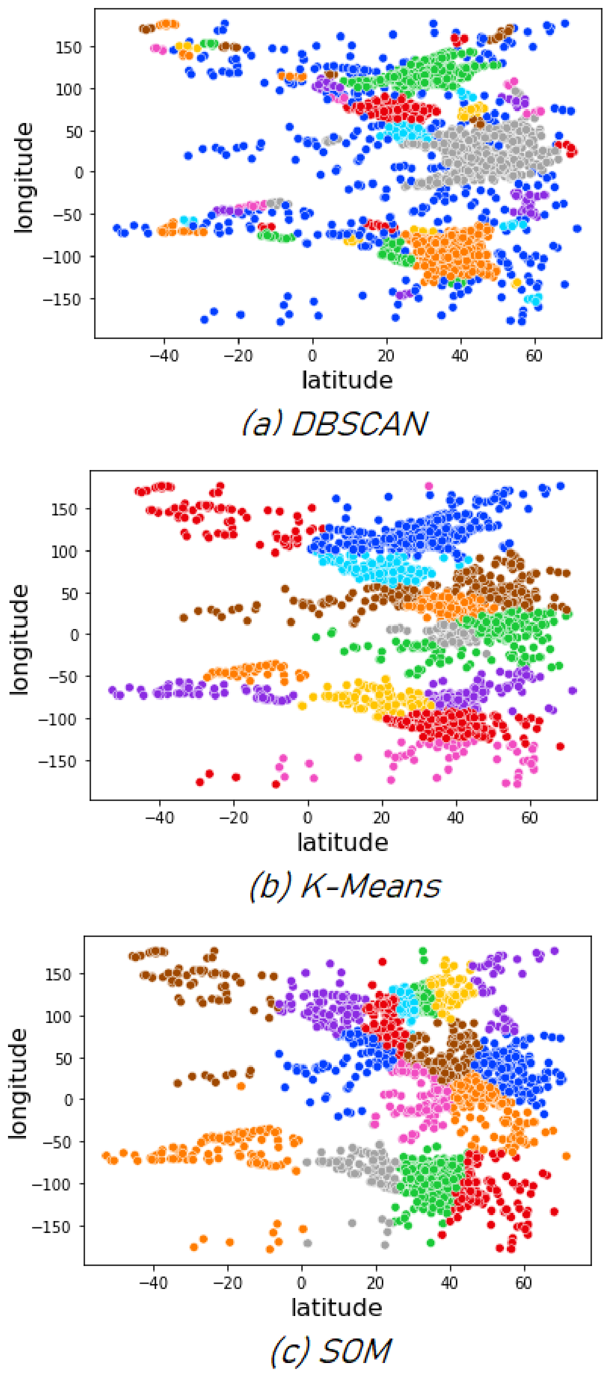


Figure 4.3: Comparison of clustering outputs under DBSCAN, K-Means, and SOM

Table 4.2: Simulation Parameters for Clustering

Simulation Parameter	**Value**
A2A Transmission Range	400 NM
Aircraft Altitude	>10000 meters
Packet Size	64 Kbytes
Link Capacity	1 Mbit/s
Timezone	UTC
Clustering algorithms	K-Means, SOM, DBSCAN, DBSCAN-Enhanced
Number of clusters	[5, 30]
Number of aircraft	[50, 350]
Packet arrival rate	[5, 40] packets/sec

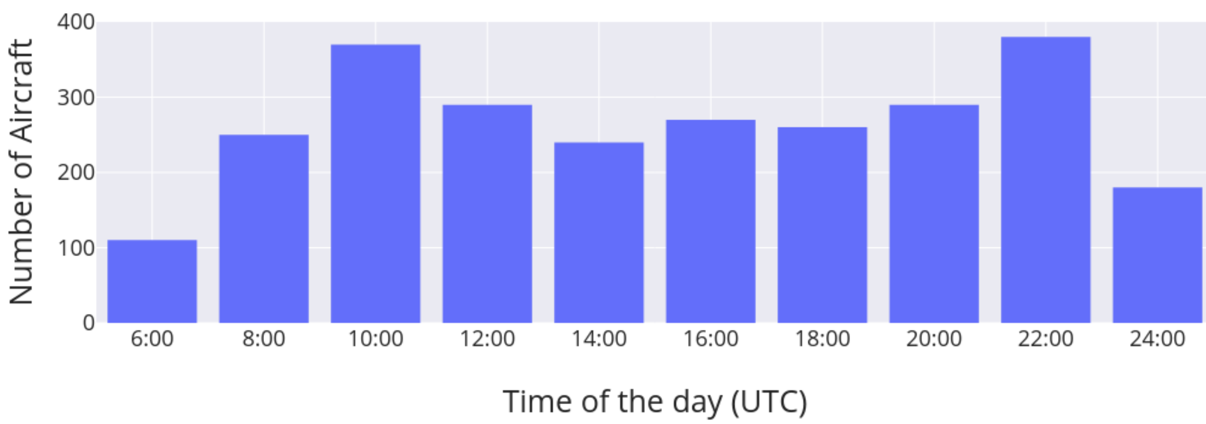


Figure 4.4: Number of aircraft vs time of the day

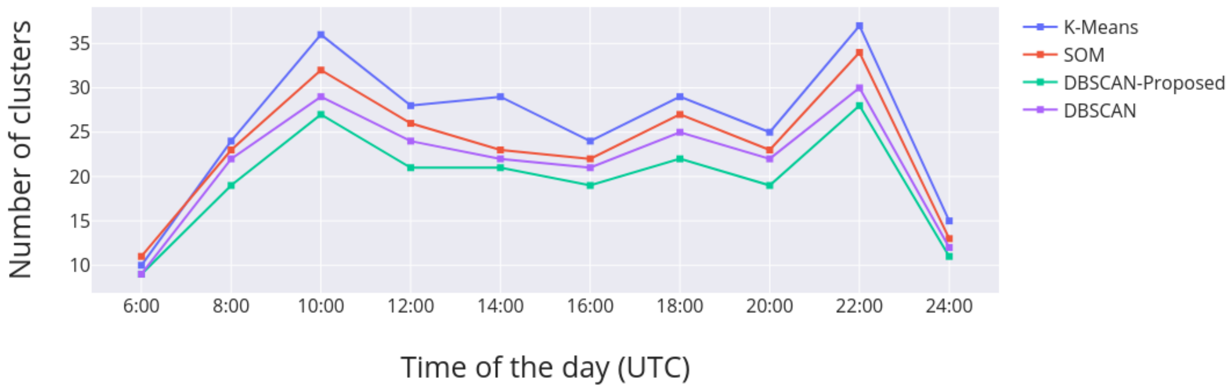


Figure 4.5: Number of clusters vs time of the day.

4.2.2 Simulation results

The proposed modification to DBSCAN-based clustering in AANET is compared to the original implementation of DBSCAN-based clustering in AANET, K-means, and Self-organizing map (SOM) [91], which is an additional unsupervised data clustering method. As indicated in [45], K-means is adjusted to utilize the Haversine distance. Through simulations, the efficacy of clustering algorithms is measured by evaluating two sets of parameters:

- 1) Clustering parameters, including the total number of cluster changes and the number of clusters at a time.
- 2) Network performance parameters, including packet delivery ratio and end-to-end transfer delay

Clustering factors

Figures 4.5 and 4.6 illustrate the outcomes of clustering algorithms based on the number of clusters throughout the day and the number of changes in the number of clusters under each algorithm during the day, respectively. It is worth noting that Figure. 4.4 depicts the number of aircraft during the day. Thus, it is straightforward to establish a relationship between those numbers and the number of clusters shown in Figure. 4.5, where the number of clusters at a specific time is depicted. Clearly, the peak time of the air traffic is between 10 a.m. and 10 p.m.; hence there are more clusters at these times. In terms of the total number of cluster changes, a smaller number suggests a more stable algorithm, as shown in Figure 4.6, which presents the total number of cluster changes that occur over the duration of the simulation. Due to the unstable nature of A2A linkages, there are numerous cluster changes, but the proposed method results in the lowest number of changes. Figure 4.3 depicts a comparison of cluster formations, highlighting how DBSCAN handles outliers in particular (blue dots).

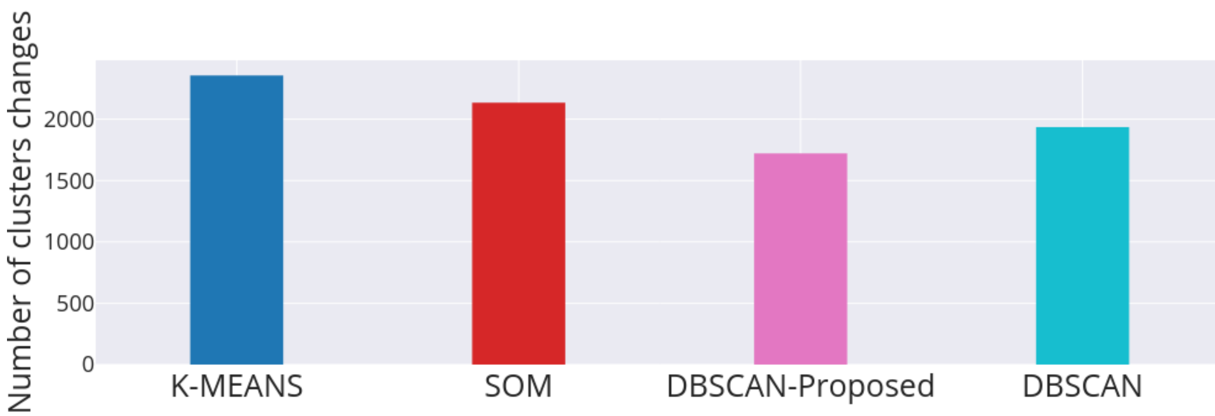
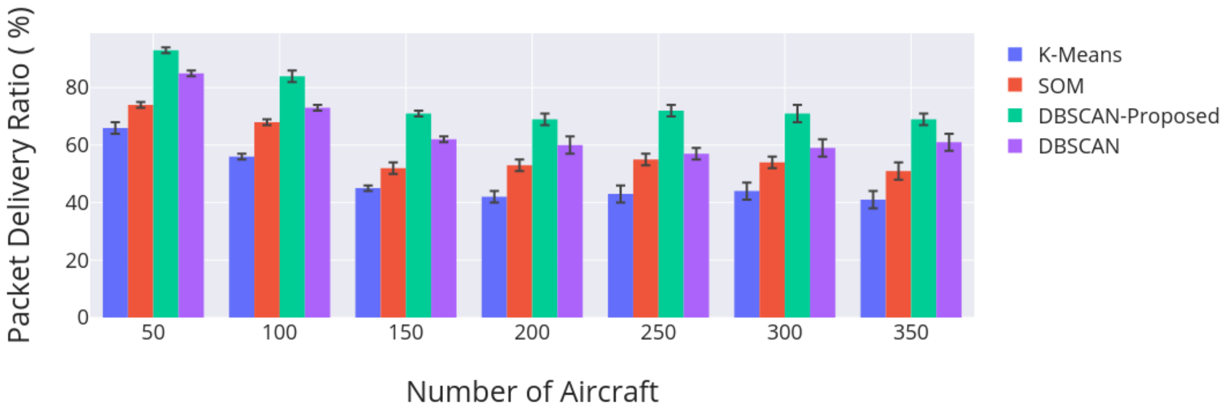
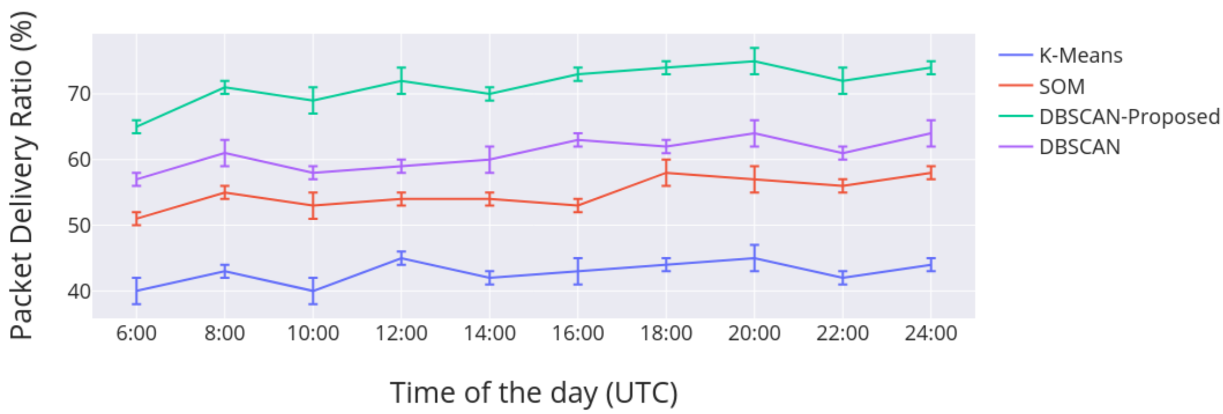


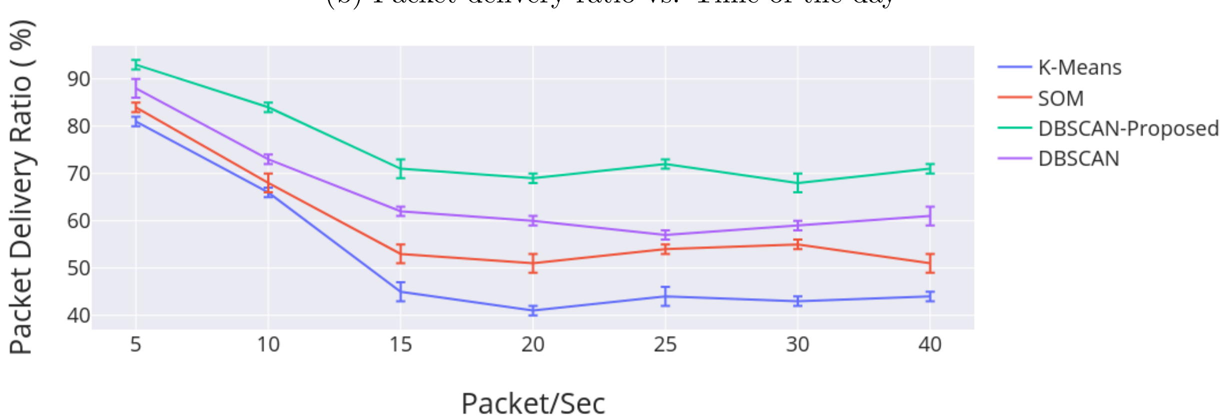
Figure 4.6: Number of cluster changes under different clustering algorithms.



