NTU GICE

Newsletter

Graduate Institute of Communication Engineering, National Taiwan University

Vol. 13, No. 3 Sep. 2022 http://www.comm.ntu.edu.tw gicenewsletter@ntu.edu.tw

GICE Honors

Prof. Kun-You Lin

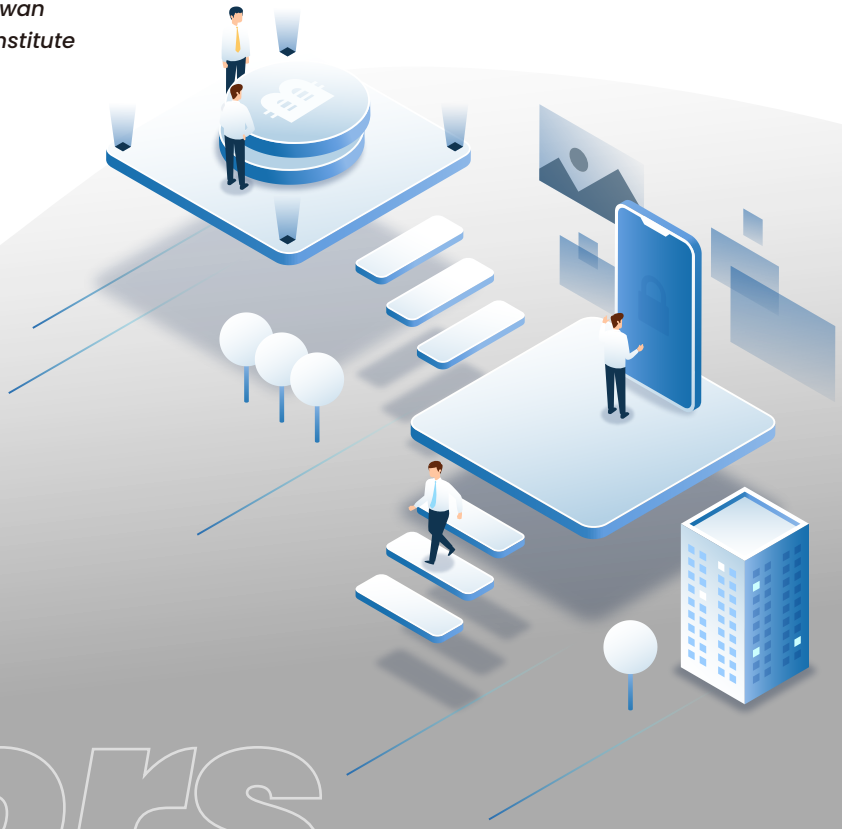
Best Chip Design Award (RF/Microwave Group), Taiwan Semiconductor Research Institute 2022

Prof. Huei Wang

Outstanding Chip Design Award (RF/Microwave Group), Taiwan Semiconductor Research Institute 2022

Prof. Chun-Hsing Li

Outstanding Chip Design Award (RF/Microwave Group), Taiwan Semiconductor Research Institute 2022



GICE Honors

Provable Verification Bounds for Convolutional Neural Networks Under Adversarial Example Attack

Prof. Pei-Yuan Wu

Associate Professor, Dept. Electrical Engineering & Graduate Inst. Communication Engineering, National Taiwan University

I. INTRODUCTION

Deep neural networks (DNN) have demonstrated great success in numerous applications. However, research community [1] has reported that neural networks are often prone to adversarial examples, where only a small amount of perturbation in data may lead the neural network to make false predictions, and seriously degrades its classification performance. As illustrated in Figure 1, a few stickers on the stop sign may fool the DNN-based traffic sign recognition model to make false predictions as speed limit sign. This in no doubt raises serious concerns if one would to rely on such a recognition model in self-driving car applications.

