IEEE COMMUNICATIONS SOCIETY
Cognitive Networks Technical Committee
http://committees.comsoc.org/tccn

TCCN Newsletter

Vol. 5, No. 2, December 2019

CONTENTS

Chair's Message

Dear Fellow TCCN Members,

2019 has been an exciting year for TCCN. Our new officers have made significant efforts to interact with the cognitive network research community. The symposium name at ICC and GLOBECON have been renovated to "Cognitive Radio and AI-Enabled Networks Symposium" which reflects the recent trend on the topics as those in this edition of the Newsletter.

TCCN has continued to select the two awards this year, which have been given in previous years:

- TCCN Recognition Award for who is deemed to have made significant and sustained contributions to cognitive network community
- TCCN Publication Award for those who are deemed to have made outstanding technical contributions to cognitive network community

The two award committees led by Professor Octavia Dobre from Memorial University and Professor Wei Zhang from University of New South Wales, respectively, evaluated many excellent nominations from the community and decided to give the awards to

- Professor Geoffrey Ye Li as the winner of 2019 IEEE TCCN Recognition Award, for his outstanding contributions to cooperative signal processing in cognitive radio.
- Professor Shuguang Cui as the winner of 2019 IEEE TCCN Publication Award, for his outstanding contributions to cognitive communications and networking.

My hearty congratulations to the winners, and many thanks to the volunteer work by the committee chairs and members. More information of the awards can be found at http://cn.committees.comsoc.org/awards/.

Last, but not least, we are always looking for more volunteers to actively engage in various aspects of the TC, including but not limited to

- Organize Special Interest Groups (SIGs) (contact: Yue Gao, Vijay Rao).
- Organize special issues for the TCCN Newsletter (contact: Daniel Benevides da Costa).
- Participate in TCCN related standardization efforts (contact: Oliver Holland).
- Contribute to the publicity efforts of TCCN (contact: Lin Gao, Yuan Ma)
- Contribute to student competition program (contact: Lucio Marcenaro, Sai Huang)
- Involve TCCN in ComSoc conference organization (contract: Lingyang Song)
- Involve TCCN in ComSoc journal special issues (contact: Yue Gao)

As always, I welcome any suggestions from TCCN members regarding how to make TCCN a better community. Please feel free to contact me at yue.gao@ieee.org if you have any suggestions.

Yue Gao
Chair, IEEE ComSoc TCCN
EPSRC Fellowship Award Holder (2018-2023)
Queen Mary University of London

Director's Message

While 5G has been deployed around the world, there exist already numerous efforts and initiatives from industry and academia to look beyond 5G and conceptualize 6G by describing its roadmap along with the emerging trends and requirements, as well as several enabling techniques and architectures. The drivers of 6G will be a confluence of past trends (i.e., densification, higher rates, and massive antennas) and of new trends that include new services and applications, such as smart wearables, implants, extended reality devices, etc. Moreover, it is expected that 6G will be able to meet strict requirements for multiterabyte per second (Tb/s) and intelligent networks.

Two promising technologies for enabling the 6G ecosystem are terahertz (THz) communications and artificial intelligence (AI). In this regard, this Newsletter will delve on these two key technologies envisioned for 6G wireless networks. In the AI area, we have interviewed Prof. Tim O'Shea, from Virginia Tech, USA, Dr. Jakon Hoydis, from Nokia Bell Labs, France, Prof. Mérouane Debbah, from Huawei, France, and Prof. Deniz Gündüz, from Imperial College London, UK, who are leading experts in this area. We have also had the pleasure to get a position paper from Prof. Deniz Gündüz. Within the context of THz communications, we have interviewed Prof. Daniel Mittleman, from Brown University, Rhode Island, Prof. Josep M. Jornet, from Northeastern University, Boston, and Dr. Onur Sahin, from InterDigital Inc., London, who provided us with their outlook on the opportunities and challenges on AI. Finally, Prof. Cyril C. Renaud, from University College London, UK, provided a position paper that discusses the advancements in demonstration of wireless bridges at THz carrier frequencies over a fiber network as an argument to the advantages of using photonic solutions compared to electronics ones. It is also discussed the potential for photonic integration to create a viable THz photonics wireless technology.

I would like to thank our two feature topic editors: Prof. F. Rafael Marques Lima, from Federal University of Ceará, Brazil, and Dr. Hadi Sarieddeen, from King Abdullah University of Science and Technology (KAUST), Saudi Arabia, for their efforts in arranging the content of this Newsletter. Moreover, we want to thank all authors and interviewees for sharing with us their experience and time. I would finally like to acknowledge the gracious support from the TCCN Chair, Dr. Yue Gao and all TCCN officers. If you have any suggestion, feel free to contact me at: danielbcosta@ieee.org. We hope that you enjoy the material of this Newsletter!

Daniel Benevides da Costa
Director, IEEE Comsoc TCCN Newsletter
Federal University of Ceará, Brazil

Feature Topic: Artificial Intelligence

Editor: Francisco Rafael Marques Lima

Computer Engineering Department, Federal University of Ceará, Sobral, Brazil
Wireless Telecommunication Research Group - GTEL
Email: rafaelm@gtel.ufc.br

Artificial intelligence (AI) is a broad term that encompasses the study and analysis related to an intelligent agent that is able to learn from data and take actions to successfully achieve objectives. Specifically, machine learning (ML), that is a subarea of AI, consists in the theory, models and algorithms that enable computer systems to perform tasks without being explicitly programmed to perform them.