(a) Packet delivery ratio vs. Number of aircraft



(b) Packet delivery ratio vs. Time of the day



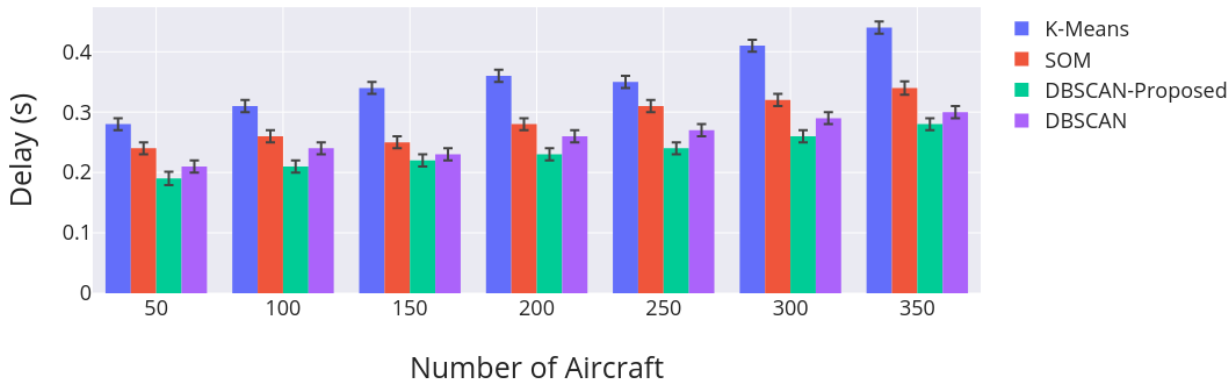
(c) Packet delivery ratio vs. Packet generation

Figure 4.7: Packet Delivery Ratio Performance

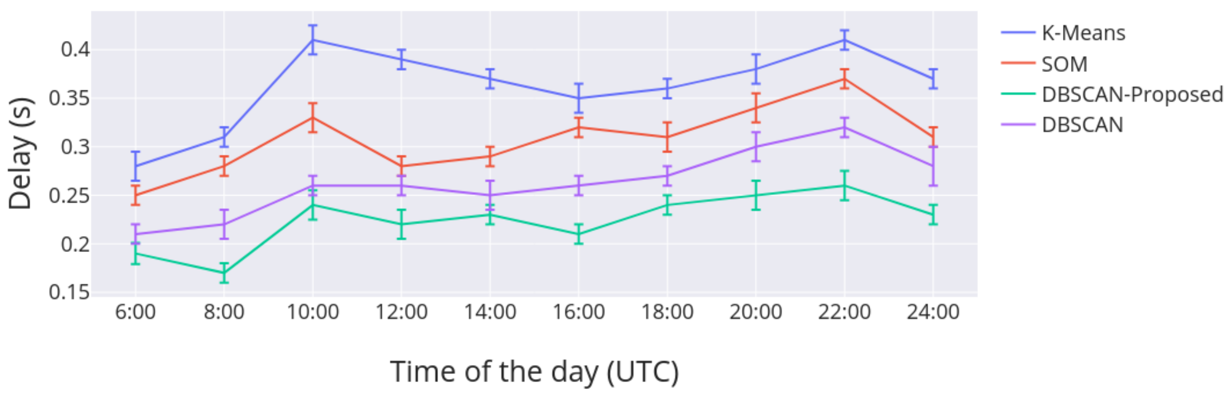
Network performance factors

To evaluate the impact of the proposed modification to the DBSCAN-driven clustering in AANET on the network performance, packet delivery ratio and end-to-end transfer delay are used as the performance indicators. Below is a detailed discussion of how these performance indicators are affected by different clustering algorithms.

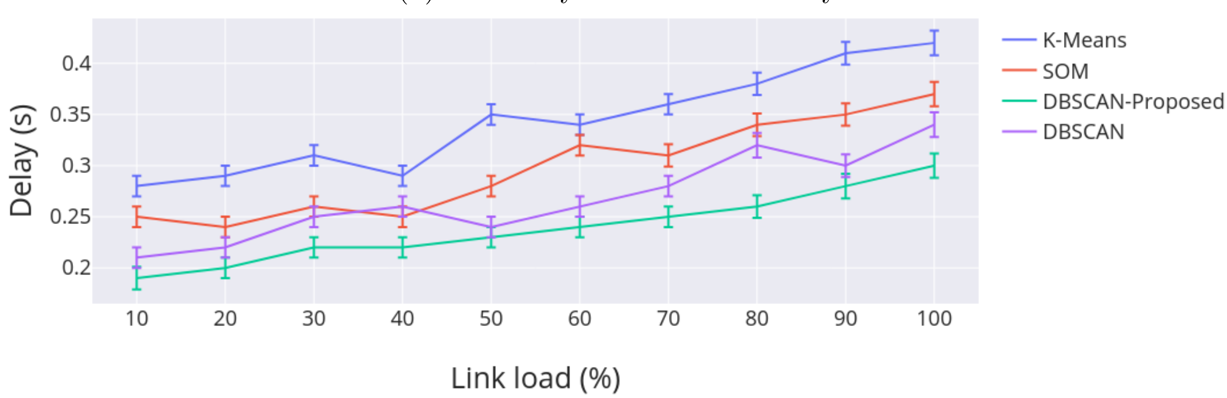
- Packet Delivery Ratio: As demonstrated in Figure 4.7a, the overall trend of packet delivery ratio (PDR) drops as the number of flights grows. Figure 4.7b depicts the PDR as a function of simulation duration. According to the AANET dynamic environment and the uncertainty of the number of clusters, 40 aircraft are randomly chosen to obtain the average results at a confidence level of 95% across 30-minute intervals via simulations repeated 10 times. Traffic is then sent from a random source to a random destination. According to the simulation, the PDR exceeds 65 percent, whereas other clustering algorithms result in a PDR of less than 60 percent. Substantial reduction is experienced during peak traffic hours. Figure 4.7c illustrates how the packet arrival rate impacts the network performance. While an increase in packet arrival rate decreases the packet delivery ratio, PDR stabilizes around 20 packets per second. Because the capacity of the link is set at 1 Mb/s, the higher amount of aircraft and packet generation will congest the link. Therefore the packets will stuck in the queuing and eventually, some will be dropped. According to the simulations around 21 % of all the PDR results are related to this issue.



(a) E2E delay vs. Number of aircraft



(b) E2E delay vs. Time of the day



(c) E2E delay vs. Link load

Figure 4.8: End-to-End Delay Performance

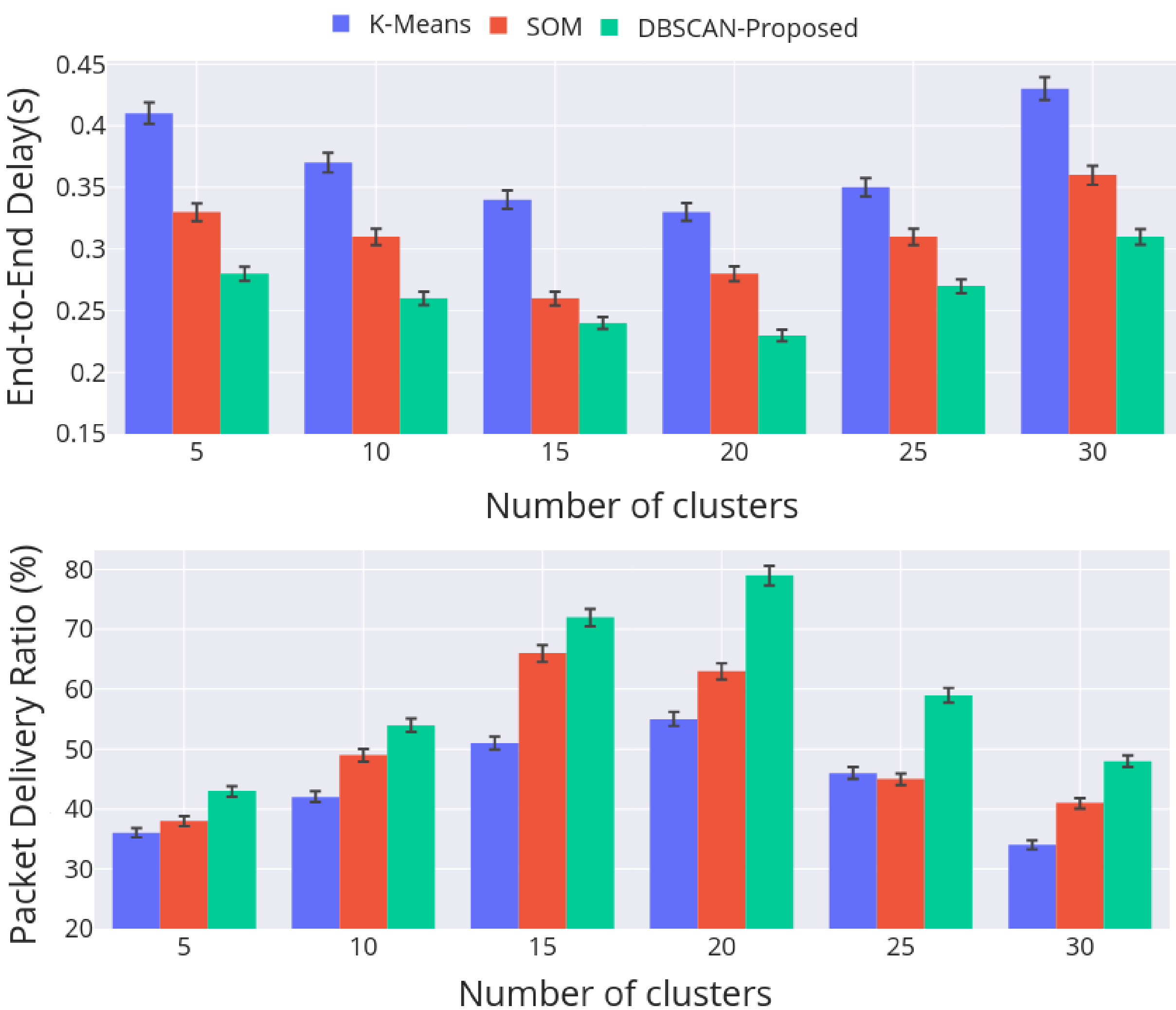


Figure 4.9: Impact of the number of clusters E2E and PDR

- End-to-End Transfer Delay: Fig. 4.8 evaluates the impact of clustering on end-to-end transmission delay with respect to varying packet arrival rates, time of the day, and the number of aircraft. As observable in Fig. 4.8a, it is possible to project an increase in end-to-end delay as the number of aircraft increases. Figure 4.8b illustrates the end-to-end latency during the course of the experiment. Despite the fact that it is evident that delay increases dramatically during peak hours, the DBSCAN remains below 0.25 seconds while others exceed 0.35 seconds. Figure 4.8c demonstrates that when the link capacity becomes more saturated, the delay increases dramatically. This is mostly due to the fact that packets must wait longer in the queue before being delivered to their final destination. At maximum capacity, a delay of fewer than 0.3 seconds is achieved while, other methods prolong the end-to-end delay by more than 0.35 seconds.

To assess the significance of the number of clusters on the network performance, we let DBSCAN determine the number of clusters and then use this value for other methods as well. Consequently, they are all normalized and include the same number of clusters. Here, 300 aircraft are chosen on a random basis to replicate the peak of traffic. Figure 4.9 confirms that the number of clusters is not the sole element impacting the algorithm performance and that the proposed algorithm performs better even with the same number of clusters. Therefore, it can be concluded that the number of clusters is a crucial process parameter; having an excessive number of clusters, and therefore small clusters, indicates that the process is inefficient and clustering's benefits will be minimal. This is due to the fact that too much communication is required to forward messages. Conversely, the existence of a small number of big clusters is undesirable since the medium is occupied by an excessive number of members within the same cluster, resulting in an increase in the delay.

4.3 Summary

Aeronautical Ad Hoc Networks (AANET) seek efficient communication and management solutions such as a clustered operation for in-flight connectivity. With this in mind, in this Chapter, after stressing sub-optimal clusters as the source of stability issues in AANETs, a multi-feature DBSCAN has been used to create aircraft clusters. Instead of utilizing a basic metric such as Haversine, the algorithm utilizes aspects of the dataset, such as location, altitude, direction, and velocity, to generate a precomputed distance matrix. According to the simulation findings, the number of cluster changes has been reduced by 22% compared to K-means and 15% compared to SOM. Moreover, there are 35% fewer clusters during peak traffic hours. These results imply that density-based clustering yields more stability. In terms of the communication metrics, the proposed method has also been shown to enhance the packet delivery ratio by 51% compared to K-means and by 28% compared to SOM. The proposed algorithm also decreases the end-to-end delay by up to 30%.

Chapter 5

A Novel Cluster Head Selection Method for Clustered AANET

This chapter expands upon the previous one, which utilized a multi-feature DBSCAN method for the clustering issue that exploited numerous parameters of real flight data, including latitude, longitude, altitude, direction, and velocity [92]. In lieu of a traditional distance metric such as Euclidean or Haversine, this approach generates a precomputed distance matrix and feeds it to DBSCAN. This approach also incorporates a weighting mechanism to represent the proportional relevance of each distance computation component.