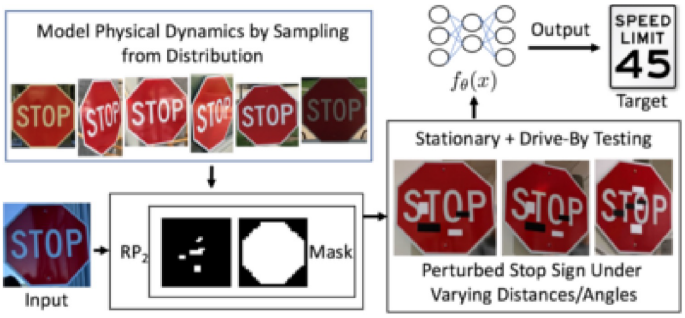


Figure 1: Adversarial example attack towards self-driving car traffic sign recognition system [1]

How can us evaluate the resilience of a DNN against adversarial example attack? More specifically, how to develop algorithm which evaluates a lower bound on the accuracy of a given DNN classifier under adversarial example attack of any kind?

Method:

It turns out that optimization theory can be applied to analyze such accuracy lower bounds [2]. To be specific, consider a classifier which predicts the class an image x belongs to as $g(x)=\text{argmax}_I f_i(x) \in \{1,...,L\}$, where $f(x)=(f_1(x),...,f_L(x)) \in \mathbb{R}^d$ is a DNN. Given that the true label for x is $y^* \in \{1,...,L\}$, the adversary may wish to fool the classifier to make a false prediction $y^{targ} \neq y^*$. Towards that end, the adversary aims to add a noise Δx into the image x, so that $f_{y^{targ}}(x+\Delta x) \geq f_{y^*}(x+\Delta x)$ for which the classifier makes a wrong prediction $g(x+\Delta x) \neq y^*$. The added noise Δx should be small, say with norm at most ϵ, so that others cannot easily detect such noise being added. On the other hand, if one can show that $f_{y^*}(x+\Delta x) - f_{y^{targ}}(x+\Delta x) \geq 0$ for each $y^{targ} \neq y^*$ whenever $\|\Delta x\| \leq \epsilon$, then one can verify that the adversary cannot fool the classifier no matter what noise Δx is added to such image x as long as Δx is within the adversary budget ϵ.

As illustrated in Figure 2, the aforementioned verification problem can be formulated as a constrained minimization problem. One thus resort to the weak duality in optimization theorem, namely each dual feasible solution yields a lower bound to the original minimization problem. Since the original problem is non-convex due to complex nature of DNN, convex relaxation tricks can be applied to transform the original minimization problem into a convex optimization problem at the cost of a loosened lower bound, but allows the strong duality theorem to work by which complementary slackness conditions can be further applied to simplify the dual problem to the form of a DNN, referred as dual network, which can be easily computed. This allows us to solve the verification problem efficiently.

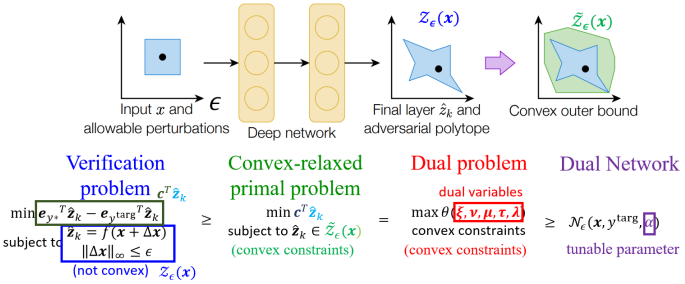


Figure 2: Evaluate accuracy lower bound under adversarial example attack with optimization theory framework.

Experiments:

We extend Wong et. al.'s work [2] originally designed for the verification of ReLU-based fully connected neural networks to convolutional neural networks (CNN) with both ReLU and maxpool activation functions. Compared to state-of-the-arts verification algorithms [3] [4] [5], our proposed verification algorithm CAPM is capable of yielding the tightest verification lower bound with significantly less computation cost under most scenarios, especially when the CNN is of larger scale when conventional verification algorithms either yields trivial bounds or is computationally infeasible. One such comparison example is illustrated in Figure 3.

