

Multiple-Objective Packet Routing Optimization for Aeronautical *ad-hoc* Networks

Jiankang Zhang, *Senior Member, IEEE*, Dong Liu, *Member, IEEE*, Sheng Chen, *Fellow, IEEE*,
Soon Xin Ng, *Senior Member, IEEE*, Robert G. Maunder, *Senior Member, IEEE*, Lajos Hanzo, *Life Fellow, IEEE*

Abstract—Providing Internet service above the clouds is of ever-increasing interest and in this context aeronautical *ad-hoc* networking (AANET) constitutes a promising solution. However, the optimization of packet routing in large ad hoc networks is quite challenging. In this paper, we develop a discrete ϵ multi-objective genetic algorithm (ϵ -DMOGA) for jointly optimizing the end-to-end latency, the end-to-end spectral efficiency (SE), and the path expiration time (PET) that specifies how long the routing path can be relied on without re-optimizing the path. More specifically, a distance-based adaptive coding and modulation (ACM) scheme specifically designed for aeronautical communications is exploited for quantifying each link's achievable SE. Furthermore, the queueing delay at each node is also incorporated into the multiple-objective optimization metric. Our ϵ -DMOGA assisted multiple-objective routing optimization is validated by real historical flight data collected over the Australian airspace on two selected representative dates.

Index Terms—Aircraft mobility model, aeronautical *ad-hoc* network, adaptive coding and modulation, routing, multiple-objective optimization.

I. INTRODUCTION

Internet access has become almost ubiquitously supported by the global terrestrial mobile networks relying on the fourth-generation (4G) and fifth-generation (5G) wireless systems. Hence, having Internet access has become virtually indispensable. The provision of Internet-above-the-clouds [1] is also of ever-increasing interest to both the civil aviation airlines and to the passengers.

In-flight WiFi relying on satellites and cellular systems has been available on the some flights of global airlines, such as British Airways, American Airlines, United Airlines, Emirates and Delta Airlines, just to name a few. However, aeronautical communications directly relying on satellites and/or cellular systems suffer from high cost, limited coverage, limited capacity, and/or high end-to-end latency. Furthermore, the cellular systems that can support ground-to-air (G2A) communications are limited to a line-of-sight range and require specially designed ground stations (GSs), which necessitate

the roll-out of an extensive ground infrastructure to cover a wide area. Intuitively, it is quite a challenge to provide ubiquitous coverage for every flight at a low cost by directly relying on satellites and/or cellular systems. As an alternative architecture, aeronautical *ad-hoc* networks (AANETs) [2]–[4] are capable of extending the coverage, whilst reducing the communication cost by hopping messages from plane to plane. Each aircraft as a network node is capable of sending, receiving and relaying messages until the messages are delivered to or fetched from a GS, so as to enable Internet access.

Hence, routing, which finds an ‘optimal’ path consisting of a sequence of relay nodes, is one of the most important challenges to be solved in support of this Internet-above-the-clouds application. Routing protocols have been intensively investigated in mobile *ad-hoc* networks (MANETs) [5] and vehicular *ad-hoc* networks (VANETs) [6] as well as in the flying *ad-hoc* networks (FANETs) [7]. However, as our analysis in [1] has revealed, AANETs have their unique features in terms of flying speed, altitude, propagation characteristics and network coverage as well as node mobility, which are different from those of MANETs, VANETs and FANETs. Therefore, the routing protocols specially developed for MANETs, VANETs and FANETs cannot be directly applied to AANETs, although their philosophies may be appropriately adopted. Hence, we will focus our attention on routing protocols specially designed for aeronautical networks.

Sakhaee and Jamalipour [8] showed that the probability of finding at least two but potentially up to dozens of aircraft capable of establishing an AANET above-the-cloud is close to 100%. It was inferred by investigating a snapshot of flight data over the United States (US). They also proposed a quality of service (QoS) based so-called multipath Doppler routing protocol by jointly considering both the QoS and the relative velocity of nodes in order to find stable routing paths. Luo *et al* [9] proposed a reliable multipath transport protocol for AANETs by exploiting the path diversity provided by heterogeneous aeronautical networks. By exploiting the geographical information, Iordanakis *et al* [10] proposed a routing protocol for aeronautical mobile ad hoc networks, which may be viewed as an evolved version of the classical *ad-hoc* on demand distance vector based routing (AODV) originally developed for MANETs. Furthermore, geographical information was intensively exploited in designing the routing protocols of [11]–[15], [19]–[24] for AANETs by considering that the locations of civil passenger aircraft are always available with

J. Zhang is with Department of Computing & Informatics, Bournemouth University, BH12 5BB, U.K. (E-mail: jzhang3@bournemouth.ac.uk).

D. Liu is with the School of Cyber Science and Technology, Beihang University, Beijing 100191, China (e-mail: dliu@buaa.edu.cn).

S. Chen, S. X. Ng, R. Maunder and L. Hanzo are with School of Electronics and Computer Science, University of Southampton, SO17 1BJ, U.K. (E-mails: {sqc, sxn, rm, lh}@ecs.soton.ac.uk).

The financial support of the Engineering and Physical Sciences Research Council projects EP/W016605/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028).

the aid of the on-board radar and the automatic dependent surveillance-broadcast (ADS-B) system [28], [29]. Explicitly, the authors of [11]–[13] developed a routing protocol termed as AeroRP, which is a highly adaptive location-aware routing algorithm exploiting the broadcast nature of the wireless medium along with the physical node location and trajectory knowledge for improving the data delivery in Mach-speed mobile scenarios. However, AeroRP ignores the delay imposed by relaying and it is prone to network congestion. Medina *et al.* [14] proposed a geographic load sharing strategy for fully exploiting the total air-to-ground capacity available at any given instant in time. In their work, the network congestion was avoided by a congestion-aware handover strategy capable of efficient load balancing among Internet Gateways. Gu *et al.* [15] proposed a delay-aware routing scheme using a joint metric relying on both the relative velocity and the expected queueing delay of nodes for selecting the next node. Wang *et al.* [30] also designed a delay-aware routing protocol for aeronautical networks, which explored the effect of dual connectivity on delay-aware resource control in heterogeneous aeronautical networks. Du *et al.* [16] aimed for minimizing the end-to-end transmission delay by jointly exploiting the direct transmissions, relayed transmissions and opportunistic transmissions. By contrast, we have also minimized the end-to-end delay of AANETs by exploiting a weighted digraph and the shortest-path algorithm in [17] and by invoking deep reinforcement learning in [18], respectively.

Both Wang *et al.* [19] as well as Mahmoud and Larrieu [20] exploited the geographical information provided by the ADS-B system. Specifically, Wang *et al.* [19] eliminated need for the traditional routing beaconing and improved the next hop selection, whilst Mahmoud and Larrieu [20] concentrated on improving the information security of routing protocols. The security issues routing were further considered by Swidan *et al.* [21] and Pang *et al.* [23]. Specifically, Swidan *et al.* [21] proposed a secure geographical routing protocol by using the GS as a trusted third party for authentication and key transport. However, their solution required an additional transceiver at each aircraft having a communication range of 150 km and wide downlink bandwidth for point-to-point communication with the GS. By contrast, Pang *et al.* [23] advocated an identity-based public key cryptosystem, which relies on the authentication of neighbor nodes and establishes a shared secret during the neighbor discovery phase, followed by the encryption of the data during the data forwarding phase. Luo and Wang [24] proposed a multiple QoS parameters-based routing protocol in order to improve the overall network performance for communication between aircraft and the ground. Explicitly, they jointly optimized the maximum path life-time period, maximum residual path load capacity, and minimum path latency from all the available paths between an aircraft node and the Internet gateways with the aid of carefully selected weighting factors for the path life-time period, residual path load capacity and path latency.

Since the overall network performance depends on multiple factors, it is unfair to optimize the overall network performance by relying on a single factor, such as the error probability, latency or capacity. As argued in [1], in contrast to conventional

single-objective optimization, multi-objective optimization is capable of finding all the global Pareto optimal solutions by potentially allowing the system to be reconfigured in any of its most desired optimal operating mode. In [25], [26], we developed a twin-component near-Pareto routing scheme by invoking the classic non-dominated Sorting Genetic Algorithm II (NSGA-II), which is capable of finding a set of trade-off solutions in terms of the total delay and the throughput. Furthermore, in [27], we extended our single objective optimization efforts of [18] to multi-objective packet routing optimization in AANETs in the north-Atlantic region. In the existing AANET literature, there is a paucity of contributions on applying multiple-objective routing optimization by jointly considering the end-to-end throughput, end-to-end latency and the corresponding path expiration time (PET). However, the nodes in AANETs are airplanes, which typically fly at a speed of 880 to 926 km/h [1], hence a path may break quite soon. Hence, the PET also becomes a much more critical metric in AANETs compare to MANETs, VANETs and FANETs. Furthermore, most of the existing routing protocols were investigated mainly based on randomly generated flight data, which cannot reveal the network performance of real-world AANETs constituted by the real flights in the air. Against this background, we propose a multiple-objective routing optimization scheme for the AANET, and we evaluate its overall network performance using large-scale real historical flight data over the Australian airspace. Explicitly, our main contributions are summarized as follows.

- 1) We propose multi-objective routing optimization for jointly optimizing the end-to-end spectral efficiency (SE), the end-to-end latency and the PET. More specifically, the latency is addressed as one of the objectives by constraining the maximum number of affordable hops, whilst the congestion is addressed by imposing a certain queueing delay at each node. Furthermore, the distance-based adaptive coding and modulation (ACM) of [31], [32], which was specifically designed for aeronautical communications is adopted for quantifying the each link's SE so as to determine the final end-to-end SE. Naturally, the lowest link-throughput limits the entire path's throughput.
- 2) At the time of writing, there is no discrete ϵ -MOGA version in the open literature and there is no application example of ϵ -MOGA in the context of routing problems. Based on the philosophy of ϵ multi-objective genetic algorithm (ϵ -MOGA) [33], which operates on a continuous parameter space for finding the Pareto-front of optimal solutions, we develop a discrete version of ϵ -MOGA by considering the specific features of the routing paths consisting of discrete aircraft IDs, which we refer to as the discrete ϵ multi-objective genetic algorithm (ϵ -DMOGA). Explicitly, in order to accommodate the unique feature of routing problem in AANETs, we have adapted the existing ϵ -MOGA to create a discrete ϵ -MOGA by considering the discrete search space of routing problems in AANETs relying on discrete aircraft IDs. This adaptation is not straightforward at all, because it in-

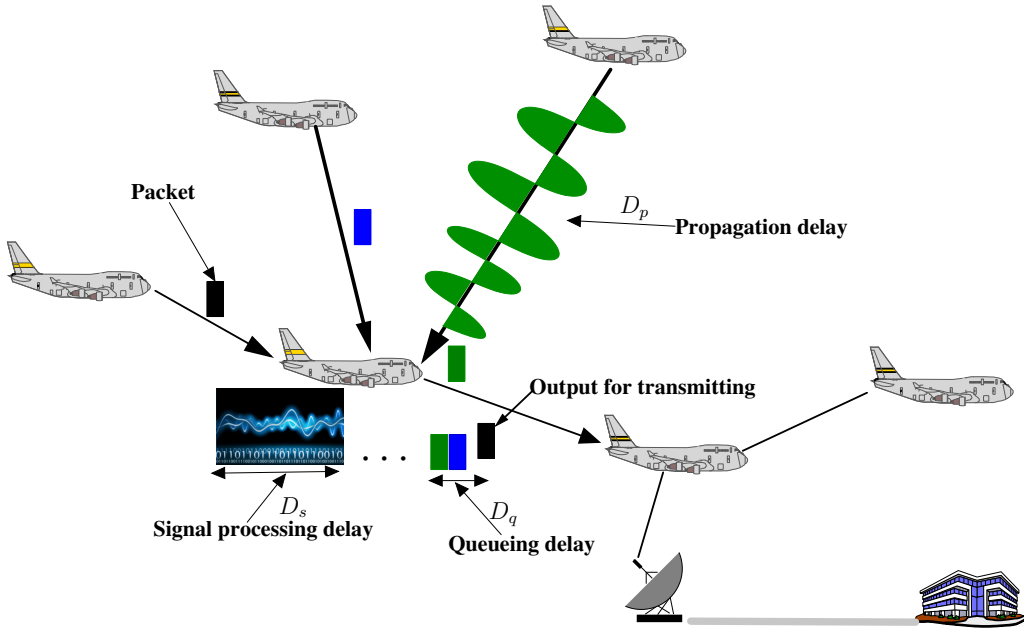


Fig. 1. Illustration for the sources of packet delay in routing

volves new operations conceiving the encoding/decoding of chromosomes, as well as new crossover and mutation operations with respect to the specific nature of discrete variables that constitute a routing path emerging from a source aircraft node to a destination ground station. We use this ϵ -DMOGA for efficiently solving the proposed multi-objective routing optimization problem.

