

Deep Neighbor Adaptation (DNA)-based Terahertz Medium Access Control (MAC) for Highly Dynamic Airborne Networks

Linsheng He¹, Fei Hu^{1*}, Zhe Chu¹, Jiamiao Zhao¹,
Nof Abuzaninab², Yalin Sagduyu², Ngwe Thawdar³, Sunil Kumar⁴

¹ Department of Electrical and Computer Engineering, University of Alabama, Tuscaloosa, AL, USA

(*: corresponding author Dr.Fei Hu, fei@eng.ua.edu)

² Intelligent Automation, Inc. (i-a-i), Rockville, MD, USA

³ Air Force Research Lab, Rome, NY, USA

⁴ . Department of Electrical and Computer Engineering, Sab Diego State University, CA, USA

Abstract—The Terahertz (THz) link provides over 100G bps of data transmission rate. However, it is extremely vulnerable to obstacles and requires very narrow (pencil-like) beams for highly focused energy radiation. The main task of THz medium access control (MAC) protocol is not channel access collision avoidance. Instead, the priority task is to coordinate all the one-hop neighbors for scheduled, antenna-aligned line-of-sight communications. This project aims to build an intelligent MAC scheme for highly mobile Terahertz airborne networks (TAN). Our TAN MAC design has three novelties as follows: (1)*Spatio-temporal TAN state learning*: We will build a predictive network state estimation model through deep learning and generative adversarial network (GAN). Based on the predicted node/link status, all the one-hop neighbors can prepare well for the antenna alignment and THz channel scheduling. (2)*Accurate, two-level MAC action generation*: we propose to use nested deep reinforcement learning (DRL) with outer/inner policy loops for high-/low-level action determination: The outer loop determines the high-level, coarse actions (such as antenna codebook selection) and the inner loop determines the low-level, fine actions (such as individual beam control) under the selected high-level action. (3)*TAN-specific, comprehensive MAC behavior control*: Based on the above deep neighbor adaptation (DNA) schemes, we finally produce a complete TAN MAC protocol that considers the routing context and dynamic network topology. Our simulations validate the smooth, high-rate THz communications in MAC layer with resilient links under high mobility.

Index Terms—TeraHertz (THz) links, MAC protocol, Airborne networks, Deep learning, Reinforcement learning

I. INTRODUCTION

Terahertz (THz) communications use the frequency beyond millimeter-wave (mmWave), i.e., 0.1T-10T Hz, to achieve a transmission rate of over 100 Gbps [1]. THz-based 6G system is at least 10 times faster than 5G [2]. By using high-gain antennas, THz links can be used for long-distance ($>500m$) airborne networks. Due to its in-the-air deployment, the THz airborne network (TAN) has much less obstacles than general ground ad hoc networks that have many radio blockages such as trees, vehicles, pedestrians, buildings, etc. Due to its large fading loss, THz node needs to use extremely narrow (pencil-like) antenna beam ($<20^\circ$) [3]. Such well-focused directional transmission can be achieved through reconfigurable antenna arrays with codebook selection to adapt to different fading/mobility environments.

In this paper we investigate the medium access control (MAC) protocol design for TANs. The goal of MAC is to guarantee collision-free, reliable communications among the dynamic one-hop neighborhood.

A. TAN MAC Requirements

THz MAC has some fundamental differences from conventional low-frequency (<15 GHz) MAC schemes. For example, today's Wi-Fi products can simply use omnidirectional antennas (OA) (Fig.1(a)) to radiate signals to 360° range for two reasons: (1) OA is relatively inexpensive compared with directional antenna (DA) (Fig.1(b)). (2) Low-frequency signals have high wavelength and can easily penetrate through many obstacles and reach the destination via different radio paths (either line-of-sight (LoS) or non-LoS). Thus, low-frequency networks do not need to concentrate the signal on a small angle as what DA does.

High-frequency networks (such as mmWave, THz) need to use DAs for two purposes: (1) The well-focused directional transmission can send the signals for a longer distance (thus overcoming the high fading loss for high-frequency signals); (2) DAs significantly reduce the radio interference from neighbors.

THz has higher frequency than mmWave and desires a pencil-like narrow beamwidth. This makes two THz nodes form a highly directional, wire-like RF channel. Such precise directionality significantly reduces the possible RF interference from any neighboring nodes. As shown in Fig.1(c), by using THz links with well-aligned antennas, nearby nodes can simultaneously establish multiple pairs of communications (A_1-B_1 , A_2-B_2 , etc.) with no or very little RF interference.

In the TAN, when both the sender and receiver are equipped with the DAs, the posture of each node (including 3 angles: yaw, pitch, and roll) as well as their relative positions, determine the orientation angle of the DA in the 3-D coordinate system. As shown in Fig.1(d), the yaw angle is used to describe the offset angle in the x-y plane, the pitch angle is used to describe the offset angle in the y-z plane, and the roll angle is used to describe the offset angle in the x-z

plane. Our MAC design will be based on the tracing of the node posture/position.

In this research, we propose an effective THz MAC design (Fig.1(e)) that shifts the design focus from (1) channel access collision avoidance in the conventional MAC schemes, to (2) the precise neighbor coordination in terms of antenna orientation and transmission (Tx)/ reception (Rx) schedule.

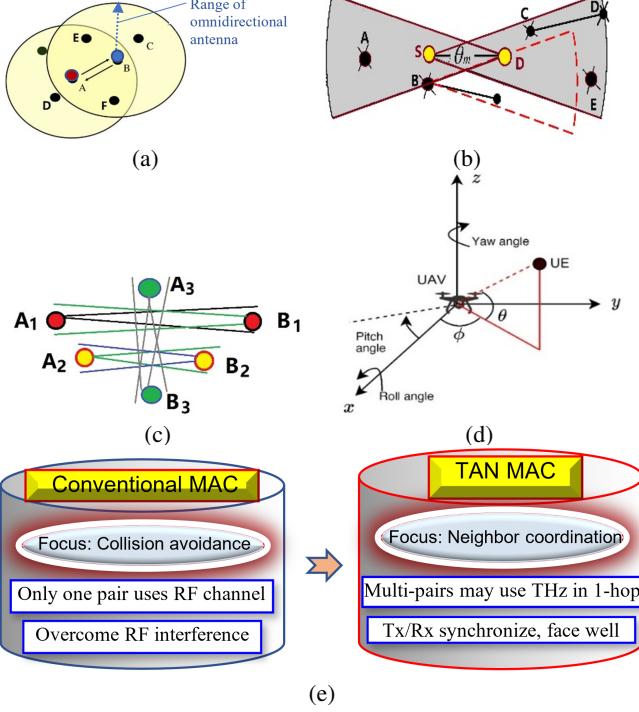


Fig. 1: (a) Low-frequency may use omnidirectional antenna; (b) Directional antenna; (c) THz pencil-like beam; little interference; (d) THz is sensitive to aircraft attitude; (e) Compare conventional MAC to TAN MAC (design focus)

In TAN, although the clear sky environment allows two airborne nodes to easily form a LoS path between them without worrying about many obstacles, there are still some challenges in the MAC design:

First, the high mobility ($>100\text{m/s}$) of an aircraft makes it difficult for a pair of pencil-like beams to align with each other accurately. THz channel is distance-selective, i.e., each distance level of a THz link favors certain THz window. This means that frequent THz channel handoff may occur under high mobility. The mobility also causes frequent network topology changes, which requires MAC protocol trigger new rounds of neighbor discovery and Tx/Rx re-scheduling.

Second, THz channel is also altitude sensitive. As shown in Fig.2 [3][4], the path fading of THz signals gets worse (see the red color) when the altitude decreases. This tells us that TANs prefer high-altitude communications. The MAC scheme must perform smooth THz channel handoff when the nodes fly across difficult altitude levels.

Third, the aircraft's body tilting/rotation actions can cause the THz beamforming to lose the original good performance. This requires that MAC protocol closely traces the aircraft's attitude/position changes and adjusts the beam orientation and data sending rate correspondingly.

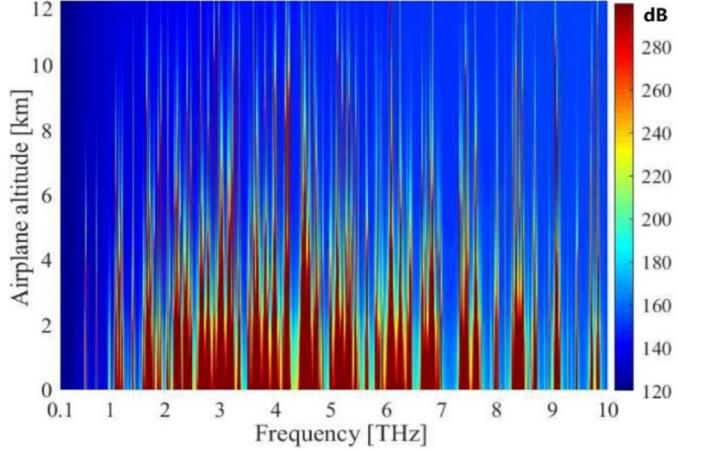


Fig. 2: Path loss between 0.1 – 1 THz.[3][4]

Moreover, since THz has a wire-like directionality, we can arrange multiple concurrent one-hop communications (even with the same THz band) in a MAC neighborhood. This requires careful MAC scheduling of data Tx/Rx, if those one-hop active links belong to different end-to-end routing flows.

B. Novelties of this Research

Novelty 1: Deep neighbor adaptation (DNA) for predictive node coordination during THz MAC control: As mentioned before, in TANs, the focus of MAC design is to coordinate well the antenna beams, Tx/Rx schedule, THz bands, data rates, and other communication parameters. Due to THz's ultra-high data rate, it cannot afford even 1ms of link outage that may cause the loss of hundreds of packets.

Therefore, we propose a predictive, spatio-temporal node state tracking based on deep learning model, called deep neighbor adaptation (DNA). Such a DNA-based tracking can accurately predict the next-time node's position/posture. As shown in Fig.3, our scheme closely tracks not only the mobility-caused altitude/distance changes (from Fig.3(a) to (b)), but also the aircraft body attitude changes (from Fig.3(b) to (c)).

