IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 21, 2018, accepted June 26, 2018, date of publication July 3, 2018, date of current version July 25, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2852707

A Scheme for Improving the Communications Efficiency Between the Control Plane and Data Plane of the SDN-Enabled Airborne Tactical Network

KEFAN CHEN, NA LV, SHANGHONG ZHAO, XIANG WANG, AND JING ZHAO
School of Information and Navigation, Air Force Engineering University, Xi'an 710077, China

Corresponding author: Shanghong Zhao (zhaoshangh@aliyun.com)

This work was supported by the National Natural Science Foundation of China under Grant 91638101 and Grant 61472443.

ABSTRACT With the battlefield becoming more challenging, the aviation swarm is considered a promising organization of air combat forces to execute both combat and non-combat air missions. As a critical component of the aviation swarm, the airborne tactical network (ATN) provides the communications capability, enabling information sharing between both manned and unmanned military aircrafts as well as surface and ground platforms. The deficiencies of legacy ATNs, particularly their inability to meet the communications demands of the aviation swarm, motivated us to employ the software-defined networking (SDN) paradigm to the ATN and to design an SDN-enabled ATN (SD-ATN) for the aviation swarm. For the SDN paradigm, the communications efficiency between the control plane and the data plane directly impacts the normal operations of the network, raising a fundamental problem of designing the SD-ATN, which is that how to ensure the QoS requirements of the non-elastic information transmitted between the control plane and the data plane (non-elastic C/D information) is met. To address this issue, a transmission framework is first proposed to make it practical to provide dedicated QoS guarantee for the non-elastic C/D information. Then, a communication protocol is designed based on the proposed framework, aiming at transmitting the non-elastic C/D information in a reliable and real-time manner. The simulation results show that, in the aviation environment, our scheme does improve the transmission efficiency of the non-elastic C/D information, which performs better than legacy solutions.

INDEX TERMS Aviation swarm, airborne tactical network, software-defined networking, transmission framework, communication protocol.

I. INTRODUCTION

With the development of information technology, network-centric warfare is a no-way-back trend in modern military operations. To enhance the adaptability and system-of-system (SoS) combat capability of the air warfare forces, based on the airborne tactical network (ATN), exploiting a swarm of manned or unmanned aircraft with various warfare capabilities to cooperatively execute air missions has recently become an important research area in the military aviation domain. We term the new organization of air warfare forces as the *aviation swarm*.

An gaviation swarm is composed of a certain number of aircraft (typically dozens of aircraft), including aviation platforms equipped with the airborne early warning and control system (AP-AEW&C), fighter jets, reconnaissance aircraft, etc. Obviously, the superiority of the aviation swarm relies on efficient collaborations among swarm members, requiring deep and extensive interactions between mission systems, weapon systems and sensors carried on different aircraft. For instance, when some aircraft in an aviation swarm use their radar in a networking manner to perform collaborative detection, they need the ATN to adaptively provide a data rate from 100's of Kbps to 10's of Mbps because of different detection accuracy requirements. Therefore, compared to legacy ATNs, an ATN with a more efficient, more appropriate and more adaptable communications capability is needed by the aviation swarm. Additionally, legacy ATNs are embraced tightly coupled and vertically integrated architecture [1]. To improve the communications capability of legacy ATNs, network engineers have to patch the network in a slow and costly way, which makes the network cumbersome, rigid and difficult to manage.

According to the above issues, legacy ATNs are ill-suited for providing communications for the aviation swarm, motivating us to design a new ATN paradigm. The emergence of the software-defined networking (SDN) provides ideas. SDN is a promising paradigm that distinctly decouples the control plane and data plane of the network. The data plane is only in charge of forwarding data flows, while in the control plane, a logically centralized network controller dictates the overall network behaviors [2], [3]. The SDN paradigm simplifies networking operations, optimizes network management and introduces innovation and flexibility compared to the legacy networking architectures. Given the advantages of the SDN, we intend to design an ATN that leverages the SDN's advantages for the aviation swarm [4]. In this article, we term the SDN-enabled ATN as the *SD-ATN*.

Since the control plane and data plane of the SDN paradigm are deployed independently, the communications efficiency between the two planes seriously impacts the benefits provided by the new network paradigm. In a SDN-enabled network, many types of emergency information must be exchanged between the control plane and data plane, such as transmission demand information, error reporting information, and network configuration information. These types of information are critical for efficient and correct network operations, and the generation of which exhibits randomness and burstiness characteristics. Accordingly, transmitting these types of information costs low bandwidth but requires a QoS guarantee in reliability and latency to allow the network to respond accurately and quickly. We term the information transmitted between the control plane and data plane, which requires a low-bandwidth and must be transmitted in a reliable and timely manner, as the *non-elastic C/D information*.

For the terrestrial networks in which the SDN paradigm is extensively discussed, the control plane and data plane can efficiently communicate with each other via a stable and high-bandwidth wired network. The communications between the two planes have mostly been incidentally investigated in the studies such as southbound interface (SBI) design, scalability improvement of the control plane and network measurement [5]–[10]. Therefore, improving the transmission efficiency of the non-elastic C/D information is currently receiving little attention. However, unlike the terrestrial networks, the SD-ATN is a network with high-speed nodes, intermittent wireless links, variable network topology and limited bandwidth, which makes it challenging for the SD-ATN's control plane to efficiently communicate with the data plane. In such a context, determining how to improve the transmission efficiency of the non-elastic C/D information in the aviation environment becomes a fundamental problem of designing the SD-ATN, which has not yet been investigated.

Considering the gap in designing the SD-ATN, we propose a communication scheme to address this issue. First, we propose a transmission framework for the SD-ATN with the aim of providing a fundamental infrastructure to improve the communications capability between the control plane and data plane. Then, based on the transmission framework proposed, we design a dedicated communication protocol, aiming at transmitting the non-elastic C/D information in both a reliable and timely manner.

The rest of this paper is structured as follows. Section II presents the background, models and assumptions related to our work. Section III describes the proposed transmission framework. Section IV presents the details of our communication protocol. Extensive experimental results are presented in Section V, and Section VI concludes this article.

II. PRELIMINARIES

The architecture of the SD-ATN is shown in fig. 1, which contains the basic planes of the SDN paradigm. The network applications that formulate operation policies for the SD-ATN constitute the application plane of the SD-ATN. The control plane of the SD-ATN includes three control hierarchies: swarm control hierarchy, platform control hierarchy and device control hierarchy. The data plane of the SD-ATN is composed of the SD-ATN transmission systems of different aircraft.

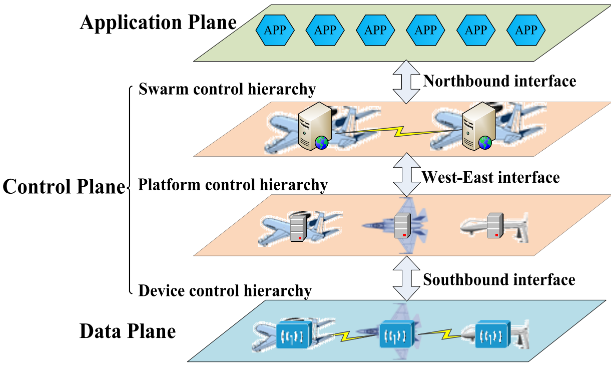


FIGURE 1. Overview of the SD-ATN architecture.

The swarm control hierarchy centrally controls the overall SD-ATN, and the controller serves this control hierarchy is called the *SD-ATN controller*. The network applications all work on the SD-ATN controller. The platform control hierarchy is responsible for controlling the SDN-enabled avionics network (employing the SDN paradigm in the avionics network is a promising way to improve the efficiency of the avionics network [11], [12]), and it also works as a proxy to configure the heterogeneous SD-ATN transmission system according to the policies given from the swarm control hierarchy, which can simplify the network control operations and protocol design. The controller serves this control hierarchy is called the *platform controller*. It is not always efficient and robust for the SD-ATN to rely on the logical centralized controller to manage the network. Therefore, the device control hierarchy embedded in the data plane is also defined in the SD-ATN, providing the SD-ATN transmission system with local network control logic to enable the SD-ATN to work in a distributed manner.

Each aircraft carries a platform controller. To improve the network robustness and facilitate the communications between the SD-ATN controller and the mission system (the latter purpose is helpful in formulating appropriate mission plans and network control strategies [13]), the SD-ATN controller is only carried on the AP-AEW&C, such as the E-3 and the KJ-2000. Each AP-AEW&C in an aviation swarm carries an SD-ATN controller (multiple SD-ATN controllers can work in cluster mode), but only one AP-AEW&C works as the active network administrator. The remaining AP-AEW&Cs work as the standby network administrators in case the active network administrator is down. The details of the election and handover mechanisms that enable a new network administrator to work are not the focus of this article; thus, these mechanisms are not discussed. For ease of description, the AP-AEW&C working as the active network administrator is called the *active control node*, and the remaining aircraft are called the *common nodes*.

Our research work focuses on the transmission behaviors between the active control node and the common nodes, since it mainly impacts the transmission efficiency of the non-elastic C/D information. Only the omnidirectional and half-duplex transmission mode is considered. When the SD-ATN is modeled as a topology graph, the protocol interference model is used to determine the transmission interference of the omnidirectional communications, in which concurrent transmission on two edges interfere with each other if and only if: 1) the two edges are adjacent; 2) a receiver is within the interference range of a non-intended transmitter transmitting at the same slot and frequency with the receiver's receiving duration.

III. FRAMEWORK FOR TRANSMITTING THE NON-ELASTIC C/D INFORMATION

As discussed above, the non-elastic C/D information is critical to the efficient and correct network operations. To provide a dedicated QoS guarantee for the non-elastic C/D information, in this section, a transmission framework is presented.