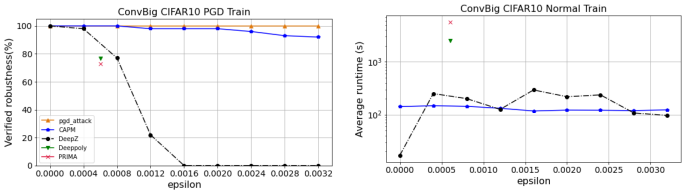


Figure 3: Comparison of the verified accuracy lower bound (left figure) and computation runtime (right figure) among various verification algorithms. The orange line indicates the accuracy under PGD adversarial example attack [3], which no lower bound should exceed.

Conclusion:

Optimization theorem can be applied to verify whether or not a DNN is resilient to adversarial example attack. The accuracy lower bound of a DNN is guaranteed through weak duality theorem, while with convex relaxation tricks and strong duality theorem one can further implement the verification algorithm efficiently. Experiment results demonstrated the effectiveness of such verification algorithms on CNNs, especially of larger scales, with both ReLU and maxpool activation functions.

References

[1] K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno and D. Song, "Robust Physical-World Attacks on Deep Learning Visual Classification," in IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018.
[2] E. Wong and J. Z. Kolter, "Provable defenses against adversarial examples via the convex outer adversarial polytope," in Proceedings of the International Conference on Machine Learning (ICML), 2018.
[3] G. Singh, T. Gehr, M. Puschel and M. Vechev, An abstract domain for certifying neural networks, Proceedings of the ACM on Programming Languages, 3(POPL), 1-30, 2019.
[4] G. Singh, T. Gehr, M. Mirman, M. Puschel and M. Vechev, Fast and effective robustness certification, Advances in neural information processing systems, 31, 2018.
[5] G. Singh, T. Gehr, M. Püschel and M. T. Vechev, Boosting robustness certification of neural networks, In International conference on learning representations, 2018, September.
[6] A. Madry, A. Makelov, L. Schmidt, D. Tsipras and A. Vladu, "Towards Deep Learning Models Resistant to Adversarial Attacks," in International Conference on Learning Representations (ICLR), Vancouver, BC, Canada, 2018.

Prof. Pei-Yuan Wu

Major Research Areas:
machine learning, active authentication, estimation theory, smart manufacturing

Abstract:
Dr. Wu, Pei-Yuan is an assistant professor at NTUEE since 2017. He received the B.S.E. degree from NTU in 2009, and the Ph.D. degree from Princeton University in 2015, all in EE department. His research interest lies in machine learning, privacy and security, computer vision, and cyber-physical system modeling.

A Compact and Low-Cost THz SoP Heterogeneously-Integrated Platform

Prof. Chun-Hsing Li

Associate Professor, Dept. Electrical Engineering & Graduate Inst. Communication Engineering, National Taiwan University

I. INTRODUCTION

Terahertz (THz) science and technologies have been considered as a potential candidate for the next sixth-generation (6G) wireless communication systems which are expected to provide a data rate higher than 100 Gb/s [1]. Moreover, THz technologies can be used for many interesting applications, including concealed weapons and explosives detection, biomedical imaging, and the failure inspection of semiconductor packages [2-3]. In this work, our developed compact and low-cost THz system-on-package (SoP) heterogeneously-integrated platform is introduced for THz system integration. With such a heterogeneous platform, different circuit modules implemented in particular technologies can be integrated into the same platform to realize high-performance, low-cost, and high-integration THz systems for the aforementioned versatile applications [4-6].

II. THZ SOP HETEROGENEOUSLY-INTEGRATED PLATFORM

Fig. 1 shows the proposed compact and low-cost THz SoP heterogeneously-integrated platform where digital baseband processors, power management circuits, analog front-ends, RF front-ends, THz front-ends, antennas, and filters realized in particular technologies are integrated on a common low-cost integrated-passive-devices (IPD) carrier through compact and low-loss chip-to-IPD interconnects. Hence, the advantages of these technologies can be exploited to carry out high-performance THz systems.