- 3) The overall network performance of our multiple-objective routing optimization quantified in terms of the end-to-end latency, the end-to-end SE and the PET, are investigated based on large-scale real historical flight data recorded over the Australian airspace. More specifically, real historical flight data of two representative dates of the top-five airlines in Australia's domestic flights are exploited for our investigations.

The remainder of this paper is organized as follows. The network architecture is presented in Section II, which includes the mobility model and the multiple-objective functions to be investigated. In Section III, we develop a discrete version of ϵ -MOGA, termed as ϵ -DMOGA, by exploiting the discrete nature of the routing paths specified by the aircraft IDs, which provides an effective tool to solve our proposed multi-objective routing optimization. Our simulation results based on real historical flight data recorded over the Australian airspace are presented in Section IV, while our conclusions are offered in Section V.

II. NETWORK ARCHITECTURE

The avionic network considered takes into account the peculiarities of aeronautical communications and exploits them for optimizing our multiple objectives. Satellites are also included into the AANET considered, which are used as the last resort for an aircraft outside all the neighbouring aircrafts' communications range. In contrast to MANETs and VANETs, the nodes of an AANET are distributed in a 3D space, and

they move at extremely high speed over long distances. The geographic information of each aircraft is available for routing optimization and network design, which can be obtained with the assistance of the global positioning system (GPS) and airborne radar carried by each aircraft [35]. Moreover, ADS-B systems have been widely deployed on commercial passenger aircraft, which can also provide an information vector including the aircraft ID, position, ground speed and heading directions [24].

A. Mobility model

In our network optimization, we consider all the aircraft during the 24 hours of both the busiest and the quietest day in a year. Based on the historical flight observation on Flightradar24¹, the busiest day was June 29th in 2018, whilst the quietest day was December 25th in 2018. The movement of a node that represents an aircraft was recorded as real historical flight data, which typically includes the flight phases of holding, takeoff/landing, taxiing, and parking that always performs at an airport or near an airport, as well as the longest en-route phase. Contrast to the traditional nodes in MANETs and VANETs as well as even the nodes in FANETs, the nodes in AANETs move at a relatively high speed during the en-route phase, typically at velocities of 800 to 1000 km/h.

As our evaluation is based on real historical flight data over the Australian airspace, the nodes tend to be sparser compared to Europe and North-America. The nodes typically fly between coastal cities, such as Sydney, Melbourne, Perth and Gold Coast. The eastern coastal airspace is much busier than the northern, western and southern coastal airspace, since most people reside in eastern coastal cities, and most international flights depart from/arrive at eastern coastal cities. The central

¹Flightradar24 is a global flight tracking service that provides real-time information of aircraft around the world, which is accessible on <https://www.flightradar24.com>

Australian airspace quite sparse, since only few people live in the central area of Australia.

B. Objective functions

In the Australian aeronautical network, there are N_G GSs at five Australian airports. An aircraft can access the Internet by relying on an optimal routing path to a GS. Hence the flights do not have to rely on costly satellites as a relay node to access the Internet.

We will consider the achievable end-to-end latency and end-to-end throughput as well as the stability of the routing path. Explicitly, the end-to-end latency is the sum of the signal propagation delays, signal processing delays and queuing delays. The end-to-end throughput is determined by the specific link having the minimum throughput. Again, the stability of the routing path is quantified in terms of the PET, which is in turn determined by the link having the minimum expiration time given a specific ACM mode.

1) *The end-to-end latency*: Let us now quantify the propagation, signal processing and queuing delays. Although there may be a ground-reflection component in the received signal of aeronautical communications, it is dominated by the Line of Sight (LoS) path for air-to-air (A2A) communications in AANETs [31]. Hence, the ground-reflected component may be neglected in A2A transmission and the propagation delay can be modelled by that of the LoS path limited by the speed of light. Explicitly, let the distance between node r_n and node r_{n+1} be denoted by $d_{r_n, r_{n+1}}$. The propagation delay $D_p(r_n \rightarrow r_{n+1})$, as illustrated in Fig. 1, between node r_n and node r_{n+1} is given by

$$D_p(r_n \rightarrow r_{n+1}) = \frac{d_{r_n, r_{n+1}}}{c}, \quad (1)$$

where c is the speed of light, $r_n \rightarrow r_{n+1}$ is the link spanning from node r_n to node r_{n+1} in the routing path $\mathbf{r} = \{r_n \rightarrow r_{n+1} \rightarrow \dots \rightarrow r_{N+n-1}\}$, which consists of $N-1$ hops from the source node r_n to the destination node r_{N+n-1} , with N being the number of nodes in the routing path \mathbf{r} .

As shown in Fig. 1, the signal processing delay is the time that a relay node takes to process a packet before it can forward it to its output queue, which includes the decoding and forwarding operations, destination lookup, packet-header updates, etc. Intuitively, when node r_n is the source node, obviously there is no decoding and forwarding operations, destination lookup, packet-header updates, hence $D_s(r_n \rightarrow r_{n+1}) = 0$ in this case. However, when node r_n is not a source node, but a relay node, it has to carry out the decoding and forwarding operations, destination lookup, packet-header updates. Then a constant signal processing delay will be imposed. Depending on the digital signal processing capability of the hardware and the detailed operations needed, the signal processing latency ranges from 0.5 to ten milliseconds with some complex designs having as much as 30 milliseconds [36]. In our investigations, we set it to 5 ms as a compromise

consideration. Without loss of generality, we formulate signal processing delay at node r_n as

$$D_s(r_n \rightarrow r_{n+1}) = \begin{cases} 0 \text{ ms,} & \text{if node } r_n \text{ is the source node,} \\ 5 \text{ ms,} & \text{if node } r_n \text{ is not a source node.} \end{cases} \quad (2)$$

The queuing delay is the time that a packet waits in a relay node after arriving at the node's queue until it can be processed plus the waiting time before the processed packet can be transmitted to the output link, as illustrated in Fig. 1. The input queuing delay is proportional to the number of packets that have already been waiting in the queue at a given time, while the output transmission delay is upper bounded by the time needed for transmitting a packet. Extensive research has been dedicated to queue theory, queue scheduling and/or minimizing the queuing delay at a node in networks [37]–[39]. We can reasonably model the input queuing delay at node r_n as follows

$$D_{q_1} = O_{r_n} D_{q_0}, \quad (3)$$

where the indicator $O_{r_n} \in \{0, 1, \dots, N_B\}$, which indicates how many packets are at the front of the queue, and N_B is the maximum number of packets that the node can have in its queue, while D_{q_0} is a fixed processing delay related to transmitting a whole packet through the node's output link. The total buffering delay, denoted as D_{q_2} , is clearly upper bounded by D_{q_0} , i.e., we have $D_{q_2} \leq D_{q_0}$. The total queuing delay of forwarding a packet from r_n to r_{n+1} , which is the sum of the input queuing delay D_{q_1} and the output buffering delay D_{q_2} , can be expressed as

$$D_q(r_n \rightarrow r_{n+1}) = D_{q_1} + D_{q_2} \approx (O_{r_n} + 1) D_{q_0}. \quad (4)$$

Note that a packet can only be routed through a node if the node's queue is not full, i.e. the number of packets waiting in the node's queue is less than N_B . This imposes a constraint on the routing decisions. Again, the delay imposed on a packet during its passage from node r_n to node r_{n+1} is the sum of the signal propagation delay, signal processing delay and queuing delay, which is given by

$$D(r_n \rightarrow r_{n+1}) = D_p(r_n \rightarrow r_{n+1}) + D_s(r_n \rightarrow r_{n+1}) + D_q(r_n \rightarrow r_{n+1}). \quad (5)$$

Therefore, the end-to-end latency of the routing path \mathbf{r} is given by

$$D(\mathbf{r}) = \sum_{n=1}^{N-1} D(r_n \rightarrow r_{n+1}). \quad (6)$$

2) *The end-to-end throughput*: The end-to-end throughput is determined by the link in the routing path \mathbf{r} , which has the minimum link throughput. Let $C(r_n \rightarrow r_{n+1})$ denote the link throughput between node r_n and node r_{n+1} . Then the end-to-end throughput is given by

$$C(\mathbf{r}) = \min_{1 \leq n \leq N-1} C(r_n \rightarrow r_{n+1}). \quad (7)$$

The achievable link throughput $C(r_n \rightarrow r_{n+1})$ is affected by the channel conditions and other factors, such as the co-channel inference. The link throughput is a function of instantaneous signal-to-interference-plus-noise ratio (SINR), where

the instantaneous SINR may be estimated using pilot signals in traditional terrestrial mobile communications. However, the problem in aeronautical communication applications is that the high speed of aircraft may result in uncorrelated small-scale fading and consequently in unreliable estimates of the instantaneous SNR or SINR, further aggravated by frequently switching among the ACM modes. Using erroneous instantaneous SNR or SINR estimates for frequently switching ACM modes may cause frequent unsuccessful transmissions, because the SINR estimates quickly become obsolete. Moreover, the instantaneous SINR does fluctuate around its average, but it has a limited range in the typical LoS scenarios. Hence, in the routing problem in AANETs, the best available distance-based ACM [31], [32] is invoked for quantifying the link throughput of the air-to-air aeronautical communications. Hence, the achievable link throughput $C(r_n \rightarrow r_{n+1})$ is a function of the communication distance between node r_n and node r_{n+1} , which is also affected by the co-channel interference imposed by the neighbour aircraft. Furthermore, given a set of K ACM modes, there are $K + 1$ distance thresholds $\{d_k\}_{k=0}^K$. The data-transmitting aircraft selects an ACM mode to transmit/relay the data according to

$$\text{if } d_k \leq d < d_{k-1} \text{ choose } k\text{-th ACM mode,} \quad (8)$$

where $d_0 = D_{\max} = D_{A2A} > d_1 > \dots > d_{K-1} > d_K = D_{\min}$. Clearly, if the distance d is outside the range of $[D_{\min}, D_{\max}]$, there will be no adequate communication link. More specifically, D_{\min} is the minimum flight-safety separation that must be obeyed according to the aviation safety regulation, whilst D_{\max} is the maximum communication range of two aircraft, which is given by the radio distance to horizon of A2A communication [1].