In general, the communication system of a military aircraft includes heterogeneous radios with different communication capabilities. Each radio is allocated an independent but limited bandwidth. Note that the communication system is mainly designed to transmit the tactical information; with the limited bandwidth, the non-elastic C/D information is easily overwhelmed if it is transmitted together with the tactical information, which severely undermines the efficiency of the network. Therefore, it is necessary to separate the transmission of the non-elastic C/D information out.

Similar to the system framework presented in [14], a SD-ATN transmission system is composed of a common SD-ATN router and different radios. The SD-ATN router communicates with the radios through a common radio-to-router interface (R2RI), which improves the interoperability by separating the radio capability (one RF hop) from the router functionality (multihop) and is helpful for routing computing. The difference of the SD-ATN transmission system is that its transmission behaviors can be managed by the SD-ATN controller through the platform controller and the SBI; the implementation details are not discussed in this article. Because the RF-integrated technology and the software-defined radio technology allow for flexible utilization of hardware resources, to separate the transmission, a new transmission path that only focuses on the transmission of the non-elastic C/D information is added to the SD-ATN transmission system, enabling the out-of-band interactions. The radio of the non-elastic C/D information transmission path works within an independent frequency band. In the SD-ATN router, we add two types of dedicated queues for the new added transmission path, one type is used for caching the information that configures the non-elastic C/D information transmission path, and another type is used for caching the non-elastic C/D information. A dedicated routing table that stores the rules of forwarding the non-elastic C/D information is also constructed in the SD-ATN router.

In particular, to cope with the abnormalities of the new added transmission path, we also define a backup transmission path that is implemented by the off-the-shelf transmission capabilities of the SD-ATN transmission system and sends the non-elastic C/D information together with the tactical flows. For instance, the transmission path implementing the tactical targeting network technology (TTNT) can be employed as the backup transmission path, with the help of which, the non-elastic C/D information can still be transmitted normally even if the non-elastic C/D information transmission path does not work.

By adding the dedicated transmission path in the SD-ATN transmission system, it is available to design a dedicated transmission solution for the non-elastic C/D information, which provides special QoS guarantee and prevents the non-elastic C/D information from being overwhelmed by the tactical flows. Meanwhile, the new added transmission path saves the bandwidth by making it easy to allocate channel resources according to the transmission demands of the non-elastic C/D information, while the needed bandwidth is difficult to learn if the non-elastic C/D information is transmitted together with the tactical information.

IV. PROTOCOL DESIGN

Based on the designed transmission framework, in this section, we design a communication protocol called the *non-elastic C/D information transmission protocol* (neCDITP) to exploit the advantages provided by the transmission framework and improve the SD-ATN's capability of transmitting the non-elastic C/D information.

In the initial state, the global network view of the SD-ATN controller is not formed. A common node utilizes the distributed communication capability of the SD-ATN to enter the network and report its network status to the active control node. The following context is presented assuming that the network initialization is completed.

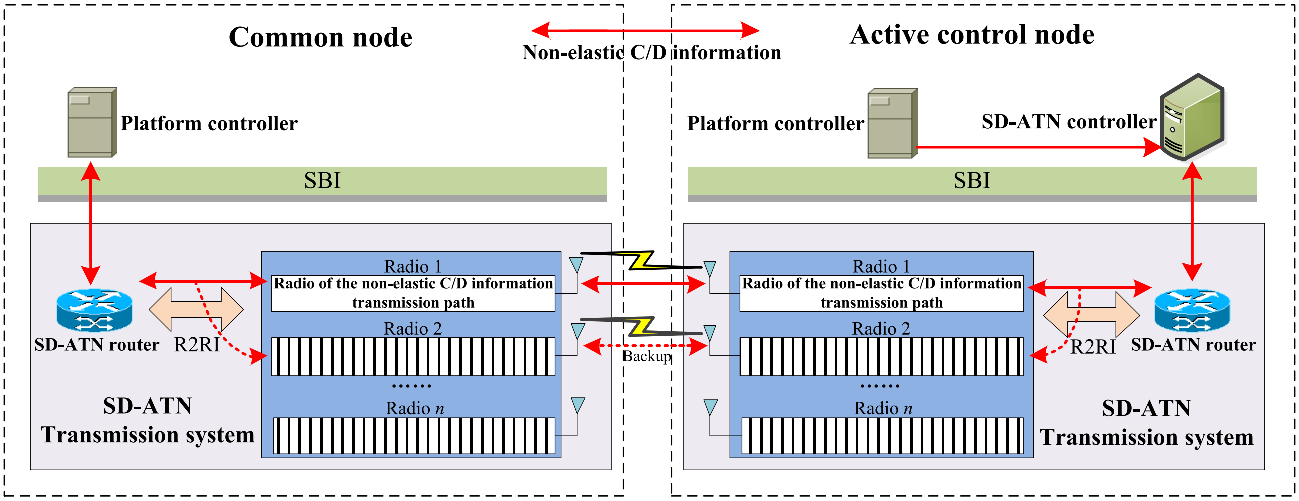


FIGURE 2. Framework for transmitting the non-elastic C/D information.

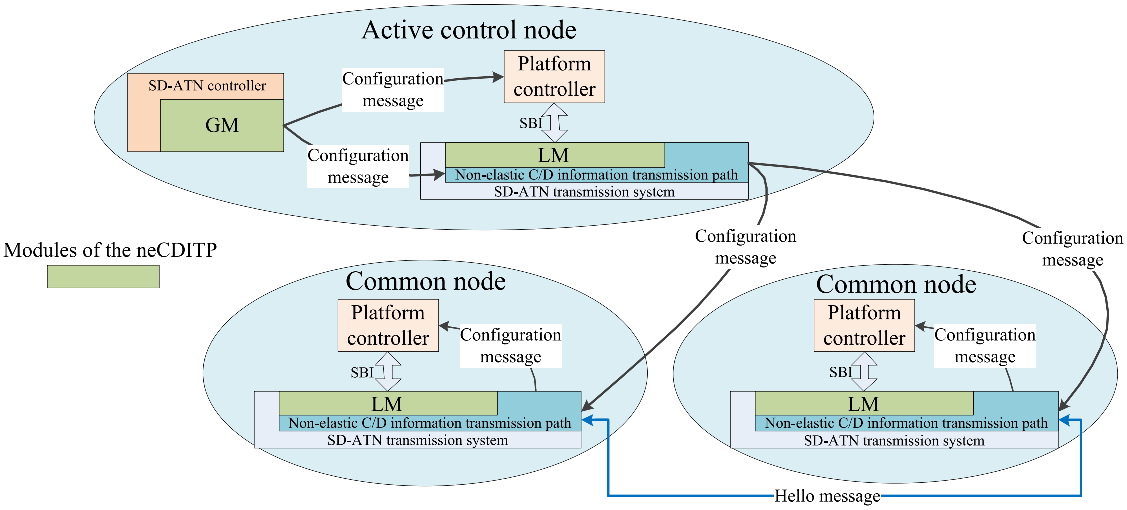


FIGURE 3. Modules and interactions of the neCDITP.

A. MODULES AND INTERACTIONS

According to the SD-ATN architecture, the SD-ATN supports both the centralized network control mode and distributed network control mode. The centralized network control mode can globally optimize the network performance with less overhead than the distributed network control mode. The distributed network control mode can perceive the local changes of the network status faster than the centralized network control mode, leading to a good adaptability. As shown in fig .3, to leverage the advantages of the two network control modes for improving the transmission efficiency of the non-elastic C/D information, the neCDITP is composed of a global module (GM) that works within the SD-ATN controller and a local module (LM) that works within the non-elastic C/D information transmission path. The GM is in charge of formulating the transmission policy for the non-elastic C/D information transmission path. The LM perceives and reports the network status and implements the transmission policy formulated by the GM to make the non-elastic C/D information transmission path work as required.

The GM indicates the operations of the LM by sending the configuration message. The configuration message is generated by the GM and is sent via the non-elastic C/D information transmission path, and it contains the formulated transmission policy of the GM. If a configuration message is generated, the configuration message is first replicated. Then, one message is sent to the platform controller of the active control node and another is broadcasted to the whole network. If a platform controller receives the configuration message, it configures the LM according to the transmission policy contained in the received message. The hello message is generated by the LM, which is used to sense the link status and is broadcast within one-hop. The hello message is also transmitted via the non-elastic C/D information transmission path. On the one hand, the sensed link status is used to indicate the operations of the LM; on the other hand, it is encapsulated in the network status message and sent to the active control node to help form the global network view. The LM periodically reports the sensed network status to the platform controller, and the platform controller transmits the reported information to the SD-ATN controller via the network status messages.

The network status message is the source of knowledge that forms the global network view of the SD-ATN controller, which contains various types of network status information sensed by each node's SD-ATN transmission system. The neCDITP does not generate the network status message but it utilizes this message type to report the status of the overlay network constructed for transmitting the non-elastic C/D information. Considering the particularity of the network status message, it is not transmitted through the non-elastic C/D information transmission path. How the active control node collects the network status messages is not the focus of this article, and the corresponding mechanism is investigated in another work.

By only considering the link layer elements or network layer elements, it is difficult to meet the QoS requirements of the non-elastic C/D information for reliability and latency. Therefore, the neCDITP is a protocol that has both link layer functions and network layer functions. Correspondingly, the GM has a link layer function unit (LLFU) and a network layer function unit (NLFU). The LLFU formulates the broadcasting policy of the configuration message and the link sensing policy of indicating how to send the hello message. The NLFU implements the routing computation that formulates the forwarding policy for the non-elastic C/D information. In particular, to utilize the global network view, an external interface is also implemented in the GM, which is used to access the SD-ATN controller's database to obtain the network view information. The LM has a local link layer function unit (L-LLFU) that is embedded in the radio of the non-elastic C/D information transmission path and a local network layer function unit (L-NLFU) that is embedded in the SD-ATN router. The L-LLFU and L-NLFU work collaboratively to make the non-elastic C/D information transmission path run according to the GM's policy. Specifically, the link sensing result is informed to the L-NLFU, making the L-NLFU respond to link status changes quickly.