This work examines the next phase following cluster formation, namely cluster head selection. Cluster heads play a crucial role in the management and effectiveness of clusters, and intra-cluster communication is largely dependent on the right selection of cluster heads [93]. A cluster head in ad hoc networks is a node that serves as the communication hub for the other nodes in its cluster. The head of the cluster is responsible for controlling and organizing communication inside the cluster and conveying information to and from other clusters. The significance of the cluster head lies in its potential to increase the overall performance and efficiency of the network by minimizing the quantity of duplicated information exchanged and by offering a more ordered communication mode. In addition, the cluster head can contribute to the network's longevity by regulating the power consumption of its cluster's nodes.

5.1 Proposed Solution

Prior to selecting the ideal cluster head, it is crucial that we have solid clusters. Sub-optimal clusters will result in stability issues and an increase in the frequency of cluster member changes. This is due to the fact that if cluster members change often, the cluster head computations will need to be performed repeatedly, which will further degrade performance. For this reason, we will develop our cluster head selection solution on top of the previously proposed modified DBSCAN [92].

The modified DBSCAN functioned by exploiting the flight dataset's numerous characteristics. Instead of supplying a basic distance metric to DBSCAN, we utilized a pre-computed distance matrix that incorporated latitude, longitude, direction, and velocity. Each of these distances was independently determined before being fed into the DBSCAN. The contribution of each distance matrix to the final distance matrix was specified by a weighting system. In this manner, more characteristics from the dataset are utilized than when location alone is employed.

Dense clusters are produced by DBSCAN, and all nodes inside a cluster may communicate with one another. This is due to the fact that no two nodes within the same cluster are separated by more than 400 nautical miles (NM), and if the distance between two aircraft is less than 400 NM, there is a probability that they can connect. Consequently, we might presume that any node in the cluster can serve as the cluster leader, which is not necessarily the case. The cluster's head must be the most connected node, capable of optimum and reliable communication with the greatest number of cluster members.

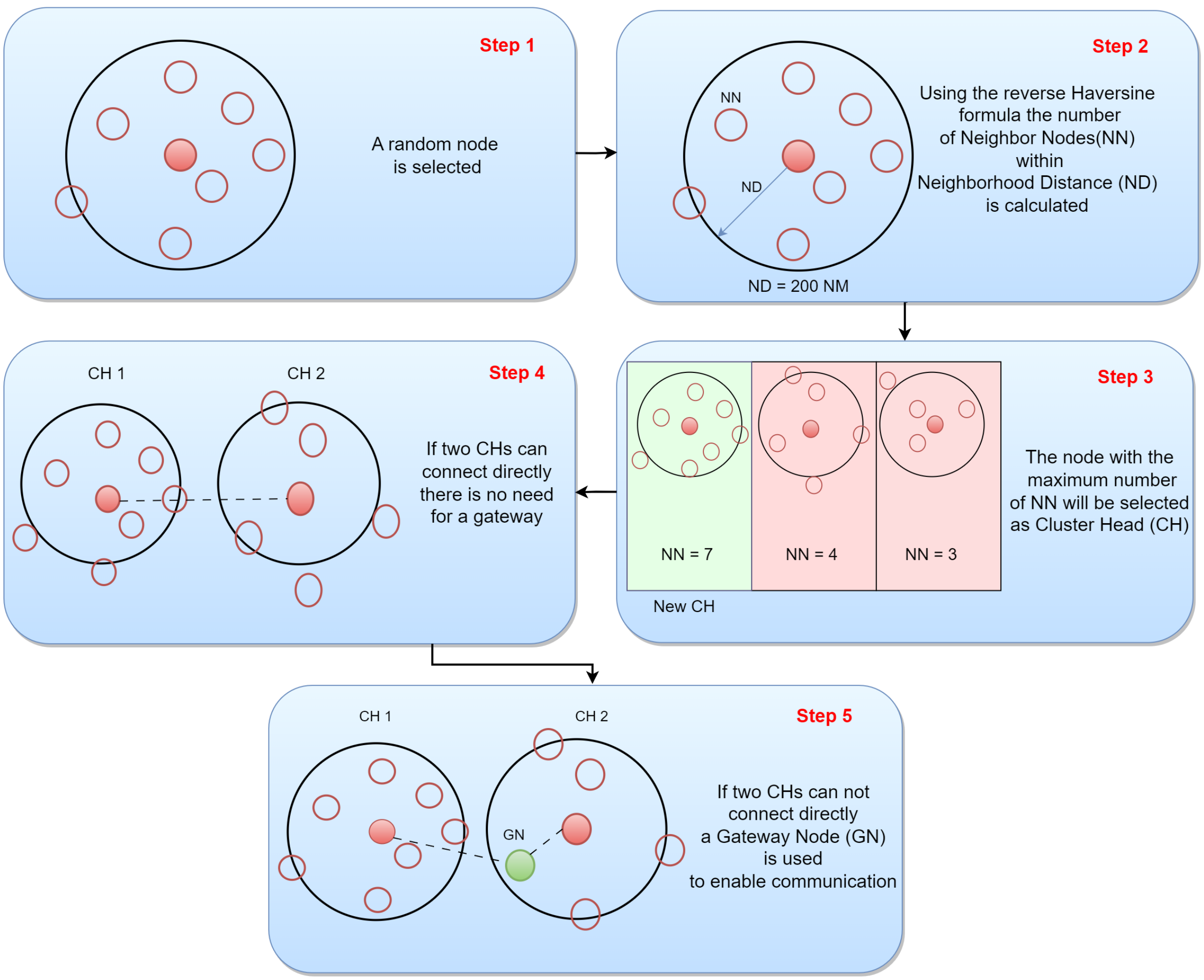


Figure 5.1: Cluster Head Selection Process

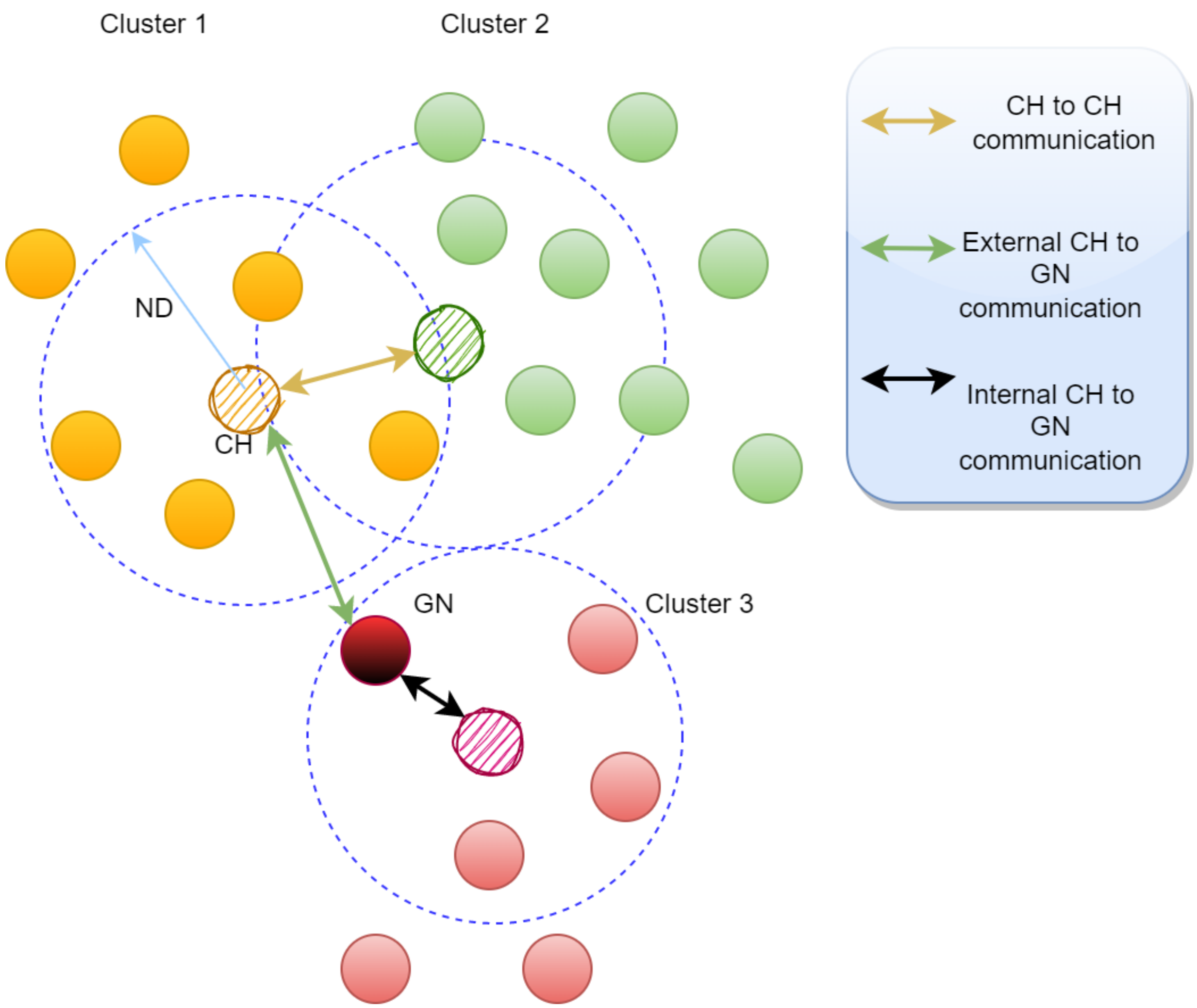


Figure 5.2: Cluster Head Selection Terminology.

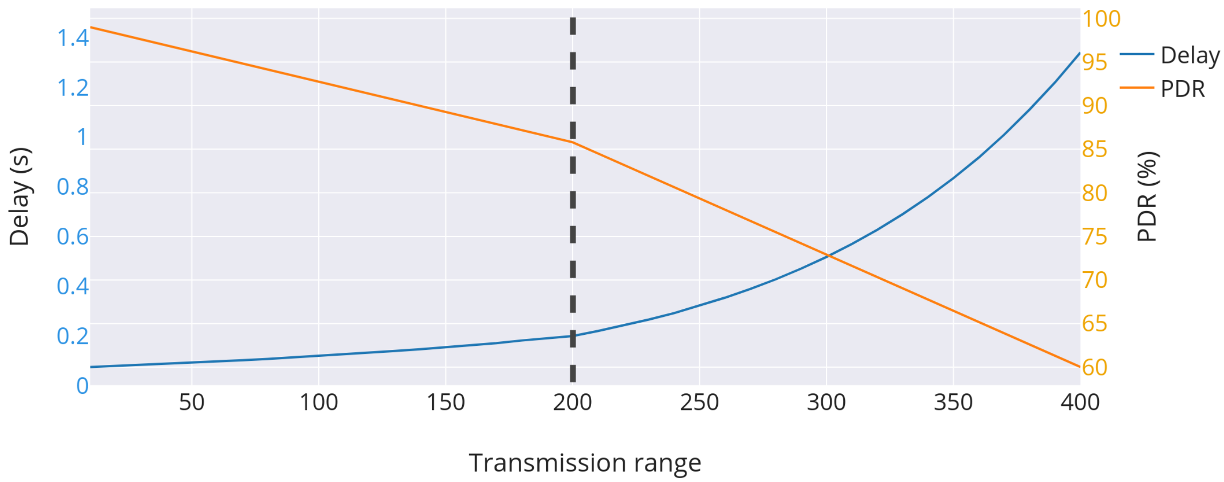


Figure 5.3: Optimal Neighborhood Distance

In order to choose a well-connected cluster head, we establish a Neighborhood Distance (ND) of 200 NM, as shown in Fig. 5.3. After 200 NM, the end-to-end delay and chance of connection failure will grow considerably. As illustrated in algorithm 1, for each node in the cluster, we calculate the number of Neighbor Nodes (NN) inside the ND-radius surrounding it. This is accomplished using the reverse haversine formula (Eq. 5.1), which takes one set of latitude/longitude and a distance as a starting point and returns the latitude/longitude pairs within that distance. The *hsine* is short for haversine. Θ_1 and Θ_2 are the latitudes of point 1 and point 2, respectively, whereas λ_1 and λ_2 are the longitudes of point 1 and point 2. r is the radius of the sphere (6,371 kilometers), and d is the distance we set at 200 NM. Afterward, the node with the greatest number of NNs would be picked as the new cluster head, i.e., the node with the most significant potential for optimum communication with the highest number of cluster members. Following this phase, the intra-cluster communication issue will be resolved.

$$d = 2r\ \mathbf{arcsin}(\sqrt{h}) \tag{5.1}$$

$$\begin{aligned} h &= hsine(\Theta_2 - \Theta_1) + cos(\Theta_1)cos(\Theta_2)hsine(\lambda_2 - \lambda_1) \\ &= sin^2(\frac{\Theta_2 - \Theta_1}{2}) + cos(\Theta_1)cos(\Theta_2)sin^2(\frac{\lambda_2 - \lambda_1}{2}) \end{aligned} \tag{5.2}$$

To enable communication across clusters, the notion of a Gateway Node (GN) is examined as illustrated in Figure 5.2. After identifying the potential cluster head, it is necessary to determine whether or not the CH may link to another CH from a different cluster. If it can, the CH can also be responsible for routing and receiving traffic to and from another cluster. For the first cluster head, this check is not necessary because there is no other cluster head. Consequently, any cluster member who can interact with two CHs, i.e., a CH from inside the cluster and a CH from an external cluster, will be chosen as the GN. Gateway is responsible for forwarding and receiving traffic from another cluster and relaying it to the cluster head or vice versa. If, after this step, there is no node with gateway capability, the entire operation, including cluster head selection, must be redone. An overview of the whole process is described in Figure 5.4.