3) *The path expiration time*: The PET is determined by the most vulnerable link of the routing path \mathbf{r} that has the shortest link expiration time (LET). Let $T_{\text{LET}}(r_n \rightarrow r_{n+1})$ be the LET of the link between node r_n and node r_{n+1} for offering ACM mode- i . Then the PET of \mathbf{r} is given by

$$T_{\text{PET}}(\mathbf{r}) = \min_{1 \leq n \leq N-1} T_{\text{LET}}(r_n \rightarrow r_{n+1}). \quad (9)$$

Since the ACM is adopted, we have to modify the formulation of calculating the LET given in [8], [24]. Specifically, given the speeds v_{r_n} and $v_{r_{n+1}}$, the heading directions θ_{r_n} and $\theta_{r_{n+1}}$ as well as the coordinates (x_{r_n}, y_{r_n}) and $(x_{r_{n+1}}, y_{r_{n+1}})$ for node r_n and node r_{n+1} , respectively, assume that the distance between r_n and r_{n+1} satisfies $d_k \leq d_{r_n, r_{n+1}} < d_{k-1}$, where d_{k-1} is the distance threshold or maximum distance that may be bridged over by the ACM mode k , as defined in Eq. (8). Then the LET of $T_{\text{LET}}(r_n \rightarrow r_{n+1})$ is calculated according to:

$$T_{\text{LET}}(r_n \rightarrow r_{n+1}) = \frac{(ab+ef) + \sqrt{(a^2+e^2)d_{k-1}^2 - (af-be)^2}}{a^2+e^2}, \quad (10)$$

where a, b, e and f are given by [8], [24]

$$a = v_{r_n} \cos \theta_{r_n} - v_{r_{n+1}} \cos \theta_{r_{n+1}}, \quad (11)$$

$$b = x_{r_n} - x_{r_{n+1}}, \quad (12)$$

$$e = v_{r_n} \sin \theta_{r_n} - v_{r_{n+1}} \sin \theta_{r_{n+1}}, \quad (13)$$

$$f = y_{r_n} - y_{r_{n+1}}. \quad (14)$$

Intuitively, when aircraft r_n and aircraft r_{n+1} have the same speed and heading direction, the LET between them is theoretically infinity. However, the associated LET is always upper-bounded by their flight time in practice. When aircraft r_n and aircraft r_{n+1} have the exact opposite heading direction, they will have the minimum LET.

C. Multi-objective routing optimization

The specific multi-objective optimization is advocated here aims for maximizing the end-to-end achievable throughput, for minimizing the end-to-end latency and for maximizing the PET, which is formulated as

$$\begin{cases} \mathcal{J}_1(\mathbf{r}) = \max C(\mathbf{r}), \\ \mathcal{J}_2(\mathbf{r}) = \min D(\mathbf{r}), \\ \mathcal{J}_3(\mathbf{r}) = \max T_{\text{PET}}(\mathbf{r}), \end{cases} \quad (15)$$

$$\text{s.t.} \begin{cases} D(\mathbf{r}) \leq 250 \text{ ms}, \\ N - 1 \leq 5. \end{cases} \quad (16)$$

The round-trip latency of geostationary satellite links is about 250 ms [14], which is imposed by the propagation delay up and down from the satellite. Hence intuitively, the end-to-end latency should be less than 250 ms, which results in the first constraint $D(\mathbf{r}) \leq 250$ ms. Naturally, when low earth-orbit (LEO) satellites are considered at say 600 Km altitude, their round-trip delay is as low as 4 ms. The second constraint of $N - 1 \leq 5$ is based on a practical consideration of the Australian scenario. Explicitly, the aircraft tend to fly over the land of Australia, where the GSs are at airports on the ground. Hence, an aircraft is typically capable of accessing the Internet with a small number of hops. Furthermore, the second constraint of $N - 1 \leq 5$ is also used for limiting the number of nodes involved, which is helpful for controlling the AANET size. Since a routing path having more hops is more vulnerable to cyber attacks, limiting the number of nodes in a routing path also helps to secure information transmission. Clearly, for geographic areas, such as the AANET over the Atlantic Ocean, we will have to set a higher value for the maximum number of hops.

III. DISCRETE ϵ -MOGA BASED PARETO-OPTIMIZATION OF AANET ROUTING PROBLEM

No closed-form solution can be derived for the multi-objective optimization problem (15) under the constraint (16). There are diverse methods of solving multi-objective optimization problems, such as the Lexicographic method [40], weighted sum method [41], elitist non-dominated sorting genetic algorithm (NSGA-II) [42], Strength Pareto Evolutionary Algorithm (SPEA) [43], SPEA-II [44], Pareto Enveloped based Selection Algorithm (PESA) [45] and PESA-II [46] as well as numerous other variations with their pros and cons. For example, the Lexicographic method is sensitive to the iteration order of objectives, while the weighted method is sensitive to the weightings and both of them suffer from high computational burden. The NSGA-II does not perform very well for more objectives, while the SPEA and SPEA2 are also exhibit high computational complexity, and the PESA as well

as PESA2 are sensitive to the size of hyperbox. By contrast, as a member of the elitist multi-objective evolutionary algorithm family based on the concept of ϵ -dominance, ϵ -MOGA has the compelling characteristics of efficient parallel computing along with the efficient control of the elitist archive, where problem solutions are stored. Hence, ϵ -MOGA outperforms the above-mentioned multi-objective optimization algorithms in terms of its convergence, diversity of solutions and computational efficiency. Hence, we apply the ϵ -DMOGA for determining the optimal Pareto-front at a moderate computational burden. Our ϵ -DMOGA is developed from the ϵ -MOGA [33], [34], which is an elitist multi-objective evolutionary algorithm based on the concept of ϵ -dominance [47], by taking into consideration the discrete nature of the routing path constituted by discrete aircraft IDs.

A. ϵ -DMOGA

In the Pareto optimal set for the multi-objective optimization problem (15), no single solution should dominate others. Explicitly, a solution routing path r_1 by definition dominates another routing path r_2 in the routing path space, if and only if all the objectives of r_1 are no worse than the objectives of r_2 and at least one objective of r_1 is better than that of r_2 , which is formulated as

$$\forall i=1, 2, 3, \mathcal{J}_i(r_1) \preceq \mathcal{J}_i(r_2) \text{ and } \exists k=1, 2, 3, \mathcal{J}_k(r_1) \prec \mathcal{J}_k(r_2). \quad (17)$$

The operator \preceq represents that the lefthand objective is no worse than the righthand one. For example, $\mathcal{J}_1(r_1) \preceq \mathcal{J}_1(r_2)$ is equal to $C(r_1) \geq C(r_2)$, whilst $\mathcal{J}_2(r_1) \preceq \mathcal{J}_2(r_2)$ is equal to $D(r_1) \leq D(r_2)$. Similarly, \prec represents that the lefthand objective is better than the righthand one. Then, the Pareto-front solution set \mathbf{R} can be formulated as

$$\mathbf{R} = \{r \in \mathbf{R} \mid \nexists \tilde{r} \in \mathbf{R} : \tilde{r} \preceq r\}, \quad (18)$$

where \nexists represents ‘does not exist’. Hence, for $r \in \mathbf{R}$, $\nexists \tilde{r} \in \mathbf{R} : \tilde{r} \preceq r$ means that there is no single \tilde{r} , which dominates r if it does not belong to \mathbf{R} .

Again, without loss of generality, we discuss the AANET over the Australian airspace. For other AANETs, similar discussions can be applied subject to minor modifications. Intuitively, the end-to-end throughput is a more dominant criterion for the Internet-above-the-clouds than the end-to-end latency and the PET. Thus, our ϵ -DMOGA based multi-objective routing optimization may start from a direct connection to any of the N_G GSs by finding the Pareto-front of optimal solutions with respect to throughput, latency and PET. Since N_G is small, we can easily find a ‘Pareto-optimal’ GS that dominates other GSs by enumerating the multiple objectives to each GS. This is effectively the single-hop solution. Then, the ϵ -DMOGA based multi-objective routing optimization can proceed to find all the Pareto-front solutions within an affordable number of hops, say $N - 1 = 2, 3, 4, 5$ with respect to the multi-component objective function (15).

The ϵ -DMOGA based multi-objective routing optimization is characterized by its initialization, individual mutation, crossover and selection operations used throughout exploring

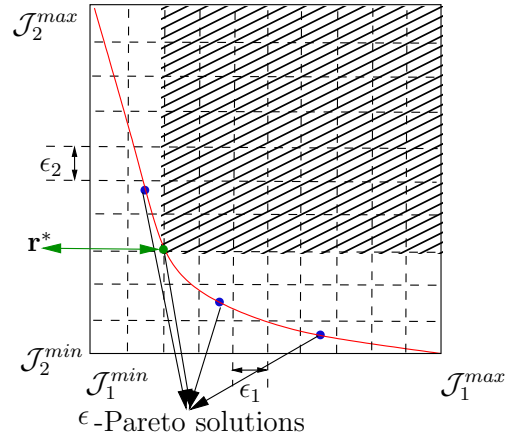


Fig. 2. The illustration of ϵ -dominance and the ϵ -Pareto-front solution.

the search space in a generation-based progression, until the termination criterion is met. We now detail this ϵ -DMOGA.

- 1) **Initialization.** At the first generation of $g = 1$, where g denotes the generation index, the ϵ -DMOGA commences its search by randomly generating an initial population of P_s N -element routing path vectors, denoted as $\mathbf{P}^{(g)}$. Explicitly, permutation encoding is invoked for generating a chromosome representing a routing path, which consists of a string of aircraft IDs from source aircraft to the destination ground station. For each individual of $\mathbf{P}^{(g)}$, the first element is the source node r_1 , and its second element is randomly selected from the node space

$$\mathbb{A} \setminus \{r_1\} = \{r \in \mathbb{A}, \sim (r \in \{r_1\})\}, \quad (19)$$

where \mathbb{A} is the node space consisting of the aircraft in air, and $\mathbb{A} \setminus \{r_1\}$ represents the node space with r_1 removed. In general, the n -th element of an individual, where $2 \leq n \leq N - 1$, is randomly selected from the node space

$$\mathbb{A} \setminus \{r_1, r_2, \dots, r_{n-1}\} = \{r \in \mathbb{A}, \sim (r \in \{r_1, r_2, \dots, r_{n-1}\})\}. \quad (20)$$

The last node in a routing path, i.e. the N -th node, is a GS, randomly selected from the GS node space \mathbb{B} , which consists of the N_G GSs at airports. The archive $\mathbf{A}^{(g)}$ that contains the elite population is initialized as the null set at the first generation $g = 1$.

The ϵ -DMOGA scheme solves the multi-objective routing optimization by evolving the main population $\mathbf{P}^{(g)}$ of P_s N -element routing path vectors from one generation to the next.

- 2) **Archive.** By calculating and comparing the multi-objective functions for the individuals of $\mathbf{P}^{(g)}$, the ϵ -Pareto-front solution set $\tilde{\mathbf{R}}$ is selected. Explicitly, the individuals in the ϵ -Pareto-front solution set $\tilde{\mathbf{R}}$ ϵ -dominate all the other individuals that are not selected into $\tilde{\mathbf{R}}$. The concept of ϵ -dominance is illustrated in Fig. 2², where the shade areas are ϵ -dominated by r^* and the other blue points on the ϵ -Pareto-front are ϵ -Pareto-front solutions.

²In Fig. 2, we only plot two dimensions in order to have a clear illustration. For our three objective dimensions as formulated in (15), the light blue areas will be cubic.