Note that the non-elastic C/D information is not the main part of the information transmitted in the SD-ATN, the system resources allocated to the new added transmission path should be limited, due to factors such as expenditures, integration difficulty, etc. Therefore, the neCDITP is designed based on the simplest case that uses one half-duplex transceiver to perform the single-channel transmission. More details of the neCDITP are presented in the following context. First, we provide the frame structure maintained by the L-LLFU, which determines the real-time performance of the neCDITP. Then, the broadcasting strategy of the configuration message is given, which indicates how the configuration message of the GM is broadcast to the whole network. Further, the routing policy that indicates the forwarding of the non-elastic C/D information and the adaptive link sensing policy that indicates the sending behaviors of the hello message are investigated, which guarantee the reliability of the non-elastic C/D information in the aviation environment.

B. FRAME STRUCTURE

As discussed above, the transmission of the non-elastic C/D information has a low cost of bandwidth. For our proposed transmission framework, the network layer transmission policy has little impact on the real-time performance of transmitting the non-elastic C/D information, because the queuing delay will be very low if the bandwidth is appropriately allocated to the dedicated transmission path. Therefore, the neCDITP intends to guarantee the real-time performance by making the non-elastic C/D information access the channel quickly, which focuses on the link layer elements.

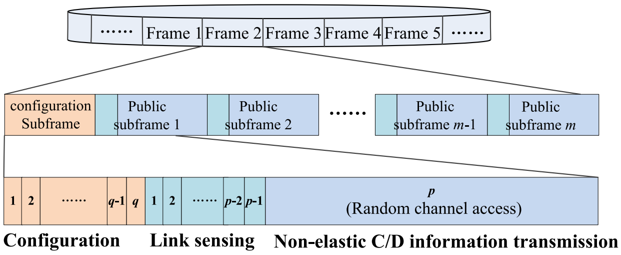


FIGURE 4. The frame structure employed by the enCDITP.

The L-LLFU of the neCDITP maintains a frame structure shown in fig. 4. The frame structure rules the channel access behaviors of the non-elastic C/D information transmission path. According to fig. 4, time is divided into periodic and equal-length frames. A frame is further divided into m public subframes and one configuration subframe. The configuration subframe is located at the beginning of each frame. There are two types of slots in the frame structure: one type is called the *short-slot*, whose duration is represented by t_s; the second type is called the *long-slot*, whose duration is represented by t_l and is much larger than t_s. Considering that the short-slot and long-slot have different durations, to simplify the slot synchronization, we fix the number of slots in a frame; thus, the length of a frame is also fixed.

The public subframe is composed of p slots, and the p slots are divided into two parts: one part is used for transmitting the non-elastic C/D information and is composed of one long-slot; another part is used for sensing the link status and is composed of $p - 1$ short-slots. During the long-slot, all nodes compete to use the channel to transmit the non-elastic C/D information. Considering the long transmission range and the complex electromagnetic environment, the statistical priority-based multiple access (SPMA) protocol [15] is used to rule the channel access behaviors during the long-slot.

Accessing the channel in a competitive manner can significantly reduce the transmission latency, which achieves a good real-time performance of transmitting the non-elastic C/D information. Although the competitive channel access manner causes collisions of packets, considering the transmission of the non-elastic C/D information costs low bandwidth, the packet collision rate can be maintained at a low level if the bandwidth is appropriately allocated. Moreover, the automatic repeat-reQuest (ARQ) mechanism and the TxRx^N waveform [15] can be used to further reduce the packet collision rate. These technologies are not discussed in this article since they are beyond the focus of our protocol.

The link sensing results are used to reflect the link status impacted by the factors such as mobility, electronic jamming, fuselage occlusion, etc. The results need not reflect the packet collision rate because it is meaningless to send special packets to perceive the packet collision rate when the SPMA protocol is employed. Based on the above considerations, the link sensing part is composed of $p-1$ independent short-slots. A short-slot is allocated to only one node for sending the hello message, which eliminates the hello messages collisions and improves the accuracy of sensing the link status. After assigning the short-slots to each node, the remaining short-slots are assigned to different nodes for transmitting the non-elastic C/D information. The length of a public subframe is represented by T_1, which is given as:

$$T_1 = (p-1)t_s + t_l. \tag{1}$$

Since the configuration message is critical in making the non-elastic C/D information transmission path operate normally and efficiently, the configuration message should be reliably transmitted to different common nodes. Thus, we define a configuration subframe in each frame to provide dedicated channel resources for broadcasting the GM's configuration message to the whole network efficiently and reliably. The configuration subframe is composed of q short-slots. There is one guard short-slot at the beginning of each configuration subframe to ensure that there is no node transmitting the non-elastic C/D information. The other short-slots in the configuration subframe are assigned to different nodes for relaying the configuration message to the whole network. The length of a configuration subframe is represented by T_2, which is given as:

$$T_2 = qt_s. \tag{2}$$

According to the length of the public subframe and configuration subframe, we can obtain the length of a frame, which is denoted by T_f:

$$T_f = mT_1 + T_2 = m(p-1)t_s + mt_l + qt_s. \tag{3}$$

C. BROADCASTING MECHANISM

The GM of the neCDITP indicates the behaviors of transmitting the non-elastic C/D information by broadcasting the configuration message to the whole network. In this section, we describe how the configuration message is broadcast during the configuration subframe. In particular, to configure the non-elastic C/D information transmission path of the active control node, the configuration message is directly sent to the platform controller via the wired network.

To configure the non-elastic C/D information transmission paths of the common nodes, the SD-ATN controller first inserts the configuration message into the dedicated queue in the SD-ATN router. This dedicated queue is only used for caching the configuration message generated by the GM. During the guard short-slot of the configuration subframe, each common node's radio of the non-elastic C/D information transmission path switches into the reception status, waiting for receiving the configuration message. When the next slot begins, the configuration message is dequeued and delivered to the channel.

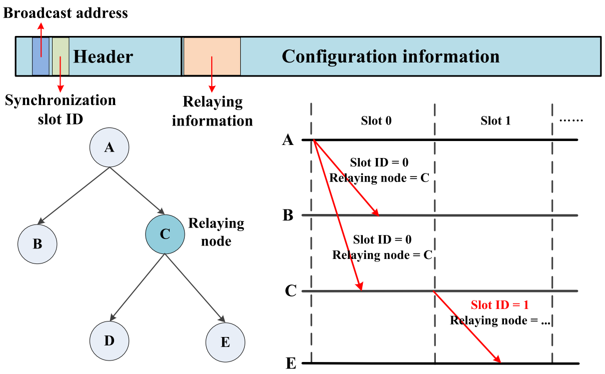


FIGURE 5. Example for broadcasting a configuration message.

The general format of the configuration message is shown in fig. 5 (the format of the configuration message is allowed to be further standardized). In the configuration message, the broadcast address in the message header ensures that the configuration message can be processed by any common node that receives it. The synchronization slot ID in the message header is changed to the ID of the slot during which the configuration message is sent, which is helpful for slot synchronization. The relaying information contained in the configuration message tells the common nodes how to broadcast the configuration message to the whole network.

To reserve more channel resources for the public subframe or to reduce the broadcast latency of the configuration message, instead of flooding the configuration message, some of the common nodes are selected out to relay the configuration message. Correspondingly, the short-slots of the configuration subframe are only assigned to the selected common nodes. Once a common node receives the configuration message, the message is temporarily cached. A replicated message is sent to the platform controller. To determine the broadcasting operations, the platform controller extracts the relaying information from the configuration message and configures the L-LLFU according to the extracted relaying information. If this common node is a relaying node, the cached configuration message is further sent at the assigned slot. If this common node is not selected as a relaying node, the cached configuration message is dropped. According to [16], if the duration of the short-slot is set appropriately, the processes above can be accomplished within one short-slot; thus, the guard slots are not needed.

The GM's LLFU is in charge of formulating the broadcasting policy of the configuration message. The relaying nodes selection processes are presented in Algorithm 1. Algorithm 2 illustrates how to assign the short-slots of the configuration subframe to the selected relaying nodes.

Algorithm 1 Relaying Nodes Selection

```
Input: $G_3$, active control node $C$
Output: $RN$
Divide the common nodes into multiple levels;
initialize $RN \leftarrow \phi, Q_1 \leftarrow \phi, Q_2 \leftarrow \phi, low\_l \leftarrow \phi, high\_l \leftarrow \phi, n_{h}igh\_l \leftarrow \phi, D \leftarrow 0, d \leftarrow \phi$;
for $i \leftarrow 1{:}h$
    Compute $low\_l, high\_l, n\_high\_l, D$;
    $Q_1 \leftarrow Q_1 \cap high\_l, Q_2 \leftarrow Q_2 \cap low\_l$;
        for each $v \in low\_l$
        if $n\_high\_l(v) \leftarrow 1$
            $RN \leftarrow RN \cup high\_l(v)$;
            Remove $low\_l(high\_l(v))$ from $Q_2$ and $high\_l(v)$ from $Q_1$;
        Update $D$;
      end if
    end for
        while $Q_1 \neq \phi$ and $Q_2 \neq \phi$
        Sort the nodes in $Q_1$ in decreasing order of $D$;
        Take out the first node in $Q_1$, say node $v$;
        $RN \leftarrow RN \cup v$;
        Remove $low\_l(v)$ from $Q_2$ and $v$ from $Q_1$;
      Update $D$;
    end while
    initialize;
end for
```

According to the neighbor nodes information obtained from the global network view, the LLFU models the overlay network that is constructed for transmitting the non-elastic C/D information as a three-dimensional undirected graph $G_3 = (V, E)$, where V denotes the set of nodes in the network and $E = \{(i, j)|i,j \in V\}$ denotes the set of edges representing the wireless links.