AI and ML have been successfully applied in many areas such as computer vision, natural language processing, robotics and autonomous systems. A branch of ML called deep learning has gained notoriety since the AlphaGo's victory in the Google DeepMind challenge match in 2016. Recently, another sophisticated software called AlphaStar from Google's Deepmind lab has mastered the real-time strategy game called StarCraft 2. The popularity of AI among scientists and researchers is highlighted by the fact that among the top 3 IEEE journals (sorted by impact factor), two journals focus on AI area (according to 2018 Journal Citation Report study in Electrical and Electronic Engineering).

Motivated by the great success of AI applications in those different areas, researchers of other traditional areas have shifted their focus on studying this powerful tool. Wireless networks is a very dynamic subarea of telecommunications that has kept a fast growing trend in the last decades and, therefore, presents many open problems and challenges. Therefore, it would be only a matter of time to the raise in the interest in AI-based solutions for wireless communication networks. Focusing on communications area, the most influential magazines of IEEE COMSOC have featured at least one special issue related to artificial intelligence in 2019.

The fifth generation (5G) networks are not only concerned with providing ultra-high transmit data rates but also assuring ultra-reliable low-latency services as well as massive machine-type access. Undoubtedly, 5G networks are the most complex communication network ever designed with focus on many vertical industries and markets. Moreover, the first initiatives on beyond 5G and sixth generation (6G) point out that these network will be both driven by and a driver of AI.

To reap the benefits of the application of AI in wireless newtorks and motivated by the plethora of technical challenges that emerge from that, one of the feature topic of this TCCN newsletter edition is devoted to AI in wireless networks. In this edition we bring together inputs from four active experts in this field from both academia and industry, and with focus on physical and system level problems: Prof. Tim O'Shea, Dr. Jakob Hoydis, Prof. Mérouane Debbah and Prof. Deniz Gündüz. All of them firstly answer a couple of questions in the interview section of this feature topic and, finally, Prof. Deniz Gündüz provides us with a position paper that presents his view on the exciting area of applying AI to wireless communications networks.

Francisco Rafael Marques Lima received the B.Sc. degree with honors in Electrical Engineering in 2005, and M.Sc. and D.Sc. degrees in Telecommunications Engineering from the Federal University of Ceará, Fortaleza, Brazil, in 2008 and 2012, respectively. In 2008, he has been in an internship at Ericsson Research in Lulea, Sweden, where he studied scheduling algorithms for LTE system. Since 2010, he has been a Professor of Computer Engineering Department of Federal University of Ceará, Sobral, Brazil. Prof. Lima is also a senior researcher at the Wireless Telecom Research Group (GTEL), Fortaleza, Brazil, where he works in projects in cooperation with Ericsson Research. He has published several conference and journal articles as well as patents in the wireless telecommunications field. His research interests include application of optimization and artificial intelligence tools in radio resource allocation and QoS/QoE guarantees in scenarios with multiple services, resources, antennas and users.

Position Paper: Machine Learning Applications for Future Wireless Networks

Prof. Deniz Gündüz
Electrical and Electronic Engineering Department - Imperial College London, UK
Email: d.gunduz@imperial.ac.uk

Recent advances in deep learning has led to remarkable progress in audio and image recognition, natural language processing, and have beaten human grandmasters in chess and Go. This progress also led to many novel applications from autonomous driving to finance, marketing, healthcare and robotics. Although it is hard to demarcate what constitutes ML research, its data-driven nature distinguishes it from more classical research methods. Many successful applications of ML stemmed from the availability of massive datasets and tremendous processing power that can efficiently train very large models on these datasets. These powerful techniques can handle highly complex datasets of image, audio, or video signals, and capture structures impossible for human experts to exploit, or even to perceive.

Almost every day we hear about a surprising new application of various machine learning (ML) tools, with results surpassing what has been possible so far with `traditional' methods. Wireless communication is no exception. The number of papers on the applications of ML to a wide variety of wireless communication problems has exploded in recent years, and I expect this trend to continue in the foreseeable future.

Not surprisingly, in wireless communications, application and network layers have been the early adopters of ML techniques as the behaviour of these layers are extremely difficult to model, often depending on complex human behaviour, such as content access or mobility patterns, while there is abundant data available to service and content providers to employ ML tools. 3GPP has already introduced the network data analytics function (NWDAF) in order to standardize the way such data is collected and communicated across various network functions. While this has limited functionality at the moment, it is widely accepted that analysis using higher layer network and user behaviour data will be an integral part of 5G and future communication network architectures, where NWDAF will orchestrate the sharing of relevant data among network functions, and ML techniques will be employed to make control and resource allocation decisions.

On the other hand, this data driven approach is in stark contrast to the model-based approaches that have long dominated physical layer communication system design. Typically, the design of a communication system is preceded by channel modelling. Only after developing highly complex and accurate models of the underlying physical communication channel, we design appropriate modulation/ demodulation techniques and error correction codes. The design is typically divided into many components, such as channel estimation, channel state information (CSI) feedback, equalization, modulation, and coding, for simplicity as well as modularity, and each of these components is individually optimized for the particular channel model.