Furthermore, ϵ_i , $i = 1, 2, 3$, is the width of a box, which is defined as

$$\epsilon_i = \frac{\mathcal{J}_i^{max} - \mathcal{J}_i^{min}}{N_{\text{box},i}}, \quad (21)$$

where $N_{\text{box},i}$ is the number of partitions in the dimension of the i -th objective, which preserves the diversity of ϵ -dominance solution in the i -th objective dimension. The Pareto-front limits \mathcal{J}_i^{min} and \mathcal{J}_i^{max} for $i = 1, 2, 3$ are calculated as follows

$$\mathcal{J}_i^{max} = \max_{\mathbf{r} \in \tilde{\mathbf{R}}} \mathcal{J}_i(\mathbf{r}), i = 1, 2, 3, \quad (22)$$

$$\mathcal{J}_i^{min} = \min_{\mathbf{r} \in \tilde{\mathbf{R}}} \mathcal{J}_i(\mathbf{r}), i = 1, 2, 3. \quad (23)$$

The individuals in $\tilde{\mathbf{R}}$ that are not ϵ -dominated by the individuals in the elite population archive $\mathbf{A}^{(g)}$ are stored into $\mathbf{A}^{(g)}$. Hence the size of the archive $\mathbf{A}^{(g)}$ may vary at different generations. Furthermore, $\tilde{\mathbf{R}}$ is an intermediate Pareto-front set at the current generation that converges towards the Pareto optimal set \mathbf{R} as the population evolves over generations.

- 3) **Variation.** A new variant is generated by the amalgamation of the ‘crossover’ and ‘mutation’ operations, which are typically two separate operations in single-objective GA optimization. However, we use the terminology ‘variant’ for the operations of crossover and mutation in our ϵ -DMOGA, since both operations are controlled by the probability of crossover/mutation $p_{c/m}$. Explicitly, two individuals, $\mathbf{r}^{(P)}$ and $\mathbf{r}^{(A)}$, are randomly selected from the main population $\mathbf{P}^{(g)}$ and the elite population $\mathbf{A}^{(g)}$, respectively. Then, a randomly generated value $\alpha \in [0, 1]$ decides which operation should be applied to $\mathbf{r}^{(g,P)}$ and $\mathbf{r}^{(g,A)}$.

- ① **Crossover.** If $\alpha > p_{c/m}$, $\mathbf{r}^{(g,P)} = \{r_1^{(g,P)}, r_2^{(g,P)}, \dots, r_N^{(g,P)}\}$ and $\mathbf{r}^{(g,A)} = \{r_1^{(g,A)}, r_2^{(g,A)}, \dots, r_N^{(g,A)}\}$ will cross over part of their elements. There exist numerous crossover mechanisms, and we opt for employing the single-point crossover due to its simplicity. Explicitly, a point on both $\mathbf{r}^{(g,P)}$ and $\mathbf{r}^{(g,A)}$ is picked randomly, which is designated as the crossover point. Then the elements to the right of the crossover point are swapped between $\mathbf{r}^{(g,P)}$ and $\mathbf{r}^{(g,A)}$, which results in two new offspring, each carrying some genetic information from both parents. Given the crossover point $n \geq 2$, the two new offspring can be expressed as

$$\begin{cases} \hat{\mathbf{r}}_1^{(g,G)} = \{r_1^{(g,P)}, r_2^{(g,P)}, \dots, r_n^{(g,P)}, r_{n+1}^{(g,A)}, \dots, r_N^{(g,A)}\}, \\ \hat{\mathbf{r}}_2^{(g,G)} = \{r_1^{(g,A)}, r_2^{(g,A)}, \dots, r_n^{(g,A)}, r_{n+1}^{(g,P)}, \dots, r_N^{(g,P)}\}, \end{cases} \quad (24)$$

where the superscript G indicates that both $\hat{\mathbf{r}}_1^{(g,G)}$ and $\hat{\mathbf{r}}_2^{(g,G)}$ are stored into an auxiliary population $\mathbf{G}^{(g)}$. Note that the same aircraft ID should be avoided both in $\hat{\mathbf{r}}_1^{(g,G)}$ and $\hat{\mathbf{r}}_2^{(g,G)}$ by applying sophisticated operations, such as for example, checking

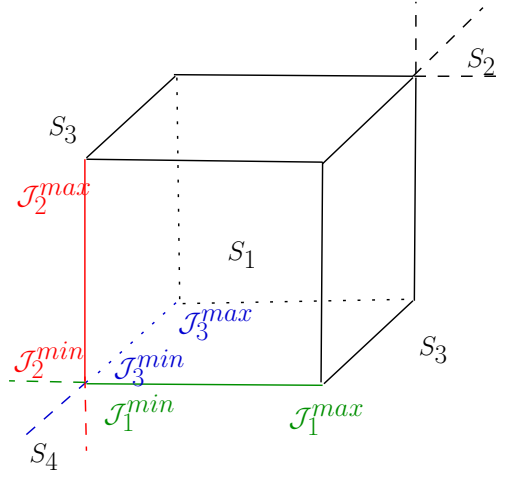


Fig. 3. The objective space can be divided into the four areas, namely, S_1 : $\mathcal{J}_i^{min} \leq \mathcal{J}_i \leq \mathcal{J}_i^{max}, \forall i = 1, 2, 3$; S_2 : $\mathcal{J}_i > \mathcal{J}_i^{max}, \forall i = 1, 2, 3$; S_4 : $\mathcal{J}_i < \mathcal{J}_i^{min}, \forall i = 1, 2, 3$; and S_3 : the rest of the objective space.

and mutating to a new non-same aircraft ID if a $r_{n+j}^{(g,A)}$ is same as a $r_j^{(g,P)}$.

- ② **Mutation.** If $\alpha \leq p_{c/m}$, $\mathbf{r}^{(g,P)} = \{r_1^{(g,P)}, r_2^{(g,P)}, \dots, r_N^{(g,P)}\}$ and $\mathbf{r}^{(g,A)} = \{r_1^{(g,A)}, r_2^{(g,A)}, \dots, r_N^{(g,A)}\}$ will mutate some of their elements. Intuitively, the mutation may occur in a single element or in multiple elements. We opt for the latter. Explicitly, an integer N_m is randomly generated in the range of $[1, N - 1]$. Then an N_m -length vector $\mathbf{l} = \{l_1, l_2, \dots, l_{N_m}\}$ is generated and each of its elements l_i , $i = 1, 2, \dots, N_m$, is selected from the integer set of $\{2, 3, \dots, N\}$ without repetition. Specifically, N_m determines the number of mutated elements with l_i , $1 \leq i \leq N_m$, specifying the positions of these elements. The pair of new offspring generated by mutation are expressed as Eq. (25), where $\hat{r}_{l_i}^{(g,P)}$ and $\hat{r}_{l_i}^{(g,A)}$, $i = 1, 2, \dots, N_m$, are the new genes generated by mutation from parent $\mathbf{r}^{(g,P)}$ and $\mathbf{r}^{(g,A)}$, respectively. These mutated elements are randomly drawn from the aircraft node set. However, the mutated elements must not be duplicated with other elements within the same individual to avoid loops in the routing path. Therefore, if a mutated element is the same as another element in the same individual, it must be mutated again until it becomes different. Similarly, both $\hat{r}_1^{(g,G)}$ and $\hat{r}_2^{(g,G)}$ are stored into the auxiliary population $\mathbf{G}^{(g)}$.

The crossover or mutation operations are operated $N_O/2$ times, which results in total of N_O new offspring in the auxiliary population $\mathbf{G}^{(g)}$.

- 4) **Selection.** The selection operation of multiple-objective optimization is much more complex than that of single-objective optimization. Explicitly, the ϵ -DMOGA calculates the multiple objective functions of the individuals in the auxiliary population $\mathbf{G}^{(g)}$ and decides, which specific individual will be selected into the elite population $\mathbf{A}^{(g)}$ on the basis of its location in the objective space, as

illustrated in Fig. 3. More specifically, there are four scenarios depending on the particular location of the individual in the objective space.

- ① *Located in S_1 .* If an individual $\hat{r}_i^{(g,G)}$, $i \in \{1, 2, \dots, N_O\}$ is located in the objective function space area S_1 and it is not ϵ -dominated by any individual of $\mathbf{A}^{(g)}$, it will be stored into the elite population $\mathbf{A}^{(g)}$, and the individuals in $\mathbf{A}^{(g)}$ that are ϵ -dominated by $\hat{r}_i^{(g,G)}$ will be removed from the elite population.
 - ② *Located in S_2 .* If an individual $\hat{r}_i^{(g,G)}$, $i \in \{1, 2, \dots, N_O\}$ is located in the objective function space area S_2 , it will not be stored into the elite population $\mathbf{A}^{(g)}$, since it is ϵ -dominated by all the individuals in $\mathbf{A}^{(g)}$.
 - ③ *Located in S_3 .* If an individual $\hat{r}_i^{(g,G)}$, $i \in \{1, 2, \dots, N_O\}$ is located in the objective function space area S_3 , the ϵ -DMOGA calculates and compares the objective functions of the individuals in $\tilde{\mathbf{P}}^{(g)} = \mathbf{A}^{(g)} \cup \hat{r}_i^{(g,G)}$. Then, the ϵ -Pareto-front set $\tilde{\mathbf{R}}^{(\tilde{\mathbf{P}})}$ is selected and the elite population $\mathbf{A}^{(g)}$ is updated as $\tilde{\mathbf{R}}^{(\tilde{\mathbf{P}})}$. Additionally, both the Pareto-front limits \mathcal{J}_i^{min} and \mathcal{J}_i^{max} as well as the box width ϵ_i are updated for all the three dimensions $i = 1, 2, 3$ according to (23), (22) and (21).
 - ④ *Located in S_4 .* If an individual $\hat{r}_i^{(g,G)}$, $i \in \{1, 2, \dots, N_O\}$ is located in the objective function space area S_4 , all the individuals in the elite population $\mathbf{A}^{(g)}$ are deleted, since all of them are ϵ -dominated by $\hat{r}_1^{(g,G)}$, and $\hat{r}_1^{(g,G)}$ is stored into $\mathbf{A}^{(g)}$. The limit of each objective function \mathcal{J}_i^{min} , $i = 1, 2, 3$, is updated as $\mathcal{J}_i(\hat{r}_1^{(g,G)})$, $i = 1, 2, 3$.
- 5) **Update.** Update the main population $\mathbf{P}^{(g)}$ by comparing its individuals and the individuals selected from the auxiliary population $\mathbf{G}^{(g)}$. Explicitly, an individual $\hat{r}_i^{(g,G)}$, $i = 1, 2, \dots, N_G$, is compared to an individual $r_j^{(g,P)}$ that is randomly selected from $\mathbf{P}^{(g)}$: if $\hat{r}_i^{(g,G)}$ dominates $r_j^{(g,P)}$ as defined by (17), $r_j^{(g,P)}$ is replaced by $\hat{r}_i^{(g,G)}$ in the main population $\mathbf{P}^{(g)}$. The updating operations are continued until all the individuals in the auxiliary population $\mathbf{G}^{(g)}$ are compared to an individual selected from the main population $\mathbf{P}^{(g)}$.
- 6) **Termination.** The ultimate stopping criterion would be that the Pareto-front solutions of the multiple-objective routing optimization problem have been found. However, we cannot offer any proof of evidence that the Pareto-optimal routing paths have indeed been found. In order to have limited and predicible computational complexity, we opt for halting the optimization procedure when the pre-defined maximum affordable number of

generations g_{\max} has been exhausted, namely, $g = g_{\max}$, and the individuals from $\mathbf{A}^{(g_{\max})}$ comprise the near-Pareto solutions. Otherwise, we set $g = g + 1$ and go to 2) **Archive**.

As a population-based nature-inspired multiple-objective optimization algorithm, the computational complexity of ϵ -DMOGA is bounded by the number of generations and the population size, with some additional complexities imposed by the crossover and mutation as well as selection operations. Hence, the computational complexity can be roughly quantified by the number of cost function (CF) evaluations, which is given by $(P_s + N_O)g_{\max}$ CF-evaluation.