Table 5.1: Simulation Settings for Cluster Head Selection **Reference?**

Parameter	**Value**
Simulation Area	2000 x 2000 NM
A2A Links Max Distance	400 NM
Min Aircraft Altitude	10000 Meters
Dataset Timezone	UTC
Size of Packets	1400 bytes
Max Link Throughput	1 Mbit/sec
Simulation Time	06:00 – 23:59
Number of Aircraft	50 - 350
MTU	1524 bytes
Layer 2	IEEE 802.11
Queue Model	Weighted Fair Queuing (WFQ)

5.2 Performance Evaluation

In the section that follows, the performance of the proposed cluster head selection approach is compared to various state-of-the-art cluster head selection techniques. In addition, several clustering techniques are investigated in order to comprehend the impacts of clustering on cluster head selection.

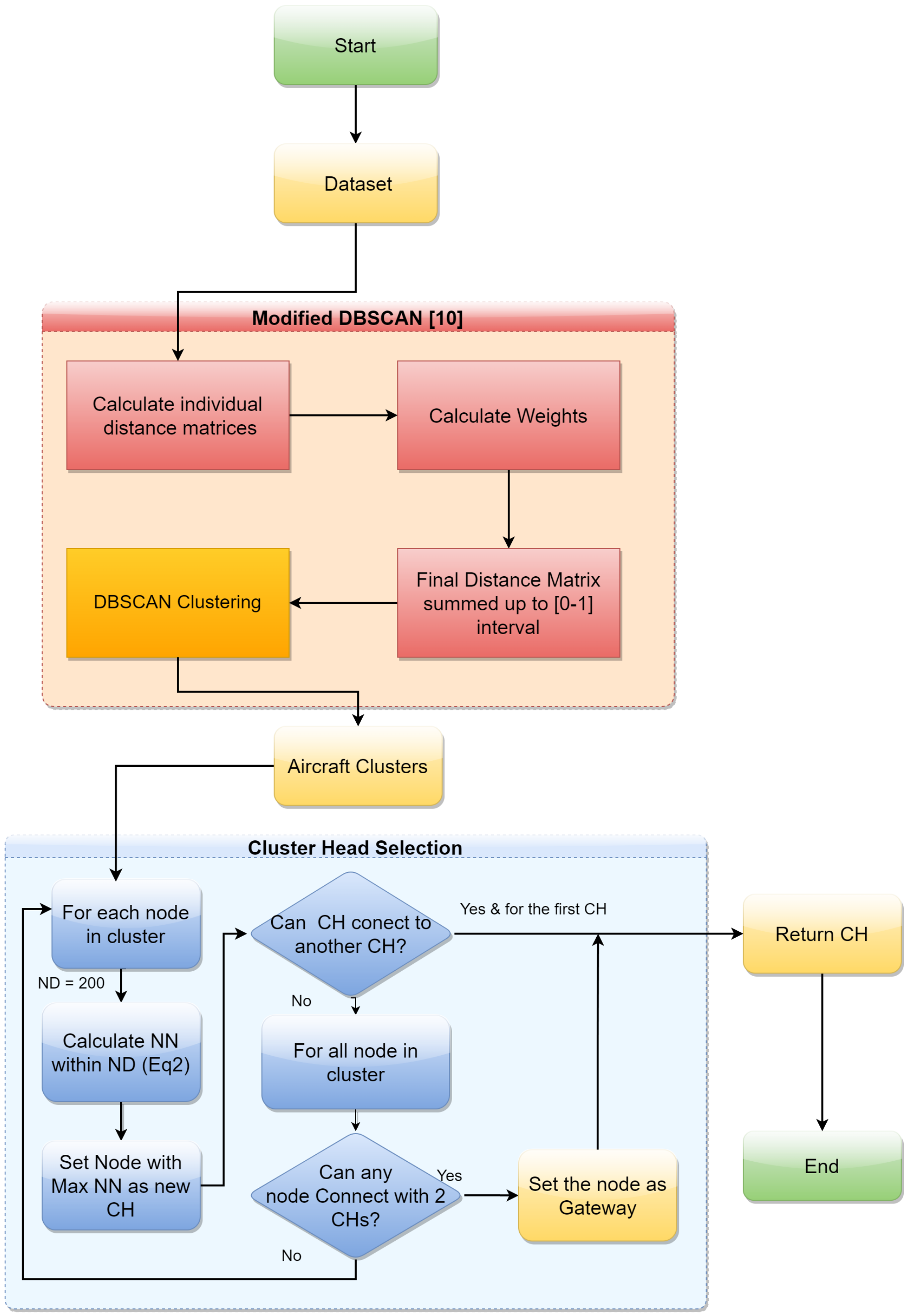


Figure 5.4: Proposed AANET topology formation framework with the new cluster head selection

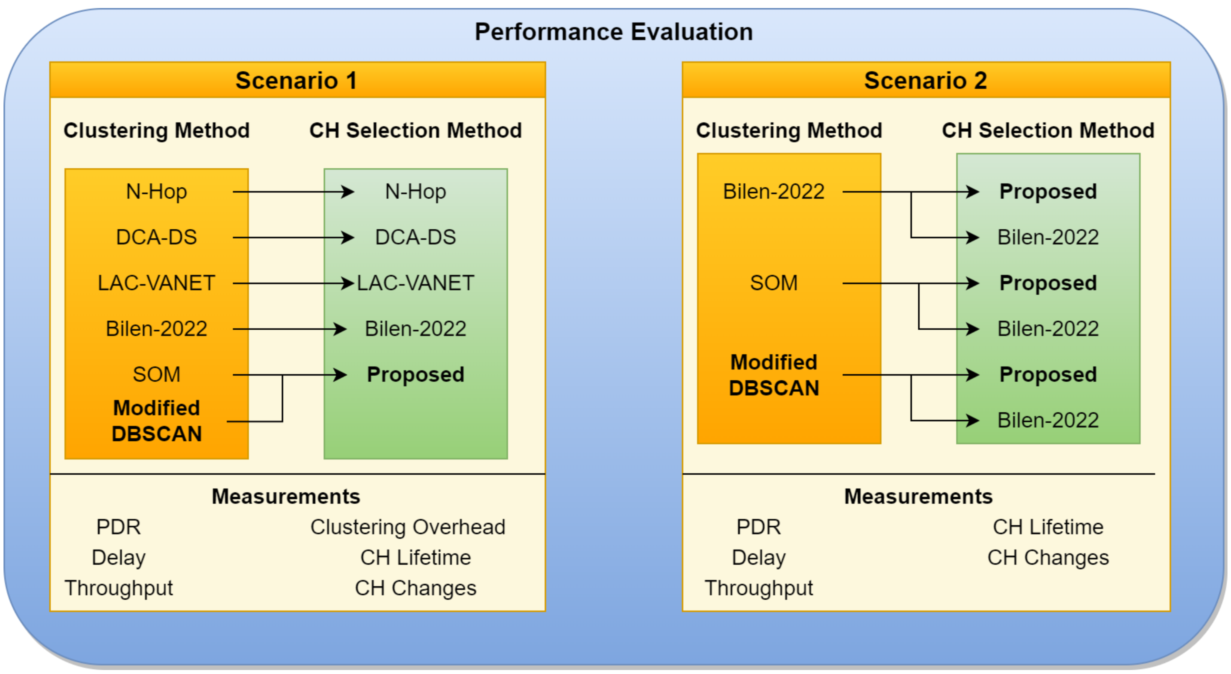


Figure 5.5: Performance Evaluation Scenarios

Algorithm 1 Cluster Head Selection

```
Input = Aircraft Clusters from DBSCAN
ND : neighborhood distance
NN : neighbor nodes
GN : gateway node
procedure FindCH
    ND ←200
    for all nodes in cluster do
        Calculate NN within ND
        Set node with max NN as CH
    end for
    return CH
end procedure
procedure FindGN
    if CH is within reach of another CH then
        Break
    else
        for all nodes in cluster do
            if node can communicate with 2 CHs then
                Set node as GN
            end if
        end for
    end if
    return GN
end procedure
```

5.2.1 Performance Metrics

To comprehend the performance of cluster head algorithms, this study evaluates two sets of metrics:

- Clustering Metrics
- Communication Metrics

Clustering metrics are those pertaining to the cluster itself. The following are examined in this thesis:

- Cluster Head Lifetime: the average lifespan of cluster heads. It denotes the duration during which an aircraft is in a CH condition before switching to another state. This can be a crucial element in comprehending the implications of clustering. A well-chosen CH will boost the cluster's stability.

- Average Cluster Head Changes: It represents the number of times a node has become or ceased to be a CH. In general, a lower value corresponds to more stable clusters and improved cluster head selection.

- Clustering Overhead: This is the percentage of packets utilized by the cluster head for clustering formation and control plane communications relative to the total amount of packets. A lower number indicates that there will be less overhead imposed on the packets.

Communication metrics are network-related and reflect how efficiently packets may be sent over the network. The most relevant communication parameters are:

- End-to-End Delay: The amount of time it takes for a packet to travel from its source to its final destination. This is significant since a large delay will render some traffic useless, such as voice traffic.

- Packet Delivery Ratio: reflects the proportion of received packets at the destination. A lower PDR results in greater packet loss, requiring the source to resend the packets, which increases network stress.

- Throughput: the maximum amount of traffic that may be carried on the connection. This is basically equivalent to link speed. The higher the speed, the quicker communication may occur.