This model-based approach has been tremendously successful, taking us from the first to the fifth generation (5G) of wireless networks. However, as we move towards the future generations of cellular networks beyond 5G, networks are becoming ever more complex, and hence, hard to model. Communication channels being used are becoming highly diverse with the introduction of new wireless spectrum, the need for seamless integration across optical or visible light links, and with the introduction of vehicular or drone terminals into the network. Moreover, many different types of traffic will be sharing this highly complicated network infrastructure, with increasingly diverse latency and reliability requirements, further augmenting the complexity.

These new challenges point to the need for a less structured network architecture and more flexible and powerful models, and I argue that ML provides the right tools and approaches to address this growing complexity. As opposed to other applications, where acquiring data to train complex learning models can be a challenge, in wireless systems data is relatively easy to generate. While rich and diverse datasets are not yet available publicly, unlike image or audio datasets, this has been mainly due to lack of interest, and an increasing number of standardised datasets are being made available as the interest grows [1].

Despite its many potentials, there are still many sceptics on the value of ML tools for the physical layer, and many claim that this “hype” in ML research in wireless will last until the next “AI winter". While trends will come and go as usual, and hopefully we will have even more exciting problems and tools to work on in the future, I believe that there is still a lot to be done with ML in wireless, and we are yet to see its impact in implementation, which needs considerably more work. In this short article I will present some examples of promising results (mainly from my own research group, see [1] for more examples), hoping to address some of the scepticisms, and attract the attention of those who have not yet had the chance to explore this exciting area.

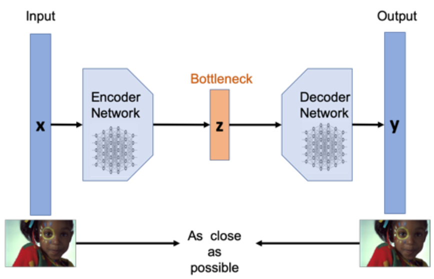

Fig. 1 Autoencoder architecture.

Learning-based data compression

Data compression is a fundamental problem in communication as it reduces the bandwidth requirements without impacting the reconstruction quality (lossless compression), or with minimal distortion (lossy compression). The traditional approach to data compression leverages expert feature knowledge for each type of information source, such as image, audio, or video, and distinct compression standards, such as MP3, JPEG and MPEG, have been developed for each source domain. These algorithms try to exploit the sparsity of the information source in a transform domain, such as discrete cosine transform in image compression, or correlations in time or space, such as motion compensation in video compression. These highly specialized algorithms are results of many decades of research and development, and in general perform quite well. However, there has been significant recent research in data compression using deep neural networks (DNNs), which, in some cases, achieve results that meet or even surpass existing standards [2].

An autoencoder architecture (see Fig. 1) is typically used for DNN-based data compression. An autoencoder consists of a pair of neural networks, the encoder and the decoder, where the original source is fed into the encoder network, and the output of the encoder network, called the bottleneck layer, is the input to the decoder network. These two networks are trained jointly with the goal of recovering the input source at the output of the decoder with minimal distortion. Typically the bottleneck layer has a lower dimension than the input data, and if the autoencoder can learn to recover the input, it can be considered as a compressed low-dimensional version of the input signal.

There are two advantages of autoencoders for data compression compared to traditional compression schemes. First, they do not require the knowledge of the underlying data distribution, or explicit identification of a certain structure; instead they learn a low-dimensional representation directly from data. Therefore, autoencoders can be optimized for specific datasets. While standard image compression algorithms follow the same procedure for all types of images, an autoencoder can learn different weights for different source domains; for example, for human faces, resulting in domain-specific and more efficient compression algorithms. Moreover, standard algorithms do not depend on the objective function; that is, the distortion measure; whereas the DNN can be trained for any loss function, resulting in objective-specific (or, "task-based") compression schemes.

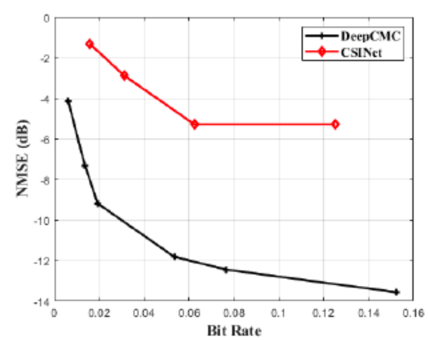

Fig. 2 - DNN-based CSI feedback schemes. DNN-based CSI feedback compression

The base station in a massive MIMO system relies on the downlink CSI. In the frequency division duplex (FDD) mode, this requires users to feedback downlink CSI to the base station. The resulting feedback overhead becomes significant due to the massive number of antennas; and hence, developing efficient CSI compression schemes is essential. DNN-based compression is particularly attractive for CSI feedback compression as it is difficult to identify and characterize the features of channel matrices, which can have quite complicated inter-dependencies through the physical environment. On the other hand, acquiring CSI data for training is easy if we have an accurate model of the physical channel, such as the 3GPP spatial channel model (SCM), WINNER, IEEE 802.16a,e, or the more advanced geometry-based COST 2100 stochastic channel model.