B. Convergence of ϵ -DMOGA

As a nature-inspired multiple-objective optimization algorithm, there is randomness in the search procedures of ϵ -DMOGA, hence it is quite challenge to definitely say whether a Pareto-optimal solution has been achieved. Nevertheless, ϵ -DMOGA tries to ensure that the elite population archive $\mathbf{A}^{(g)}$ converge toward an ϵ -Pareto set $\tilde{\mathbf{R}}^{(\tilde{\mathbf{P}})}$ in a smart distributed manner along the Pareto front. The convergence of ϵ -DMOGA can be studied in a similar manner to [48] by the probability of convergence, which is defined as

$$\lim_{g \rightarrow \infty} Pr \left(d \left(\mathbf{A}^{(g)}, \mathbf{R} \right) \rightarrow 0 \right) = 1, \quad (26)$$

where $d(\mathbf{A}^{(g)}, \mathbf{R})$ is a distance function between the g -th generation's elite population archive and the Pareto optimal \mathbf{R} . Additionally, the convergence of ϵ -DMOGA may be also studied in a manner similar to *Theorem 1: Almost sure convergence* in [48]. Hence motivated readers are referred to [48] for a detailed study of convergence in multiple-objective evolutionary algorithms.

IV. SIMULATION RESULTS

In this section, we investigate the achievable network performance of our proposed ϵ -DMOGA based multi-objective routing optimization scheme.

A. Simulated AANET

The mobility characteristics of the nodes are critical for designing and analysing AANETs. In [49], we developed a semi-stochastic aircraft mobility model, which is capable of generating an arbitrary number of flights. However, "it would be ideal to use actual node position", as stated by Kingsbury [50]. from Massachusetts Institute of Technology. Hence, in contrast to relying on a mobility model, which generates artificial flights and their trajectories for approximating aircraft movement, we simulate a realistic AANET in the Australian airspace based on a large-scale real historical flight data on both the busiest and quietest day of 2018. Specifically, June

$$\begin{cases} \hat{r}_1^{(g,G)} &= \left\{ r_1^{(g,P)}, \dots, \hat{r}_{l_1}^{(g,P)}, \dots, \hat{r}_{l_2}^{(g,P)}, \dots, \hat{r}_{l_{N_m}}^{(g,P)}, \dots, r_N^{(g,P)} \right\}, \\ \hat{r}_2^{(g,G)} &= \left\{ r_1^{(g,A)}, \dots, \hat{r}_{l_1}^{(g,A)}, \dots, \hat{r}_{l_2}^{(g,A)}, \dots, \hat{r}_{l_{N_m}}^{(g,A)}, \dots, r_N^{(g,A)} \right\}, \end{cases} \quad (25)$$

TABLE I
DISTANCE-BASED ADAPTIVE CODING AND MODULATION SCHEME FOR AERONAUTICAL COMMUNICATIONS.

Mode k	Mode color	Modulation	Code rate	Spectral efficiency (bps/Hz)	Switching threshold d_k (km)
0	None	None	None	< 0.459	> 740.8
1	Black	BPSK	0.488	0.459	500
2	Magenta	QPSK	0.533	1.000	350
3	Green	QPSK	0.706	1.322	200
4	Yellow	8-QAM	0.642	1.809	110
5	Blue	8-QAM	0.780	2.194	40
6	Cyan	16-QAM	0.731	2.747	25
7	Red	16-QAM	0.853	3.197	5.56

29, 2018, which represents the busiest day, and December 25, 2018, which represents the quietest day. The busiest/quietest day is determined by the number of flights in air on the day, which indicates the traffic in airspace. We assume that there are $N_G = 5$ GSs, namely these at Perth airport (PER), Melbourne airport (MEL), Sydney airport (SYD), Brisbane airport (BNE), and Darwin International airport (DRW). The selection of these five representative airports has jointly considered the geographical distribution and flight handling capacity. The flights considered for our investigation are real historical flights of the top-5 domestic airlines scheduled on June 29, 2018 and December 25, 2018. The top-5 domestic airlines in Australia were Qantas, Jetstar, Tigerair, Virgin Australia and Rex (Regional Express) in 2018.

The AANET employs the time division duplexing (TDD) protocol, which has already been standardized by existing aeronautical communication systems, such as the Automatic Dependent Surveillance-Broadcast (ADS-B) [51], L-band digital aeronautical communications system (L-DACS) [52] and the aeronautical mobile airport communication system (AeroMACS) [53]. Following the physical layer design in [31], orthogonal frequency-division multiplexing (OFDM) is adopted as the transmission technique of broadband aeronautical communications. Each aircraft has 32 transmit antennas and 4 receiver antennas. The network is allocated $B_{\text{total}} = 6$ MHz bandwidth at the carrier frequency of 5 GHz. This bandwidth is divided into 512 subcarriers. The number of cyclic prefix (CP) samples is $N_{\text{cp}} = 32$. The transmit power per antenna is $P_t = 1$ Watt. Furthermore, the distance-based ACM scheme of [31], [32] designed for aeronautical communications is employed for quantifying the link quality between a pair of communicating aircraft, in which the transmit aircraft activates a specific ACM mode based on its distance from the receiver aircraft. Explicitly, our distance-based ACM scheme using $K = 8$ modes is given in Table I, where an ACM mode is represented by a color. If the distance of two aircraft is longer than 740.8 km, there exists no adequate communication link between these two aircraft, which is marked as ‘None’ in Table I. The default parameters of the AANET used in our simulations are summarised in Table II.

B. A specific example of the flight TT589

First, we investigate the achievable multiple objectives of the network layer performance via the specific example of the flight TT589 on June 29, 2018, which may be extrapolated to other flights and other dates. As shown in Fig. 4(a), there is no available direct link to any GS at the five airports considered. However, if we increase the number of the affordable relay nodes and tolerate a higher delay, there are available routing paths to the GS at Brisbane airport. As shown in Fig. 4(b) and Fig. 4(c), there are two Pareto-optimal 2-hop routing paths to the GS deployed at Brisbane airport relying on the relay nodes JQ935 and QF2366, respectively. Both routing paths have a spectrum efficiency (SE) of 0.459 bps/Hz according to the SE of the ‘black’ link of Table I. The multi-objective functions of these two solutions are compared in the left half of Table III. It is clear that the two routing paths shown in Fig. 4(b) and Fig. 4(c) do not dominate each other, and hence both are Pareto-optimal routing paths, provided that the affordable number of relay nodes is one.

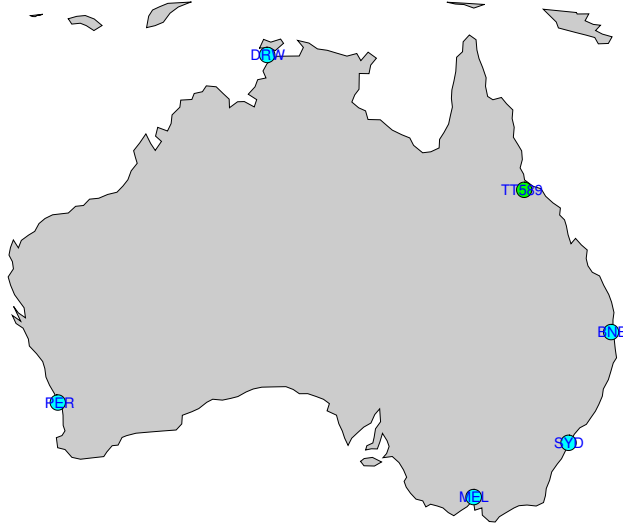
If the affordable number of relay nodes is two, we can find routing paths having a higher SE. As shown in Fig. 4(d) and Fig. 4(e), there are two Pareto-optimal routing paths to the GS at Brisbane airport via three hops, namely, the routing path TT589→QF2407→QF974→BNE and the routing path TT589→QF2407→QF935→BNE, respectively. Both the routing paths have a SE of 1.000 bps/Hz according to the SE of the ‘magenta’ link, which determines the end-to-end SE. As confirmed in the right half of Table III, these two 3-hop routing paths are also Pareto-optimal, as they do not dominate each other.

C. Overall multi-objective network-layer performance

We now investigate the overall network-layer performance, including how many flights in the air over a period of 24 hours can access the Internet through our ϵ -DMOGA based multi-objective routing as well as the average end-to-end SE, the average end-to-end latency and the routing paths’ average PET, on June 29, 2018 and December 25, 2018, respectively.

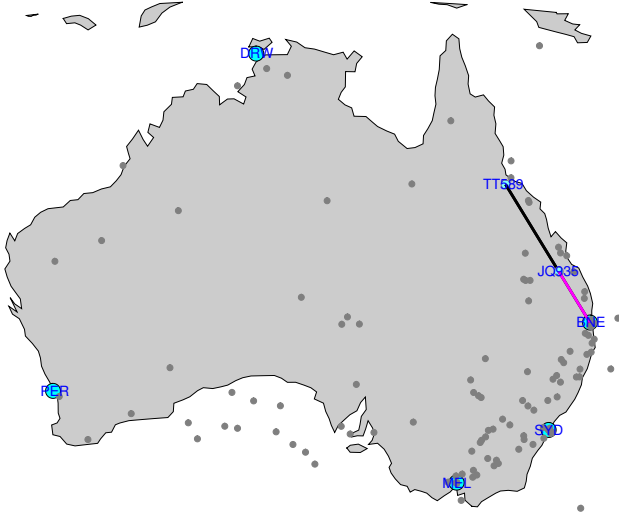
1) *The number of flights that can access the Internet:* Clearly, the number of flights in air changes over the 24 hours of the day, as shown in Fig. 5(a) and Fig. 5(b) for June 29, 2018 and December 25, 2018, respectively. The peak number

1 hop routing with SE = 0.000 bps/Hz and delay = 100000.00000361 s,
 path expiration time = 0.00000 s



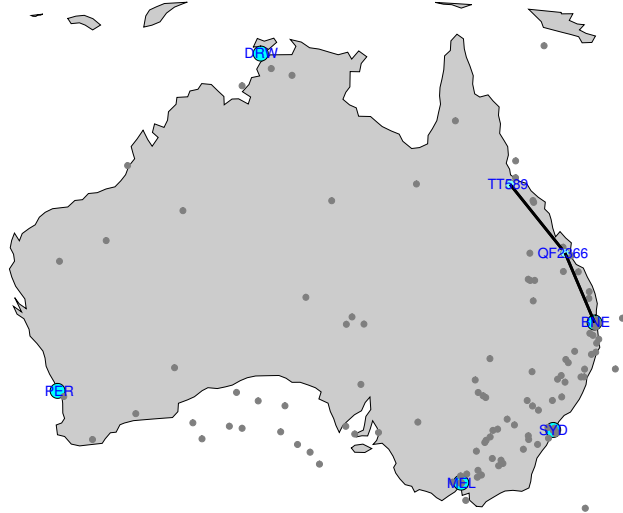
(a) No 1 hop to a GS

2 hops routing with SE = 0.459 bps/Hz and Delay = 0.01500361 s,
 path expiration time = 1902.86767 s



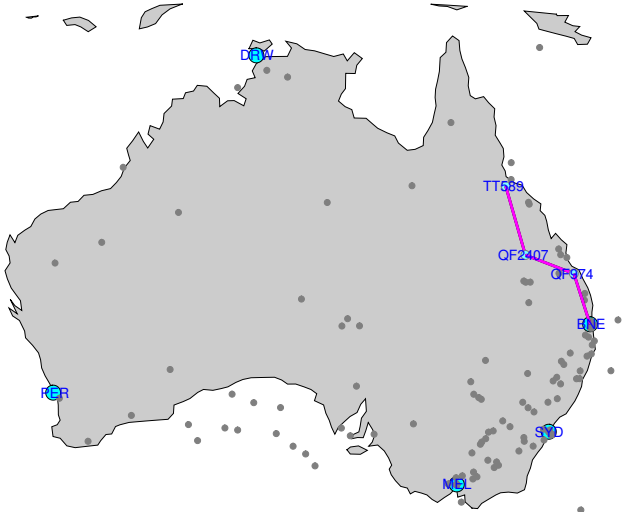
(b) 2 hops to a GS – Solution-1

2 hops routing with SE = 0.459 bps/Hz and Delay = 0.01500365 s,
 path expiration time = 2586.56434 s