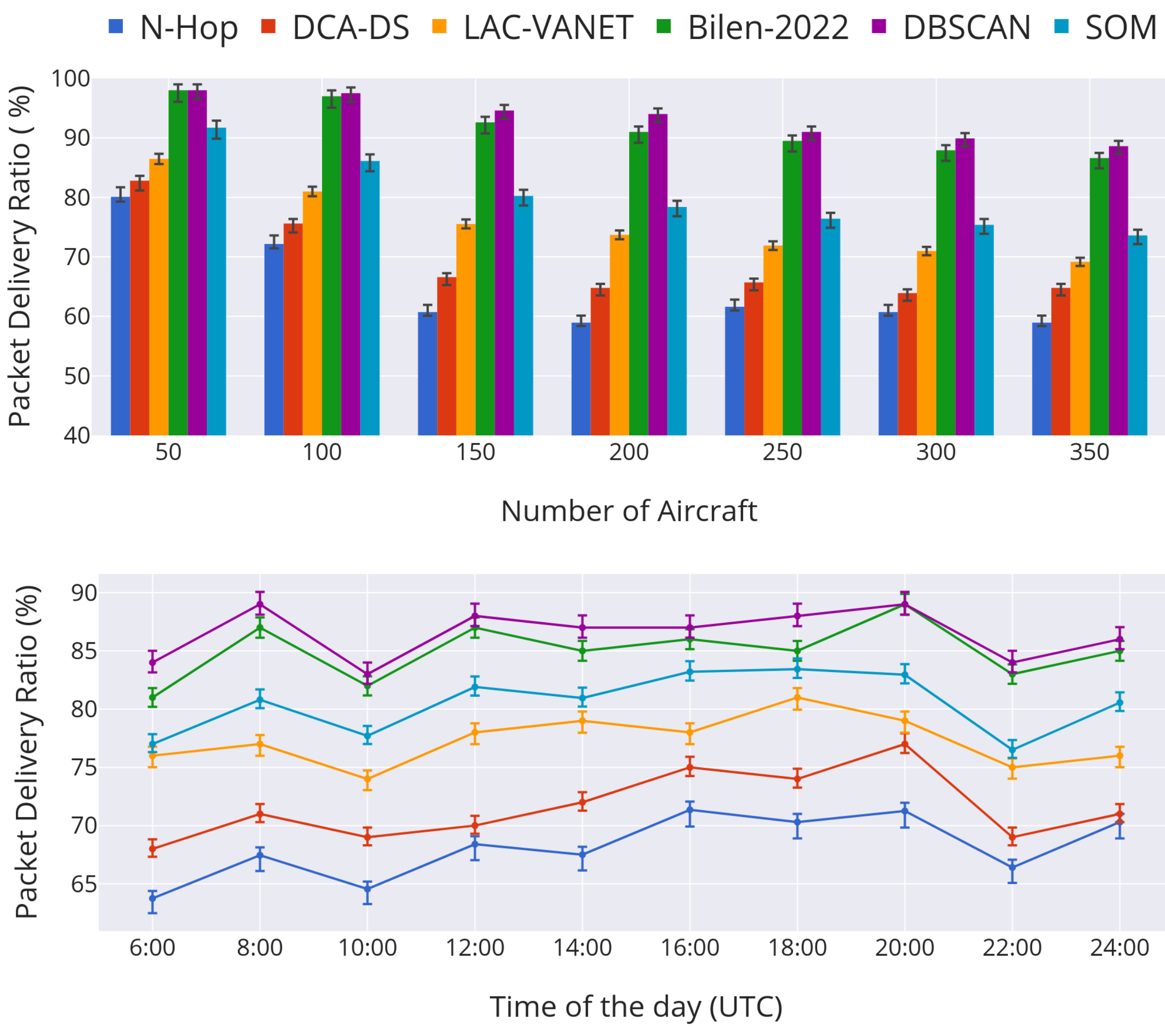


Figure 5.6: PDR vs. Number of aircraft and the simulation time in scenario 1

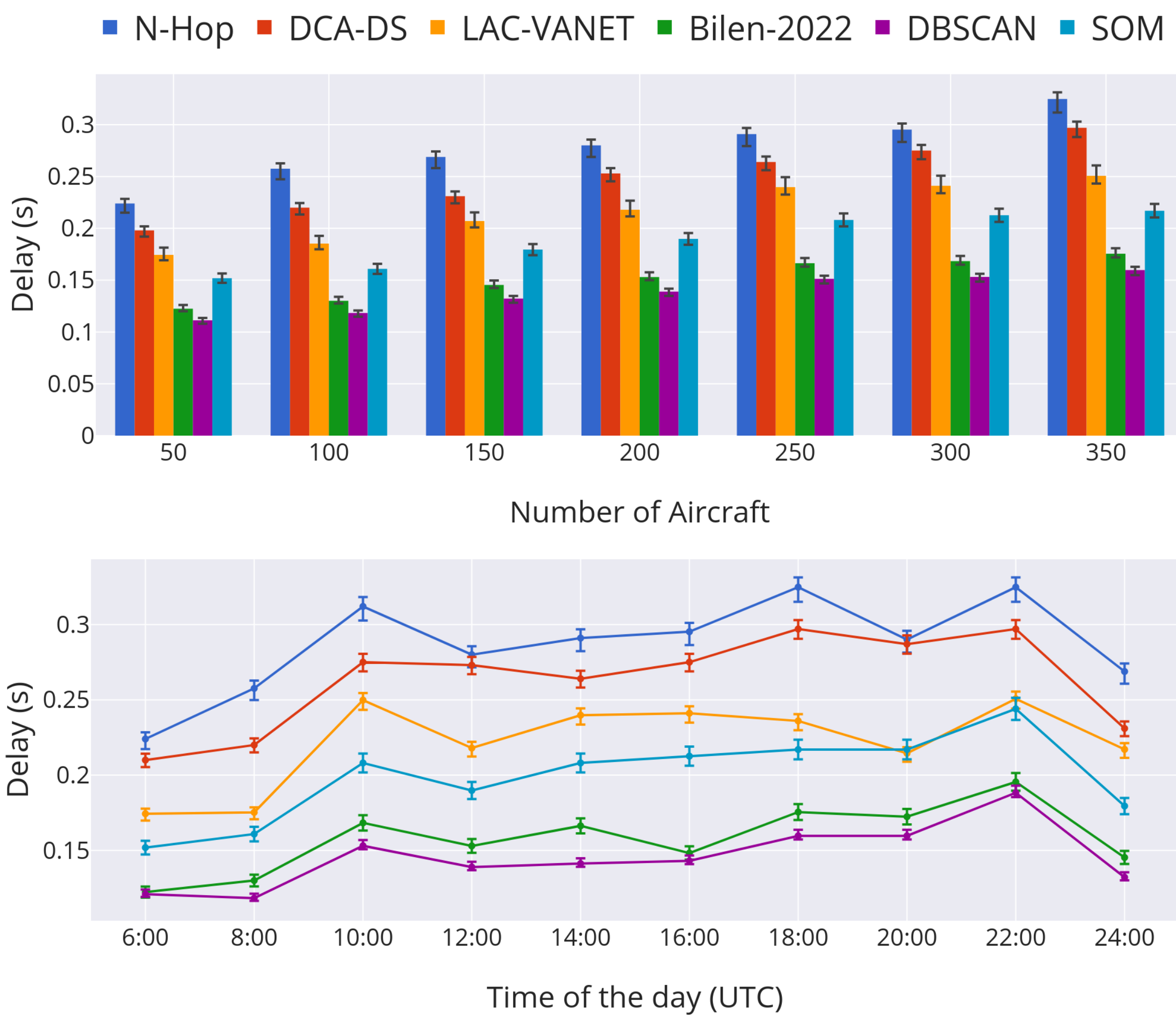


Figure 5.7: Delay vs. Number of aircraft and the simulation time in scenario 1

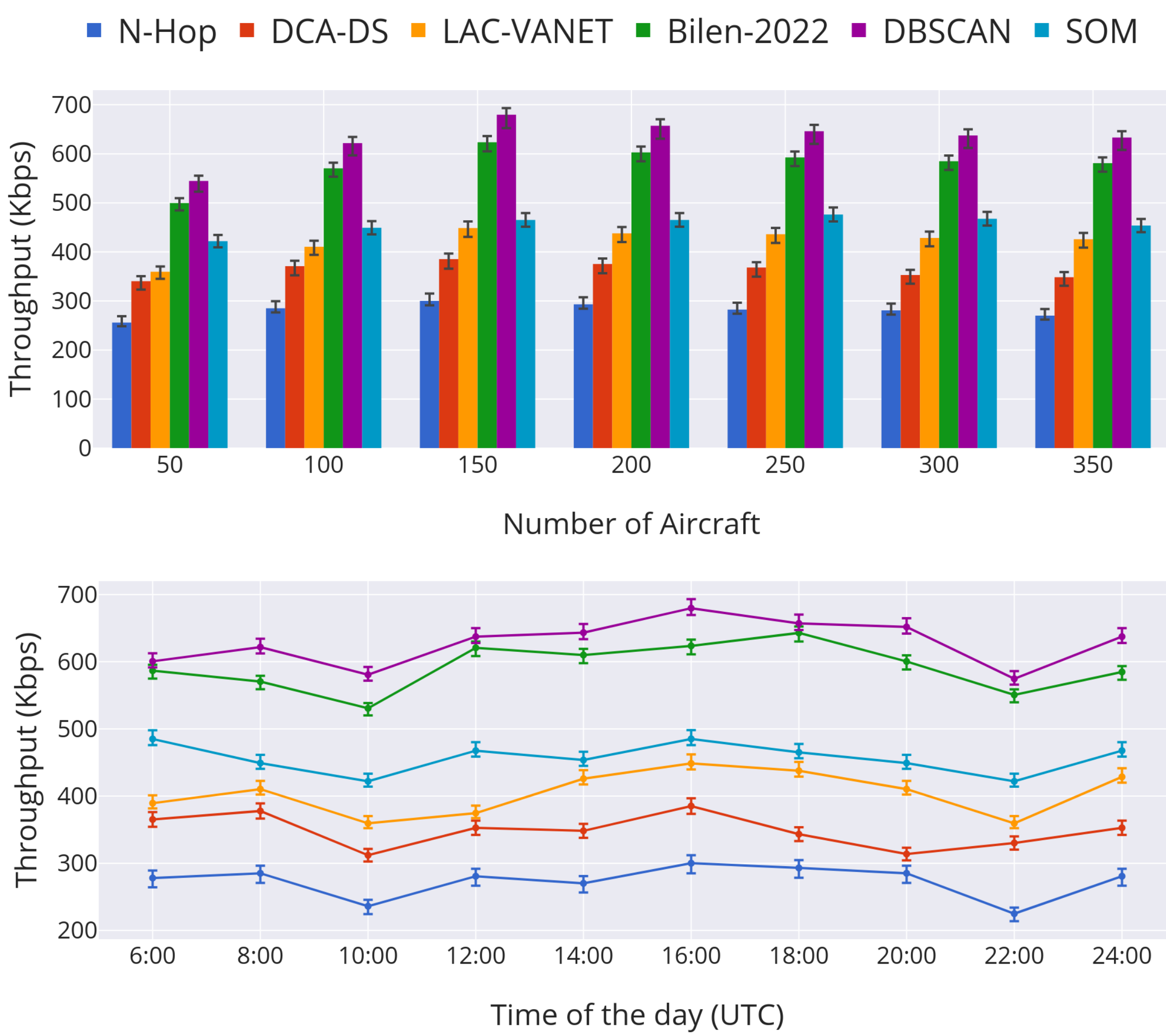


Figure 5.8: Throughput vs. Number of aircraft and the simulation time in scenario 1

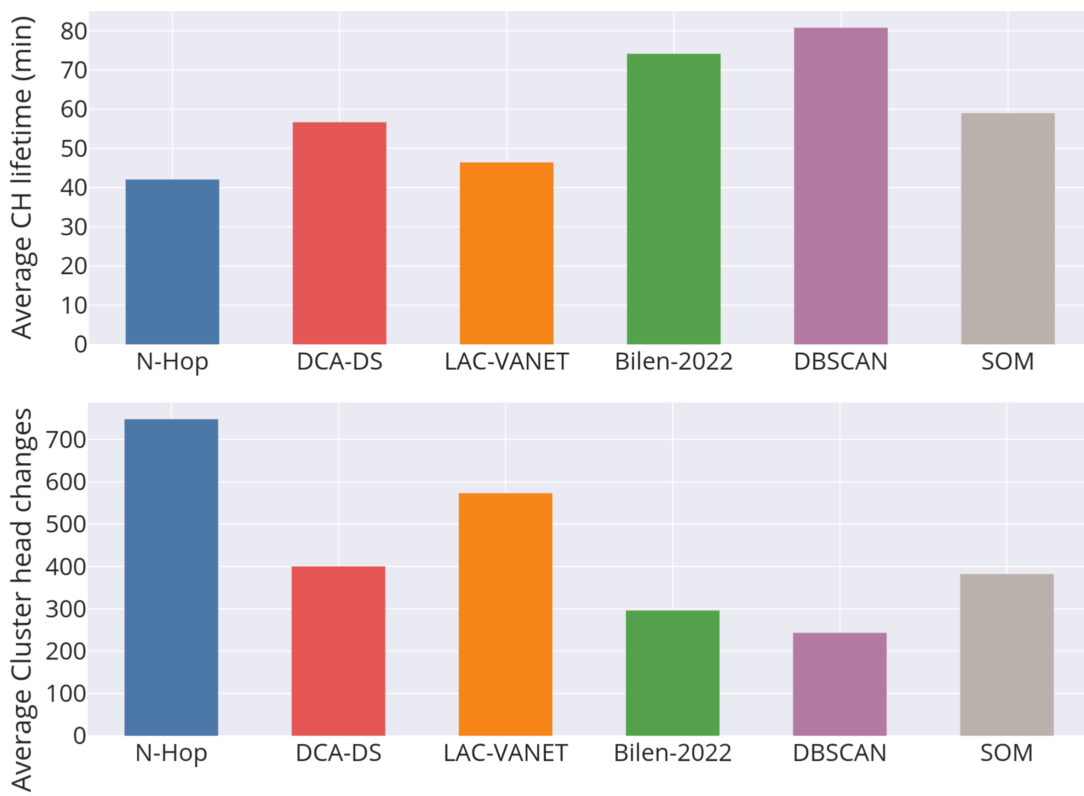
5.2.2 Simulation setup

Having the aircraft clusters is crucial in order to evaluate the effectiveness of the cluster head selection and the significance of this factor, and for this reason, we employ most of the same emulation settings as the previous chapter for clustering the aircraft. The frameworks OMNET++ and INET [90] are used to mimic mobility and traffic data. We utilize a pre-existing dataset [92], including actual flight data from historical flight data collected using the Aviationstack API [87]. This data is retrieved on January 4, 2022, between 6:00 a.m. and 11:59 p.m. UTC Time. Due to the magnitude of the dataset, only current flights with altitudes above 10,000 meters are included in the dataset used for the simulations. The dataset contains numerous flight characteristics, such as latitude, longitude, direction, velocity, origin, and destination, which are used by modified DBSCAN to generate optimum clusters. Table. 5.1 describes the simulation settings in full.

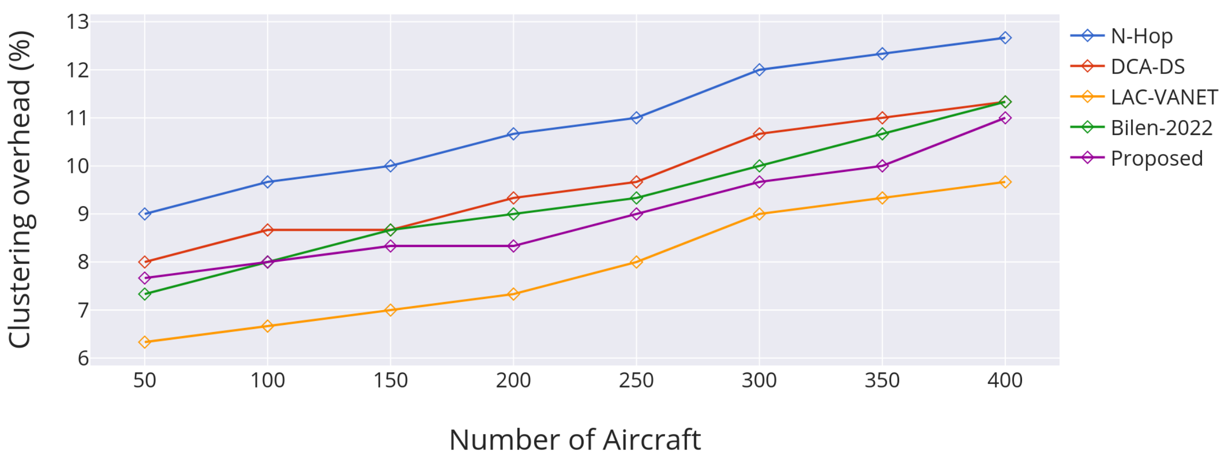
AANET's most common use is providing Internet connection to rural places. For the simulation, a 2000-by-2000-nautical-mile region is examined over the north Atlantic Ocean [89]. Even though cluster heads are essential for packet routing, they cannot do so without a routing algorithm. Because greedy routing algorithms are typically employed in ad hoc contexts, we choose Greedy Perimeter Stateless Routing (GPSR) [94]. 400NM is assumed to be the transmission range of air-to-air connections at the height of 10,000 meters [63]. In principle, if the distance between two aircraft is less than 400NM, they may be able to communicate; nevertheless, this gap may be shorter in practice. The maximum link capacity is considered to be 1 Mbit per second, whereas the packet size is 1400 bytes. Additionally, the maximum transmission unit (MTU) is set at 1524 in order to prevent fragmentation [95]. Weighted Fair Queuing (WFQ) was chosen as the queue model for the simulations because, with the emergence of real-time sensitive traffic such as voice traffic, a simple queue model is no longer enough [96]. WFQ can prioritize important real-time data above typical traffic.