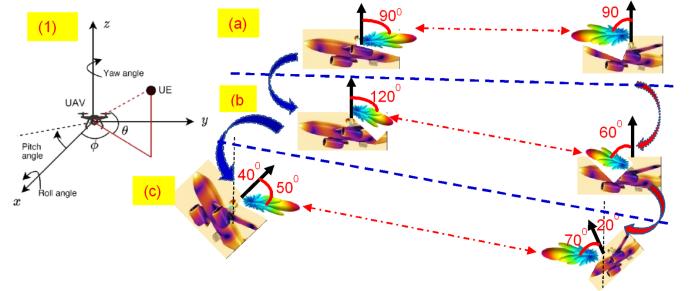


Fig. 3: Impacts of aircraft mobility/tilting on antenna alignment.

Our DNA-based TAN state learning has the following four advantages:

- (1) Comprehensive state vector prediction: Many factors influence the setup of THz communication parameters. Our prediction algorithm can trace a vector of node/link parameters

(through the use of deep neural network (DNN)), including mobility modes, aircraft yaw/pitch angles, link quality, etc.

(2) Scalable network topology pattern identification: With the maturity and popularity of software-defined network (SDN) [5], the control panel of a TAN could easily collect the entire network's parameters on each node/link. Thus, it is desired that the state tracking/prediction algorithms are scalable to any large network with hundreds of (or more) nodes. Our deep tracking model is scalable to a large network.

(3) Spatio-temporal evolutional tracking: DNA-based tracking can capture the spatial patterns (by using graph model to describe the spatial relationship among all nodes) as well as the temporal patterns (by using deep learning to analyze a series of snapshots in different times to capture the evolutional patterns).

(4) Tolerable to limited TAN training dataset: The training set of a deep learning network (DNN)-based TAN may not be complete, considering there are many network scenarios (in terms of node density, mobility modes, mission-based grouping behaviors, THz dynamics in different altitudes, etc. We propose to use generative adversarial network (GAN) [6] to synthesize various TAN communication scenarios.

Based on the above DNA-based network "state" tracking, we further propose to use the nested deep reinforcement learning (DRL) to generate optimal network control "actions" (Fig.4):

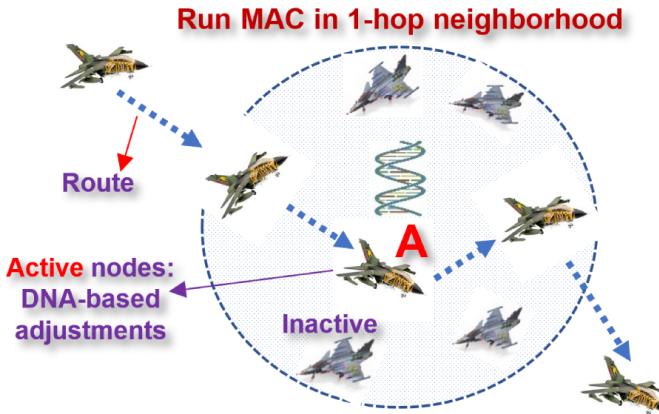


Fig. 4: DNA-based MAC in 1-hop neighborhood.

(1) High-level action (implemented through the outer loop of nested DRL model): Because the THz window is sensitive to altitude, distance, velocity, etc., a high-level action that we need to make is to select the suitable codebook for optimal antenna array structure, which brings the approximate antenna coverage/orientation for a pair of active communication nodes.

(2) Low-level action: Once the high-level (outer) action is made, we can further determine low-level actions by using the inner policy of the nested DRL model. Particularly, since the high-level action only tells the codebook to be used, what gain level should be used for each antenna element? What is the control vector for the antenna array? These settings will be determined by the inner DRL policy.

Our above nested-DRL actions will be applied in each one-hop neighborhood (Fig.4). They should be the high-priority task of TAN MAC layer since they are used for accurate node

coordination to ensure the super fast THz channels to always stay in a ready-for-transmission status.

Novelty 2: Our THz protocol design places MAC parameter control under the context of the end-to-end routing: To reduce MAC protocol overhead and achieve seamless node coordination, we solve the MAC node coordination issue under the context of the end-to-end routing behavior. As shown in Fig.5, the nodes participate in MAC operations in two routing phases:

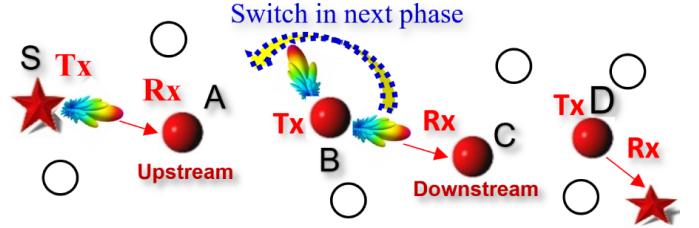


Fig. 5: TAN MAC control depends on routing context.

(1) Initial route seeking phase: During the initial path establishment, all the nodes that received route request (RREQ) will be activated and blindly broadcast RREQ to its one-hop neighbors. Here strict antenna alignment is not required since official THz data transmission does not begin yet and we can use the out-of-band channels (such as Ka/Ku-bands) for long-distance, omnidirectional protocol message exchange.

(2) Data transmission phase: Once the path is established, the routing operations mainly includes (a) THz data delivery, and (b) path repair (due to node mobility).

Operation (a) requires the coordination of neighborhood transmissions and the entire path's Tx/Rx schedules. As shown in Fig.5, due to the half-duplex nature of RF links, we need to use a Tx/Rx-alternating schedule. For example, when the sender (S) is in Tx mode, A is in Rx mode (to receive data from S). However, A cannot talk with B at this time. To avoid resource waste, B can simultaneously send data to C. In the next schedule cycle, A and C need to switch to Tx mode. Such path-wide Tx/Rx schedule coordination should be carefully arranged among each one-hop neighborhood.

Operation (b) requires that MAC protocol always maintains one or multiple 'backup' nodes in 1-hop range, in case the downstream node is not available due to mobility. This means that the DNA-based node coordination should also be performed between active nodes and backup nodes.

Novelty 3: Comprehensive MAC protocol design for highly dynamic, large-scale TANs. Fig.6 shows our proposed TAN MAC framework. As we can see, the long-/short-term memory (LSTM)-based DNN [7] is used to capture the large-scale TAN's spatio-temporal dynamics. The GAN is further used to handle sparsity nature of the network feature matrix and overcome the limited TAN training datasets. The learned TAN node/link "states" are used for node coordination "action" determination (via nested DRL model). Finally, such actions will be combined with other MAC operations (such as Tx/Rx schedule control) to form a complete TAN MAC protocol.

Paper organization: In Section 2, we will first compare our work to existing THz MAC schemes. Then the TAN system

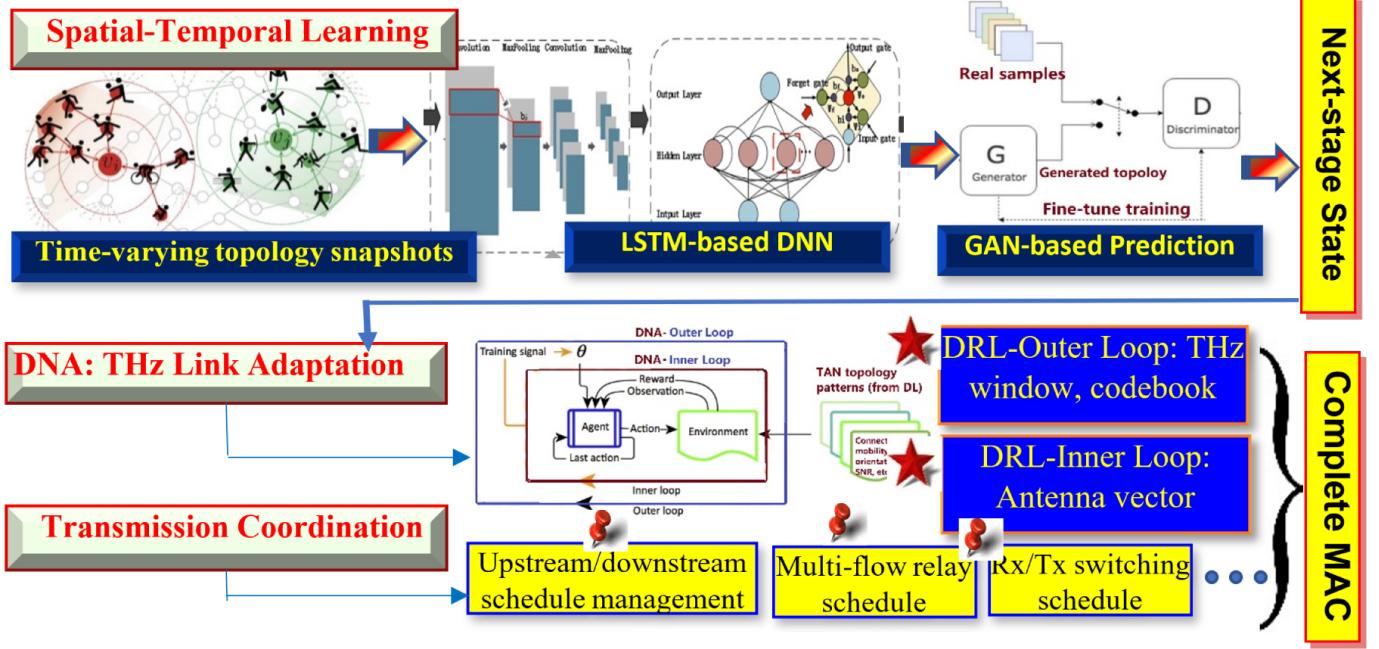


Fig. 6: Proposed TAN MAC Implementation Framework.

model and some assumptions are given in Section 3. The DRL-based action generation model is described in Section 4. The complete MAC protocol will be detailed in Section 5. Performance analysis results are provided in Section 6, followed by the concluding marks in Section 7.

II. RELATED WORKS

We review existing THz MAC schemes based on three taxonomy methods as below:

(1) Centralized vs. distributed: Centralized MAC scheme assumes the existence of a central controller that controls the channel access schedule. A typical scheme is the access point (AP)-based MAC [8] where the AP allocates the channels based on the neighboring nodes' priorities and channel conditions. However, centralized MAC may have difficulties to achieve real-time AP-nodes schedule sharing if the link is over 1km, since the beam alignment becomes difficult during official data transmission.