In Algorithm 1, the common nodes in the network are divided into multiple levels according to their hop-count from the active control node. L_i denotes the set of nodes at level i, and h denotes the number of levels. The common nodes that are only one hop away from the active control node belong to the highest level which is represented by level 1. RN represents the set of selected relaying nodes, $low\_l(v)$ denotes the set of nodes that belong to the lower level and are connected with node v. $high\_l(v)$ denotes the set of nodes that belong to the higher level and are connected with node v. $n\_high\_l(u)$ denotes the size of $high\_l(v)$. Algorithm 1 is iterated every two adjacent levels. At the beginning of an iteration, the nodes in the higher level are temporarily stored in a buffer Q_1, and the nodes in the lower level are temporarily stored in a buffer Q_2. During the iteration, some nodes in Q_1 are selected as the relaying nodes, for the purpose of making any node in Q_2 have coverage by the selected relaying nodes. The nodes in Q_2 that have only one connected node in Q_1 are found first. These nodes' connected nodes in Q_1 are chosen as the relaying nodes. Then, considering better communication quality can be achieved by shortening the distance between two nodes, a node in Q_1, say node v, which minimizes $D(v)$, is selected as the relaying node, where $D(v)$ is represented as:

Algorithm 2 Slot Assignment for Broadcasting

```
Input: $G_3$, $RN$, active control node $C$
Output: slot assignment result
As in algorithm 1, divide all the common nodes into multiple levels;
Initialize $low\_l\_r \leftarrow \phi, Q \leftarrow Q \cap C$;
Compute $low\_l\_r$ according to $G_3$ and $RN$;
Schedule all the common nodes to be idle;
while $Q \neq \phi$
    Take the first node from $Q$, say node $v$.
    Schedule $v$ to send according to the broadcast constraints
    Schedule the nodes in $low\_l(v)$ to receive;
    Remove $v$ from $Q$;
    $Q \leftarrow Q \cup low\_l\_r(v)$;
end while
```

$$D(v) = \min_{u \in Q_2} \max d(v, u). \tag{4}$$

$d(v,u)$ denotes the distance between node v and node u, which is represented as:

$$d(v, u) = \sqrt{(x_u - x_v)^2 + (y_u - y_v)^2 + (z_u - z_v)^2}, \tag{5}$$

where x, y and z constitute the three-dimensional Cartesian coordinate that identifies a node's location. For each time of iteration, Algorithm 1 stops selecting relaying nodes from Q_1 until Q_1 is empty or all nodes in Q_2 are covered by the selected relaying nodes.

Before describing the details of Algorithm 2, the constraints that ensure the correct broadcast operation are presented:

1) A relaying node can be scheduled to send only if it has received a message.

2) A relaying node can be scheduled to send during a slot if it is not scheduled to receive during the same slot.

3) A relaying node can be scheduled to send if it does not interfere with the reception of another node according to the protocol interference model.

In Algorithm 2, the common nodes in the network are also divided into multiple levels as in Algorithm 1. Then, low_l_r is computed according to G_3 and L_i, where $low\_l\_r(v)$

denotes the set of lower level relaying nodes that are connected with the relaying node v. Initially, the active control node is stored in a buffer Q. As long as Q is not empty, a node in Q is found and scheduled to send abiding by the broadcast constraints. Correspondingly, all the sending node's connected nodes in the lower level are scheduled to receive. The node that is scheduled to send is removed from Q, and then all the relaying nodes that are in the lower level and connected with the removed node are stored in Q. Each node can only be stored in Q once.

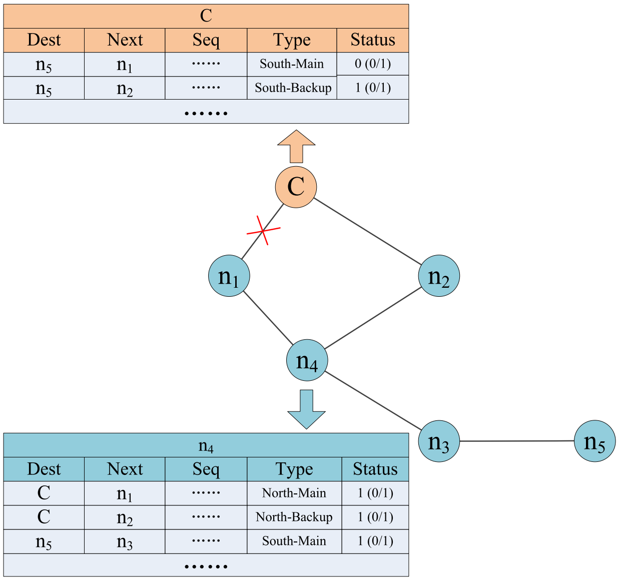
According to Algorithm 1 and Algorithm 2, we give an upper bound of the short-slot number allocated for transmitting the configuration message, which is equal to $\max(\frac{2N+h-2}{3}, h)$. N denotes the number of nodes in the network and h denotes the maximum hop-count of the network. According to the upper bound, the number of short-slots in the configuration subframe can be determined in advance, which is helpful for pre-network planning to reserve more channel resources for the public subframe or to reduce the broadcast latency of the configuration message. The proof of the upper bound is given in the appendix of this article.

D. ROUTING MECHANISM

The configuration message contains the information of the formulated routing policy and link sensing policy. The routing policy indicates the nodes' behaviors of forwarding the non-elastic C/D information. In this section, we present the routing mechanism of the neCDITP.

Considering the link quality of the SD-ATN is unstable due to the mobility, electronic jamming, and other factors, different routing paths have different reliability performance. Therefore, the routing mechanism of the neCDITP must consider the reliability requirement of the non-elastic C/D information. Since the transmission policy of the network layer has little impact on the real-time performance, the routing mechanism of the neCDITP only focuses on ensuring the reliability performance.

When the platform controller receives the configuration message, it extracts the routing information and accordingly configures the L-NLFU in the SD-ATN router. Thus, the routing table (the structure of the routing table is allowed to be further standardized), which is dedicatedly constructed for forwarding the non-elastic C/D information, is updated. There are 4 types of route entries in the dedicated routing table: north main route entry, north backup route entry, south main route entry, and south backup route entry. The north route entries record the routing paths that instruct the common nodes to forward the non-elastic C/D information to the active control node. The south route entries record the reversed routing paths. Considering the intermittent links of the SD-ATN and the reliability requirement of the non-elastic C/D information, two types of routing paths are maintained: one type is called the *main routing path* and another type is called the *backup routing path*. A packet is first forwarded according to the main routing path. If the main routing path is unavailable, the packet and its following packets are then forwarded according to the corresponding backup routing path. If there is no available routing path in the routing table, the backup transmission path is enabled to transmit the non-elastic C/D information. To avoid routing loops, there is an indicator in the packet's header indicating the routing path type that should be looked up. If the main routing path is unavailable, the indicator is changed and will not be changed back during the following transmission, even when there is a new available main routing path. In particular, the L-NLFU of each node independently determines the availability of a routing path according to the link sensing result. The link sensing mechanism of the neCDITP is presented in the following section.



C				
Dest	Next	Seq	Type	Status
n_5	n_1	······	South-Main	0 (0/1)
n_5	n_2	······	South-Backup	1 (0/1)
······				

n_4				
Dest	Next	Seq	Type	Status
C	n_1	······	North-Main	1 (0/1)
C	n_2	······	North-Backup	1 (0/1)
n_5	n_3	······	South-Main	1 (0/1)
······				

FIGURE 6. Example for the routing table.

As shown in fig. 6, when n_4 needs to transmit its non-elastic C/D information to the active control node, the north main route entry is first looked up. If the north main route entry is available (the status of the route entry equals to 1), the packet that carries the non-elastic C/D information of n_4 is sent to n_1 which is the next-hop node recorded in the route entry. After n_1 receives the packet sent from n_4, n_1 continues to forward the received packet in the same way as n_4 until the packet is forwarded to the active control node. When the active control node needs to send its non-elastic C/D information to n_5, the south main route entry that records n_5 as the destination is first looked up. If this south main route entry is unavailable, instead, the active control node looks up the south backup route entry that records n_5 as the destination. If the south backup route entry is available, the active control node sends the packet that carries its non-elastic C/D information to n_2 which is the next-hop node recorded in the route entry. During the following transmission of this packet, only the south backup routing entries are looked up since the indicator in the packet's header is changed.

To select a reliable routing path, the following two metrics are considered:

(1) Path maintenance time. Path maintenance time is equal to the minimum link availability period between intermediate nodes in the routing path. The link availability period l_T is computed as follows. Assuming node m and node n are neighbors. Let (x_m, y_m, z_m), v_m and θ_m be the three-dimensional coordinate, moving speed and moving direction of node m, respectively. The parameters of node n are represented in the same way. The above parameters of each node are part of the global network view. First, the three-dimensional parameters defined above are projected to a two-dimensional plane. The obtained two-dimensional parameters include (x'_m, y'_m), (x'_n, y'_n), v'_m, v'_n, θ'_m, and θ'_m. Then, according to the above two-dimensional parameters, the l_T is computed as:

$$l_T = \frac{-(ab+cd) \pm \sqrt{(a^2+c^2)R^2-(ad-bc)^2}}{a^2+c^2}, \tag{6}$$

where

$$\begin{aligned} a &= v'_m \cos\theta'_m - v'_n \cos\theta'_n \\ b &= x'_m - x'_n \\ c &= v'_m \sin\theta'_m - v'_n \sin\theta'_n \\ d &= y'_m - y'_n. \end{aligned} \tag{7}$$

According to the l_T of different links, the path maintenance time is given as follow:

$$P_T = \min_{l \in P}\{l_T\}. \tag{8}$$

(2) Path signal-to-interference-plus-noise ratio. Path signal-to-interference-plus-noise ratio (SINR) is equal to the minimum link SINR between intermediate nodes in the routing path. For any two neighbor nodes m and n, the link SINR l_S is equal to the value of $S(m, n)$ or $S(n, m)$, where $S(m, n)$ represents the SINR when node m receives the hello message from node n. The minor value of $S(m, n)$ and $S(n, m)$ is taken as the link SINR.