5.2.3 Results

This section compares the proposed CH selection method to DCA-DS, LAC-VANET, N-Hop [97], and [51]'s three-phased framework model for AANETs. For clustering along the modified DBSCAN, self-organizing maps (SOM) are also employed. Moreover, in order



(a) Average cluster head Lifetime and Average cluster head changes



(b) Clustering Overhead

Figure 5.9: Clustering metrics results from scenario 1

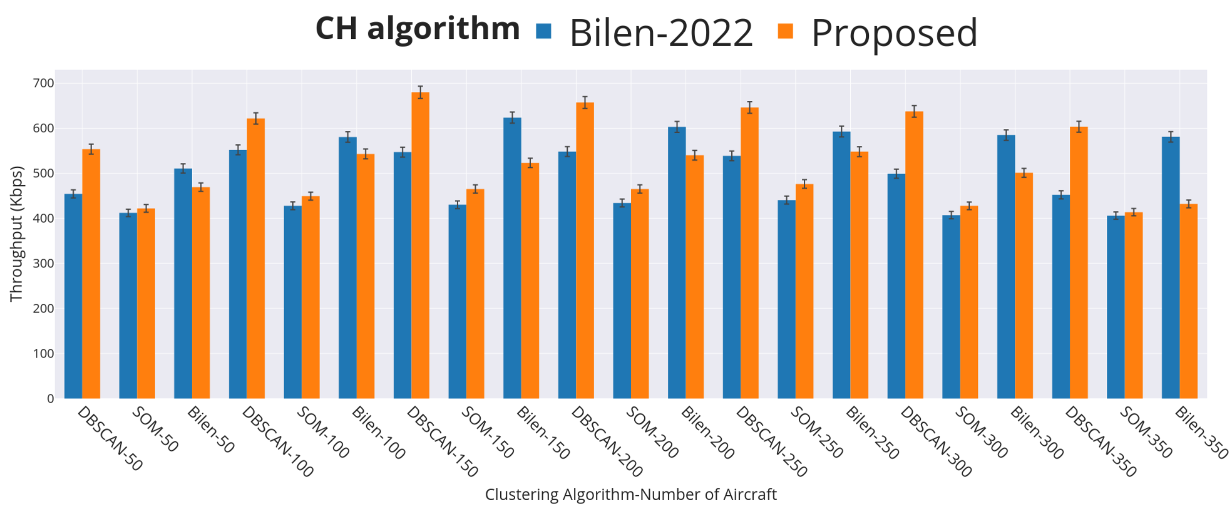


Figure 5.10: Throughput vs. Number of aircraft in scenario 2

to assess the implications of optimum clustering on cluster head selection, comprehensive simulations in the form of two scenarios are conducted.

Scenario 1

Under this case, all the clustering methods are used together with their respective cluster head selection component. Hence, clustering and cluster head selection is not distinct. We employ our proposed cluster head selection for SOM and modified DBSCAN. For findings that are dependent on the number of aircraft, aircraft are selected at random from different clusters, and the result is reported as the average of 30-minute intervals. As they are derived from historical flight data, the findings that are a function of simulation time contain the fluctuating number of aircraft according to the time of day, with a figure near 400 during the traffic peak. All results are collected with a confidence level of 95% and replicated ten times. These specifications are met by every outcome in the subsequent sections.

Figures 5.6 to 5.8 depict communication data; the general trend for packet delivery ratio (PDR) is falling as the number of aircraft increases, as seen in Fig.5.6. This is primarily due to the fact that as the number of nodes rises, so does the network's load and the likelihood of congestion and packet stacking in the queue. Nevertheless, the proposed

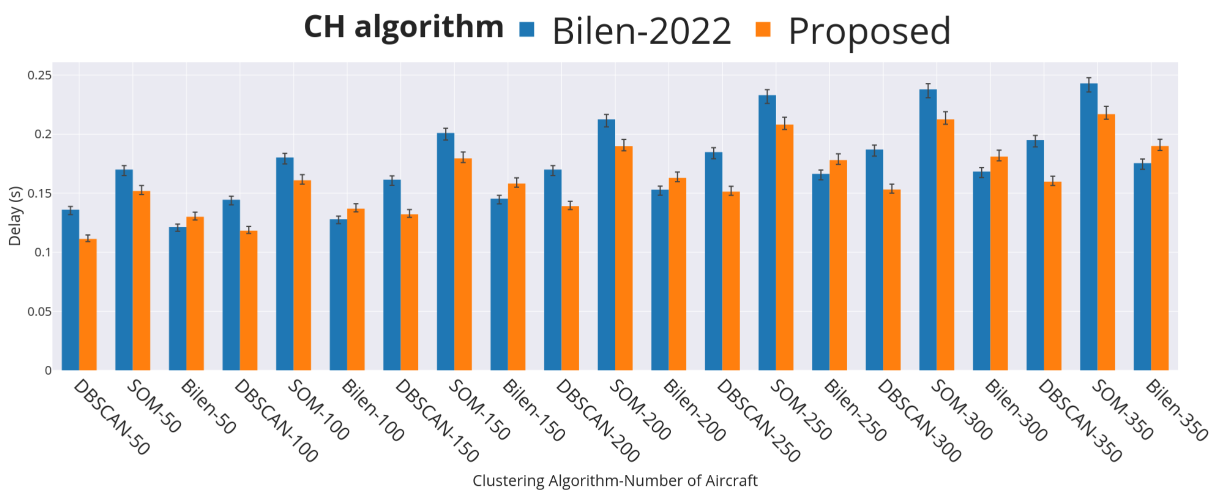


Figure 5.11: Delay vs. Number of aircraft in scenario 2

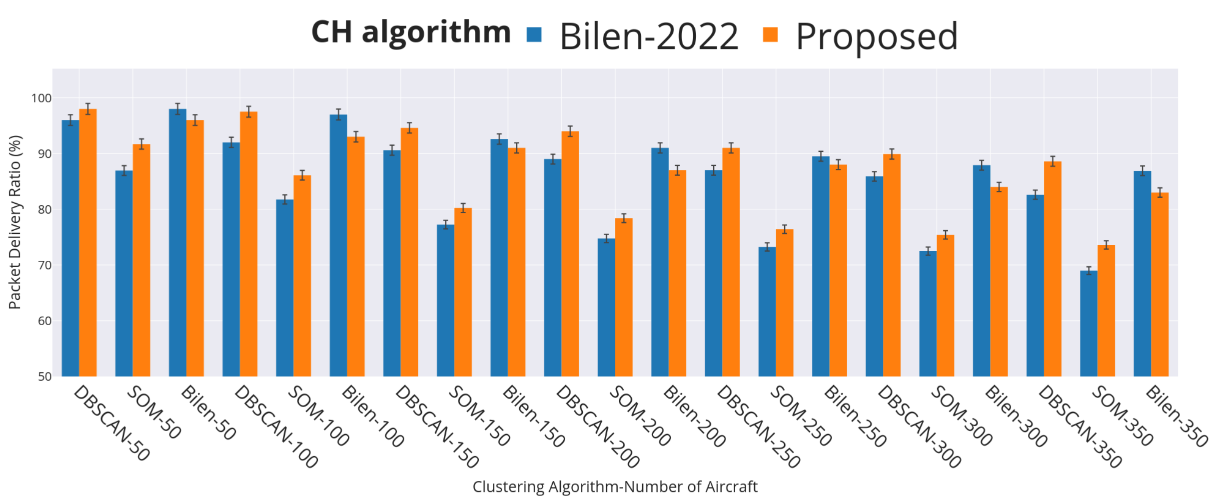


Figure 5.12: PDR vs. Number of aircraft in scenario 2

method maintains a PDR above 85% with 350 aircraft. It is obvious that three algorithms, including LAC-VANET, DCA-DS, and N-Hop, do not have ideal performance, as their respective percentages remain below 70%, 65%, and 60%. This is to be anticipated, given that these algorithms are not optimized for AANETs. Also, if we consider simulation time, the highest performance reduction occurs between 10 a.m. and 10 p.m. due to heavy traffic. Even during heavy traffic, the suggested technique maintains a PDR of over 80 percent. Compared to Bilen-2022 [51], the proposed method achieves around a 5 percent improvement overall.

In the case of end-to-end delivery delay, as observed in Fig. 5.7, the latency will rise as the number of nodes increases. The greater the number of nodes, the longer it takes for a packet to go from source to destination and the more hops it must make in between. At a maximum of 350 aircraft, the proposed method produces a delay of around 0.15 seconds, compared to 0.17 for Bilen-2022. The subpar performance of the other algorithms, which remain above 0.25 seconds, is primarily due to the fact that the criteria they use for cluster head selection are not suitable for use cases involving aircraft; as a result, routing decisions cannot be made in a timely manner, thereby increasing the delay. In terms of simulation time, we notice a performance drop around 10 a.m. and 8 p.m., similar to PDR, with the suggested technique reaching a low of 0.188 seconds compared to Bilen-2022, which achieves a low of 0.195.

The trending behavior of Throughput differs from that of delay and PDR, as illustrated in Fig5.8. As the network converges, the throughput improves between 50 to 150 aircraft but then falls as the number of aircraft grows. Due to the limited number of aircraft, the topology has not yet been established and stabilized, resulting in low throughput. After convergence, we reach the top performance at about 150 nodes, with the proposed approach achieving a maximum of 679 Kbps, whereas Bilen-2022 achieves 623 Kbps, an 8% increase. N-Hop, DCA-DS, and LAC-VANET achieve maximum speeds of 300, 385, and 448 kbps, respectively, compared to top performers. The trend is similar when compared to simulation time.

Fig. 5.9 displays the performance of clustering metrics; Fig. 5.9a depicts the frequency with which a node is elected as the cluster head (CH) or ceases to be the CH. The suggested cluster head selection using modified DBSCAN clustering achieves 243, which is a 21 percent improvement over the current state-of-the-art Bilen-2022, which records 296. In comparison to the other algorithms, N-Hop has a comparatively large number of changes,

Figure 5.13: Clustering metrics results from scenario 2

with 748. These methods fail to choose the ideal cluster head for the AANET situation, resulting in a large number of cluster head changes and an unstable topology. The proposed approach achieves a CH lifespan of around 80 minutes, indicating a highly stable cluster and good cluster head selection. When a node remains in the CH state for an extended period of time, there is less disruption in the topology and a lower probability of connection failure. While the Bilen-2022 also earns a score above 70, the other algorithms fail to meet the AANET requirements and fall behind.

As indicated in Fig. 5.9b, the final measure to evaluate is clustering overhead. The general trend of overhead increases as the number of nodes increases, mostly because more packets are required for control plane messages sent by the cluster head to maintain the topology. For 400 nodes, the LAC-VANET has the lowest overhead at around 10%. The proposed method and Bilen-2022 both fall short of 11%, while N-Hop achieves the lowest performance at 13 percent. While this metric cannot determine the performance of CH selection algorithms on its own, it is vital to consider it while selecting an appropriate method since it provides insight into the trade-off between stability and performance.

Based on the findings of Scenario 1, it can be concluded that the proposed method has the best performance in relation to the AANET Specifications, followed by Bilen-2022. Regarding the other algorithms, it is clear that while LAC-VANET has superior performance in terms of PDR, latency, and throughput, DCA-DS is the more stable method due to its lower number of cluster head changes and longer average CH lifetime. Choosing between these two methods is, therefore, a compromise between stability and communication performance. Another important takeaway is SOM's performance. It outperforms DCA-DS and LAC-VANET even without additional modifications and a cluster head selection component designed expressly for it. This is to be expected given that the SOM is naturally efficient at processing unknown inputs and unstructured data and is a suitable fit for AANETs.

Scenario 2

In the second scenario, clustering algorithms are decoupled from their counterpart cluster head algorithm in order to assess the effects of clustering on cluster head selection. The previous instance showed that VANET algorithms are not suited for AANETs; thus, we do not analyze their performance in this situation. For clustering, DBSCAN, SOM, and

Bilen-2022 are considered, and for cluster head selection, the proposed method is compared to Bilen-2022. Thus, we may gain a solid idea of the capabilities of each combination.

Figures 5.10, 5.11, and 5.12 depict the communication metrics findings for scenario 2. Similar to the last experiment, the end-to-end latency grows as the number of aircraft increases. The combination of improved DBSCAN and the proposed cluster head selection algorithm outperforms Bilen-2022 at 350 aircraft. The combination of DBSCAN + Bilen-2022 cluster head selection and Bilen-2022 + proposed CH selection method performs poorly. This is due to the fact that each of these cluster head selection techniques is suited to the specifications of its clustering equivalents, and when they are decoupled, we can see a decrease in performance compared to the results of scenario 1. The ratio of delivered packets and throughput follows the same trend. As neither the Bilen-2022 nor the proposed approach is created for SOM, we can also pay special attention to SOM. So it will be a useful indicator of the success of the CH selection process. In this case, the combination of SOM and Proposed CH performs better by around 3%, 12%, and 4% with respect to PDR, latency, and throughput, respectively.

Fig. 5.13 depicts the clustering outcomes of scenario 2's measurements. There is no need to reevaluate the clustering overhead because it is independent of the clustering algorithm, and the performance of the clustering overhead is unaffected by adopting a different clustering algorithm. In terms of the average cluster head changes, it can be noted that the suggested CH selection method performs roughly 21 percent better when paired with modified DBSCAN than Bilen-2022, indicating more stability. Taking into account the neutral clustering algorithm SOM, the performance of the suggested CH selection method is superior. In terms of Cluster head lifespan, the performance of the suggested CH selection is again enhanced when paired with the modified DBSCAN. In terms of average CH changes and CH lifespan, it is obvious that the combination of DBSCAN + Bilen-2022 and Bilen + Proposed CH selection is not optimal and performs poorly.

From this scenario, we can deduce that the clustering algorithm does impact the performance of the cluster head method, and it is crucial that the CH selection algorithm be constructed on top of a strong clustering algorithm in order to enhance stability and communication reliability.

5.3 Summary

To improve the performance of cluster-based Aeronautical Ad Hoc Networks (AANETs) topologies, it is essential to select a robust cluster head that improves connection and stability. Keeping this in mind, this chapter proposes a novel CH selection technique that employs a previously suggested modified DBSCAN clustering for AANETs, after evaluating the different state of the art regarding selecting the cluster head. The notion of Neighbor Nodes was established in order to compute and select the most linked node as the cluster leader. In addition to ensuring connectivity with other clusters, the concept of a Gateway Node is studied. Extensive simulations demonstrate that the proposed CH method outperforms previous state-of-the-art algorithms in terms of PDR, throughput, and latency by 3, 9, and 10 percent, respectively. Moreover, Cluster head lifetime has increased by 8%, while average cluster head changes have lowered by 17%, indicating greater stability.