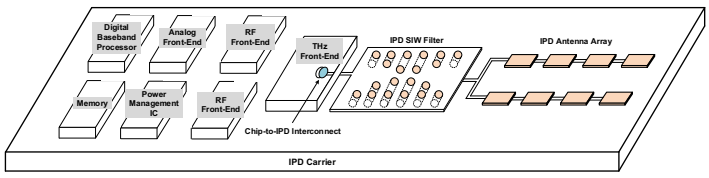


Fig. 1. Proposed THz SoP heterogeneously-integrated platform.

To make the heterogeneous integration feasible, we must have low-loss THz interconnects to connect the chip module to the carrier. Fig. 2 illustrates the proposed ac-coupled and directly-connected THz interconnects supporting single-band and dual-band, and broadband operations, respectively [4]. The broadband interconnect directly connects the chip's ground-signal-ground (GSG) pads to the carrier's GSG pads by using an Au-Au flip-chip thermo-compressive packaging technique with gold stud bumps. In contrast, single-band and dual-band THz interconnects are realized by ac coupling two transmission lines deployed on the chip and the carrier, respectively. By appropriately designing the transmission line length, single-band and dual-band operations can be acquired. The proposed THz interconnects are implemented in 0.18-μm CMOS and IPD technologies. As shown in Fig. 2, the measured insertion loss of the proposed single-band, dual-band, and broadband THz interconnects is 1.7, 2.7 and 2.9 dB, and lower than 3.3 dB at 306.5, 140 and 324.5 GHz, and from 140 to 330 GHz, while ensuring the return loss better than 10 dB. Fig. 3 illustrates the proposed substrate-integrated waveguides (SIW) and fourth-order Chebyshev SIW filter realized in a GaAs IPD technology which can give the measured insertion of 0.7 and 3.6 dB at 327 GHz, respectively [6]. Such a low-cost and compact heterogeneously-integrated THz platform is very suitable to realize THz systems for the next-generation THz sensing and communication applications.

References

[1] K. Katayama et al., "A 300 GHz CMOS transmitter with 32-QAM 17.5 Gb/s/ch capacity over six channels," IEEE J. Solid-State Circuits, vol. 51, no. 12, pp. 3037–3048, Dec. 2016.

[2] H. Siegel, "Terahertz technology," IEEE Trans. Microw. Theory Techn., vol. 50, no. 3, pp. 910–928, Mar. 2002.

[3] C.-H. Li, C.-L. Ko, M.-C. Kuo, and D.-C. Chang, "A 340-GHz heterodyne receiver front end in 40-nm CMOS for THz biomedical imaging applications," IEEE Trans. THz Sci. Technol., vol. 6, no. 4, pp. 625–636, Jul. 2016.

[4] C.-H. Li and T.-Y. Chiu, "Low-loss single-band, dual-band, and broadband mm-wave and (sub-)THz interconnects for THz SoP heterogeneous system integration," IEEE Trans. THz Sci. Technol., vol. 12, no. 2, pp. 130-143, Mar. 2022.

[5] T.-Y. Chiu and C.-H. Li, "340-GHz heterogeneously-integrated THz imager with 4°-beamwidth 16 × 16 IPD antenna array for lensless terahertz imaging applications," IEEE Access, vol. 9, pp. 102195-102206, Jul. 2021.

[6] T.-Y. Chiu and C.-H. Li, "Low-loss low-cost substrate-integrated waveguide and filter in GaAs IPD technology for terahertz applications," IEEE Access, vol. 9, pp. 86346-86357, Jun. 2021.