Distributed MAC avoids the single-point-failure issue by requesting all neighbors to coordinate their transmissions in a distributed style. A typical example is the scheme presented in [9]. The distributed scheme produces more channel access collisions than the centralized scheme.

(2) Random vs. scheduled: Random channel access schemes may use CSMA for THz link access. To reduce the collisions, in [10] a one-way handshake protocol is used to achieve initial coordination. However, when nodes are mobile, RTS/CTS involves much protocol overhead since repeated messages need to be sent out to detect the neighbors' status. Therefore, in [11] it uses a geometry model to estimate the angle-of-arrival (AoA) and then adjusts the beam to face the sender. But this work did not consider high path and absorption loss.

Scheduled channel access requires the use of time slots and priority-aware resource allocation among nodes. The calculation of optimal time slot allocation is often NP-hard.

In [12] a scheme called FTDMA is proposed to use both time slots and sub-bands to achieve resource allocation among THz nodes. A TDMA-based scheme is proposed in [13]. It relies on pulse-level beam-switching and energy control.

A hybrid scheme is proposed in [14] that overcomes the shortcomings of pure random access schemes (which cause many collisions) and pure scheduled schemes (which may not utilize the resources efficiently since some time slots may be wasted).

(3) Transmitter vs. Receiver-Initiated: In [15], a transmitter-initiated communication model is proposed to achieve high-speed vehicle networking. It has considered the channel capacity, autonomous relaying, and network establishment. Although transmitter-initiated schemes are widely used and have lower complexity in distributed applications, the receiver-initiated schemes avoid the tight synchronization requirement between the sender and receiver, and also overcome the deafness issue. A receiver-initiated MAC scheme is proposed in [10]. It uses a sliding window flow control scheme with one-way handshake to increase the channel utilization.

However, all the above THz MAC schemes have the following shortcomings: First, they target either nano-networks or networks with very short link distance (<10m). However, TAN has long link distance (>500m). Such an airborne network requires suitable THz channel model that considers long distance (distance-selective) and various altitude levels (altitude-selective); Second, those MAC schemes still put channel access control as the first priority. However, node coordination is a more important task in THz networks. In Table 1 we compare conventional and our MAC designs.

TABLE I: Compare conventional MAC and our DNA-based MAC scheme.

Comparison items	Conventional MAC	DNA-based MAC
Main MAC design goal	Collision-free, low-energy THz channel access control	Neighborhood coordination management for well-aligned THz Tx/Rx
Assumed antenna models	Omni-directional or directional antennas with certain beamwidth	Highly directional, extremely narrow beamwidth (pencil-like)
Main concerns	Hidden/exposed terminal problems; overheat interference from neighbors	Antenna misalignment; May not use in-band (THz) for protocol message exchange
Intelligent algorithms used	Seldomly use AI algorithms; Reactive response to special events	Use scalable, high-precise deep learning and DRL algorithms for proactive response
Performance Metrics	Throughput, channel collision rate	THz channel outage probability; THz window switching delay; Data relay delay

III. TAN SYSTEM MODEL AND ASSUMPTIONS

A. Out-of-band and In-band Communications

As mentioned before, THz has strict requirements for antenna alignment. In a particular communication task, most TAN data (video, sensor data, etc.) can be transmitted within seconds. This means that as long as two nodes maintain DNA-based coordination, THz transmission can be finished without outage. In this project we propose to run DNA algorithms continuously in the background to keep track of the nodes' states in real-time to support the end-to-end route maintenance.

Unlike the data transmission, the protocol message exchange has three main features: (1) It occurs frequently in the background, such as the broadcasting of RTS (request-to-send) and CTS (clear-to-send), neighbor discovery messages, etc. It is used to closely monitor the neighbors' states such that the system is always ready for any incoming data transmission task. (2) Its traffic amount is small (the protocol messages have very short packet size, typically less than 100 bytes each). If using THz channel for such tiny traffic load, we can significantly waste the huge THz bandwidth. (3) Unlike THz data transmission which only occurs between a pair of nodes, the protocol messages are often exchanged among multiple neighbors. For example, a node needs to run neighbor discovery protocol to detect all one-hop neighbors; during route seeking process, a node may help to flood RREQ messages; etc. (4) It may require long-distance ($>10\text{km}$) communications in a large airborne network. For example,

in Software defined network (SDN)-based [18] or other hierarchical airborne networks, an aircraft may need to report its status to a center/satellite

Therefore, a reasonable strategy is: Do Not rely on vulnerable THz links for network protocol operations. Only use THz (in-band) for data transmissions and use other lower frequency bands (such as Ka-band) for protocol message exchanges. Those lower-frequency bands (out of THz band) can use omnidirectional antennas to easily reach all surrounding neighbors.

B. Network architecture

Our proposed DNA scheme can handle large-scale network state learning. This is especially useful to the SDN that is an important 5G management model due to its easy-to-reprogram nature, which allows the entire network easily to reconfigure its protocols based on different network conditions. SDN-based architecture (Fig.7) is especially suitable to TAN dynamics adaptation by using (1) global node/link state sensing/collection (see Fig.7, '1. Monitoring'), and (2) AI and machine learning in the control plane (which can be run in a satellite or a central node).

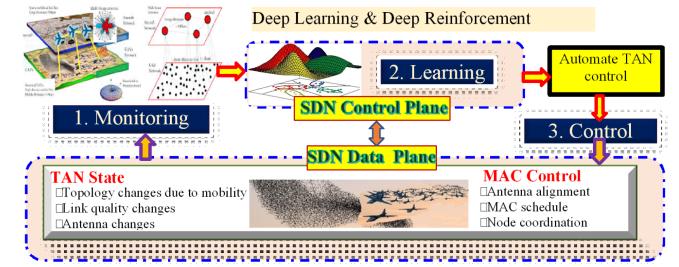


Fig. 7: (Global mode Architecture 1) If using SDN, the control panel can run all AI algorithms.

If the TAN does not have a global SDN-based management model, it may adopt the distributed, cluster-based architecture (Fig.8(a)), which selects some nodes as the cluster heads (CHs). Some border nodes may serve as the gateways. The communications may occur between CH-to-CH, members-to-CH, or gateway-to-CH. In fact, the above cluster-based architecture has been popularly adopted in many protocols. For example, OLSR [16] uses some special nodes as multi-point relays (MPRs). Those MPRs have similar role as CHs. In cluster-based architecture, our proposed DAN algorithms can be run in the CHs. Each CH can collect the state parameters from all of its cluster members.

If the cluster-based architecture is not available, a mini-version of our DNA algorithm can be run only in a node, which closely traces the status of each neighbor (Fig.8(b)).

On the collection of aircraft status parameters: Today an aircraft is equipped with many sensors (such as GPS receiver, accelerometers, compass, RF detectors, etc.) to collect the status information, including the 3D position, velocity, tilting angle, RF signal quality, distance to each neighbor, etc.

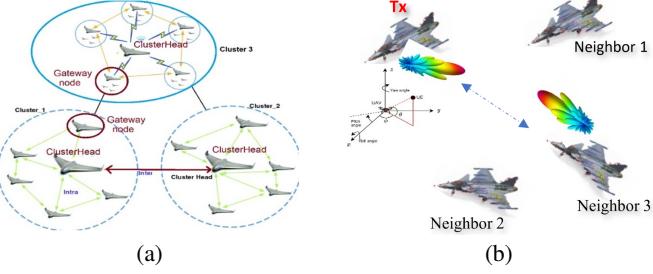


Fig. 8: (a) Cluster mode: Cluster-based distributed management; (b) Local mode: Runs AI in neighborhood

IV. DEEP NEIGHBOR ADAPTATION (DNA)

A. GAN-enhanced DL for Predictive Situational Awareness (PSA)

As mentioned before, THz channel is extremely sensitive to antenna misalignment. If we wait for the detection of THz link outage and then adjust the transmission rate, a large number of packets may have been lost. Therefore, the accurate prediction of node position/posture (both determine the antenna orientation changes) is important to the earlier preparation for the link degradation. Particularly, spatio-temporal prediction is beneficial to TAN MAC control due to two reasons:

(1) **Spatial** learning/prediction: The entire network can be represented as a graph with nodes (vertexes) and links (edges). The spatial changes directly impact on the MAC performance. As shown in Fig.9, the spatial changes could be node position changes. In Fig.9(1) to (2), node D slightly moves down, which causes the antenna misalignment. Thus, THz antennas must be adjusted.

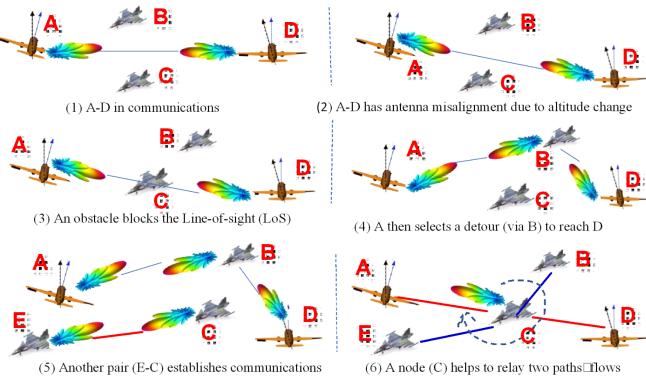


Fig. 9: Snapshots of graphs for different situations of a dynamic TAN.

The spatial change can also cause the change of communication nodes. For example, from Fig.9(2) to (3), the movement of C causes the loss of line-of-sight (LoS) between A and D. Thus, A has to resort to another node to reach D. Here MAC layer must activate a detour link, for example, A→B→D.

Apparently, we should not just wait for the occurrence of spatial changes and then perform antenna adjustments. Using Fig.9(2)→(3) as an example. The loss of LoS between A and D cannot be detected until a timer-out event comes when A is waiting for D's ACK feedback. By that time a large amount

of data has been lost (>100 packets/ms in THz channel). Therefore, it is critical to predict the spatial changes. For example, if we predict that A-D link will be blocked by another node, we can gradually decrease the data rate in the link A-D, meanwhile a detour link can be established via a backup node (here it is B).

(2) **Temporal** learning/prediction: Single spatial snapshot can only tell us the current network status without the evolutionary patterns. If we take a series of snapshots (i.e., graphs) of the TAN network states, we will be able to learn the continuous network changes. Based on the temporal patterns, we can further predict the next-phase network snapshot.