According to the l_S of different links, the path SINR is given as follow:

$$P_S = \min_{l \in P}\{l_S\}. \tag{9}$$

According to the above metrics, it is prone to select the routing path with larger P_T and P_S. Therefore, we provide a joint metric$P_{weight,i}$ for the routing path selection:

$$P_{weight,i} = \alpha \frac{P_{T,i}}{P_T^{max}} + \beta \frac{P_{S,i}}{P_S^{max}}. \tag{10}$$

$P_{T,i}$ and $P_{S,i}$ denote the maintenance time and the SINR of routing path i, respectively. P_T^{max} and P_S^{max} denote the maximum maintenance time and the maximum SINR of all available routing paths, respectively. α and β are the weights corresponding to the significance of the two parameters in the joint metric. The values of α and β lie between 0 and 1 and satisfy $\alpha + \beta = 1$.

Algorithm 3 Route Selection

```
Input: $G_3$, $l_T$ of each link, $l_S$ of each link
Output: Routing paths for forwarding the non-elastic C/D information
Compute each link's weight.
Take the active control node as the start point and compute the north main routing paths using the *Dijkstra* algorithm;
Take reverse search for each common node's north main routing path to get the south main routing paths;
Sort the links of all north main paths in decreasing order of the weight, and begin to traverse all links in order;
while there is a link that has not been traversed
    Use *Tarjan* algorithm to compute the cut edges set $CE$ of $G_3$;
    if the traversed link is not included in $CE$
        Remove the traversed link from $G_3$;
    end if
end while
Based on the updated $G_3$, compute the north backup routing paths using the *Dijkstra* algorithm;
Take reverse search for each north backup routing path to obtain the south backup routing paths;
```

The GM's routing path selection algorithm is shown in Algorithm 3. In Algorithm 3, the weight of each link is first computed referring to the joint metric $P_{weight,i}$, which is given as follow:

$$l_{weight,i} = \frac{l_T^{max} l_S^{max}}{\alpha l_{T,i} l_S^{max} + \beta l_{S,i} l_T^{max}}, \tag{11}$$

where $l_{T,i}$ and $l_{S,i}$ denote the maintenance time and the SINR of link i, respectively. l_T^{max} and l_S^{max} denote the maximum maintenance time and the maximum SINR of all links, respectively. Based on the computed links' weight, the routing paths that are adaptive to the high mobility and have high SINR can be selected out, which can improve the reliability of transmitting the non-elastic C/D information. In Algorithm 3, according to the weight of each link, the *Dijkstra* algorithm is used to compute the north main routing paths. Afterwards, the south main routing paths are computed by taking a reverse search. To compute the north backup routing paths, we first delete the links that belong to the main routing paths and are not the cut edges. A link with a larger weight is more prone to be deleted, since it is be included in the main routing paths more frequently. Then, the north backup routing paths and south backup routing paths are computed according to the updated network topology.

After computing all routing paths, the GM converts the computed routing paths to the routing information and encapsulates the routing information in the configuration message.

E. ADAPTIVE LINK SENSING MECHANISM

For a wireless network, sensing the link status is the basic step of forming the network view. Additionally, it is necessary for the nodes to sense the link status changes in a timely manner for responding to the network changes quickly. The link status mentioned in this article includes the neighbor node information and the corresponding link quality. The neCDITP utilizes the hello message to sense the link status. The hello messages are generated by the L-LLFU and are broadcast within one-hop range. The format of the hello message is shown in fig .7 (the format of the hello message is allowed to be further standardized). A node obtains its neighbor nodes information according to the sender address contained in the received hello messages. Meanwhile, a node perceives the quality of a link based on the SINR of receiving the hello messages sent from the corresponding neighbor node.

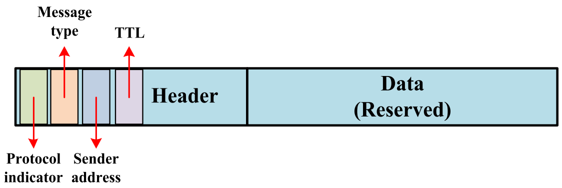


FIGURE 7. Format of the hello message.

Considering a short-slot of a frame's link sensing part is only assigned to one node, the total number of link sensing times during one frame is limited to $m(p-1)$, which is equal to the times of broadcasting the hello messages. To sense the link status changes more accurately and quickly, the link sensing frequency should be improved. However, the total number of link sensing times during one frame is limited; thus, the neCDITP employs the adaptive link sensing mechanism. Given the benefits of the frame structure of the neCDITP, it is practical to change the link sensing frequency of a node by dynamically assigning short-slots of each frame's link sensing part, which in fact rules different nodes' frequency of broadcasting the hello messages. The neCDITP stipulates that a node should broadcast two hello messages during one frame to provide a relatively accurate and timely link sensing result. The remaining short-slots of the frame's link sensing part are dynamically assigned to different nodes to sense the status of the links that break easily. The GM of the neCDITP is responsible for assigning the short-slots for link sensing, and the assignment result is encapsulated in the configuration message.

For each frame's link sensing part, the number of short-slots that can be dynamically assigned is very limited after subtracting the basic number of short-slots assigned to each node. Therefore, when adaptively adjusting the link sensing frequency, 1) only the links that belong to the main routing paths are considered. Moreover, considering that the link with a smaller distance is less easily broken, 2) the link with an availability period less than a frame's duration is prone to be sensed more frequently. For a link that meets the two conditions listed above, say e_i, the recommended number of short-slots assigned to the end node of e_i, which is denoted by $S_a(e_i)$, is computed as follow:

$$S_a(e_i) = \begin{cases} \left[\dfrac{S_r}{l_{T,e_i} \sum\limits_{e_j \in L_s} \frac{1}{l_{T,e_j}}}\right] & \text{if } S_a(e_i) < S_a^{max} \\ S_a^{max} & \text{otherwise.} \end{cases} \tag{12}$$

The above formula guarantees that a link with lower availability period is allocated more slots. L_s denotes the set of links that belong to the main routing paths and have an availability period that is less than a frame's duration. S_r denotes the total number of short-slots that can be assigned. S_a^{max} represents the maximum number of short-slots that can be assigned to one node for broadcasting hello messages. If a node simultaneously belongs to multiple links that meets the conditions, the maximum number of assigned short-slots is taken.

Algorithm 4 Slot Assignment for Link Sensing

```
Input: $G_3$, routing information
Output: assignment result of the short-slots for link
        sensing
Number the short-slots of a frame's link sensing part;
while there is a node being not assigned any short-slot
  for $t \leftarrow 1{:}k$
    Find a short-slot which is not assigned and is within
     the number from $n_{min}(t)$ to $n_{max}(t)$.
    Assign the short-slot to the node;
  end for
end while
Find the links that meet the conditions and store the
 found links into set $L_s$;
while $L_s \neq \phi$
  Find the link that has the minimum availability
   period, say $e_i$;
  if there is an end node of $e_i$ being not assigned
   any short-slot
    Compute $S_a(e_i)$;
    for $t \leftarrow 1{:}S_a(e_i)$
      Find a short-slot which is not assigned and is
       within the number from $n_{min}(t)$ to $n_{max}(t)$;
      Assign the found short-slot to the end node;
    end for
  end if
  Remove the link from $L_s$;
end while
```

Algorithm 4 shows the process of assigning the short-slots of each frame's link sensing part. First, a same number of short-slots is assigned to each node to guarantee the basic link sensing frequency which is represented by k (the value of k can be larger than 2). Then, the remaining short-slots are assigned to the end nodes of the links that meet the conditions mentioned above. To evenly distribute the time points at which a node broadcasts the hello messages, the short-slots are divided into k or $S_a(e_i)$ sets before they are assigned. Then, one non-occupied short-slot from one set is assigned to a node. For example, if it is required to assign m short-slots to a node, the short-slots in the link sensing part are divided into m sets. One non-occupied short-slot from one set is assigned to the node. $n_{min}(t)$ and $n_{max}(t)$ denote the minimum short-slot number and the maximum short-slot number of set t, respectively.

The formulated link sensing policy is also encapsulated in the configuration message. After receiving the configuration message, the L-LLFU makes the radio of the non-elastic C/D information transmission path sense the link status according to the received link sensing policy.

V. PERFORMANCE EVALUATION

In this section, simulation results are given to illustrate the performance of our scheme. The experiment is constructed based on the network simulator called EXata 5.1.

A. SIMULATION SCENARIO

As shown in fig. 8, we build a simplified avionics network in EXata 5.1 for every aircraft and regard an avionics network as a node of the SD-ATN. The simplified avionics network, which is regarded as the common node, includes two devices that implement the functions of the platform controller and SD-ATN transmission system, respectively. The avionics network that represents the active control node includes one more device that implements the functions of the SD-ATN controller. The devices within an aircraft are connected through wired links. The 802.3 MAC protocol and the open shortest path first routing protocol (OSPF) are used to ensure efficient communications between the different devices. The GM of the neCDITP is implemented in the device that implements the functions of the SD-ATN controller and the LM is implemented in the device that implements the functions of the SD-ATN transmission system.