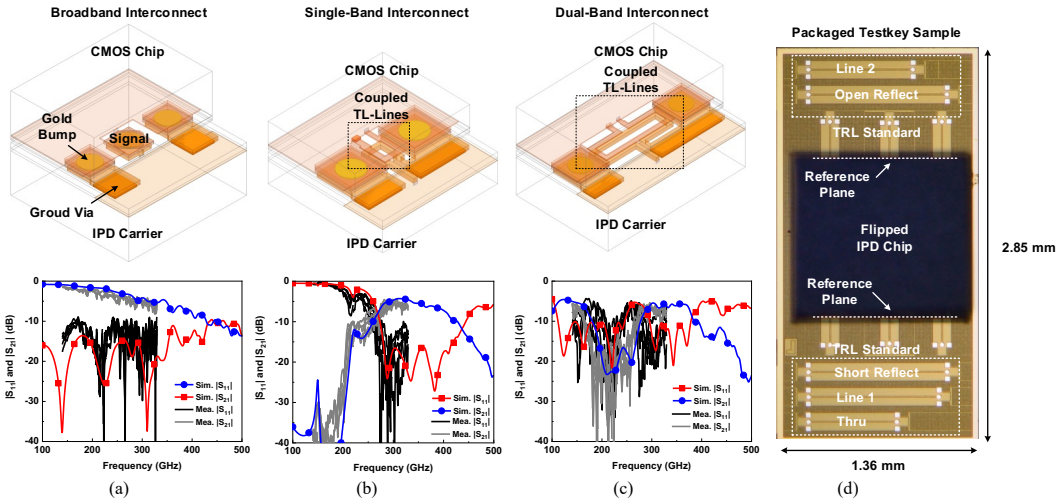


Fig. 2. Measured S-parameters of (a) broadband THz interconnects, (b) single-band THz interconnects, and (c) dual-band THz interconnects. (d) Chip photo of the packaged testkey sample and TRL calibration standards.

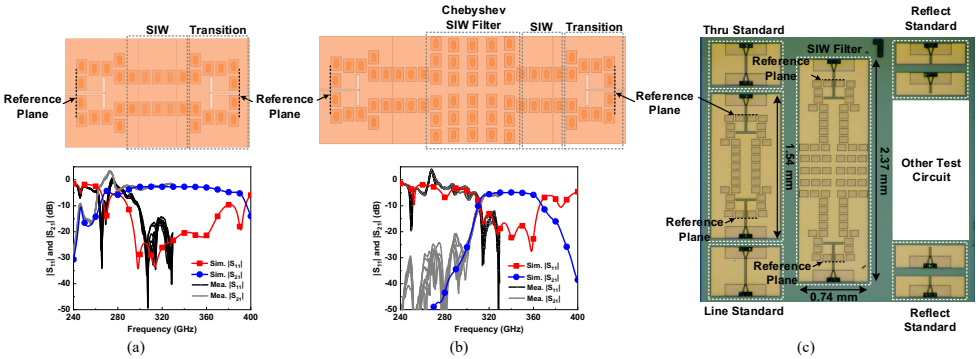


Fig. 3. Measured S-parameters of (a) 220-μm SIW with a CPW-to-SIW transition and (b) fourth-order Chebyshev SIW filter with a CPW-to-SIW transition. (c) IPD chip photo of the CPW-to-SIW transition, SIW filter with a CPW-to-SIW transition, and TRL calibration standards.

Prof. Chun-Hsing Li

Associate Professor, Dept. Electrical Engineering & Graduate Inst. Communication Engineering, National Taiwan University

Major Research Areas:

RF, millimeter-wave, and THz integrated circuits and systems, RF energy harvesting systems, and THz imaging and radar systems

Abstract:

Prof. Li's current research interests include RF, millimeter-wave, and terahertz integrated circuit and system design. He has been serving as an Associate Editor for IEEE Access since 2021 and a member of the TC-21 Terahertz Technology and Applications Committee of the IEEE Microwave Theory and Technology Society since 2022.

Sharing of MediaTek 6G Vision

Chun-Ying Wu
M.S. degree from the Graduate Institute of Communication Engineering, NTU

5G aims to support three use cases: eMBB (enhanced Mobile Broadband), URLLC (Ultra-Reliable and Low-Latency Communications), and mMTC (massive Machine-Type Communications). However, due to various reasons, the first version of 5G (Rel-15) puts most of its focus on eMBB, including features like BWP (Bandwidth Part, see [1] for more details), UL enhancements [2], and Dynamic Spectrum Sharing ([3]). Later on, more features of URLLC and mMTC are introduced in Rel-16 and Rel-17.