In this research, we aim to predict the spatio-temporal features based on the historical snapshots. Particularly, we focus on the following 6 parameters that are critical to DNA operations (Fig.10): (1) 3D position information (X, Y, Z). This can be easily obtained via GPS. If GPS is not available, other methods such as received signal strength (RSS) or angle-of-arrival (AoA) information [17], may be used to obtain the approximate position information. (2) Yaw angle (ψ), pitch angle (θ), and roll angle (φ). Those 3 angles reflect the aircraft body attitude changes. For example, the roll angle reflects the tilting scope of two wings (up and down), while pitch angle can change when the aircraft's head is elevated. The change of the yaw angle means that the aircraft is rotating in X-Y plane. Any change of the above 3 angles can cause the antenna misorientation. (3) Link quality information: while the above (1) and (2) indicate an individual node's state, (3) shows the link characteristics between a pair of nodes. It consists of BER (bit error rate), and load (traffic rate). The link load measures how much capacity of the link is utilized.

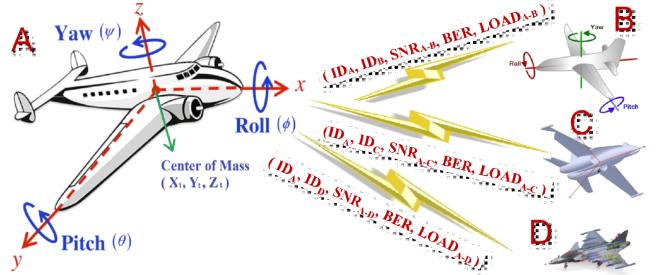


Fig. 10: Graph parameters (Position, tilt angles, link quality, etc.).

If using a graph model, we can mark the above (1) & (2) as the node attribute, and (3) as the edge attribute (i.e., weights) between two nodes. If we use a feature matrix to describe a network snapshot, each row of the feature matrix has the information of an edge in the graph, which includes (ID₁, ID₂, edge attribute, ID₁'s attribute (position/attitude), ID₂'s attribute (position/attitude)).

The feature matrix (\mathbf{F}) of each snapshot captures the **spatial** information of the TAN topology at a particular time. The past T snapshots' graph models (**temporal** information) become the input of the deep learning (DL) model, which is enhanced with long-/short-term memory (LSTM) units (Fig.11).

Note that although some conventional matrix-based methods (such as ARMA models [18]) have been used for network state prediction, they are mostly based on linear computations and

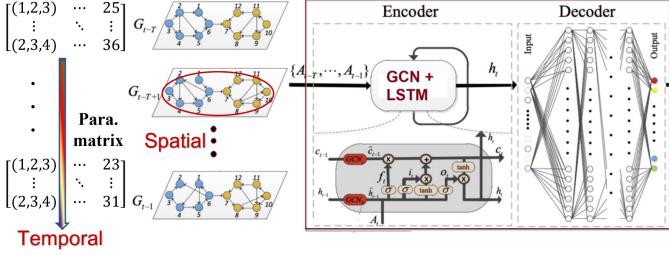


Fig. 11: Spatio-temporal Deep Learning for TAN evolutionary pattern extraction.

thus cannot capture the nonlinear characteristics of topology dynamics. Moreover, when the SDN is used for the entire TAN management, the conventional prediction algorithms has poor scalability and stability. LSTM-based DL can overcome those drawbacks. Particularly, LSTM can handle sequential data (here we use a series of snapshots) with different time intervals and long-term dependency. This helps to capture the time-evolving network patterns.

The prediction problem can be defined as: Given the previous T snapshots, $\{A_{(\tau-T)}, A_{(\tau-T+1)}, \dots, A_{(\tau-1)}\}$, as well as the current snapshot A_τ , our goal is to predict the next snapshot:

$$\overline{A}_\tau = f(A_{(\tau-T)}, A_{(\tau-T+1)}, \dots, A_{(\tau-1)}), \quad (1)$$

here $f(\cdot)$ is a nonlinear function of DL model.

For the convenience of calculation, we use a connectivity matrix $\mathbf{C} \in \mathbb{R}^{N \times N}$ to show the existence of the direct link or not. If two nodes are too far away from each other, there is no connectivity there ('0' in the matrix position). Both matrices \mathbf{C} and \mathbf{F} serve as inputs to LSTM-GCN (graphical convolutional network). The fully connected output layer generates the following output (\mathbf{Y}) [19]:

$$Y = \text{LSTM} - \text{GCN}(\mathbf{C}, \mathbf{F}) \quad (2)$$

$$= g\left(\mathbf{W} \cdot \mathbf{F} \cdot \left(\widehat{\mathbf{D}}^{-\frac{1}{2}} (\mathbf{C} + \mathbf{I}_N) \widehat{\mathbf{D}}^{-\frac{1}{2}}\right)\right) \quad (3)$$

Here LSTM-GCN is the DL model, $g(\cdot)$ is the activation function, \mathbf{W} is the weight matrix of the DNN, each element of \mathbf{D} has the format of $\widehat{\mathbf{D}}_{ii} = \sum_{j=1}^N \widehat{\mathbf{C}}_{ij}$, and $\widehat{\mathbf{D}}^{-\frac{1}{2}} (\mathbf{C} + \mathbf{I}_N) \widehat{\mathbf{D}}^{-\frac{1}{2}}$ is the approximated filter of GCN. Here \mathbf{I}_N is the N -dimensional identify matrix.

Note that we use one LSTM-GCN for any snapshot. The learned outcomes (from the output layer) $Y = \{Y_{(\tau-T)}, Y_{(\tau-T+1)}, \dots, Y_{(\tau-1)}\}$ will be fed into the LSTM modules for evolving patterns capture. The LSTM architecture encapsulates a cell with multiplicative gate units. For LSTM model we assume that the gates and the memory cell (m) has their corresponding parameter sets (weights, bias) as follows:

The input gate i has (W_i, b_i) , output gate i has (W_o, b_o) , forget gate f has (W_f, b_f) , the memory cell \mathbf{m} has (W_m, b_m) . Also let h_{t-1} represent the output embedding matrix at timestamp (t-1), and h_t is the obtained output embedding matrix at the current time t . Denote $\sigma(\cdot)$ as the sigmoid activation function, which limits the gate output within $[0, 1]$.

1]. Let x_t mean the input vector at time t . Then we have [7]:

$$\mathbf{i}_t = \sigma(W_i [h_{t-1}, x_t] + b_i) \quad (4)$$

$$\mathbf{f}_t = \sigma(W_f [h_{t-1}, x_t] + b_f) \quad (5)$$

$$\mathbf{o}_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \quad (6)$$

Note that in LSTM the gates i and f provide partial information for current state cell \mathbf{m}_t . Likewise, the output gate \mathbf{o} and \mathbf{m}_t provide partial information for output embedding matrix \mathbf{h}_t . Let m' represent an intermediate variable. We have [7][20]:

$$m' = \tanh(W_m [h_{t-1}, x_t] + b_m) \quad (7)$$

$$\mathbf{m} = \mathbf{f}_t \times \mathbf{m}_{t-1} + \mathbf{m}' \quad (8)$$

$$\mathbf{h}_t = \mathbf{o}_t \times \tanh(\mathbf{m}_t) \quad (9)$$

Although the above LSTM-GCN model can capture the evolutionary long-term patterns, it has three shortcomings: (1) It can only reflect the features of the historical snapshots. To predict the next-time snapshot, we need to form an iterative process: using historical evolutions to make a prediction of the next snapshot, then the ground-truth snapshot can be used to further correct the prediction error. In other words, we need to form a two-player game, one is the generated snapshot, the other one is the ground-truth-based discriminator, which is used to ‘push’ the generator to keep improving its prediction accuracy. (2) LSTM is typically trained by using the mean square error (MSE) loss function. However, MSE cannot handle the matrix sparsity issue [20]. The connection matrix \mathbf{C} may have sparsity nature since THz is more suitable to short-distance (one-hop) communication (the nodes which are not neighbors cannot establish THz link directly). (3) MSE cannot well handle the wide value range issue [20]. The feature matrix \mathbf{F} may have wide-scope values.

To overcome the above three issues, here we use GAN with adversarial training to ‘push’ the generator (\mathbf{G}) (from LSTM-GCN) to keep improving its prediction accuracy by using the discriminator (\mathbf{D}), as shown in Fig.12. Here \mathbf{D} alternatively takes \mathbf{G} ’s output (\overline{A}_τ) or the ground-truth A_τ as the input. To further improve the GAN training stability, we use Wasserstein GAN (WGAN) [21] which can effectively avoid ‘useless’ \mathbf{G} improvement[22].

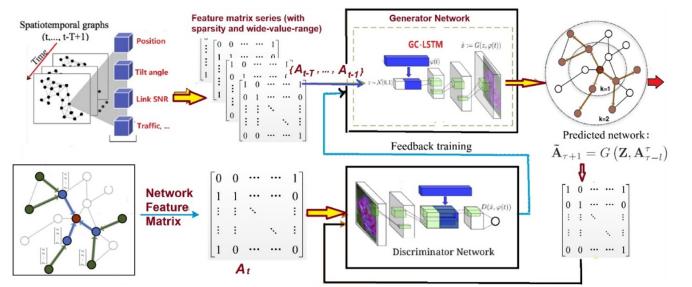


Fig. 12: Generative Adversarial Network (GAN) for Next-Phase TAN topology prediction.

B. Nested DRL for Precise Neighbor Adaptation (PNA)

To achieve precise node coordination, we need to distinguish between two types of actions:

(1) **High-level (coarse) action:** Some actions can determine the coarse behavior of the THz node. For example, a specific antenna codebook may only tell the approximate main lobe shape and the number of on/off antenna elements. However, the detailed adjustment of each antenna element (orientation, gain, etc.) is unknown yet. (2) **Low-level (fine) action:** Once the above high-level action is determined, we need to further make some fine-level adjustments, such as the detailed setup of each antenna element. This action can make the antenna reach the proper combination of gain, which can not only reduce the antenna power loss, but also minimize the exclusive region (ER) [23] generated by the sender.

We thus propose a nested DRL (Fig.13) to determine the above two-level actions. In outer policy, we use codebook to set up the “aggregated single-direction coverage” without much side lobe effect (Fig.14(a)). Moreover, we can easily adjust the beamwidth/orientation by changing codebook. In neighbor discovery or 1:M multicast communications, we can use antenna array to achieve a multi-point communication effect (Fig.14(b)). Typically, when two aircraft are far away from each other, the beam will be more pencil-like. Antenna array can be easily reconfigured to achieve pencil-like gain coverage (Fig.14(c)).