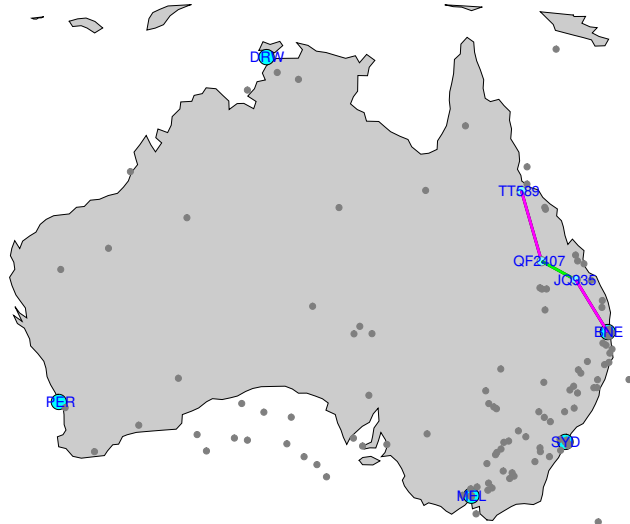
(c) 2 hops to a GS – Solution-2

3 hops routing with SE = 1.000 bps/Hz and Delay = 0.03000395 s,
 path expiration time = 1900.42115 s



(d) 3 hops to a GS – Solution-1

3 hops routing with SE = 1.000 bps/Hz and Delay = 0.03000379 s,
 path expiration time = 1132.96596 s



(e) 3 hops to a GS – Solution-2

Fig. 4. A specific example of the flight-TT589's routing path to a GS, all of which are Pareto-optimal compared to any of other routing paths.

TABLE II
PARAMETERS USED IN SIMULATING AANET

AANET environment	Airspace	Australian airspace
	Airlines considered	Top-5 domestic airlines of Qantas, Jetstar, Tigerair, Virgin Australia and Rex
	Location of GSs	PER, MEL, SYD, BNE, and DRW
	Representative dates investigated	December 25th, 2018 June 29th, 2018
	Time period observed	00:00 ~ 24:00
	Total number of flights on December 25th, 2018	802
	Total number of flights on June 29th, 2018	1007
	Latitude	Determined by each aircraft
	Longitude	Determined by each aircraft
	Altitude	Determined by each aircraft
Communication parameters	Carrier frequency	5 GHz
	Bandwidth B_{total}	6 MHz
	Number of CPs N_{cp}	32
	Number of subcarrier N_{cc}	512
	Rice factor K_{Rice}	5 dB
	ACM	As detailed in Table I
	Maximum A2A communication distance	740.8 km
	Maximum A2G communication distance	370.4 km

TABLE III
COMPARING THE MULTI-OBJECTIVES OF THE PARETO-OPTIMAL ROUTING PATHS FOR FLIGHT TT589 ON JUNE 29, 2018.

	2-hop solution 1	2-hop solution 2	3-hop solution 1	3-hop solution 2
\mathcal{J}_1 SE (bps/Hz)	0.459	0.459	1.000	1.000
\mathcal{J}_2 delay (s)	0.01500361	0.01500365	0.03000395	0.03000379
\mathcal{J}_3 PET (s)	1902.86767	2586.56434	1900.42115	1132.96596

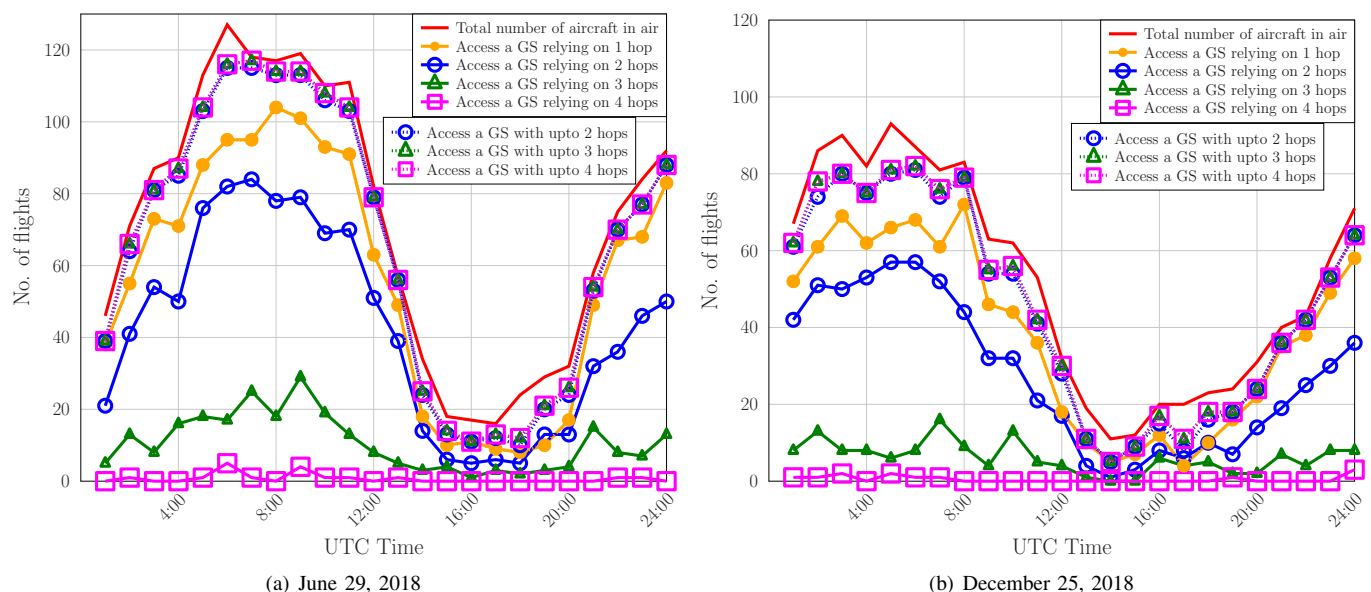


Fig. 5. The number of flights in air over 24 hours capable of accessing the Internet using the ϵ -DMOGA multi-objective routing optimization.

TABLE IV
COMPARING END-TO-END SPECTRUM EFFICIENCY OF ONE-HOP PARETO-OPTIMAL ROUTING SOLUTIONS AND PARETO-OPTIMAL ROUTING SOLUTIONS WITH UP TO TWO HOPS AT UTC TIME 18:00 ON JUNE 29, 2018.

	Individual flights' SE (bps/Hz)								Average SE (bps/Hz)
8 flights with 1 hop	1.8090	1.3220	2.1940	2.1940	1.8090	1.0000	1.0000	1.3220	1.58125
10 flights with up to 2 hops	1.8090	1.3220	2.1940	2.1940	2.1940	1.3220	1.3220	1.3220	1.52282
	0.7295	0.8197							

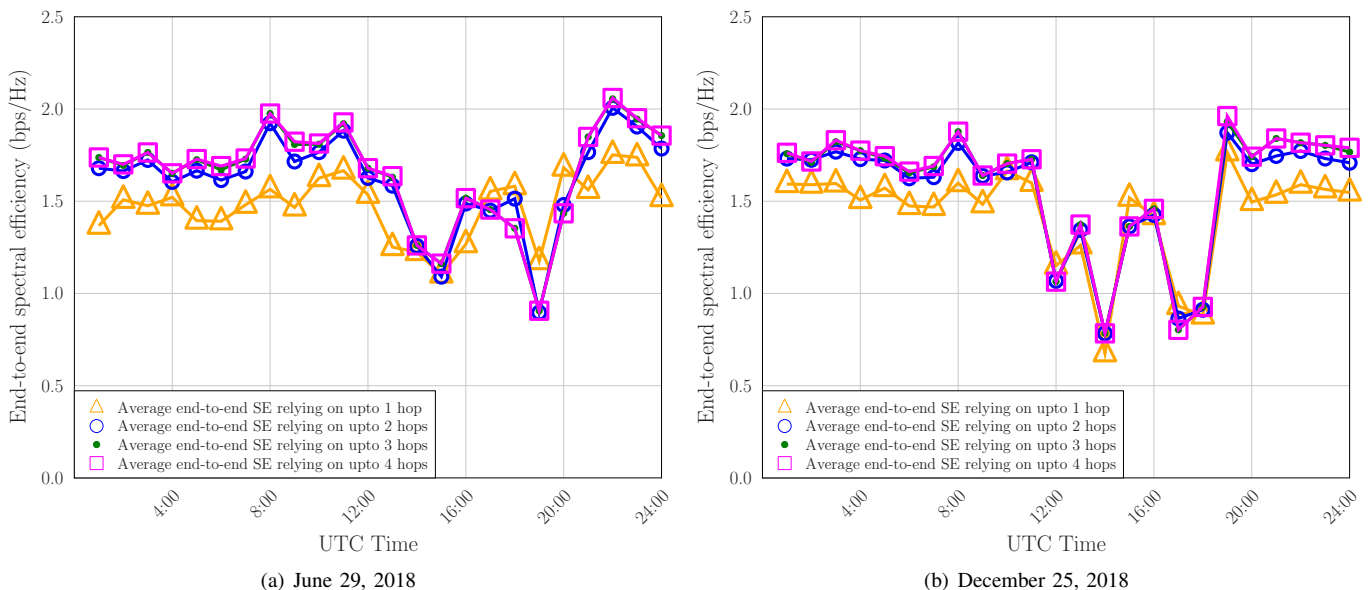


Fig. 6. The average end-to-end spectral efficiency achieved by the ϵ -DMOGA based multi-objective routing optimization.

of flights in air occurs at UTC time 06:00, namely 127 flights, whilst the lowest number of the flights in air occurs at UTC time 17:00, namely 16 flights. We start by investigating the number of flights that can access a GS at any one of the five airports considered. Explicitly, in Fig. 5, we investigate how many flights can access a GS relying on 1 hop, 2 hops, 3 hops and 4 hops, respectively, using our ϵ -DMOGA based multi-objective routing optimization scheme. Additionally, we also provides the numbers of flights that can access a GS with up to 2 hops, up to 3 hops and up to 4 hops, respectively, obtained by the ϵ -DMOGA based multi-objective routing optimization.

As shown by the solid lines of Fig. 5(a) and Fig. 5(b), most of the flights in air can access a GS relying on 1 hop, i.e. via a direct link. Furthermore, by analyzing the results obtained using the ϵ -DMOGA scheme, it can be found that there is a significant number of flights which can achieve higher end-to-end SE than the single-hop paths by relying on 2 hops. By contrast, seldom flights can achieve higher end-to-end SE relying on 3 hops and 4 hops than those relying on 2 hops. Although it is not explicitly shown by the figures, we found that almost all flights can access a GS with up to 2 hops. Hence, increasing the affordable number of hops to 3 and 4 contributes little to the total number of flights that can access a GS, regardless of the hour of the day.