There are reasons why 5G is called New Radio. The whole new design of air interface, RAN, and Core Network is posing new challenges to the industry. In MediaTek, our 6G vision is a global standardized technology that can overcome the current limitations of 5G and deliver 10x to 100x performance. In order to achieve this target, we have envisioned the design principles for the upcoming 6G technologies. Here are some examples from MediaTek 6G White Paper [4] that I'd like to share with you in this article. For more details, please refer to the [4].

Wireless Access Convergence

We think 7-24GHz and sub-THz frequencies will be candidates for 6G spectrum. In order to provide better coverage, a new kind of "Hybrid Node" will arise. A 6G Hybrid Node can communicate with any other Hybrid Node, where it can play the role of either device or base station or both. In addition to the Uu interface between device and base station, Sidelink communication between devices could also facilitate the coverage enhancement required in the high-frequency short-range spectrum environment.

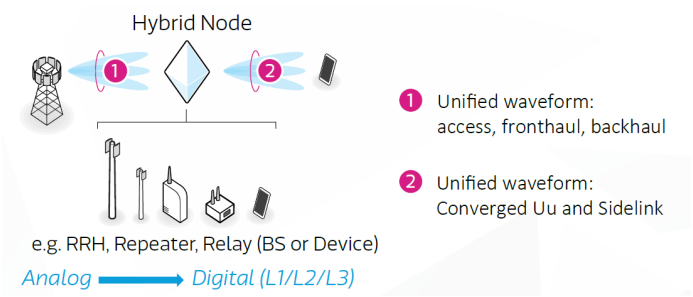


Fig.1 Hybrid Node [4]

MIMO evolution towards True Edge-less Experience

We expect MIMO and multi-antenna technology to keep playing an important role as they did in 4G and 5G. To evolve in 6G, we envision a distributed MIMO deployment where Tx/Rx is no longer bounded to a single site, will provide a revolutionary "cell-free" architecture to overcome the limitations in the traditional cell site design.

Towards Extreme and Predictable QoS – Lean User Plane Protocol Stack

Today's protocol stack design puts the focus on lossless data delivery. The layered design is aimed at more efficient collaboration between layers. However, the drawback is the underlying layers lack the knowledge of upper-layer applications. For example, when the application is loss-tolerant but delay-sensitive, such as video streaming or immersive AR/VR, the lossless data delivery becomes the bottleneck: the overhead to ensure the in-order delivery is not appreciated. In 6G, we think dynamic mutual awareness between Radio/Transport and Application layers is crucial for the user experiences.

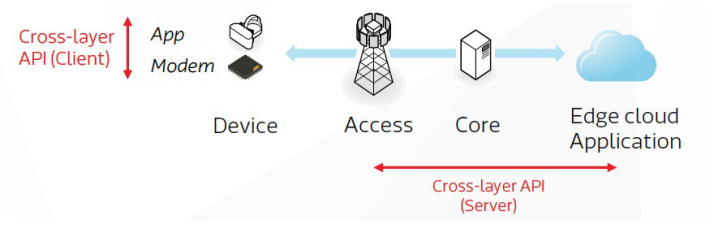


Fig.2 Cross-layer API concept [4]

Terrestrial and Non-Terrestrial Convergence

The convergence of Terrestrial Network (TN) and Non-Terrestrial Network (NTN) is a long anticipated which might happen in the 6G development. The advantage of NTN, such as cost-effective way of coverage for unpopulated area, could be a good complimentary technology to the TN. We expect that there will be native integration of NTN and TN in the network architecture, spectrum re-use, and one single device to serve all.