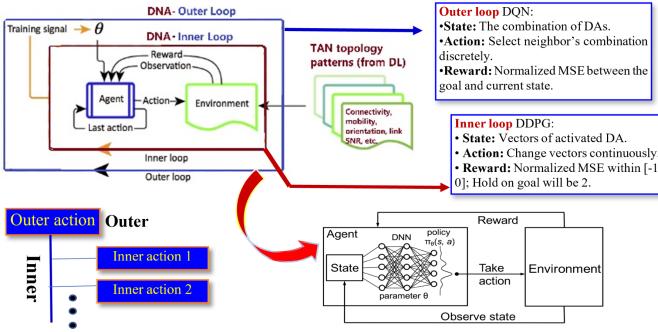


Fig. 13: Nested DRL for big/small action determination.

1) **Outer DRL Policy:** The main task of the outer policy is to distinguish among different combinations of activated antennas. We have implemented Deep Q-learning Network (DQN) that consists of convolutional neural network (CNN) and Q-learning. DQN uses stochastic gradient descent method with a value function to evaluate the future returns. Assume that a node is equipped with an antenna array with at least four DA elements, each of which has at least 4 states, i.e., facing the East, North, West, or South. The DRL can generate at least four actions: a node can make its antenna go Up, Left, Down or Right (assume a 2-D case).

An example environment of the outer DRL policy is shown in Fig.15. The 4 x 4 grids represent all cases of DA combinations, and four rows and four columns represent the antenna elements 1~4, separately. The black blocks represent the invalid combinations of antennas; the yellow circle represents the goal of the combination. In the example of Fig.15, it means that the North antenna and West antenna should be used in the next step. The red block represents the current state of combination. In this figure, it means that only the South antenna is active.

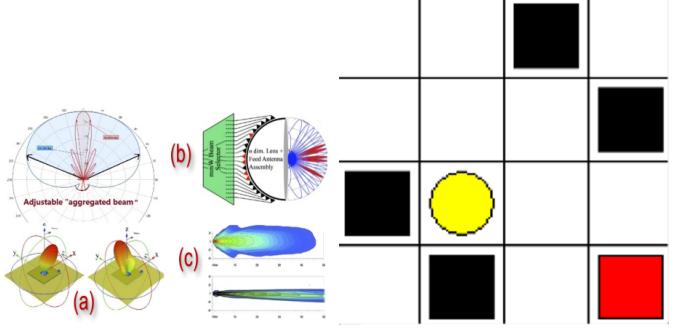


Fig. 14: Benefits and impacts of codebook selection. Fig. 15: Demo of Outer DRL Policy.

The loss function of the outer DRL policy is as follows,

$$\text{Loss}(\theta) = E \left[(Q_t \text{target} - Q(s, a|\theta))^2 \right] \quad (10)$$

$$= E \left[\left(r + \gamma \max_{a'} (Q(s', a'|\theta^-)) - Q(s, a|\theta) \right)^2 \right] \quad (11)$$

where r is the old reward, γ is the reward discount per time-step, θ^- is the parameter of the target network in DQN, whereas θ is the parameter of the evaluation network in DQN. This loss function is determined based on Q-Learning model. Our goal is to make the difference between the q-target and q-eval as small as possible, and then update the parameters based on the methods such as stochastic gradient descent. However, this loss function cannot intuitively reflect the optimization performance of the model, and we propose a new parameter to reflect the ability of the model to reflect the distance between the current state and the goal state. The cost of the outer DRL is the accumulated, reduced average value of the squared difference between two normalized tensors (i.e., the goal and current state). Its equation is as below:

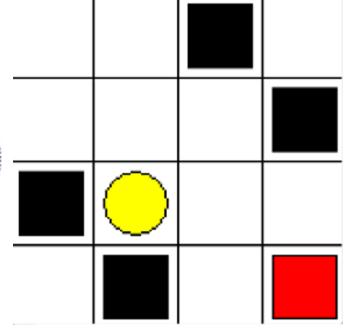
$$\text{Cost} = \frac{\sum (N_{\text{current}} - N_{\text{goal}})^2}{s} \quad (12)$$

where the s is the total number of steps of movements before achieving the goal, N_{current} and N_{goal} are two tensors with the information of antenna combinations.

In Fig.15, we set the reward value to -10 when the current state is on the black block; in other cases, the reward value is 0 except when the current state is the goal (for which the reward value is 1). The specific equation of the reward function is as below:

$$\text{Reward} = \begin{cases} 1, & \text{current state reaches the goal} \\ -10, & \text{reaches an unavailable state} \\ 0, & \text{for all other cases} \end{cases} \quad (13)$$

2) **Inner DRL Policy:** The inner policy loop is conducted after the outer policy is determined. While the Outer DRL tells how many numbers of antennas should be activate, the inner policy further determines exactly how much gain that each beam should use. Because the task of inner DRL is to control the gain of the sender and receiver's antennas continuously until finally it reaches the goal, this task could have infinite actions. The discrete action format used by the outer policy



cannot be used here, since it assumes that the probability of n finite actions is generated from the Softmax output layer in the DNN.

To overcome the above issue, we use deep deterministic policy gradient (DDPG) [24] algorithm for the inner policy model. DDPG is a classical DRL model for continuous action prediction. It is not based on the probabilistic updates of behaviors, but on specific behaviors. DDPG also uses the DQN structure, but it uses a dual network – the actor network and critic network. The former aims to maximize the expected reward, and the latter updates the error between the reality and the actor's estimate.

For the Critic network, the loss function is similar to the DQN in terms of mean square error:

$$\text{Critic Loss}(\theta) = \frac{1}{m} \sum_{i=1}^m (y_i - Q(S_i, A_i, \theta))^2 \quad (14)$$

where θ is the critic parameters, S_i , A_i are the state and action from Actor network, y_i is the Q-value calculated by critic target network, Q is the evaluation function in critic. The actor's loss is defined below. The larger the obtained feedback Q value is, the smaller the loss in critic will be, vice versa.

$$\text{Actor Loss}(\theta') = -\frac{1}{m} \sum_{i=1}^m Q(S_i, A_i, \theta') \quad (15)$$

where θ' is the actor parameters, S_i , A_i are the state and action, Q is the evaluation function in the actor. An example of the inner RL environment is shown in Fig.16. The red stick represents the current state of the gain of combined antenna vectors. Regardless of which DAs are active, we restrict that the blue dot (8x8 pixels of square) is located in quadrant I, which symbolizes the combined gain of the two antennas. Because the DA's beamwidth is less than 20° in THz networks, we further restrict the blue point can change from 10° to 80°. This is because at 0°-10° and 80°-90°, only one DA is activated; between 10° and 80°, the combined gain of two antennas is available. The original state of the combined vector is in the center of the green sector (its radius is 150 pixels). Fig.16(a) and (b) illustrate a case that the blue point is far away from the original state. The DDPG model can be trained to reach the goal area by changing the angle. In (c) and (d), the situation is opposite to the above case.

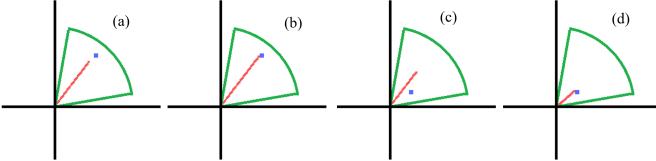


Fig. 16: Example settings of inner RL environment.

The reward function depends on the accumulated normalized MSE of the goal and current antenna vectors (-1-0). If the current vector reaches the goal, the accumulated reward will increase by 2. The equation of the reward function is shown below:

$$\text{Reward} = \begin{cases} \sum -\gamma \times \sqrt[2]{\Delta d_N^2}, & \text{out of the goal area} \\ 2, & \text{in the goal area} \end{cases} \quad (16)$$

where d_N is the normalized distance from the end point of the current state to the center of the goal, γ is the reward discount value.

V. TAN MAC PROTOCOL

In this section, we will describe the MAC operations, which consist of PSA and PNA protocols.

A. PSA Protocol

Our proposed PSA (Predictive Situational Awareness) protocol can be executed in global, distributed, or individual modes, depending on the network routing architecture. Since this paper focuses on MAC layer design, we assume that the routing layer can run either SDN-based global management mode (in which the control panel can collect the node/link information from the data panel), or cluster-based protocols (in which the CH collects each member node's state), or just uses flatten topology (a node just runs DL-based PSA for its one-hop neighbors). The PSA protocol is shown in Protocol 1.

Protocol 1: PSA (Predictive Situational Awareness)

INPUT: Network management modes, node parameters (position (x, y, z), yaw (ψ), pitch (θ), roll(φ)). Link parameters (SNR or BER). Need those parameters for the past T time slices, as well as *ground truth* of the snapshot in each time slice.

OUTPUT: The predicted snapshot of the next time slice.

Protocol operations: Check network management mode first:

- (1) **IF SDN mode** (Mode = '*Global*'):
 - a Each node periodically reports its state parameters (via ka/ku-band etc.) to the SDN controller.
 - b The control plane builds the graph for each time slice, with connectivity matrix C and feature matrix F .
 - c Controller runs LSTM-GCN to obtain the pattern vector (P) of each snapshot.
 - d Controller collects the latest ground-truth snapshot S and configure GAN discriminator (D).
 - e Controller uses P (as the generator G) and S to run GAN to output prediction result.
 - f D 's output is fed back to G input for feedback training purpose; S is also fed back to D .
 - g Controller uses two ways to timely tell nodes the prediction results: (i) Directly set up flow table rules for each node based on topology changes; (ii) sends the needed partial (1-hop) snapshot to each node.
- (2) **IF Cluster mode** (Mode = '*Distributed*'):
 - a Each node is notified its role (clusterhead (CH), members, gateway). CH manages its own cluster. Members/gateway report states to their own CH. CH maintains C & F & ground truth for its own cluster.
 - b CH runs LSTM-GCN-GAN to predict its own next-time snapshot and tell its members.
 - c Gateway nodes help to forward prediction results between CHs for inter-cluster communications.
- (3) **IF flat mode** (Mode = '*Individual*'):
 - a Each node runs LSTM-GCN-GAN to predict its own next-time snapshot.