An autoencoder based compression scheme, called CSINet, is introduced in [3], and shown to provide significant improvement compared to existing schemes exploiting sparsity. However, CSINet does not take into account quantization during training. In [4], we have considered quantization and an entropy coding as part of the autoencoder architecture, to further improve the compression efficiency. This architecture, called deepCMC, is trained end-to-end with an objective function that combines the reconstruction distortion and the entropy of the quantized features, which corresponds to the average number of compressed bits. Note that the quantization operation is non-differentiable. We have replaced it with additive uniform noise during training to enable back propagation. In Fig. 2 the normalized mean square error (NMSE) achieved by CSINet and DeepCMC schemes are compared as a function of the bit rate (bits per channel gain) for 32 transmit antennas and 256 subcarriers. We observe that deepCMC provides an impressive 7dB reduction in NMSE, compared to CSINet, which already outperforms existing techniques with a huge margin.

I would like to highlight that the entropy coder is not trained, as codes that approach the entropy of a data source already exist (e.g., arithmetic coding). This is an example of an architecture where ML techniques are combined with a known structured code for improved end-to-end performance.

DNN-based detection and decoding

Data detection is an essential component of any communication system, and is also the quintessential classification problem, where the goal is to classify received noisy vectors into transmitted symbol sequences. Current model-based solutions employ a mathematical model describing the underlying communication channel, whose parameters are estimated through training. The detector employed is typically the theoretically optimal one (i.e., the Viterbi decoder) assuming perfectly known channel model; yet, falls short of the optimal performance in practice due to lack of perfect channel model and state information.

DNNs can also be used to recover coded symbols. Decoding of codewords from a certain channel code is another classification problem. However, the number of classes grows exponentially with the blocklength, leading to unmanageable training complexity. Therefore, most of the current approaches incorporate DNNs into the existing decoder structures.

A fully DNN-based channel decoder is considered in [5]. To keep the complexity reasonable, codelength is limited to 16 with a code rate of 1/2. The authors trained the decoder NN both for a polar code and a random code. While a performance close to a maximum aposteriori (MAP) decoder is achieved for the polar code, the gap to the MAP decoder performance is much larger for the random code. Although this gap can potentially be closed by more training, the result highlights the point that NNs are most effective when the data has an underlying structure to be learned. The authors also considered training with only a subset of the codebook, to see whether decoder can generalize to the rest of the codebook. They observed that the decoder for the polar code was able to generalize, while this was not the case for the random code. This shows that the NN-based decoder is able to learn the structure of the decoding algorithm, and apply it to unseen codewords, while no such structure exists in the case of random codes.

DNN-based channel code design

Similarly to autoencoder based compression schemes, a channel code can also be obtained through training a pair of DNNs, by treating the noisy communication channel that connects the output of the encoder NN to the input of the decoder NN as an untrainable layer with a fixed transformation. In channel coding, instead of the entropy constraint on the bottleneck layer, we impose an average power constraint. End-to-end training of the physical layer can bypass the modular structure of conventional communication systems that consists of separate blocks for data compression, channel coding, modulation, channel estimation and equalization, each of which can be individually optimized. While this modular structure has advantages in terms of complexity and ease of implementation, it is suboptimal. An autoencoder is trained for coding and modulation over an additive white Gaussian noise (AWGN) channel in [6], and it is shown to perform similarly to conventional coding and modulation schemes in short blocklengths. However, it is challenging to extend these results to even moderate blocklengths as the number of messages grows exponentially with the blocklength.

Deep joint source-channel coding (JSCC)
Conventional communication systems employ separate modules for compression and channel coding. I have discussed DNN-based architectures for both of these modules separately. Shannon's separation theorem proves that this two-step approach is optimal theoretically in the asymptotic limit of infinitely long source and channel blocks. However, in practical scenarios, that is, for finite blocklengths and for non-ergodic source and channel distributions, or sources and channels with memory, JSCC is known to outperform the separate approach. The delay and complexity constraints of many emerging applications, such as Internet-of-things (IoT), require operating over strictly limited blocklengths and in the low power regime. Despite its strict suboptimality, almost all communication systems employ a separate coding architecture due to its modularity and the lack of high-performance practical JSCC schemes.

To overcome these limitations, in [7] we have designed a DNN-based JSCC architecture to map the underlying signal samples directly to channel inputs, and vice versa at the receiver, particularly focusing on wireless image transmission. This can be considered as an "analog" JSCC scheme as the input image is never converted into bits, and the channel input signal is not limited to take values from discrete constellation. A fully convolutional architecture is employed, which allows to transmit images of any size, by simply specifying the bandwidth ratio, i.e., the number of channel uses available per pixel.

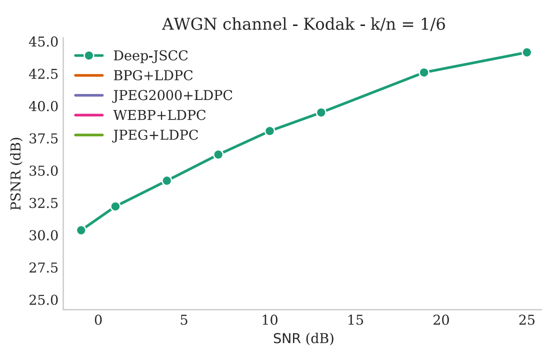

Fig. 3 - PSNR vs. channel SNR for deepJSCC and various digital transmission schemes.