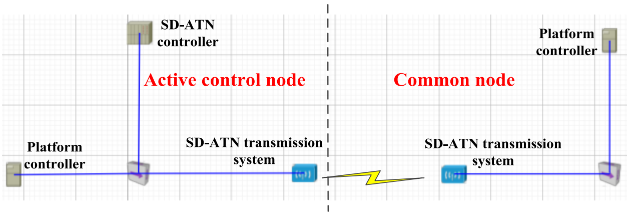


FIGURE 8. Example for simulation scenario.

The device that implements the functions of the platform controller is in charge of generating and processing the non-elastic C/D information, and the device is also responsible for generating the network status message. The device that implements the functions of the SD-ATN transmission system uses different interfaces to simulate different radios. The network status message is reported to the active control node every 2 seconds via a dedicated interface. A simple SBI that implements the configuration functions needed by the platform controller is also designed. Since an aircraft is represented by its avionics network in our experiment, a new random waypoint mobility model is developed to simulate the mobility of different aircraft. This mobility mode makes the aircraft move in a three-dimensional random way; meanwhile, the devices that belong to the same aircraft remain relatively stationary.

The dimension of the experimental scenario is set in the Cartesian coordinate system, where X and Y are both equal to 3×10^5 meters. The altitude of the simulation scenario is equal to 1×10^4 meters above sea level. In the aviation swarm deployed in our simulation scenario, there is one aircraft that serves as the active control node, and the remaining aircraft serve as the common nodes. The minimum speed of an aircraft is set to 200 m/s. The short-slot duration is set to 2.5 ms, the long-slot duration is set to 200 ms, and the frame duration is set to 5.7 s. The generation interval of the configuration message is set to 6 s. The data rate of the radio, which belongs to the non-elastic C/D information transmission path, is set to 2 Mbps. The aeronautical channel model [17] is used to simulate the aviation channel environment.

B. SIMULATION RESULTS

Based on the proposed transmission framework, the reliability performance and real-time performance of the designed communication protocol neCDITP are investigated. The simulation results of the neCDITP are compared with the performance of the combinatorial applications of classic routing and MAC protocols. Four combinations are employed for comparison:

- C1: Carrier sense multiple access protocol (CSMA) and optimized link state routing protocol (OLSR).
- C2: CSMA and ad hoc on-demand distance vector routing protocol (AODV).
- C3: Time division multiple access protocol (TDMA) and OLSR.
- C4: TDMA and AODV.

Afterwards, the comprehensive performance of our scheme is simulated, which is compared with the performance of the widely used military communication systems of legacy ATNs.

Fig. 9 shows the reliability performance of different protocols. With the network size increases (the number of nodes in the network is used to characterize the network size), the packet delivery ratio of all simulation objects decreases, since the bandwidth becomes more limited. The reliability performance of the neCDITP is better than that of most combinations (C1, C2, and C3), except C4. However, the packet delivery ratio of C4 decreases significantly when the number of nodes becomes larger than 60. In contrast, the neCDITP achieves the best packet delivery ratio when there are more than 70 nodes in the network. With the packet generation interval increases, the packet delivery ratio of all simulation objects increases because the packet collision rate is reduced. C4 achieves the highest packet delivery ratio for all cases. The packet delivery ratio of the neCDITP is only lower than that of C4, which is higher than 85% when the packet generation interval is larger than 1 s. With the increase of the aircraft's speed, the network topology changes

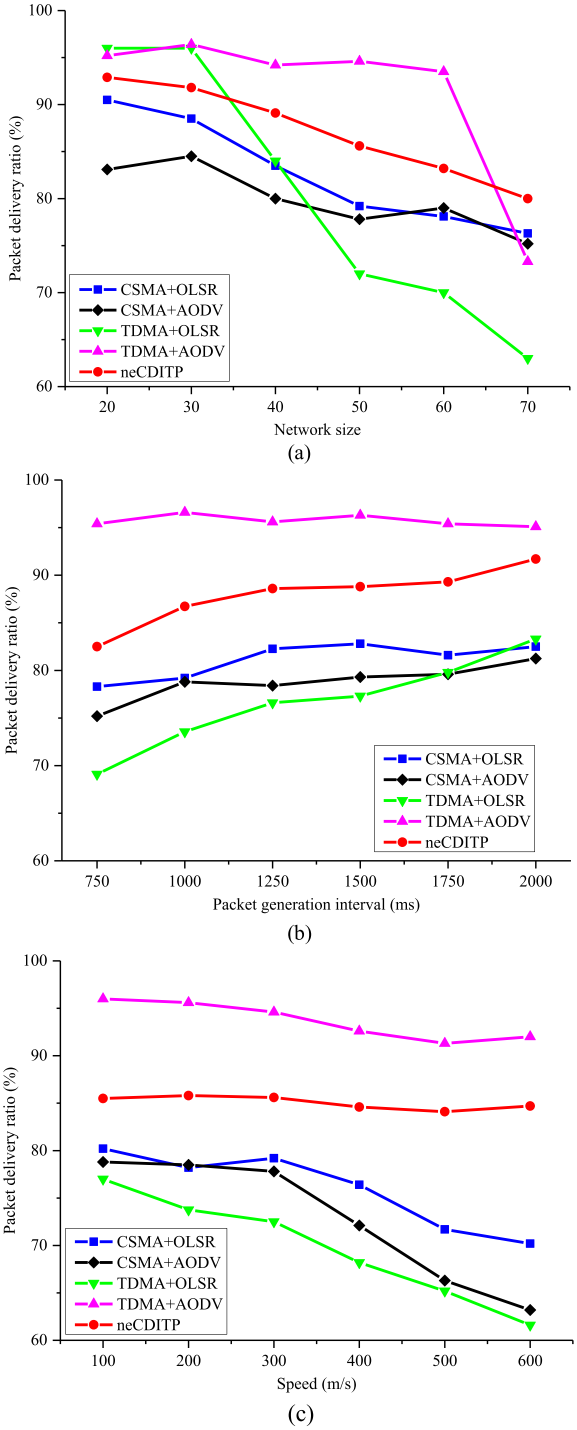


FIGURE 9. **Packet delivery ratio of different protocols. (a) Packet delivery ratio with respect to the network size. (b) Packet delivery ratio with respect to the packet generation interval. (c) Packet delivery ratio with respect to the aircraft's speed.**

more frequently. The reliability performance of the neCDITP is only lower than that of C4. Moreover, the reliability performance of the neCDITP is nearly unfluctuating, while the reliability performance of the compared objects is more easily affected. In general, the average packet delivery ratio of the neCDITP is higher than 85% in most cases, which is a better performance than that of most compared objects.

For C3, the OLSR produces large network control overhead, which causes queue overflow because the packets have to wait a long time before they are sent to the channel when the TDMA is employed. Therefore, the reliability performance of C3 is the worst when the network size becomes large. C2 employs the CSMA, which makes the packets transmitted to the channel quickly. Although there are collisions of packets, the possibility of queue overflow is drastically reduced. Therefore, the reliability performance of C2 is better than that of C3 in most cases. When the CSMA is employed, the significant network control overhead of the OLSR improves the packet collision rate. Although the packet collision rate is improved, the OLSR is more adaptable to the link status changes than the AODV, since the OLSR can compute a new route faster than the AODV whose RREQ message and RREP message cannot be reliably transmitted. Consequently, C1 performs better than C2. The AODV has much less network control overhead than the OLSR. Therefore, even though the TDMA is employed by C4, there is no overflowed queue. Leveraged by the low network control overhead of the AODV, the reliability advantage of the TDMA is fully exploited; thus, the reliability performance of C4 is the best in most cases. However, when the network size becomes larger, the network control overhead of the AODV also becomes large; thus, the reliability of C4 decreases significantly.

For the neCDITP, the network control overhead of the neCDITP is maintained at a low level, which is much less than that of the compared objects. This is because the neCDITP leverages the low overhead advantage of the centralized network control mode, and the network status messages are collected independent of the transmission of the non-elastic C/D information. Therefore, the packet collision rate and the possibility of queue overflow are both reduced. Meanwhile, the routing mechanism of the neCDITP always tries to select a routing path that can reliably forward the non-elastic C/D information to the destination. Moreover, the neCDITP's adaptive link sensing mechanism makes it sensitive to the link status changes. Based on the adaptive link sensing mechanism, the neCDITP further exploits the backup routing path and the back transmission path to transmit the non-elastic C/D information, which improves the adaptability to the link status changes. Additionally, since the random channel access technology is also employed by the neCDITP, the packets can be quickly transmitted to the channel; thus, the possibility of queue overflow is drastically reduced. Due to the above factors, the reliability performance of the neCDITP is better than that of C1, C2, and C3, and this superiority is maintained under different circumstances.