2) *Average end-to-end SE*: Since the routes of the same number of hops for different flights have different end-to-end

SEs, Fig. 6(a) and Fig. 6(b) portray the achievable average end-to-end SE over 24 hours on June 29, 2018 and December 25, 2018, respectively. As shown in Fig. 6(a) and Fig. 6(b), the routing paths relying on up to 2, 3 and 4 hops are capable of achieving somewhat higher average end-to-end SE than those relying on 1 hop most times of the day except for UTC times 17:00, 18:00, 19:00 and 20:00 on June 29 as well as UTC time at 15:00 on December 25, mainly in the scenarios of having low flight times. But naturally, this SE improvement is achieved at the cost of a higher delay. To elaborate a little further, it is unexpected to encounter a lower SE for 2 hops than for 1 hop, because a single hop tends to be longer, which results in a lower SE. The issue here is that there are more Pareto-optimal routing paths of up to 2 hops than the number of Pareto-optimal 1-hop routes. We illustrate this point using the example at UTC time 18:00 on June 29, 2018. There are 8 flights having direct links to a GS, while there are 10 flights having routing paths to a GS relying on up to 2 hops – some relying on direct link and some relying on 2 hops to access a GS. Table IV compares the individual flights' SEs of these two groups as well as their average SEs. Although the average SE of the second group is lower than that of the first group, observe that eight of the ten Pareto-optimal routing paths with up to 2 hops have the same or higher SEs than those of the eight 1-hop routing paths.

Extrapolating from the results of the busiest day and the

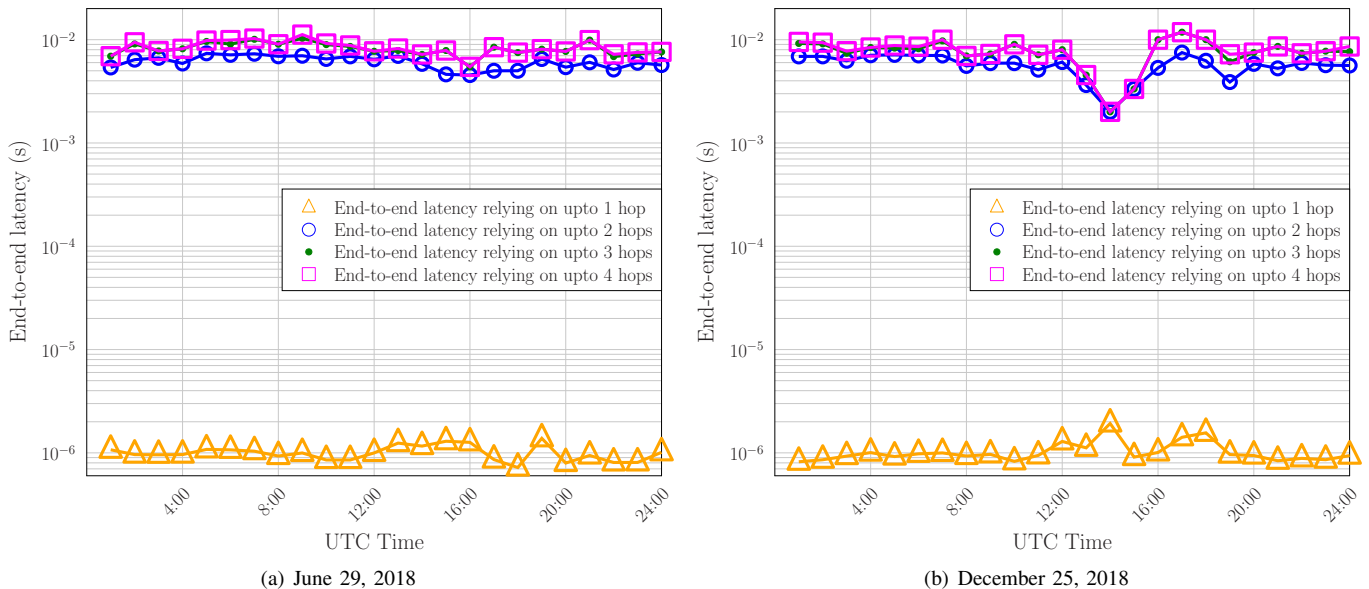


Fig. 7. The average end-to-end latency achieved by the ϵ -DMOGA based multi-objective routing optimization.

quietest day depicted in Fig. 6(a) and Fig. 6(b), we may draw the conclusion that the AANET over the Australian airspace can achieve an average SE of about 1.2 bps/Hz for low-flight-times during the period of UTC 12:00-20:00, while its average SE increases to around 1.7 bps/Hz for high-flight-times.

3) *Average end-to-end latency*: The average end-to-end latency imposed by the Pareto-optimal routing paths over 24 hours on June 29, 2018 and December 25, 2018 are depicted in Fig. 7(a) and Fig. 7(b), respectively. As expected, the direct links have the lowest latency, since they only have single-hop propagation delay upon accessing a GS. As seen in Fig. 7(a) and Fig. 7(b), the average end-to-end latency relying on up to 2 hops is significantly higher than that relying on direct links. This is because the end-to-end latency of a 2-hop routing path also includes the signal processing delay and queuing delay, which are significantly higher than the propagation delay. Furthermore, the average end-to-end latency relying on up to 3 hops and up to 4 hops is higher than that relying on up to 2 hops. But the difference between the average end-to-end latency relying on up to 3 hops and up to 4 hops is small, because the number of routing paths relying on 4 hops is relatively small, as shown in Fig. 5(a) and Fig. 5(b).

As observed from Fig. 7(a) and Fig. 7(b), the variation of the average end-to-end latency over 24 hours of a day is small. Furthermore, by extrapolating from the results of Fig. 7(a) and Fig. 7(b), we may conclude that the average end-to-end latency may be as low as 0.01 s in the AANET over the Australian airspace, provided that a link is available.

4) *Average path expiration time*: The average PETs over 24 hours on June 29, 2018 and December 25, 2018 are portrayed in Fig. 8(a) and Fig. 8(b), respectively. Intuitively, the routing path may become vulnerable to potential breakage upon increasing the number of hops. Our investigations based on real historical flight data both on June 29, 2018 and on December 25, 2018 confirm this intuition. Explicitly, as shown in Fig. 8(a) and Fig. 8(b), the routing paths relying on a direct

link to a GS have considerably longer average PET than those relying on 2, 3 and 4 hops, except around UTC time 14:00 on June 29, 2018. Furthermore, the routing paths relying on 2 hops have longer average PET with a high probability than those relying on 3 and 4 hops, but the difference between the average PETs relying on 3 and 4 hops are hardly noticeable, since the number of routing paths relying on 4 hops is very low.

It can be seen that the average PET varies over the 24 hours of a day. On average, we may extrapolate that the PET is around 800 s i.e. over minutes in the AANET over the Australian airspace.

5) *Summary*: Based on the results obtained using real historical flight data on two representative dates in 2018, namely on the busiest day of June 29 and on the quietest day of December 25, we may extrapolate the achievable network layer performance for the AANET over the Australian airspace using our ϵ -DMOGA based multi-objective routing optimization. The overall network performance is summarized in Table V at a glance.

V. CONCLUSIONS

In order to provide Internet service above the clouds, an ϵ -DMOGA based multiple-objective routing optimization has been developed by taking into account the unique features of routing problem in AANETs. Explicitly, the end-to-end SE, the end-to-end latency and the PET have been jointly optimized by the proposed ϵ -DMOGA for determining the Pareto-optimal routing paths. The achievable end-to-end SE, end-to-end latency and PET performance using our ϵ -DMOGA based multiple-objective routing optimization have been investigated based on the top-5 Australia domestic airlines' real historical flight data on two representative dates in 2018, namely the busiest day of June 29 and the quietest day of December 25, in term of the number of flights in air. Our simulation results have quantified the networking capability of the AANET over

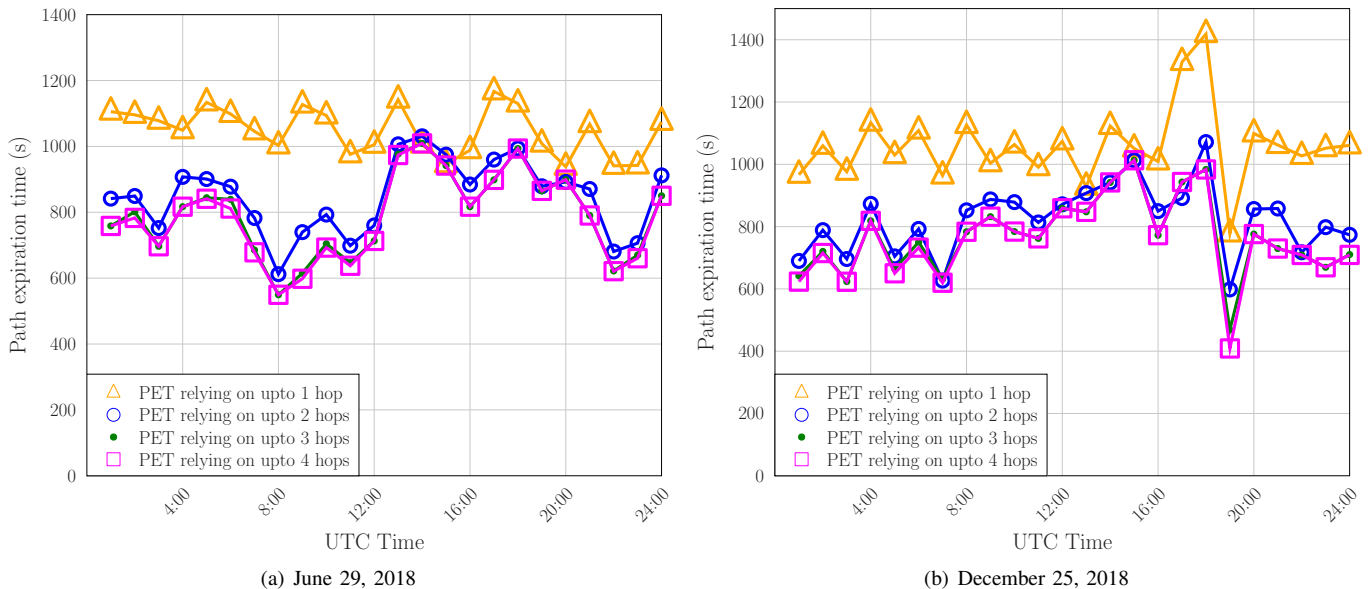


Fig. 8. The average path expiration time achieved by the ϵ -DMOGA based multi-objective routing optimization.

TABLE V
THE ACHIEVABLE NETWORK PERFORMANCE OF THE AANET OVER AUSTRALIA AIRSPACE

Average spectrum efficiency	Average end-to-end latency	Average path expiration time
1.2 bps/Hz at low-flight times	0.01 s	800 s \approx 13.3 min
1.7 bps/Hz at high-flight times		

the Australian airspace. Furthermore, our investigations have also offered useful design considerations for the AANETs in other parts of the worlds.