In Fig. 3 above we show the results achieved by deepJSCC compared to state-of-the-art digital image transmission schemes, e.g., JPEG/ WEBP/ JPEG2000/ BPG image compression codecs followed by LDPC channel codes. Here, we set the bandwidth ratio to k/n = 1/6, and train a different network for each channel SNR value. We observe that deepJSCC performs better or similarly to BPG followed by LDPC, while clearly outperforming all other image compression schemes. This is quite impressive considering that these compression and channel coding algorithms are products of decades-long intensive research, while deepJSCC is obtained only after several hours of training. Note also that, while DNN-based channel coding schemes are limited to very short blocklengths, we improve the state-of-the-art in JSCC while transmitting approximately 200K channel symbols for each Kodak image of size 768 x 512 pixel.

Another striking property of deepJSCC worth mentioning is the graceful degradation it exhibits with channel SNR. A deep JSCC architecture trained for a particular target channel SNR gracefully degrades if the channel SNR falls below this value, and its performance improves gradually if the channel SNR goes above the target value [7]. This is unlike the 'cliff effect' observed in digital systems, where the performance saturates at a certain target value dictated by the compression rate, and sharply deteriorates if the channel capacity falls below the channel code rate. This "analog" behaviour of deepJSCC is particularly attractive when multicasting to multiple receivers with different channel qualities, or when transmitting over a time-varying channel. Indeed, it is shown in [7] that, over fading channels, deepJSCC outperforms digital schemes with a much larger margin. We have later demonstrated that deepJSCC also provides adaptivity to channel bandwith [8].

While these results are mainly obtained through simulations using a fixed channel model, we have also implemented deepJSCC on software defined radio, and observed that it performs similarly to the performance promised by the numerical simulations.

Conclusions
These few examples, and many others in the recent literature show a great potential in applying ML techniques to the physical layer communication systems. Other application ranges from DNN-based channel estimation, channel equalization, beamforming design, or resource allocation across distributed terminals.

A major challenge is to evaluate these techniques in real channels and to implement them on mobile devices. As I have mentioned above, we have already implemented the deepJSCC architecture, and showed that practical performance is not far from the numerical results. This is mainly because the tested environment exhibits behaviour sufficiently similar to the AWGN channel used for training. This is not very different from why structured codes designed for AWGN channels work in practice. However, more research and trials are needed how much this extends to more general and complex channel models. More advance ML techniques can also be employed to learn the channel together with the communication scheme. Another important and essential research direction is to explore low-complexity implementation of DNN-based encoders and decoders on complexity and power-limited mobile devices.

References

[1] D. Gunduz, P. de Kerret, N. Sidiroupoulos, D. Gesbert, C. Murthy, M. van der Schaar, "Machine learning in the air," *IEEE Journal Selected Areas in Comms.*, Oct. 2019.
[2] D. Minnen, J. Ballé, G. Toderici, "Joint autoregressive and hierarchical priors for learned image compression," NeurIPS, 2018.
[3] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep learning for massive MIMO CSI feedback," IEEE Wireless Comm. Letters, May 2018.
[4] Q. Yang, M. Boloursaz Mashhadi, and D. Gunduz, "Deep convolutional compression for massive MIMO CSI feedback," IEEE MLSP, Oct. 2019.
[5] T. Gruber, S. Cammerer, J. Hoydis, and S. ten Brink, "On deep learning-based channel decoding," in Proc. Conf. on Information Sciences and Systems (CISS), 2017.
[6] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," IEEE Trans. on Cognitive Comms. and Networking, Dec. 2017.
[7] E. Bourtsoulatze, D. Burth Kurka and D. Gunduz, "Deep joint source-channel coding for wireless image transmission,"IEEE Trans. on Cognitive Comms. and Networking, Sep. 2019.
[8] D. Burth Kurka and D. Gunduz, "Successive refinement of images with deep joint source-channel coding," *IEEE SPAWC, Jul. 2019.*

Deniz Gündüz received his M.S. and Ph.D. degrees from NYU Tandon School of Engineering (formerly Polytechnic University) in 2004 and 2007, respectively. After postdoctoral positions at Princeton and Stanford Universities, se served as a research associate at CTTC in Spain for three years. In 2012 he joined the Electrical and Electronic Engineering Department of Imperial College London, UK, where he is currently a Reader (Associate Professor) in information theory and communications, serves as the deputy head of the Intelligent Systems and Networks Group, and leads the Information Processing and Communications Laboratory (IPC-Lab). His research interests lie in the areas of communications, information theory, machine learning, and privacy. Dr. Gündüz is an Editor of the IEEE Transactions on Wireless Communications and IEEE Transactions on Green Communications and Networking. He served as a Guest Editor of the IEEE JSAC Special Issue on Machine Learning in Wireless Communication (2019). He is a Distinguished Lecturer for the IEEE Information Theory Society (2020-21). He is the recipient of the IEEE Communications Society - Communication Theory Technical Committee (CTTC) Early Achievement Award in 2017, a Starting Grant of the European Research Council (ERC) in 2016, and best paper awards at several conferences.