The real-time performance of different protocols is shown in fig. 10. With the network size increases, the average end-to-end delay of C3 and C4 increases drastically; this is because more slots are needed in the network. In contrast, the end-to-end delay of C1 and C2 is maintained lower

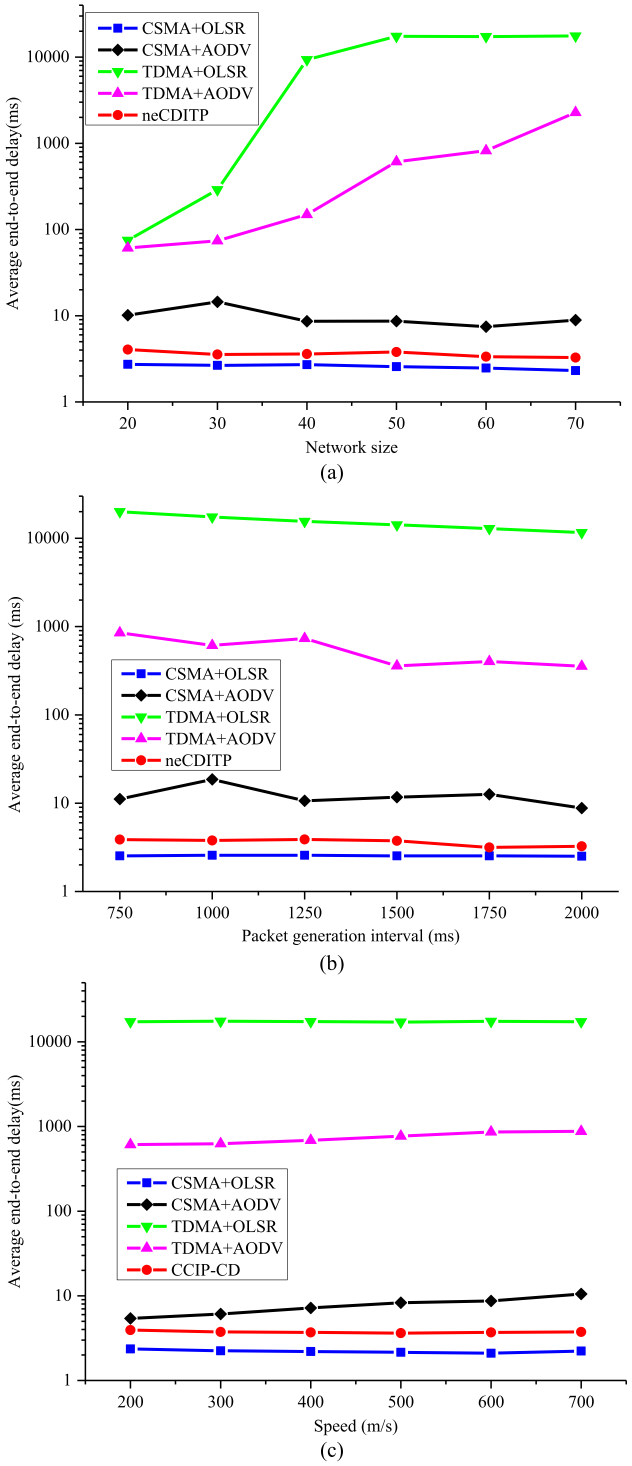


FIGURE 10. End-to-end delay of different protocols. (a) End-to-end delay with respect to the network size. (b) End-to-end delay with respect to the packet generation interval. (c) End-to-end delay with respect to the aircraft's speed.

than 3 ms and 20 ms, respectively. The average end-to-end delay of the neCDITP is maintained lower than 5ms, which is only higher than C1. With the packet generation interval increases, the average end-to-end delay of C3 and C4 decreases several seconds and hundreds of milliseconds, respectively, while the real-time performance of C1 and C2 has no obvious change. The real-time performance of the neCDITP also has no obvious change, which is maintained slightly worse than C1. Since more time is needed by the AODV to find a new route with the increase in the aircraft's speed, the average end-to-end delay of C2 and C4 increases. The real-time performance of C1, C3, and the neCDITP is not affected.

Due to the characteristic of the TDMA, the packets have to wait a long time before they are sent to the channel. Therefore, the real-time performance of C3 and C4 is not good. In particular, the OLSR produces larger network control overhead than the AODV; therefore, the real-time performance of C3 is the worst. For the compared objects that employ the CSMA, the real-time performance is greatly improved because the CSMA ensures that the packets are quickly sent to the channel. Since the packets have to wait for the AODV to temporarily find a route if there is no available routing path, the end-to-end delay of C2 is higher than that of C1 which employs the OLSR.

The neCDITP uses the SPMA to send the packets to the channel. The SPMA is also a random channel access protocol like the CSMA. Therefore, the packets can be quickly sent to the channel, which leads to a good real-time performance. The routing mechanism of the neCDITP provides more options than the AODV to route a packet to the destination; thus, an available routing path can be found more quickly than the AODV in the dynamic environment. Therefore, the end-to-end delay of the neCDITP is lower than that of C2. According to the frame structure of the neCDITP, a period of time is allocated to broadcast the configuration message and to sense the link status, the non-elastic C/D information can only be transmitted during the long-slot. Therefore, the real-time performance of the neCDITP is worse than that of C1.

In general, compared to the combinatorial applications of classic routing and MAC protocols, the neCDITP not only performs well in transmitting the non-elastic C/D information reliably, but also has good real-time performance. Moreover, the performance of the neCDITP is stable under different circumstances.

To synthetically illustrate the superiority of our scheme in transmitting the non-elastic C/D information, we simulate the proposed scheme's performance from a system perspective. The performance of the communication systems, which are widely used in legacy ATNs, is simulated for comparison. Because the two communication systems used for comparison are employed in the military domain, only the general characteristics of the two communication systems are listed:

- System-1: A communication system that uses the TDMA and achieves multihop transmission by configuring fixed relaying nodes (system-1 is similar to Link-16).
- System-2: A communication system that uses the TxRx^N waveform, random access technology, and proactive routing protocol (system-2 is similar to the TTNT).

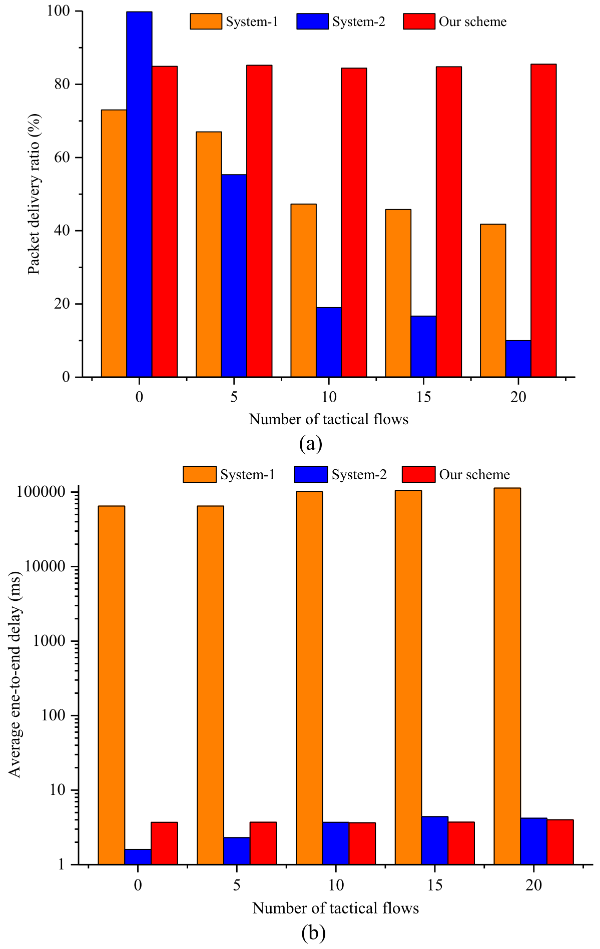


FIGURE 11. Performance of different communication schemes. (a) Reliability performance. (b) Real-time performance.

The experiment is constructed for a 2-hop network. For system-1 and system-2, the non-elastic C/D information is transmitted mixed with the tactical information. The priorities of different information types are not considered. The number of tactical flows is taken as the variable. Under the premise of the same bandwidth consumption, fig. 11 shows the simulation results.

According to fig .11, the best packet delivery ratio of the system-1 is lower than 80%, and it drops with the number of tactical flows increases. The end-to-end delay of system-1 is much higher than that of the other simulation objects. This is because system-1 makes a node wait a long time to transmit or relay its packets to the destination. When there is no transmission demand for tactical information, system-2 has the highest packet delivery ratio and the lowest end-to-end delay. With the number of tactical flows increases, the packet delivery ratio of system-2 drops significantly. This is because the packets that carry the non-elastic C/D information are subject to severe interference from the packets that carry the tactical information. Compared to system-1 and system-2, our scheme transmits the non-elastic C/D information in a relatively reliable and real-time way. This is because our scheme separates the transmission of the non-elastic C/D information from the other types of information and provides dedicated QoS guarantees for the non-elastic C/D information transmitted in the SD-ATN.

VI. CONCLUSION

An aviation swarm needs an ATN that provides a more flexible and efficient communications capability. To design an ATN that satisfies the communications demands of the aviation swarm, applying the SDN paradigm to the ATN is a promising solution because of the SDN's flexibility, programmability and openness features. Remarkably, one of the fundamental problems of designing an SDN-enabled airborne tactical network (SD-ATN) is determining how to ensure that the QoS requirements of the non-elastic information transmitted between the control plane and data plane (non-elastic C/D information) in the aviation environment are met. In this article, we propose a scheme for improving the communications capability between the control plane and data plane of the SD-ATN, aimed at making a reliable and real-time transmission of the non-elastic C/D information. The proposed scheme includes a new transmission framework and a communication protocol that is designed based on the transmission framework. The transmission framework makes it practical to provide dedicated QoS guarantees for transmitting the non-elastic C/D information. The communication protocol, which is called the neCDITP, has good reliability and real-time performance under the transmission framework.

In particular, our scheme exchanges hardware resources for bandwidth resources, needing the support of advanced hardware technologies and improves the difficulty of hardware design. Meanwhile, the anti-jamming capability of our scheme needs to be enhanced to ensure the reliable transmission in the complex electromagnetic environment. In future work, we will investigate the hardware implementation of our scheme and discuss the application of the anti-jamming technologies.

APPENDIX

For a network that can be modeled as a connected graph, when Algorithm 1 and Algorithm 2 are used to determine the relaying policy, we obtain the following lemmas:

Lemma 1: If there are m relaying nodes in level L_i, there are at least m nodes in level L_{i+1}, and each of the m nodes is connected to one different relaying node in L_i.