- a Each node only collects node/link states from MAC neighborhood (1-hop). Or, in the minimum version, a sender only runs LSTM-GCN-GAN for its receiver(s).
 - b Sender & receiver share the prediction results for antenna alignment.
 - c Sender is aware of 1-hop neighbors' states for backup node selection purpose (if existing link is broken).
-

B. PNA Protocol

After the above PSA protocol is used to predict the next-time network topology and each node/link's *state*, eventually certain *action* needs to be adopted. Actions are generated from Precise Neighbor Adaptation (PNA) protocol (see *Protocol 2*), which uses nested DRL algorithm to determine the antenna codebook and beam orientation. Such a protocol can be easily extended to include more actions, such as THz band selection.

Protocol 2: PNA (Precise Neighbor Adaptation)

INPUT: PSA outcomes (predicted next-time-slice snapshot), including each node's status.

OUTPUT: THz window, antenna codebook, antenna vector.

Protocol operations: Both sender and receiver obtain the PSA results on each other's states.

* Offline training of Deep Reinforcement Learning (DRL) – assume done before protocol execution. The training dataset includes various scenarios with different parameters (such as distance d , link SNR, altitudes, etc.) and the corresponding optimal actions (*selecting THz channels, codebooks, etc.*);

* Online DRL (real-time) for action determination.

(1) Outer policy:

- a Set up reward function as the antenna activation combinations, measured by feature matrix.
- b Generate the activation combination by the prediction results of node position and antenna attitude.
- c Use the results to update the accumulative reward values in the entire DRL policy.
- d Run DRL outer state/action to select the optimal combination that minimize the selection time.

(2) Inner policy:

- a Set up reward function as the MSE of the one or two antenna gain vectors.
 - b Revise reward as normalized SNR loss in terms of the number of antennas (not full beamforming gain). Thus, inner policy is independent of the THz path loss and the number of active antenna elements.
 - c Using the PSA predicted snapshot and outer actions to determine the exact beam steering angle of each antenna element. The adjustment of antennas orientation can maximize the normalized reward.
 - d Update the beam vector based on the adjusted beam angle; Also update the normalized reward value.
-

C. Comprehensive TAN MAC Scheme

The comprehensive TAN MAC should not only meet general MAC functionality requirements (such as channel

access control, flow control, etc.), but also need to consider the special features of THz mobile networks, such as the initial 'activation' of THz links when receiving a new ROUTE_SEEK request.

1) **MAC for Upstream/Downstream switching:** A node in the path (except the source and destination) needs to periodically switch its communication destinations between the upstream (Up) and downstream (Down) node due to its half-duplex nature. It is critical for a node to coordinate well with its Up/Down nodes. Here we use one example to illustrate this point. Suppose each packet has a size of 10K bytes (THz's super-fast link favors a larger packet size instead of typical 1,500 bytes). Assume the THz link provides 100Gbps of data rate. Suppose a node (V in Fig.17) switches from Down node W to Up node A. However, A is still communicating with S, and it may take 1ms to switch to V. Without such a time waste, V could have used the 1ms of waiting time to send >10,000 packets to W.

Therefore, in an end-to-end THz path, all the nodes should follow a strict transmission (Tx) / reception (Rx) switching schedule. The tide-like alternating Rx/Tx modes should be carefully coordinated. As shown in Fig.17, the sender is always in Tx mode. Note that S could still use omnidirectional antenna to receive any protocol-related messages via non-THz channel from each path node, in case any path node wants to report a special situation (such as link outage). The destination D is always in Rx mode. But the middle nodes should follow the Tx/Rx switching schedule that can be sent from the source. Once a node switches to a Up or Down node, the previous PSA/PNA protocols will be executed to ensure the proper antenna alignment.

To achieve seamless Tx/Rx coordination, each node exchanges the UP_DOWN_INFO message with its Up/Down nodes. Such a message has the following packet format (Table 2): (1) *Up/down*: Indicate the receiver of this message, which can be the upstream node, downstream node, or even a backup node (in case that the path needs to be repaired). (2) *Tx/Rx sequence position*: Its value is determined by the source. It tells when each path node switches to Rx or Tx mode. A node may be assigned with "odd or even" schedule: 'odd' means "going to Tx in odd time windows (#1, #3, #5, ...); (3) *Tx/Rx duration*: It indicates the stay time length for each Tx/Rx phase. It should be a fixed value in each routing task. Suppose the antenna's switching time is 1ms. The Tx/Rx duration may choose to be 10ms. If it is too long, it may cause each node to hold too many packets in the queue; if it is too short, it could cause frequent antenna switching. (4) *THz band ID (a channel in [0.1T, 10THz])*: It comes from PNA protocol and will be used for the channel ID when a node communicates with a Up or a Down node. (5) *Alignment (Predicted)*: This field stores the predicted snapshot (the node state) from PSA protocol. (6) *Alignment (Ground-truth)*: This field stores the latest available ground-truth alignment status, which will be used by the PSA algorithm (serving as the input to the GAN discriminator) to predict the future snapshot. (7) *Other fields*: Such as link outage report, node queue status (for MAC flow control purpose), etc.

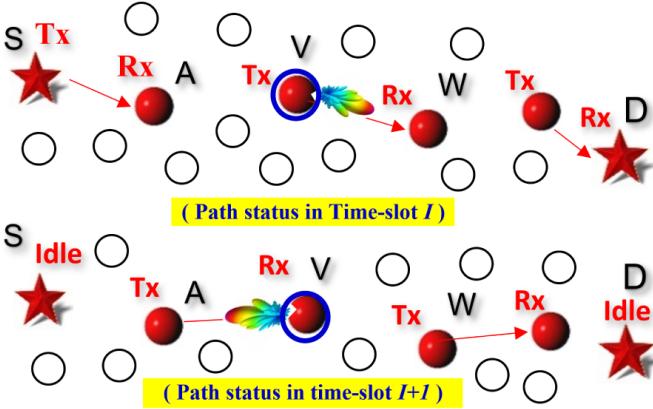


Fig. 17: Coordinated Tx/Rx control.

TABLE II: UP_DOWN_INFO message format.

Up / Down	TX / RX Seq.	Tx / Rx duration	THz window	Alignment (Predicted)	Alignment (Ground-truth)	Other fields
2 bits	8 bits	16 bits	4 bits	16 bytes	8 bytes	...

2) **Flow control:** Although in the transport layer the congestion control function such as TCP can effectively avoid the situation such as traffic being stuck in a specific node, it relies on end-to-end, multi-hop feedback and performs sending rate control in the *source node only*. This may not be suitable to TAN conditions due to two reasons: (1) TAN has super-fast date rate (>100 Gbps). Any single link can get congested easily. If we rely on end-to-end, cloud-like congestion control (TCP sees the entire network as a ‘cloud’ and only relies on source-destination interactions for congestion control), it is difficult to respond to the congestion in a timely manner. (2) TAN nodes have high mobility and could trigger frequent end-to-end congestion control scheme. A more effective way is to control the flow pace in each individual THz link. Such “nip congestion in the bud” solution could avoid avalanche effect by performing flow control in the MAC layer. In fact, flow control is one of the main functions of MAC protocols and is particularly important to THz links. TAN requires *proactive* flow control between each upstream-downstream pair since it cannot afford even 1ms of packet loss (>1000 packets with the size 10K bytes) after queue overflow occurs. The queue status of each node can be easily added to the previous DNA feature matrix (F) to achieve accurate prediction of Up/Down nodes’ queuing status.

“Queue + Link” evolution-based flow control model: Our MAC layer flow control scheme not just uses the queue status as the direct indication of congestion status, but also closely monitors the link quality degradation trends: If the link deteriorates at a fast rate, the sending rate should be decreased more aggressively. Particularly, we propose the following flow rate control model:

$$r_{t+1} = \begin{cases} r_t - \Delta r, & \rho(t) \text{ is relatively constant or shows Gaussian changes} \\ \frac{r_t}{\kappa \xi(t)}, & \xi(t) > 1 \text{ in the past but with stable value, here } \kappa > 1 \\ \frac{r_t}{e^{\kappa \xi(t)}}, & \text{exponential decrease when } \xi(t) > 1 \text{ and has an increasing trend} \end{cases} \quad (17)$$

Here r_{t+1} is the rate at time $(t+1)$, $\xi(t)$ is the link quality decrease rate, $\rho(t)$ is the measure of link quality (here we use weighted sum of SNR and BER), κ is the flow control sensitivity. Note that in (17) we adopt *linear* rate decrease when the queue of the receiver is filling up but the link quality is still satisfactory. This case occurs when the sender’s traffic rate goes high in a short term. In the second case, we use *multiplicative* decrease if the link quality shows a linear decrease. In the third case, the exponential decrease is used to handle the fast decay of link quality, which likely indicates an incoming link failure (due to node mobility or complete antenna misalignment).

Notifying the source: Since congestion control is never a single-hop task due to the “hop-to-hop” congestion accumulation effect from the source to the congestion location, the flow control state should be fed back to the source along the reverse path. The anti-congestion capability of the entire path depends on the weakest point since the congestion in any node can slow down the entire path. Therefore, once heard about the congestion in any node, the source should trigger the ultimate rate throttling by controlling its sliding window, or it launches a new path with less congestion possibility. Since the end-to-end congestion control belongs to the *transport* layer issue, we will leave it for our future research.

3) Overall MAC Protocol: The overall TAN MAC operations are described in Protocol 3.

Protocol 3: Complete THz MAC Protocol in Mobile Airborne Networks

INPUT: (1) Past T snapshots of the TAN topology (either SDN-global, cluster-based, or Tx/Rx only). Each snapshot includes node status (position, 3 angles, mobility mode, queue, antenna attributes, etc.) and link status (SNR, BER, etc.). (2) The ground-truth snapshot in the latest time slice. (3) Routing protocol specifications.

OUTPUT: MAC behavior control specifications.

Protocol operations: Assume out-of-band is used for protocol operations. In-band (THz) is used for data only.