REFERENCES

- [1] 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," *Proc. IEEE*, vol. 107, no. 5, pp. 868–911, May 2019.
- [2] D. Medina and F. Hoffmann, "The airborne Internet," in S. Plass (ed.), *Future Aeronautical Communications*. InTech, 2011, pp. 349–374.
- [3] Q. Vey, A. Pirovano, J. Radzik, and F. Garcia, "Aeronautical ad hoc network for civil aviation," in *Proc. 6th Int. Workshop Communi. Technologies for Vehicles* (Offenburg, Germany), May 6–7, 2014, pp. 81–93.
- [4] E. Sakhaee, A. Jamalipour, and N. Kato, "Aeronautical ad hoc networks," in *Proc. WCNC 2006* (Las Vegas, NV, USA), Apr. 3–6, 2006, pp. 246–251.
- [5] M. Mauve, J. Widmer, and H. Hartenstein, "A survey on position-based routing in mobile ad hoc networks," *IEEE Network*, vol. 15, no. 6, pp. 30–39, Nov.–Dec. 2001.
- [6] F. Li and Y. Wang, "Routing in vehicular ad hoc networks: A survey," *IEEE Vehicular Technology Mag.*, vol. 2, no. 2, pp. 12–22, Jun. 2007.
- [7] D. S. Lakew, U. Sa'ad, N.-N. Dao, W. Na, and S. Cho, "Routing in flying ad hoc networks: A comprehensive survey," *IEEE Communi. Surveys & Tutorials*, vol. 22, no. 2, pp. 1071–1120, Secondquarter 2020.
- [8] E. Sakhaee and A. Jamalipour, "The global in-flight Internet," *IEEE J. Sel. Areas Communi.*, vol. 24, no. 9, pp. 1748–1757, Sep. 2006.
- [9] Q. Luo, J. Wang, S. Liu, "AeroMRP: A multipath reliable transport protocol for aeronautical ad hoc networks," *IEEE Internet of Things J.*, vol. 6, no. 2, pp. 3399–3410, Apr. 2019.
- [10] M. Iordanakis, D. Yannis, K. Karras, G. Bogdos, G. Dilintas, M. Amirfeiz, G. Colangelo, and S. Baiotti, "ad-hoc routing protocol for aeronautical mobile ad-hoc networks," in *Proc. CSNDSP 2006* (Patras, Greece), Jul. 19–21, 2006, pp. 1–5.
- [11] A. Jabbar and J. P. G. Sterbenz, "AeroRP: A geolocation assisted aeronautical routing protocol for highly dynamic telemetry environments," in *Proc. ITC 2009* (Las Vegas, NV, USA), Oct. 26–29, 2009, pp. 1–9.
- [12] K. Peters, A. Jabbar, E. K. Cetinkaya, and J. P. Sterbenz, "A geographical routing protocol for highly-dynamic aeronautical networks," in *Proc. WCNC 2011* (Cancun, Mexico), Mar. 28–31, 2011, pp. 492–497.
- [13] J. P. Rohrer, E. K. Cetinkaya, H. Narra, D. Broyles, K. Peters, and J. P. Sterbenz, "AeroRP performance in highly-dynamic airborne networks using 3D Gauss-Markov mobility model," in *Proc. MILCOM 2011* (Baltimore, MD, USA), Nov. 7–10, 2011, pp. 834–841.
- [14] D. Medina, F. Hoffmann, F. Rossetto, and C.-H. Rokitansky, "A geographic routing strategy for north Atlantic in-flight Internet access via airborne mesh networking," *IEEE/ACM Trans. Networking*, vol. 20, no. 4, pp. 1231–1244, Aug. 2012.
- [15] W. Gu, J. Li, M. Lv, Q. Sun, and F. Yang, "Delay aware multipath Doppler routing in aeronautical ad hoc networks," in *Proc. 14th Int. Conf. Computational Science and Engineering* (Dalian, China), Aug. 24–26, 2011, pp. 251–255.
- [16] B. Du, X. Di, D. Liu, H. Zhang, "Dynamic graph optimization and performance evaluation for delay-tolerant aeronautical ad hoc network," *IEEE Trans. Communi.*, vol. 69, no. 9, pp. 6018–6036, Sep. 2021.
- [17] J. Cui, D. Liu, J. Zhang, H. Yetgin, S. X. Ng, R. G. Maunder, L. Hanzo, "Minimum-delay routing for integrated aeronautical ad hoc networks relying on real flight data in the north-Atlantic region," *IEEE Open Journal of Vehicular Technology*, vol. 2, pp. 310–320, Jun. 2021.
- [18] D. Liu, J. Cui, J. Zhang, C. Yang, L. Hanzo, "Deep reinforcement learning aided packet-routing for aeronautical ad-hoc networks formed by passenger planes," *IEEE Trans. Vehicular Technology*, vol. 70, no. 5, pp. 5166–5171, May 2021.
- [19] S. Wang, C. Fan, C. Deng, W. Gu, Q. Sun, and F. Yang, "A-GR: A novel geographical routing protocol for AANETs," *J. Systems Architecture*, vol. 59, no. 10, pp. 931–937, Nov. 2013.
- [20] M. S. B. Mahmoud and N. Larriau, "An ADS-B based secure geographical routing protocol for aeronautical ad hoc networks," in *Proc. COMPSACW 2013* (Kyoto, Japan), Jul. 22–26, 2013, pp. 556–562.
- [21] A. Swidan, S. Khattab, Y. Abouelseoud, and H. Elkamchouchi, "A secure geographical routing protocol for highly-dynamic aeronautical networks," in *Proc. MILCOM 2015* (Tampa, FL, USA), Oct. 26–28, 2015, pp. 708–713.
- [22] B. Zheng, H.-Y. Zhang, B.-L. Wang, and W. Zhao, "Load balancing geographic routing strategy for aeronautical ad hoc networks," *J. Communi.*, vol. 37, no. 12, pp. 67–76, Dec. 2016. (in Chinese)

- [23] S.-C. Pang, C.-Y. Luo, and H. Guan, "Secure geographic based routing protocol of aeronautical ad hoc networks," in *Proc. ITITS 2015* (Xi'an China), Dec. 12-13, 2015, pp. 295–306.
- [24] Q. Luo and J. Wang, "Multiple QoS parameters-based routing for civil aeronautical ad hoc networks," *IEEE Internet of Things J.*, vol. 4, no. 3, pp. 804–814, Jun. 2017.
- [25] J. Cui, H. Yetgin, D. Liu, J. Zhang, S. X. Ng, L. Hanzo, "Twin-component near-Pareto routing optimization for AANETs in the north-Atlantic region relying on real flight statistics," *IEEE Open Journal of Vehicular Technology*, vol. 2, pp. 346–364, Jul. 2021.
- [26] J. Cui, S. X. Ng, D. Liu, J. Zhang, A. Nallanathan, L. Hanzo, "Multi-objective optimization for integrated ground-air-space networks: current research and future challenges," *IEEE Vehicular Technology Magazine*, vol. 16, No. 3, pp. 88–98, Jun. 2021.
- [27] D. Liu, J. Zhang, J. Cui, S. X. Ng, R. G. Maunder, L. Hanzo, "Deep Learning Aided Packet Routing in Aeronautical Ad-Hoc Networks Relying on Real Flight Data: From Single-Objective to Near-Pareto Multi-Objective Optimization," *IEEE Internet of Things J.*, Early Access, Aug. 2021.
- [28] <https://www.faa.gov/nextgen/programs/adsb/>
- [29] [https://www.skybrary.aero/index.php/Automatic_Dependent_Surveillance_Broadcast_\(ADS-B\)](https://www.skybrary.aero/index.php/Automatic_Dependent_Surveillance_Broadcast_(ADS-B))
- [30] D. Wang, Y. Wang, S. Dong, G. Huang, J. Liu, and W. Gao, "On delay-aware resource control with statistical QoS provisioning by dual connectivity in heterogeneous aeronautical network," *IEEE Trans. Vehicular Technology*, vol. 69, no. 3, pp. 2915–2927, Mar. 2020.
- [31] J. Zhang, S. Chen, R. G. Maunder, R. Zhang, and L. Hanzo, "Adaptive coding and modulation for large-scale antenna array-based aeronautical communications in the presence of co-channel interference," *IEEE Trans. Wireless Communi.*, vol. 17, no. 2, pp. 1343–1357, Feb. 2018.
- [32] —, "Regularized zero-forcing precoding-aided adaptive coding and modulation for large-scale antenna array-based air-to-air communications," *IEEE J. Sel. Areas Communi.*, vol. 36, no. 9, pp. 2087–2103, Sep. 2018.
- [33] J. Herrero, M. Martínez, J. Sanchis, and X. Blasco, "Well-distributed Pareto front by using the ϵ -MOGA evolutionary algorithm," in: F. Sandoval, A. Prieto, J. Cabestany, and M. Graña (eds.) *Computational and Ambient Intelligence*, LNCS 4507. Springer-Verlag: Berlin Heidelberg, 2007, pp. 292–299.
- [34] M. Martínez, J. M. Herrero, J. Sanchis, X. Blasco and S. Garcia-Nieto, Applied Pareto multi-objective optimization by stochastic solvers. *Engineering Applications of Artificial Intelligence*. Vol. 22 pp. 455 - 465, 2009.
- [35] S. Batabyal and P. Bhaumik, "Mobility models, traces and impact of mobility on opportunistic routing algorithms: A survey," *IEEE Communi. Surveys & Tutorials*, vol. 17, no. 3, pp. 1679–1707, Thirdquarter 2015.
- [36] D. McNell, Networked Audio Transport: Looking at the methods and factors Accessed on Dec. 22, 2021, at the ProSoundWeb.com.
- [37] D. Das and A. A. Abouzeid, "Spatial-temporal queuing theoretic modeling of opportunistic multihop wireless networks with and without cooperation," *IEEE Trans. Wireless Communi.*, vol. 14, no. 9, pp. 5209–5224, Sep. 2015.
- [38] M. Baz, P. D. Mitchell, and D. A. J. Pearce, "Analysis of queuing delay and medium access distribution over wireless multihop PANs," *IEEE Trans. Vehicular Technology*, vol. 64, no. 7, pp. 2972–2990, Jul. 2014.
- [39] A. Al-Halafi, A. Alghadhban, and B. Shihada, "Queuing delay model for video transmission over multi-channel underwater wireless optical networks," *IEEE Access*, vol. 7, pp. 10515–10522, Jan. 2019.
- [40] H. Isermann, Linear lexicographic optimization, *Operations-Research-Spektrum*, Vol. 4, pp. 223-228, 1982.
- [41] L. Zadeh, Optimality and Non-Scalar-Valued Performance Criteria, *IEEE Transactions on Automatic Control*, vol. 8, no. 1, pp. 59–60, 1963.
- [42] E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach, *IEEE Transactions on Evolutionary Computation*, Vol. 3, No. 4, pp. 257-271, 1999.
- [43] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 2, 2002.
- [44] E. Zitzler, M. Laumanns, and L. Thiele, SPEA2: improving the strength pareto evolutionary algorithm for multiobjective optimization, in *Evolutionary Methods for Design, Optimisation and Control*, vol. 3242, pp. 95–100, CIMNE, Barcelona, Spain, 2002.
- [45] D. W. Corne, J. D. Knowles, M. J. Oates, The Pareto envelope-based selection algorithm for multiobjective optimization, *Parallel Problem Solving from Nature PPSN VI*. Springer Berlin Heidelberg, 2000.
- [46] D. W. Corne, et al. PESA-II: Region-based selection in evolutionary multiobjective optimization, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2001)*. 2001.
- [47] G. Reynoso Meza, *Controller Tuning by Means of Evolutionary Multiobjective Optimization: A Holistic Multiobjective Optimization Design Procedure*, Ph.D. dissertation, Editorial Universitat Politècnica de València, May 2014.
- [48] T. Hanne, On the convergence of multiobjective evolutionary algorithms, *European Journal of Operational Research*, no. 117, pp. 553-564, 1999.
- [49] J. Zhang, et al, Semi-stochastic aircraft mobility modelling for aeronautical networks: An Australian case-study based on real flight data, *IEEE Transactions on Vehicular Technology*, Vol. 70, No. 10, pp. 10763 - 10779, Oct. 2021.
- [50] R. W. Kingsbury, "Mobile ad hoc networks for oceanic aircraft communications," M.S. thesis, Dept. Aeronaut. Astronaut. Massachusetts Inst. Technol., 2009.
- [51] Minimum Aviation System Performance Standards for Automatic Dependent Surveillance-Broadcast (ADS-B), document DO-242A, RTCA, Jun. 2002.
- [52] M. Schnell, U. Epple, D. Shutin, and N. Schneckenburger, "LDACS: Future aeronautical communications for air-traffic management," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 104–110, May 2014.
- [53] J. M. Budinger and E. Hall, Aeronautical Mobile Airport Communications System (AeroMACS), document NASA/TM-2011-217236, NASA Glenn Research Center, Oct. 2011.