Interview with Prof. Tim O'Shea

DeepSig Inc. Arlington, USA
Virginia Tech, Arlington, USA
Email: toshea@deepsig.io

Q1: Artificial intelligence has been successfully applied in many areas such as voice/video recognition and biomedical sciences. Nowadays, we are witnessing increasingly interest in applying artificial intelligence in wireless communication problems. Do you think that artificial intelligence will experience the same success as in other areas? In other words, do you think artificial intelligence in wireless communications networks is just a hype or it will sustain its seemingly revolutionary role in the next decades? Why?

A1: We will absolutely see significant, sustained, and transformative change in wireless communications algorithm approaches from the application of machine learning to communications just as we have in computer vision (CV), voice, and natural language processing (NLP). While the hype surrounding AI and ML is currently extremely high and it can sometimes be difficult to filter through the noise, there are many applications and approaches that make sense, are realistic, and provide benefit when considering real world datasets, assumptions, and conditions.

Communications systems have, from the beginning, generally been centered around optimization techniques for obtaining the best performance in our signal processing algorithm chains, and have progressively considered more and more factors over which to optimize as we got better and better at it. This spans from information theoretic optimal bounds on compact statistical formulations of communications systems to optimal estimation techniques for channel state information and symbol detection, to design techniques to optimize performance metrics on various modulation and encoding schemes.

Deep learning brought about an incredibly powerful new set of optimization tools which allow us to take the next natural step in greatly advancing the way we approach almost all of these optimization problems by considering more information and optimizing for more realistic sets of data, constraints, and measurements rather than simplified analytic models. This step is exciting because we now have optimization tools which can scale to optimize for factors ranging from propagation effects, to hardware impairments, to structured interference, to traffic composition and distribution, to a wide variety of other information sources and factors which impact our communications systems. Moreso, we can now do so with the ability to trade off between model performance in various dimensions such as computational complexity, accuracy, and generalization to fit system requirements. For numerous problems within communications this combination gives us excellent state of the art solutions to intractable optimization problems which are computationally efficient.

The degree to which I believe this will be widely transformative within communications signal processing systems reflects why the topic fascinated me during my dissertation work, after watching how the fields of CV and NLP transformed so rapidly in the 2013-2015 timeframe, and why I personally redirected essentially my entire life to focus on building out an incredible team, production level software capabilities, and intelectual property (IP) focusing on rapidly embracing this approach to key wireless problems at DeepSig since then.

Q2: Do you think that wireless networks will be fully controlled/designed by artificial intelligence tools with none or minimal intervention of humans in the future? If so, how far we are from that? If not, what are the limitations of artificial intelligence that prevent it from achieving that?

A2: We'll definitely see increasing automation and autonomy in wireless networks over time as people are comfortable with it and approaches mature and are proven out, in fact, we've already seen this begin to happen significantly. In general

I think operators will retain control over the things they want and automate the things they can and which make sense to offload. Today many cellular networks already adopt AI/ML ideas such as self organizing networks (SONs) in order to adapt basic operating settings such as transmit powers and antenna tilts in a distributed way to help improve network performance. With 4G, 5G-NR, and Beyond 5G the number of configurable parameters and operating modes and resource allocation choices on base stations is growing out of control such that there's no way a person could ever hope to hand-tune all of these settings optimally. We will continue to see AI and ML used increasingly to optimize these settings, scheduling decisions, resource allocation, and other such factors within cellular and other wireless networks based on performance metrics and other optimization objectives. Many such resource allocation applications of AI/ML are rolling out today nearly immediately, as this has been the major focus of many carriers and original equipment manufacturers (OEMs) recently and provides the most immediate payoff for them in terms of leveraging and optimizing existing finite resources. There are now many powerful AI/ML tools to help accomplish this, and really the primary barriers and limitations lie in engineering, integration, testing, and access to data and systems. Fundamental limitations in this area include speed of adaptation from limited data, stability of underlying processes, generalization of learned models, the ability to feed back performance metrics in interoperable ways with other devices, the ability to manage privacy of users and traffic with the need for metrics for adaptation.

The most exciting areas, to me, lay within learning and adaptation within physical layer processing rather than just tuning parameters and knobs on top of existing methods and algorithms. This is an exciting area where we can greatly enhance our estimation, awareness, spectrum access, and resource sharing methods using information and techniques we're not leveraging today, and provide fundamental improvements in comms system performance. These are areas where we could never have managed tuning with a human in the loop because the dimensionality and time-scale is unmanageable without automated optimization schemes, but now with AI/ML in the loop we absolutely can improve them based on data across many dimensions on short time scales. There are numerous practical things we can do in this space rapidly such as channel estimation, compensation for hardware and spatial properties of specific radio systems, and which I expect we will see employed over the next several years. This sort of adaptation has numerous limitations as well and I suspect we will see a large initial wave of successes based on feasibility using current AI/ML methods over the next 1-5 years followed by a slowdown as we cope with fundamental limits of optimization of combinatorially complex systems, difficulty of generalization from few examples, etc. Many of these challenges and limitations mirror the challenges that machine learning in general has faced in the deployment of systems for computer vision, natural language, voice and other applications. As these barriers become more clear over time, and the limits of various applications using existing state of the art data-driven methods are solidified, we will see slower and more incremental progress in the long term by researchers incorporating radio and communications specific model enhancements as well as leveraging fundamental machine learning techniques applicable across a wide range of applications. Ultimately, I believe we will arrive at more fundamental information theoretic limitations for what is feasible to achieve in learning and signal processing systems based on stability, availability, and complexity of data distributions in the underlying process.