- (1) **Initial MAC operations** - occurs during the route-seeking phase
 - a If requested, a node helps to blindly broadcast RREQ message, without worrying THz node coordination issue;
 - b If notified (by the routing source) as one of the path nodes, it records upstream (Up)/Downstream (Down) IDs;

- c Run any wireless neighbor discovery scheme (such as [26]) to find 1-hop neighbors (may serve as backup nodes);
- d If in SDN-global or cluster-based architecture, each node automatically receives the info. on their neighbors;
- i Otherwise, the node runs PSA and PNA protocols with both Up and Down nodes to maintain node coordination;
 - ii Only runs PSA with its backup nodes (1-hop away), no PNA is needed at this time.
- e The node establishes and updates the “Neighbor Table” based on PSA/PNA outcomes.
- (2) **Route maintenance for official THz data transmission phase:** Note: The protocol always runs in the background via out-of-band. It monitors the path status to prepare link repair in MAC layer.
- a Up/Down switching:
 - i The source sends wave-like Tx/Rx schedule and rate/queue suggestions to each node.
 - ii The node performs THz channel switching if necessary, during Up/Down switching.
 - iii The node also notifies the Up node about its queue status and Downlink quality dynamics.
 - b Flow control: MAC layer uses the “traffic pacing” in each single hop to help with whole-path congestion control.
 - i Each path node uses queue status, predicted link quality from PSA protocol and ground-truth (current) quality to adjust its flow rate via equation (17).
 - ii To avoid congestion accumulation effect, the congested node notifies the source on rate pacing.
 - c To avoid high-cost entire-path repair, the node always tries to execute local repair first.
 - i If a node predicts (via PSA) its downlink will go down soon, it issues BACKUP_SEEK message.
 - ii If a proper backup node is found, it will play the role of the original Down node.
 - iii To avoid high packet loss during transfer, the node slowly decreases its sending rate.
 - iv The backup node runs PSA/PNA with its Up and Down nodes, and sets up proper queue size and rate.
 - v The node completely switches to backup node, and notifies the source on such a switch.
 - d Collision Avoidance: Although node coordinate is #1 task in THz MAC, collisions can still happen.
 - i Suppose a pair of THz link has been built between Tx1 and Rx1. Another new pair (Tx2 and Rx2) needs to be carefully set up to avoid interference with existing Rx. Mainly we want to make sure that Tx2 do not interfere with Rx1. Tx2 needs to avoid the “interference zone” where it can interfere with Rx1.
 - ii By using PSA/PNA results broadcasted by the SDN center, Tx2 gets to know Rx1/Tx1’s state parameters.
 - iii Tx2 calculates Rx1’s “interference area” and avoids such a range.
- (3) **Other MAC operations during route re-establishment and tear-down.**
- a If too many local repairs occurred, the entire path may not be optimal anymore. The source then broadcasts a RE_ROUTE message, which will be heard of by all path nodes and their neighbors.
- b The nodes that receive the RE_ROUTE in the earliest time (among all nodes with the same no. of hops from source) will be the path nodes in the new route.
- c Then each new path node runs PSA/PNA with neighbors.
- d If ROUTE_DONE message is received, all nodes stop PSA/PNA and other MAC operations.
-

VI. PERFORMANCE ANALYSIS

A. (GCN+GAN)-based Prediction of Aircraft Position/Posture

To validate the accuracy of GCN+GAN based aircraft state prediction, we have synthesized the dataset including aircraft location, acceleration, attitude, and RF links. Table 3 shows the prediction model’s parameters used in the simulations.

TABLE III: Parameters of GDL-GAN for PSA algorithm.

LSTM window size	Input feature	LSTM features	GCN features	Output features
10	1	128	512	64
Batch size	High mobility	Low mobility	GAN optimizer	
64	300 m/s	200 m/s	RMSprop	

Fig.18(a) shows the time-continuous trajectories of different aircraft and their attitudes in 3D space. The realistic aircraft trajectories can be described by smooth-turn or Gaussian-Markov model [25], and thus look smooth (no fast velocity changes). Fig.18(a) shows our simulated aircraft trajectories; Fig.18(b) is a zoomed-in version of (a). As one can see, we also simulated the posture of each aircraft (two wings).

The prediction results of the GCN+GAN for PSA protocol are shown in Fig.18(c) and (d). In the figures, the five curves represent the time-continuous aircraft trajectory data. The discrete points at the end of each curve represent the predicted positions. As we can see, they match with the ground-truth data very well.

B. Antenna Adjustment based on Nested DRL algorithms

Decision cost: Fig.19 shows the convergence curve of the algorithm cost in the Outer DRL model. In the cost model (mentioned in Section 4.2), the normalized MSE demonstrates the performance of the DRL training. Here Y-axis is the cost, X-axis represents the total number of learning steps. In one episode, the training model may learn once or several times. Thus the total number of learning steps is larger than that of episodes. The DRL parameters are shown in Table 4.

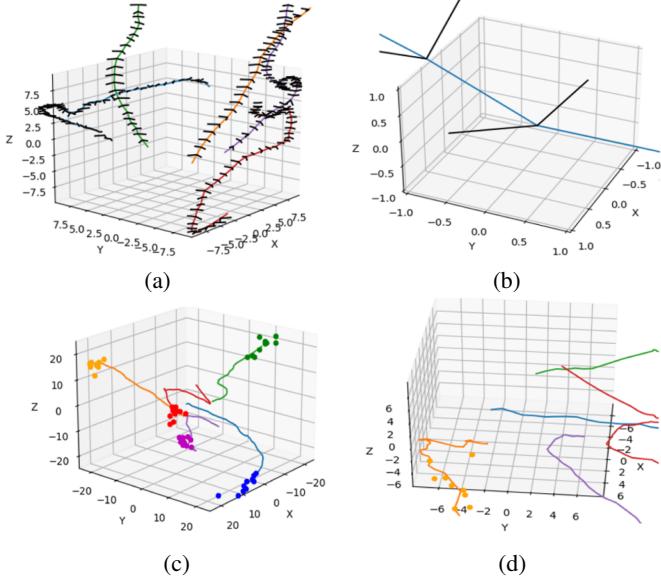


Fig. 18: (a) Synthesized aircraft trajectory/posture; (b) Zoomed-in trajectory/posture; (c) The predicted aircraft positions (dots); (d) Zoomed-in predictive trajectory

TABLE IV: Parameters of nested DRL algorithms.

Outer	Learning rate	Reward decay	Greedy rate	Max learning steps
RL	0.01	0.9	0.9	50
	Training episodes	Testing episodes	Optimizer	
Inner	Actor learning rate	Critic learning rate	Reward decay	Batch size
	0.001	0.001	0.9	32
RL	Max steps per episode	Training episodes	Testing episodes	Optimizer
	300	1000	100	Adam

In the example shown in Fig.19, we can see that after step 3300, the training model reaches the close-to-minimum cost, and there is a growing trend in the subsequent steps. But our simulation still continues to train more episodes. This helps to eliminate overfitting issue and increase model robustness.

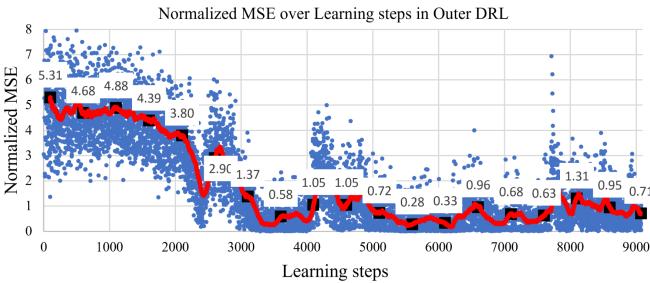


Fig. 19: Cost (MSE of actions) over Training Episodes in Outer DRL.

If we use general traversal method to reach the optimal

DA open/closed state, on average we need to try at least 3.5 actions. But in our nested DRL model, we only need an average of 3.0 action changes to find the optimal antenna element on/off combination during the DRL testing phase. Such a gain of nearly 15% allows the nodes with antenna array to more quickly automate the task of turning on/off antenna elements.

Fig.20 shows the reward change trend in the Inner DRL algorithm. The reward should converge to the value of 20 since the best action of combined vectors should reach the goal area within 10 steps, with 2 reward points per step. From Fig.20 we can see that for the first 150 epochs, the reward is negative, which means that the model did not reach the goal yet. After epoch 300, the reward reaches around 20.

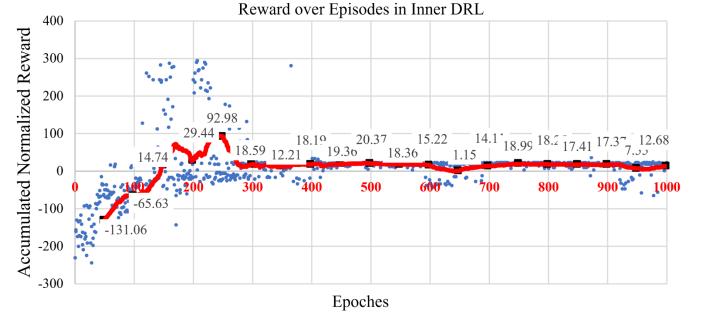


Fig. 20: Accumulated Normalized Reward over Episodes in Inner DRL.

C. QOS Performance

The prediction algorithm contains the GCN-GAN for PSA and the nested DRL for PNA. The high and low mobility are specified in Table 3. Other simulation parameters are shown in Table 5.

TABLE V: Parameters of MAC layer simulation.

Data size	Tx/Rx period	Time resolution	Buffer size	Buffer Tx rate
10 Gb	2 ms	0.1 ms	0.5 Gb	10000 Gbps
Link capacity	Batch size	Sending rate incremental	Numbers of trials	
			20	

We have simulated 4 cases. Case (I) is the best scenario in which all nodes have low-to-medium mobility levels (50-100m/s) and use (GCN+GAN)-based prediction algorithms and nested DRL for precise antenna control. Case II is similar to Case I, except that we also considered the RF interference from neighboring THz communication links. Such interference typically comes from the partial antenna coverage overlap between two nearby THz links, although such interference occurs rarely due to THz link's semi-wire characteristics. Case III also adopts prediction algorithm; but the nodes have high mobility (100-200m/s). In this case it becomes more difficult to predict the aircraft posture and adjust the antennas. Case IV

is the worst scenario because it does not use prediction model and nodes have high mobility.