Proof: Suppose there are n nodes in level L_{i+1}, and each of the n nodes is connected to one different node in level L_i. If $n < m$, only n relaying nodes are selected out in level L_i. If $n < m$, n nodes must be selected as the relaying nodes in level L_i since all the n nodes in level L_{i+1} are required to be covered. Thus, n can only be equal to m.

Lemma 2: For level L_i, if m relaying nodes in level L_i transmit during m different short-slots, there must be at least c_m^2 nodes in level L_{i+1}.

Proof: According to the broadcast constraints described in section IV-C, if any two of the m nodes in level L_i cannot transmit during the same short-slot, any two of the m nodes in level L_i have at least one same covered node in level L_{i+1}, and the covered nodes in level L_{i+1} are different. Thus, there are at least C_m^2 nodes in level L_{i+1}.

Lemma 3: The nodes that are in level L_{i+1} and are derived from *Lemma* 1 are different from the nodes that are in level L_{i+1} and are *derived* from *Lemma* 2.

Proof: For a node that is in level L_{i+1} and is *derived* from *Lemma* 2, it is covered by at least two nodes in level L_i. If this node can also be *derived* from *Lemma* 1, this node is connected to only one node in L_i; thus, the contradiction arises.

Theorem 1: For the network described above, when Algorithm 1 and Algorithm 2 are employed, the upper bound of the allocated short-slots number, which is needed for broadcasting a configuration packet to the entire network, equals to $(\frac{2N+h-2}{3}, h)$, where N denotes the number of nodes in the network and h denotes the maximum hop-count of the network.

Proof: For any two adjacent levels L_i and L_{i+1}, assuming that there are m relaying nodes in level L_i. If the number of short-slots needed by the nodes in level L_i is maximum, according to *Lemma* 1, *Lemma* 2 and *Lemma* 3, there are at least $C_m^2 + m$ nodes in level L_{i+1}. Assuming that there are m_k relaying nodes in level L_k and that the number of short-slots needed for relaying between any two adjacent levels is maximum. Ignoring the case that a short-slot is multiplexed by the nodes belonging to different levels, the required number of nodes in the network can be given as:

$$n = 1 + m_1 + \frac{1}{2}(m_1 + m_1^2) + \ldots + \frac{1}{2}(m_{h-1} + m_{h-1}^2), \quad (13)$$

where h denotes the number of total levels.

Assuming that the total number of nodes in the network is equal to N, we can thus get

$$1 + m_1 + \frac{1}{2}(m_1 + m_1^2) + \ldots + \frac{1}{2}(m_{h-1} + m_{h-1}^2) \leq N. \quad (14)$$

Since $m_k - 1 \geq 0$, it is true that

$$\frac{1}{2}(m_k + m_k^2) \geq \frac{3m_k - 1}{2}. \quad (15)$$

Then, we obtain

$$1 + m_1 + \frac{3m_1 - 1}{2} + \ldots + \frac{3m_{h-1} - 1}{2} \leq n \leq N. \quad (16)$$

The active control node only requires one short-slot, and there is no relaying node in the last level. Ignoring the case that a short-slot is multiplexed by the nodes belonging to different levels, the maximum number of short-slots required for relaying can be given as:

$$AS'_{max} = 1 + m_1 + m_2 + \ldots + m_{h-1}, \quad (17)$$

thus, it is true that

$$AS'_{max} \leq \frac{2N + h - 2 - 2m_1}{3} < \frac{2N + h - 2}{3}. \quad (18)$$

When the case that a short-slot is multiplexed by the nodes belonging to different levels is considered, the actual number of short-slots needed is smaller than AS'_{max}, which can be given as:

$$AS_{max} \leq AS'_{max} < \frac{2N + h - 2}{3}. \quad (19)$$

Considering the special case that there is only one node in each level, which requires h short-slots for relaying, we can finally give an upper bound of the needed short-slots for relaying:

$$AS_{max} \leq \max(\frac{2N + h - 2}{3}, h). \quad (20)$$

REFERENCES

[1] B.-N. Cheng *et al.*, "Design considerations for next-generation airborne tactical networks," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 138–145, May 2014.

[2] F. Bannour, S. Souihi, and A. Mellouk, "Distributed SDN control: Survey, taxonomy, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 333–354, 1st Quart., 2018.

[3] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proc. IEEE*, vol. 103, no. 1, pp. 14–76, Jan. 2015.

[4] Z. Shanghong, C. Kefan, L. Na, W. Xiang, and Z. Jing, "Software defined airborne tactical network for aeronautic swarm," *J. Commun.*, vol. 38, no. 8, pp. 140–155, Aug. 2017.

[5] A. Tootoonchian, M. Ghobadi, and Y. Ganjali, "OpenTM: Traffic matrix estimator for OpenFlow networks," in *Proc. PAM*, 2010, pp. 201–210.

[6] Z. Su, T. Wang, Y. Xia, and M. Hamdi, "FlowCover: Low-cost flow monitoring scheme in software defined networks," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 1956–1961.

[7] N. L. M. V. Adrichem, C. Doerr, and F. A. Kuipers, "OpenNetMon: Network monitoring in OpenFlow software-defined networks," in *Proc. IEEE/IFTP NOMS*, May 2014, pp. 1–8.

[8] S. R. Chowdhury, M. F. Bari, R. Ahmed, and R. Boutaba, "PayLess: A low cost network monitoring framework for software defined networks," in *Proc. IEEE/IFTP NOMS*, May 2014, pp. 1–9.

[9] A. Yassine, H. Rahimi, and S. Shirmohammadi, "Software defined network traffic measurement: Current trends and challenges," *IEEE Instrum. Meas. Mag.*, vol. 18, no. 2, pp. 42–50, Apr. 2015.

[10] M. Moshref, M. Yu, R. Govindan, and A. Vahdat, "SCREAM: Sketch resource allocation for software-defined measurement," in *Proc. ACM CoNEXT*, 2015, Art. no. 14.

[11] Z. Li, Q. Li, L. Zhao, and H. Xiong, "Openflow channel deployment algorithm for software-defined AFDX," in *Proc. IEEE DASC*, Oct. 2014, pp. 4A6-1–4A6-10.

[12] P. Heise, F. Geyer, and R. Obermaisser, "Deterministic OpenFlow: Performance evaluation of SDN hardware for avionic networks," in *Proc. IEEE CNSM*, Nov. 2015, pp. 372–377.

[13] A. Betances, K. M. Hopkinson, and M. Silvius, "Context aware routing management architecture for airborne networks," *IET Netw.*, vol. 5, no. 4, pp. 85–92, Jul. 2016.

[14] B.-N. Cheng, R. Charland, P. Christensen, L. Veytser, and J. Wheeler, "Evaluation of a multihop airborne IP backbone with heterogeneous radio technologies," *IEEE Trans. Mobile Comput.*, vol. 13, no. 2, pp. 299–310, Feb. 2014.

[15] S. M. Clark, K. A. Hoback, and S. J. F. Zogg, "Statistical priority-based multiple access system and method," U.S. Patent 7 680 077 B1, Mar. 16, 2010.

[16] P. De Mil, B. Jooris, L. Tytgat, J. Hoebeke, I. Moerman, and P. Demeester, "snapMac: A generic MAC/PHY architecture enabling flexible MAC design," *Ad Hoc Netw.*, vol. 17, no. 3, pp. 37–59, Jun. 2014.

[17] E. Haas, "Aeronautical channel modeling," *IEEE Trans. Veh. Technol.*, vol. 51, no. 2, pp. 254–264, Mar. 2002.

KEFAN CHEN received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2013, and the M.S. degree from Air Force Engineering University, Xi'an, China, in 2016, where he is currently pursuing the Ph.D. degree in aerospace network.

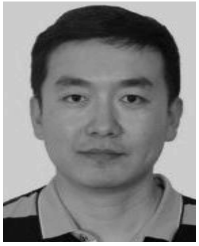
His research interests include airborne tactical network, software-defined networking, and aviation data link system.

NA LV received the B.S. degree in testing technology and instrumentation, the M.S. degree in control theory and applications, and the Ph.D. degree in armament science and technology from Northwestern Polytechnical University, Xi'an, China, in 1992, 1995, and 2010, respectively.

She is currently a Full Professor with Air Force Engineering University, Xi'an. Her current research interests include aviation data link system, military air communications, and software-defined networking.

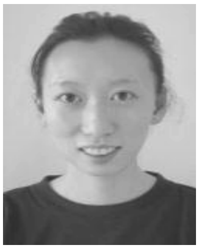
SHANGHONG ZHAO received the B.S. and M.S. degree from the Physics Department, Lanzhou University, Lanzhou, China, in 1984 and 1989, respectively, and the Ph.D. degree from the Xi'an Institute of Optics and Precision Mechanics, CAS, in 1998.

He is currently a Full Professor with Air Force Engineering University, Xi'an, China. He is also the Head of the Space-Based Information Network Research Center. He has authored over 300 academic articles and four books. His research interests include aerospace network, airborne network, quantum key distribution, microwave photonics, and optoelectronics.



XIANG WANG received the B.S., M.S., and Ph.D. degrees in communication and information system from Air Force Engineering University, Xi'an, China, in 2006, 2009, and 2013, respectively.

He is currently a Lecturer with Air Force Engineering University. His research interests include airborne network and software-defined networking.



JING ZHAO received the B.S. and M.S. degrees from Air Force Engineering University, Xi'an, China, in 2011 and 2014, respectively, where she is currently pursuing the Ph.D. degree in communication and information system.

Her research interests include software-defined networking and atmospheric optical communication.