Chapter 6

Conclusion

Aeronautical Ad Hoc Networks (AANET) are a viable solution to fix the shortcomings of traditional fixed networks such as terrestrial or satellite networks. AANET seeks efficient communication and management solutions such as a clustered topology for providing in-flight connectivity, especially for remote regions like oceans where an air to ground connections is not possible. With this in mind, in this thesis, after reviewing the current state of the art and stressing sub-optimal clusters as the source of stability issues in AANETs, as the first contribution, a multi-feature DBSCAN has been proposed to create aircraft clusters. Instead of utilizing a basic metric such as Haversine, the algorithm utilizes aspects of the dataset, such as location, altitude, direction, and velocity, to generate a precomputed distance matrix. According to the simulation findings, the number of cluster changes has been reduced by 22% compared to K-means and 15% compared to SOM. Moreover, there are 35% fewer clusters during peak traffic hours. These results imply that density-based clustering yields more stability. In terms of the communication metrics, the proposed method has also been shown to enhance the packet delivery ratio by 51% compared to K-means and by 28% compared to SOM. The proposed algorithm also decreases the end-to-end delay by up to 30%.

In the second part of the dissertation, the challenges of selecting a cluster head are tackled, which is rarely considered for AANET in the literature. To increase the performance of cluster-based Aeronautical Ad Hoc Networks (AANETs) topologies, it is crucial to choose a cluster head that enhances connectivity and stability. Keeping this in mind, this thesis proposed a novel CH selection technique that employed a previously suggested modified DBSCAN clustering for AANETs, after evaluating the different state of the art regarding selecting the cluster head. The notion of Neighbor Nodes was established in order to compute and select the most linked node as the cluster leader. In addition to ensuring connectivity with other clusters, the concept of a Gateway Node is studied. Extensive simulations demonstrate that the proposed CH selection method outperforms previous state-of-the-art algorithms in terms of PDR, throughput, and latency by 3, 9, and 10 percent, respectively. Moreover, Cluster head lifetime has increased by 8%, while average cluster head changes have lowered by 17%, indicating greater stability.

6.1 Future Work And Open Issues

We have obtained relatively good communication performance and stable clusters by using the proposed density-based clustering method and building a cluster head selection mechanism on top of it. However, there is still room for improvement in the formation, management, and sustainability of aeronautical ad hoc networks. Moreover, the cluster head selection method that is proposed in this thesis is designed to work over the specific clustering solution we utilized and will not have optimal performance when coupled with other clustering approaches. While we explored the tasks of clustering and cluster head selection, the important challenge of routing still remains, which is the limitation of this thesis.

Modified Self-Organizing Map for AANETs

Self-Organizing Map was used in this thesis to compare the evaluate the results of the proposed density-based approach. However, the implementation of SOM that we used was a generic one, and SOM can be modified to AANET specifications to improve performance. SOMs' ability to reduce the dimension of input data and form structure out of unknown inputs is extremely valuable in the formation of AANETs.

Routing Protocols for AANETs

Routing protocols are crucial for transporting packets from their source to their destination. Although there are several routing algorithms built for ad hoc networks, they function poorly in an AANET architecture. In the literature [98], different types of studies are conducted on routing, but there is still potential for improvement, particularly in the use of Reinforcement Learning and Q-Learning techniques for this task, which might be the continuation of this work.

Effects of External Elements on The Performance And Stability of AANETs

In this study, we hypothesized that the short lifetime of air-to-air linkages could be the fundamental cause of stability issues that were solved by optimum clustering. However, external environmental factors might also alter the quality of air-to-air communication. This might include weather and atmosphere-related occurrences, including rain, storm, clouds, and other abnormalities. Therefore, the impact of these factors on communication quality should be investigated further.

Security Challenges in AANETs

The characteristics of ad hoc networks necessitate strict security requirements. Given the nature of airplanes, a malicious assault on AANETs can have disastrous consequences. Whether AANET infrastructure is used to connect passengers to the internet or for aircraft-to-ground communication, security is of the highest significance. To maintain the security of AANETs, authentication, integrity, and encryption mechanisms must be implemented.

Real World Implementation of AANET

A real-world implementation of the AANET may differ from the hypothesis presented in this thesis. In contrast to our emphasis on preliminary tasks such as clustering and cluster head selection, real-world situations present additional difficulties. For instance, the link capacity we specified for the simulation is insufficient to meet the streaming requirements of passengers in an actual network. In addition, the entire calculation of clusters and CHs requires a substantial amount of time, so it is possible that the calculations could be completed before the flights. A further consideration is the positioning of the CH on nodes that can have satellite or ground connections; in this case, there is no need to recalculate the CH, as it would make sense for these nodes to be the CH, as they have an uplink.

References

[1] MarketsandMarkets, "In-flight Entertainment & Connectivity Market Size & Share | Industry Report, 2021-2026 |."

[2] SITA, "Passenger IT Insights 2020."

[3] Insmarsat, "Passenger Confidence Tracker 2021."

[4] J. Zhang, T. Chen, S. Zhong, J. Wang, W. Zhang, X. Zuo, R. G. Maunder, and L. Hanzo, "Aeronautical $Ad Hoc$ Networking for the Internet-Above-the-Clouds," *Proceedings of the IEEE*, vol. 107, pp. 868–911, May 2019. Conference Name: Proceedings of the IEEE.

[5] A. Numani, S. M. Gulfam, M. A. Javed, B. Muhammad, R. Prasad, and S. J. Nawaz, "Network Architecture and Wireless Ad Hoc Routing for Airborne Internet Services," *Wireless Personal Communications*, vol. 122, pp. 3529–3543, Feb. 2022.

[6] J. Cui, D. Liu, J. Zhang, H. Yetgin, S. X. Ng, and L. Hanzo, "Minimum-Delay Routing for Integrated Aeronautical Ad Hoc Networks Relying on Passenger-Planes in the North-Atlantic Region," Oct. 2020. arXiv:2010.08315 [cs, math].

[7] S. Ghosh and A. Nayak, "Multi-dimensional clustering and network monitoring system for aeronautical ad hoc networks," in *2015 Seventh International Conference on Ubiquitous and Future Networks*, pp. 772–777, July 2015. ISSN: 2165-8536.

[8] M. Fowlie, E. Rubin, and R. Walker, "Bringing Satellite-Based Air Quality Estimates Down to Earth," *AEA Papers and Proceedings*, vol. 109, pp. 283–288, May 2019.

[9] A. Khan, S. Khan, A. S. Fazal, Z. Zhang, and A. O. Abuassba, "Intelligent cluster routing scheme for flying ad hoc networks," *Science China Information Sciences*, vol. 64, p. 182305, June 2021.

[10] S. Al-Sultan, M. M. Al-Doori, A. H. Al-Bayatti, and H. Zedan, "A comprehensive survey on vehicular Ad Hoc network," *Journal of Network and Computer Applications*, vol. 37, pp. 380–392, Jan. 2014.

[11] M. Lee and T. Atkison, "VANET applications: Past, present, and future," *Vehicular Communications*, vol. 28, p. 100310, Apr. 2021.

[12] R. Ramanathan and J. Redi, "A brief overview of ad hoc networks: challenges and directions," *IEEE Communications Magazine*, vol. 40, pp. 20–22, May 2002.

[13] P. Mohapatra and S. Krishnamurthy, *AD HOC NETWORKS: Technologies and Protocols*. Springer Science & Business Media, Sept. 2004. Google-Books-ID: izNUbXbK7e4C.

[14] "The Handbook of Ad Hoc Wireless Networks | Mohammad Ilyas, Richard C.."

[15] S. Mirza and S. Z. Bakshi, "INTRODUCTION TO MANET," vol. 05, no. 1.

[16] K. Xu, X. Hong, and M. Gerla, "An ad hoc network with mobile backbones," in *2002 IEEE International Conference on Communications. Conference Proceedings. ICC 2002 (Cat. No.02CH37333)*, vol. 5, pp. 3138–3143 vol.5, Apr. 2002.

[17] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing*, vol. 2, no. 5, pp. 483–502, 2002. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcm.72.

[18] "Mobile ad hoc networking and the IETF | ACM SIGMOBILE Mobile Computing and Communications Review."

[19] S. Dhar, "MANET: Applications, Issues, and Challenges for the Future," *International Journal of Business Data Communications and Networking (IJBDCN)*, vol. 1, pp. 66–92, Apr. 2005. Publisher: IGI Global.

[20] J. Hoebeke, I. Moerman, B. Dhoedt, and P. Demeester, "An Overview of Mobile Ad Hoc Networks:Applications and Challenges,"

[21] M. Gavhale and P. D. Saraf, "Survey on Algorithms for Efficient Cluster Formation and Cluster Head Selection in MANET," *Procedia Computer Science*, vol. 78, pp. 477–482, Jan. 2016.

[22] C. S. Raghavendra, K. M. Sivalingam, and T. Znati, *Wireless Sensor Networks*. Springer, Dec. 2006. Google-Books-ID: 4zyDBwAAQBAJ.

[23] I. F. Akyildiz and M. C. Vuran, *Wireless Sensor Networks*. John Wiley & Sons, June 2010. Google-Books-ID: 7YBHYJsSmS8C.

[24] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, pp. 393–422, Mar. 2002.

[25] D. Kandris, C. Nakas, D. Vomvas, and G. Koulouras, "Applications of Wireless Sensor Networks: An Up-to-Date Survey," *Applied System Innovation*, vol. 3, p. 14, Mar. 2020. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

[26] O. I. Khalaf and B. M. Sabbar, "An overview on wireless sensor networks and finding optimal location of nodes," *Periodicals of Engineering and Natural Sciences*, vol. 7, pp. 1096–1101, Aug. 2019. Number: 3.

[27] M. J. N. Mahi, S. Chaki, S. Ahmed, M. Biswas, M. S. Kaiser, M. S. Islam, M. Sookhak, A. Barros, and M. Whaiduzzaman, "A Review on VANET Research: Perspective of Recent Emerging Technologies," *IEEE Access*, vol. 10, pp. 65760–65783, 2022. Conference Name: IEEE Access.

[28] M. S. Anwer and C. Guy, "A Survey of VANET Technologies," 2014.

[29] C. Cooper, D. Franklin, M. Ros, F. Safaei, and M. Abolhasan, "A Comparative Survey of VANET Clustering Techniques," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 657–681, 2017. Conference Name: IEEE Communications Surveys & Tutorials.

[30] H. Hasrouny, A. E. Samhat, C. Bassil, and A. Laouiti, "VANet security challenges and solutions: A survey," *Vehicular Communications*, vol. 7, pp. 7–20, Jan. 2017.

[31] A. Lamssaggad, N. Benamar, A. S. Hafid, and M. Msahli, "A Survey on the Current Security Landscape of Intelligent Transportation Systems," *IEEE Access*, vol. 9, pp. 9180–9208, 2021. Conference Name: IEEE Access.

[32] N. Taherifard, M. Simsek, C. Lascelles, and B. Kantarci, "Attention-Based Event Characterization for Scarce Vehicular Sensing Data," *IEEE Open Journal of Vehicular Technology*, vol. 1, pp. 317–330, 2020. Conference Name: IEEE Open Journal of Vehicular Technology.

[33] N. Taherifard, M. Simsek, and B. Kantarci, "Bridging Connected Vehicles with Artificial Intelligence for Smart First Responder Services," in *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1–5, Nov. 2019.

[34] N. Taherifard, M. Simsek, C. Lascelles, and B. Kantarci, "Machine Learning-Driven Event Characterization under Scarce Vehicular Sensing Data," in *2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 1–6, Sept. 2020. ISSN: 2378-4873.

[35] N. Taherifard, M. Simsek, C. Lascelles, and B. Kantarci, "Prior Knowledge Input to Improve LSTM Auto-encoder-based Characterization of Vehicular Sensing Data," in *ICC 2021 - IEEE International Conference on Communications*, pp. 1–6, June 2021. ISSN: 1938-1883.

[36] M. Simsek, A. Boukerche, B. Kantarci, and S. Khan, "AI-driven autonomous vehicles as COVID-19 assessment centers: A novel crowdsensing-enabled strategy," *Pervasive and Mobile Computing*, vol. 75, p. 101426, Aug. 2021.

[37] E. Sakhaee, A. Jamalipour, and N. Kato, "Aeronautical ad hoc networks," in *IEEE Wireless Communications and Networking Conference, 2006. WCNC 2006.*, vol. 1, pp. 246–251, Apr. 2006. ISSN: 1558-2612.

[38] T. S. Madhulatha, "An Overview on Clustering Methods," May 2012. arXiv:1205.1117 [cs].

[39] M. G. H. Omran, A. P. Engelbrecht, and A. Salman, "An overview of clustering methods," *Intelligent Data Analysis*, vol. 11, pp. 583–605, Jan. 2007. Publisher: IOS Press.

[40] O. Senouci, S. Harous, and Z. Aliouat, "Survey on vehicular ad hoc networks clustering algorithms: Overview, taxonomy, challenges, and open research issues," *International Journal of Communication Systems*, vol. 33, no. 11, p. e4402, 2020. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/dac.4402.

[41] M. Ahmad, A. Hameed, A. A. Ikram, and I. Wahid, "State-of-the-Art Clustering Schemes in Mobile Ad Hoc Networks: Objectives, Challenges, and Future Directions," *IEEE Access*, vol. 7, pp. 17067–17081, 2019. Conference Name: IEEE Access.

[42] L. A. Maglaras and D. Katsaros, "Social Clustering of Vehicles Based on Semi-Markov Processes," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 318–332, Jan. 2016. Conference Name: IEEE Transactions on Vehicular Technology.