We will consider three QoS metrics, that is, delay, BER (bit error rate), and throughput. The delay refers to the end-to-end latency that includes all transmission delay (= data amount / link rate), queuing delay and Tx/Rx switching time. Here we consider 5 7 hops from the source to destination, which reflects the path lengths of most practical applications. The packet loss mainly comes from two sources: (1) Queue overflow: This is the main source of packet loss. Due to the large data rate of THz channel, any stuck in a queue of the path can cause a big data loss. Certainly we can set up a longer queue to avoid data loss; but that strategy could cause long queuing delay. (2) Bit errors: Due to the large THz channel fading (the THz signal quickly fades away with the distance getting longer), some bits can be damaged (which can be detected by error detection codes). Those damaged packets need to be discarded in the receiver side. Moreover, the high mobility and antenna misalignment may also significantly worsen the channel quality and cause more bit errors. Our simulations have considered the above factors.

Note that the antenna operates in half-duplex mode. Thus even though we set up the maximum capacity of a THz channel as 100 Gbps, in reality the entire path can only achieve a throughput between 0-50 Gbps. When we calculate the throughput, we only collect the data amount received in the destination. Since this is MAC layer test, we do not consider the cases of packet re-transmissions, which typically occurs in the transport layer.

The delay performance is shown in Fig.21. As we can see, the best scenario (Case I) has almost 6 times less delay than the worst scenario (Case IV). The use of deep-learning-based aircraft position/posture prediction provides the nodes with accurate antenna adjustment, which significantly speeds up the data transmission and clears up all queues more quickly. Thus the delay is greatly reduced.

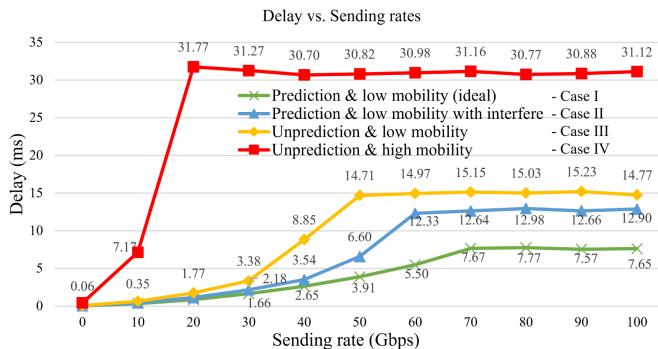


Fig. 21: Delay in MAC simulation.

Fig.22 shows the BER performance. The worst scenario (Case IV) does not adopt the prediction algorithms. Thus it responds to the antenna misalignment only when the ACKs from the receiver are not received. Often it is too late to perform antenna adjustment when a large number of packet loss events have occurred. As we can see, Case IV has over 10 times larger data damage rate than Case I. When the sending rate is low (less than 30Gbps), the first 3 cases have similar

low BER, which is due to the light traffic load (thus no queue overflow).

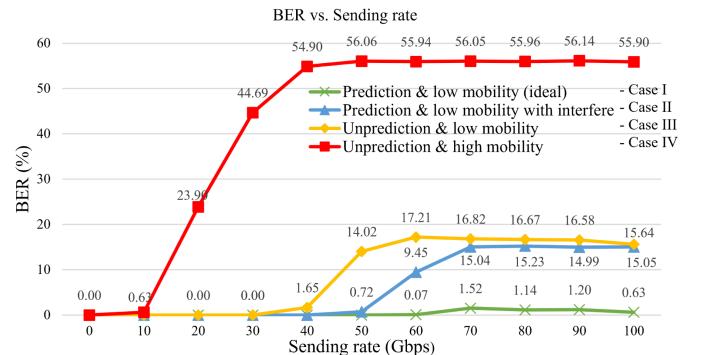


Fig. 22: BER in MAC simulation.

Fig.23 shows the throughput performance. It matches the above two results well. Due to half-duplex and unavoidable fading loss in each hop (the total path has 5 7 hops), even the best scenario (Case I) can only reach 33 Gbps, which is 66% of the theoretical value (50Gbps). Without the prediction model Case IV can only reach 8 Gbps after entering saturation phase.

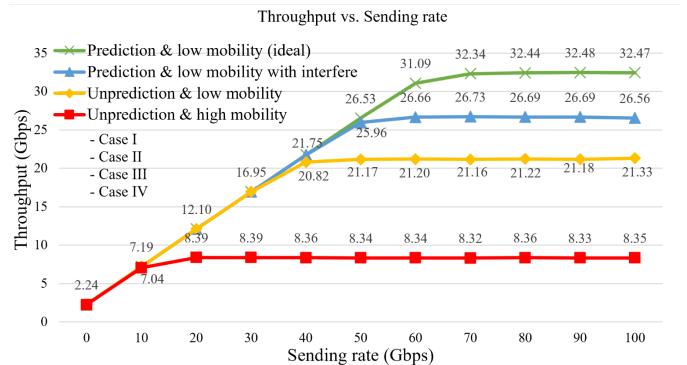


Fig. 23: Throughput in MAC simulation.

VII. CONCLUDING MARKS

In this research, we have conducted the pioneer studies on the THz airborne network MAC layer design, based on the DL prediction model for tracing the node posture/trajectory changes. Moreover, we place the MAC design under the context of the entire route setup because the THz links require coordinated Tx/Rx activities among all path nodes. Our simulations show that, with prediction and low mobility and no interference, the TAN has the best QoS performance. This is because all the hops in the end-to-end path can operate in nearly full capacity status, under the timely node state prediction and accurate antenna adjustment. In the future work, we will further investigate the application of distributed federated learning for the intelligent THz MAC/Routing across-layer design.

REFERENCES

- [1] Z. Chen, X. Ma, B. Zhang, Y. Zhang, Z. Niu, N. Kuang, W. Chen, L. Li, and S. Li, "A survey on terahertz communications," *China Communications*, vol. 16, no. 2, pp. 1–35, 2019.

- [2] S. Koenig, D. Lopez-Diaz, J. Antes, F. Boes, R. Henneberger, A. Leuther, A. Tessmann, R. Schmogrow, D. Hillerkuss, R. Palmer *et al.*, "Wireless sub-thz communication system with high data rate enabled by rf photonics and active mmic technology," in *2014 IEEE Photonics Conference*. IEEE, 2014, pp. 414–415.
- [3] K. Guan, H. Yi, D. He, B. Ai, and Z. Zhong, "Towards 6g: Paradigm of realistic terahertz channel modeling," *China Communications*, vol. 18, no. 5, pp. 1–18, 2021.
- [4] J. Kokkonieniemi, J. M. Jornet, V. Petrov, Y. Koucheryavy, and M. Juntti, "Channel modeling and performance analysis of airplane-satellite terahertz band communications," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2047–2061, 2021.
- [5] D. Kreutz, F. M. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, 2014.
- [6] M. Durgadevi *et al.*, "Generative adversarial network (gan): a general review on different variants of gan and applications," in *2021 6th International Conference on Communication and Electronics Systems (ICCES)*. IEEE, 2021, pp. 1–8.
- [7] T. He and J. Droppo, "Exploiting lstm structure in deep neural networks for speech recognition," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 5445–5449.
- [8] C. Han, W. Tong, and X.-W. Yao, "Ma-adm: A memory-assisted angular-division-multiplexing mac protocol in terahertz communication networks," *Nano communication networks*, vol. 13, pp. 51–59, 2017.
- [9] X.-W. Yao and J. M. Jornet, "Tab-mac: Assisted beamforming mac protocol for terahertz communication networks," *Nano Communication Networks*, vol. 9, pp. 36–42, 2016.
- [10] Q. Xia, Z. Hossain, M. Medley, and J. M. Jornet, "A link-layer synchronization and medium access control protocol for terahertz-band communication networks," in *2015 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2015, pp. 1–7.
- [11] W. Tong and C. Han, "Mra-mac: A multi-radio assisted medium access control in terahertz communication networks," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.
- [12] Z. Li, L. Guan, C. Li, and A. Radwan, "A secure intelligent spectrum control strategy for future thz mobile heterogeneous networks," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 116–123, 2018.
- [13] J. Lin and M. A. Weitnauer, "Pulse-level beam-switching mac with energy control in picocell terahertz networks," in *2014 IEEE Global Communications Conference*. IEEE, 2014, pp. 4460–4465.
- [14] L. You, Z. Ren, C. Chen, and Y.-H. Lv, "An improved high throughput and low delay access protocol for terahertz wireless personal area networks," *Journal of Computers*, vol. 28, no. 3, pp. 147–158, 2017.
- [15] C. Zhang, K. Ota, J. Jia, and M. Dong, "Breaking the blockage for big data transmission: Gigabit road communication in autonomous vehicles," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 152–157, 2018.
- [16] F. De Rango, M. Fotino, and S. Marano, "Ee-olsr: Energy efficient olsr routing protocol for mobile ad-hoc networks," in *MILCOM 2008-2008 IEEE Military Communications Conference*. IEEE, 2008, pp. 1–7.
- [17] S. Wang, B. R. Jackson, and R. Inkol, "Hybrid rss/aoa emitter location estimation based on least squares and maximum likelihood criteria," in *2012 26th Biennial Symposium on Communications (QBSC)*. IEEE, 2012, pp. 24–29.
- [18] Y. Sun, R. Wang, B. Sun, W. Li, and F. Jiang, "Prediction about time series based on updated prediction arma model," in *2013 10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*. IEEE, 2013, pp. 680–684.
- [19] Z. Li, G. Xiong, Y. Chen, Y. Lv, B. Hu, F. Zhu, and F.-Y. Wang, "A hybrid deep learning approach with gcn and lstm for traffic flow prediction," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2019, pp. 1929–1933.
- [20] X. Lv, Z.-L. Wang, Y. Ren, D.-Z. Yang, Q. Feng, B. Sun, and D. Liu, "Traffic network resilience analysis based on the gcn-rnn prediction model," in *2019 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE)*. IEEE, 2019, pp. 96–103.
- [21] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *International conference on machine learning*. PMLR, 2017, pp. 214–223.
- [22] G. Bouallegue and R. Djemal, "Eeg data augmentation using wasserstein gan," in *2020 20th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*. IEEE, 2020, pp. 40–45.
- [23] L. X. Cai, L. Cai, X. Shen, and J. W. Mark, "Rex: A randomized exclusive region based scheduling scheme for mmwave wpans with directional antenna," *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 113–121, 2010.
- [24] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.
- [25] M. Zhao and W. Wang, "Wsn03-4: A novel semi-markov smooth mobility model for mobile ad hoc networks," in *IEEE Globecom 2006*. IEEE, 2006, pp. 1–5.