Q3: Artificial intelligence and its branch of machine learning are able to tackle many difficult problems in wireless communications. However, most of the problems in this area have been studied by researchers and engineers over the past decades using well-established techniques with strong mathematical background such as optimization, statistics and game theory. On the other hand, a common criticism is the difficulty to guarantee that machine learning solutions will always work in general scenarios or converge to the optimal solutions. Another common criticism to machine learning solutions that rely on neural networks is that they are seen as black boxes whose outputs cannot be completely explained, thus raising doubts about reliability and biases. What is your view about these aspects? Moreover, do you think that the classical solutions for wireless communications problems will be still useful in the future or they will be completely replaced by machine learning-based solutions?

A3: Classical solutions are great when model assumptions are correct. George Box's timeless remark that "All models are wrong, [but some are useful]." could never be more relevant than today

-- machine learning and deep learning will outperform well-established statistical models where they are less-wrong given real world data, where they can reduce computational complexity, and where statistical models and traditional optimization approaches are intractable for joint optimization of end-to-end communications schemes.

Classical analytic performance guarantees and bounds are often given for statistically convenient system models such as Gaussian or Rayleigh fading channels, single user channel access, etc. These can be powerful tools for gaining intuition and expectations for real world systems, but at some point there is no pristine AWGN channel in the real world. Similarly, deep learning models can be easily evaluated against a wide range of test and evaluation environments to provide empirical performance measures in both simulated and real world measurement environments, but providing strong analytic performance guarantees can be difficult. Both of these approaches have drawbacks so there is no magic bullet here, analysis can provide invaluable insight, while ultimately performance of systems in the real world is the metric that we all care about. There has actually been a lot of recent work in computer vision focused on trying to provide stronger performance bounds and guarantees. One of the fundamental difficulties with this is that in a data-driven system, performance bounds are a function of the data-distribution which is often complex and hard to represent in any compact and easy to analyse way. This has been explored within the contexts of adversarial attacks as well as generative adversarial network (GAN) training, where the notion of mode-collapse has been shown to be a function of the dataset -- that is to say that, without enough data to represent your target distribution, generalization and performance will be degraded regardless of the model. Then there's simply that degrees of freedom in the model which are difficult to cope with in a traditional analytic scheme -- but not impossible. Numerous works (including Feinman et al, “Detecting adversarial samples from artifacts”) have explored a Bayesian analysis for tracking neural network uncertainty through high dimensional datasets and models with interesting results. Ultimately, this kind of robust statistical analysis is possible within the context of Deep Learning, but it needs to be done programmatically at scale and will not be easily done by hand -- There are many exciting applications of this to the communications space which have yet to be explored. So the answer here is mixed, we will absolutely continue to need all of the fundamental tools of probability and analysis here, but we may not be able to use some of the optimization tools in the way we do now to solve sufficiently complex problems.

The criticism that ML is a black box is indeed also a common one, and is mixed in its truth and impact. Again much of this problem is due to the target distribution or dataset being hard to represent in a compact and understandable way. Its true that learned solutions are sometimes hard to understand or explain what is going on, but there are also countless counter-examples to that where we can understand quite clearly what is going on. For example in the autoencoder case, we can learn 2-dimensional I/Q representations and directly visualize constellation points to see how they spread out from each other to minimize codeword/class confusion. In the case of classifiers, numerous works have inspected filter weights, activations due to certain stimulus, generative models, and a wide range of other tools which allow us to gain an understanding into what is being learned within each network. While this is not a satisfying answer for some, it can provide significant valuable insight and intuition which helps in the design, engineering, and performance analysis process.

Beyond this, there have been an increasing number of model-driven or domain-knowledge-driven network architectures leveraged within the context of data-driven machine learning models in recent years. Our original work on the radio transformer network (RTN) was one example of this which showed how incorporating network structure (e.g. equalization, synchronization, etc) into the network allows for reduced complexity, faster training, and better generalization, while still allowing the benefits of a data driven approach and forcing certain intermediate values to become very understandable or equivalent to traditional estimation values. Beyond that we have seen enormous value in the ideas behind deep unfolding, or incorporating the structure from analytic models such as belief propagation/message passing into end-to-end machine learning architectures and then allowing a data-driven approach to fine-tune or improve their performance under real world data, providing architectures where we can gain significant understanding of how they work.

So to summarize, domain knowledge and Bayesian probability analysis are not going away -- they remain important and central tools, and new ways to use them within the context of machine learning have proven to be valuable for performance guarantees, understanding, and improving model performance -- and letting us scale to solve much larger and more realistic problems.

Q4: Machine learning has many techniques/algorithms that can be classified in supervised learning, unsupervised learning, and reinforcement learning. The neural networks are very relevant building blocks of many machine learning solutions. In your opinion, what is(are) the most promising algorithm(s)/architecture(s)/framework(s) from machine learning area to be applied in wireless communications problems?