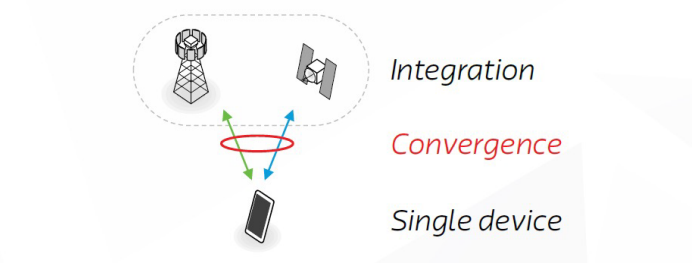


Fig.3 Native TN/NTN integration and convergence [4]

Native AI-integrated System – Communication and Computing Convergence

AI and Machine Learning can bring new opportunities to optimize system performance for networks and devices. In 3GPP Rel-18, many Study Items (SI) and Work Items (WI) are introduced for AI/ML. The industry has the consensus to explore the possibilities to use AI/ML to improve the overall performance, not only the of the air interface, but also of services and network automation, to name a few.

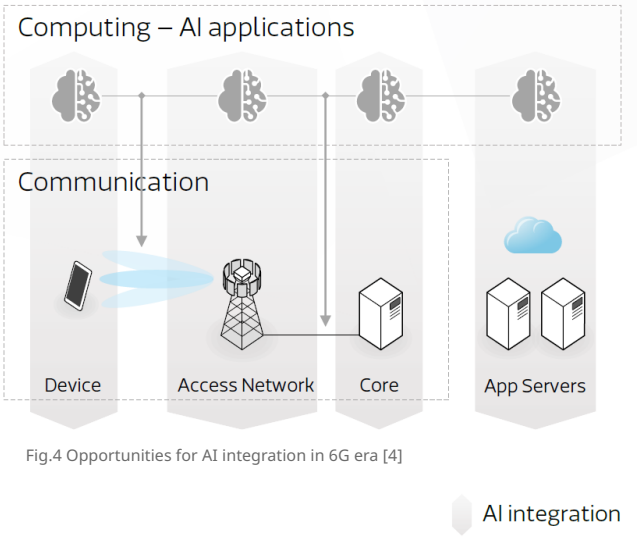


Fig.4 Opportunities for AI integration in 6G era [4]

Conclusions

According to the timeline of ITU-R IMT2030, 6G roll-out will start in 2030 and the pre-commercialization activities might happen in a year earlier. In this pace, we should start to lay the foundation of 6G technologies now, both in academics and industries. In MediaTek, we believe that the 6G design should keep the "S.O.C" principles: Simplexity, Optimization, and Convergence. With these principles in mind, 6G could be more than just another G to the world.

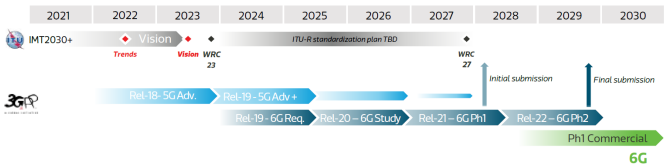


Fig.5 6G Timeline [4]

References

[1] Bandwidth Part Adaptation: 5G NR User Experience & Power Consumption Enhancements. White Paper, MediaTek Inc., Feb 2019
[2] 5G NR and 4G LTE Coexistence: A Comprehensive Deployment Guide to Dynamic Spectrum Sharing. White Paper, MediaTek Inc., Feb 2020
[3] 5G NR Uplink Enhancements: Better Cell Coverage & User Experience. White Paper, MediaTek Inc., Feb 2019
[4] 6G Vision. White Paper, MediaTek Inc., Jan 2022 [Online]. Available: https://www.mediatek.tw/whitepapers

Acknowledgement

Thanks for the invitation from Associate Professor Shih-Chun Lin. The solid study and training in the Graduate Institute of Communication Engineering, NTU, has been one of the drivers of my careers in the wireless communication industry. The communication technology has never stopped innovation and you are already in a very good position for the 6G and beyond!

Author. Chun-Ying Wu

Chun-Ying Wu received M.S. degree from the Graduate Institute of Communication Engineering, NTU, in 2002. He has been working in MediaTek since 2004 and is currently a Senior Technical Manager.

National Taiwan University
Graduate Institute of Communication Engineering

No.1, Sec.4, Roosevelt Road,
Taipei 10617, Taiwan

Editor in Chief
Prof. Shih-Chun Lin

Visit us at:
http://www.comm.ntu.edu.tw

E-mail
gicenewsletter@ntu.edu.tw