[43] M. Royer, F. Garcia, and A. Pirovano, "An enhanced 1-hop clustering algorithm for Publish / Subscribe systems in AANETS," in *2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC)*, pp. 2D2–1–2D2–6, Sept. 2015. ISSN: 2155-7209.

[44] F. Shi, Y. Shi, and L. Lai, "A clustering algorithm of Ad-Hoc network based on honeycomb division," in *2011 IEEE International Conference on Granular Computing*, pp. 863–866, Nov. 2011.

[45] T. Bilen, P. J. Aydemir, A. E. Konu, and B. Canberk, "Customized K-Means Based Topology Clustering for Aeronautical Ad-hoc Networks," in *2021 IEEE 26th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 1–5, Oct. 2021. ISSN: 2378-4873.

[46] *Advances in Self-Organizing Maps, Learning Vector Quantization, Clustering and Data Visualization.*

[47] Y. Wang, M. C. Ertürk, J. Liu, I.-h. Ra, R. Sankar, and S. Morgera, "Throughput and delay of single-hop and two-hop aeronautical communication networks," *Journal of Communications and Networks*, vol. 17, pp. 58–66, Feb. 2015. Conference Name: Journal of Communications and Networks.

[48] E. Sakhaee and A. Jamalipour, "Stable Clustering and Communications in Pseudolinear Highly Mobile Ad Hoc Networks," *IEEE Transactions on Vehicular Technology*, vol. 57, pp. 3769–3777, Nov. 2008. Conference Name: IEEE Transactions on Vehicular Technology.

[49] V. Megas, S. Hoppe, M. Ozger, D. Schupke, and C. Cavdar, "A Combined Topology Formation and Rate Allocation Algorithm for Aeronautical Ad Hoc Networks," *IEEE Transactions on Mobile Computing*, pp. 1–17, 2022. Conference Name: IEEE Transactions on Mobile Computing.

[50] S. Bharany, S. Sharma, S. Badotra, O. I. Khalaf, Y. Alotaibi, S. Alghamdi, and F. Alassery, "Energy-Efficient Clustering Scheme for Flying Ad-Hoc Networks Using an Optimized LEACH Protocol," *Energies*, vol. 14, p. 6016, Jan. 2021. Number: 19 Publisher: Multidisciplinary Digital Publishing Institute.

[51] T. Bilen and B. Canberk, "Three-phased clustered topology formation for Aeronautical Ad-Hoc Networks," *Pervasive and Mobile Computing*, vol. 79, p. 101513, Jan. 2022.

[52] R. Kaur, R. K. Ramachandran, R. Doss, and L. Pan, "The importance of selecting clustering parameters in VANETs: A survey," *Computer Science Review*, vol. 40, p. 100392, May 2021.

[53] O. Senouci, Z. Aliouat, and S. Harous, "DCA-DS: A Distributed Clustering Algorithm Based on Dominating Set for Internet of Vehicles," *Wireless Personal Communications*, vol. 115, pp. 401–413, Nov. 2020.

[54] S. Beheshti, S. Adabi, and A. Rezaee, "Location-Aware Distributed Clustering with Eliminating GPS in Vehicular Ad-hoc Networks," Jan. 2020.

[55] W. Ahsan, M. F. Khan, F. Aadil, M. Maqsood, S. Ashraf, Y. Nam, and S. Rho, "Optimized Node Clustering in VANETs by Using Meta-Heuristic Algorithms," *Electronics*, vol. 9, p. 394, Mar. 2020. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.

[56] Y. Xu, K. Nash, A. Acharjee, and G. V. Gkoutos, "CACONET: a novel classification framework for microbial correlation networks," *Bioinformatics*, vol. 38, pp. 1639–1647, Mar. 2022.

[57] L. Ye, Y. Zhang, Y. Li, and S. Han, "A Dynamic Cluster Head Selecting Algorithm for UAV Ad Hoc Networks," in *2020 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 225–228, June 2020. ISSN: 2376-6506.

[58] I. Hussain and C. Bingcai, "Cluster Formation and Cluster Head Selection Approach for Vehicle Ad-Hoc Network (VANETs) using K-Means and Floyd-Warshall Technique," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 12, 2017.

[59] X. Duan, X. Wang, Y. Liu, and K. Zheng, "SDN Enabled Dual Cluster Head Selection and Adaptive Clustering in 5G-VANET," in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, Sept. 2016.

[60] S. Ucar, S. C. Ergen, and O. Ozkasap, "VMaSC: Vehicular multi-hop algorithm for stable clustering in Vehicular Ad Hoc Networks," in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2381–2386, Apr. 2013. ISSN: 1558-2612.

[61] X. Cheng and B. Huang, "A Center-Based Secure and Stable Clustering Algorithm for VANETs on Highways," *Wireless Communications and Mobile Computing*, vol. 2019, p. e8415234, Jan. 2019. Publisher: Hindawi.

[62] J. Yang, K. Sun, H. He, X. Jiang, and S. Chen, "Dynamic Virtual Topology Aided Networking and Routing for Aeronautical Ad-Hoc Networks," *IEEE Transactions on Communications*, vol. 70, pp. 4702–4716, July 2022. Conference Name: IEEE Transactions on Communications.

[63] K. Fuger, C. Petersen, and A. Timm-Giel, "AODV-LD: Link Duration Based Routing for Multi-Hop Aircraft-to-Ground Communication," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, pp. 1–5, June 2022. ISSN: 2577-2465.

[64] B. Du, X. Di, D. Liu, and H. Zhang, "Dynamic Graph Optimization and Performance Evaluation for Delay-Tolerant Aeronautical Ad Hoc Network," *IEEE Transactions on Communications*, vol. 69, pp. 6018–6036, Sept. 2021. Conference Name: IEEE Transactions on Communications.

[65] B. Du, X. Di, D. Liu, and H. Ning, "Coalitional Dynamic Graph Game for Aeronautical Ad Hoc Network Formation," *IEEE Internet of Things Journal*, vol. 9, pp. 5773–5784, Apr. 2022. Conference Name: IEEE Internet of Things Journal.

[66] T. Bilen and B. Canberk, "Q-Learning Driven Routing for Aeronautical Ad-Hoc Networks," *Pervasive and Mobile Computing*, vol. 87, p. 101724, Dec. 2022.

[67] A. Mchergui, T. Moulahi, and S. Zeadally, "Survey on Artificial Intelligence (AI) techniques for Vehicular Ad-hoc Networks (VANETs)," *Vehicular Communications*, vol. 34, p. 100403, Apr. 2022.

[68] A. Forster, “Machine Learning Techniques Applied to Wireless Ad-Hoc Networks: Guide and Survey,” in *Sensor Networks and Information 2007 3rd International Conference on Intelligent Sensors*, pp. 365–370, Dec. 2007.

[69] M. Qassab and Q. Ali, “A Comprehensive Survey on Current Literature, Standards, Applications and Projects of Self-Organizing Aerial Ad Hoc Network (AANET) in Smart Cities,” *Current Chinese Computer Science*, vol. 01, Feb. 2021.

[70] M. W. Berry, A. Mohamed, and B. W. Yap, eds., *Supervised and Unsupervised Learning for Data Science*. Unsupervised and Semi-Supervised Learning, Cham: Springer International Publishing, 2020.

[71] L. Lorenzo and J. Arroyo, “Analysis of the cryptocurrency market using different prototype-based clustering techniques,” *Financial Innovation*, vol. 8, p. 7, Jan. 2022.

[72] K. P. Sinaga and M.-S. Yang, “Unsupervised K-Means Clustering Algorithm,” *IEEE Access*, vol. 8, pp. 80716–80727, 2020. Conference Name: IEEE Access.

[73] H. Deniz, M. Simsek, and B. Kantarci, “On Delay Sensitivity Clusters of Microgrid Data Aggregation Under LTE-A Links,” in *2021 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, pp. 1–6, May 2021.

[74] P. Govender and V. Sivakumar, “Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019),” *Atmospheric Pollution Research*, vol. 11, pp. 40–56, Jan. 2020.

[75] D. Miljković, “Brief review of self-organizing maps,” in *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 1061–1066, May 2017.

[76] M. Simsek, B. Kantarci, and A. Boukerche, “Collaborative Self Organizing Map with DeepNNs for Fake Task Prevention in Mobile Crowdsensing,” in *ICC 2022 - IEEE International Conference on Communications*, pp. 4794–4799, May 2022. arXiv:2203.12434 [cs].

[77] M. Simsek, B. Kantarci, and A. Boukerche, “Utility-Aware Legitimacy Detection of Mobile Crowdsensing Tasks via Knowledge-Based Self Organizing Feature Map,”

IEEE Transactions on Mobile Computing, pp. 1–1, 2021. Conference Name: IEEE Transactions on Mobile Computing.

[78] M. Simsek, B. Kantarci, A. Boukerche, and S. Khan, "Machine Learning-Backed Planning of Rapid COVID-19 Tests With Autonomous Vehicles With Zero-Day Considerations," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, pp. 41–52, Feb. 2022. Conference Name: IEEE Transactions on Emerging Topics in Computational Intelligence.

[79] M. Simsek, B. Kantarci, and A. Boukerche, "Self Organizing Feature Map-Integrated Knowledge-Based Deep Network Against Fake Crowdsensing Tasks," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, pp. 1–6, Dec. 2020. ISSN: 2576-6813.

[80] X. Chen, M. Simsek, and B. Kantarci, "Locally reconfigurable Self Organizing Feature Map for high impact malicious tasks submission in Mobile Crowdsensing," *Internet of Things*, vol. 12, p. 100297, Dec. 2020.

[81] Y. Zhang, M. Simsek, and B. Kantarci, "Empowering Self-Organized Feature Maps for AI-Enabled Modeling of Fake Task Submissions to Mobile Crowdsensing Platforms," *IEEE Internet of Things Journal*, vol. 8, pp. 1334–1346, Feb. 2021. Conference Name: IEEE Internet of Things Journal.

[82] M. Simsek and B. Kantarci, "Artificial Intelligence-Empowered Mobilization of Assessments in COVID-19-like Pandemics: A Case Study for Early Flattening of the Curve," *International Journal of Environmental Research and Public Health*, vol. 17, p. 3437, Jan. 2020. Number: 10 Publisher: Multidisciplinary Digital Publishing Institute.

[83] Y. Zhang, M. Simsek, and B. Kantarci, "Self Organizing Feature Map for Fake Task Attack Modelling in Mobile Crowdsensing," in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Dec. 2019. ISSN: 2576-6813.

[84] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "OPTICS: ordering points to identify the clustering structure," *ACM SIGMOD Record*, vol. 28, pp. 49–60, June 1999.

[85] P. Bhattacharjee and P. Mitra, "A survey of density based clustering algorithms," *Frontiers of Computer Science*, vol. 15, p. 151308, Sept. 2020.

[86] T. Bilen, H. Ahmadi, B. Canberk, and T. Q. Duong, "Aeronautical Networks for In-Flight Connectivity: A Tutorial of the State-of-the-Art and Survey of Research Challenges," *IEEE Access*, vol. 10, pp. 20053–20079, 2022. Conference Name: IEEE Access.

[87] Aviationstack, "Real-Time Flight Tracker & Status API."

[88] P. Liashchynskyi and P. Liashchynskyi, "Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS," Dec. 2019. Number: arXiv:1912.06059 arXiv:1912.06059 [cs, stat].

[89] D. Medina, F. Hoffmann, S. Ayaz, and C.-H. Rokitansky, "Feasibility of an Aeronautical Mobile Ad Hoc Network Over the North Atlantic Corridor," in *2008 5th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pp. 109–116, June 2008. ISSN: 2155-5494.

[90] R. Nagel and S. Eichler, "Efficient and Realistic Mobility and Channel Modeling for VANET Scenarios using OMNeT++ and INET-Framework," in *Proceedings of the First International ICST Conference on Simulation Tools and Techniques for Communications Networks and Systems*, (Marseille, France), ICST, 2008.

[91] X. Zhang and Y. Li, "Self-organizing map as a new method for clustering and data analysis," in *Proceedings of 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan)*, vol. 3, pp. 2448–2451 vol.3, Oct. 1993.

[92] M. Shahbazi, M. Simsek, and B. Kantarci, "Density-Based Clustering and Performance Enhancement of Aeronautical Ad Hoc Networks," in *2022 International Balkan Conference on Communications and Networking (BalkanCom)*, pp. 51–56, Aug. 2022.

[93] M. Ayyub, A. Oracevic, R. Hussain, A. A. Khan, and Z. Zhang, "A comprehensive survey on clustering in vehicular networks: Current solutions and future challenges," *Ad Hoc Networks*, vol. 124, p. 102729, Jan. 2022.

[94] B. Karp and H. T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in *Proceedings of the 6th annual international conference on Mobile computing and networking*, MobiCom '00, (New York, NY, USA), pp. 243–254, Association for Computing Machinery, Aug. 2000.

[95] B. Bakshi, P. Krishna, N. Vaidya, and D. Pradhan, "Improving performance of TCP over wireless networks," in *Proceedings of 17th International Conference on Distributed Computing Systems*, pp. 365–373, May 1997. ISSN: 1063-6927.

[96] A. Kherani, R. El-Azouzi, and E. Altman, "Stability-Throughput Tradeoff and Routing in Multi-hop Wireless Ad-Hoc Networks.," pp. 25–40, Jan. 2006.

[97] Z. Zhang, A. Boukerche, and R. Pazzi, "A novel multi-hop clustering scheme for vehicular ad-hoc networks," in *Proceedings of the 9th ACM international symposium on Mobility management and wireless access*, MobiWac '11, (New York, NY, USA), pp. 19–26, Association for Computing Machinery, Oct. 2011.

[98] P. Kumar and S. Verma, "Implementation of modified OLSR protocol in AANETs for UDP and TCP environment," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, pp. 1305–1311, Apr. 2022.