A4: I believe all of these classes of algorithms hold enormous potential in different aspects of communications systems. Supervised learning has proven to be a powerful tool for end-to-end feature learning, unsupervised and semi-supervised learning have proven to be extremely powerful tools for coping with unlabeled data and representation and reconstruction learning tasks, while reinforcement learning has shown to be a powerful tool for exploring very large state spaces where action-reward mappings can be extremely difficult to model or represent in other tractable ways. In general, I believe compact deep neural networks (DNN)/ convolutional neural networks (CNN) architectures with numerous variations will be powerful tools in features extraction from raw data, function approximation, estimation, and data transformation for a lot of applications, especially while optimizing for small architectures and reduced precision. Meanwhile, reinforcement learning is an enormously powerful building block which I think will be key in numerous resource allocation problems and state-space exploration problems in communications systems (e.g. resource scheduling, control of RF front-ends, parameters and hardware, and transmit adaptations). We have been focused at DeepSig on building upon powerful mature software frameworks such as Torch and TensorFlow at DeepSig as well as upon powerful software radio frameworks such as USRP Hardware Driver (UHD), GNU Radio, Liquid and others in order to tightly couple high performance ML into baseband algorithms for rapid measurement, learning, adaptation and iteration -- I think enabling software and synthesis tools which is tailored for the communications system domain is really a key enabler across all of these algorithm, application, and architecture domains.

Q5: In your opinion, what are the most important problems to be faced by artificial intelligence in physical layer? And in a system level? What are your own short-term and long-term research plans in artificial intelligence for wireless communication?

A5: We've sort of broken this down into two classes of problems and perhaps a third which joins them together.

1) Awareness: Simply ingesting information from sensors to boil this down into highly accurate descriptions of what's going on in the spectrum around you quickly. This is a core enabler for orchestrating dense unplanned radio frequency (RF) deployments, unlicensed and shared spectrum bands, understanding of interference, impairments, and malicious emitters as well as an enormous enabler of new emitter analytics and the underlying physical processes they represent. We already have RF edge sensors all over the place, but being able to leverage and act on this information given limited bandwidth, computation, and development time at scale is game changing proposition in numerous industries.
2) Communications Efficiency: Improving the density, energy efficiency, accuracy, and performance of how we encode and decode information for wireless systems and for specific components of wireless systems. Taking advantage of all of the dimensions we have been leaving on the table during wireless system design for so many years --

actual distributions of channel statistics after deployment, effects of hardware and imperfections in real systems, multi-user and multi-antenna strategies to improve density, and data and experience fueled methods for resource allocation and system operation. We have simply not had the optimization tools powerful enough to do so in a convenient, tractable, and efficient way until recently.

3) Finally, bringing these together -- combining rapid spectrum awareness with optimized communications system spectrum access, encoding and processing offers to give us a whole new level of visibility, performance, power efficiency, spectral efficiency, multi-user efficiency, spatial efficiency, density, and multi-dimensional optimization and cost reduction in how we share and and monitor spectrum access, and detect impairments and anomalies. Longer term, both of these fields coming together to allow truly dynamic spectrum access

In the short term I'm focused on building software and real world realizations of #1 and #2 above, which leverage data-driven approaches to providing awareness and leverage a data-driven approach to optimizing modem performance and inserting these into the communications ecosystem where they can add significant value versus how we do things today. Longer term, I believe we will see these come together to enable smarter spectrum re-use and allocation schemes on a larger scale.

Q6: Could you please briefly introduce the most recent research project(s) that you have done in this area? (Please explain the key idea(s) and interesting findings)?

A6: My major focus recently have been working with our team at DeepSig building out mature data-driven wideband RF sensing and modem learning capabilities that work in the real world to mirror my answers from above. This includes)

1) Building out the OmniSIG Sensor software and software development kit (SDK) which extends basic modulation recognition to the wideband detection and classification of numerous emitters using machine learning. This has led to RF sensing which is 700x-1000x faster than traditional energy detection and feature based methods, which generalizes well to a wide range of emitter types, and is data-driven, allowing users to train up new wideband recognition models in minutes or hours which would previously take months or years of engineering and implementation time. This is our spectrum awareness engine, which consumes gigabits of raw RF data and outputs a structured JSON/SIGMF representation of all of the RF emissions occurring in the spectrum which can be used to enable interference detection, decision making, analytics, threat detection, and spectrum access. At VT we're also now looking at how to use this stream of structured data to do wireless intrusion detection and cyber threat detection on this datastream in a scalable and generalizable way which mirrors wired-network threat detection today.
2) Building out learning enabled modem demonstrations and software for standards-based and non-standards-based contexts to enhance communications systems by leveraging more information in the environment and more efficiency processing. In the case of standards based technologies, we're building out a 5G-NR base transceiver station (BTS) demonstrators showing where learning based algorithms can improve channel estimation, equalization, multiple-input multiple-output (MIMO) performance within the baseband unit (BBU) by taking advantage of online learning after deployment. And in the context of non-standards based modems we're building the OmniPHY Modem runtime and SDK which is based on extensions of the autoencoder approach as a full usable system for point-to-point, backhaul, satcom, and mesh deployment scenarios. This has allowed us to test these ideas in the real world, carrying video, internet protocol (IP) traffic, and encryption all over a completely adaptive physical layer which can be customized for a wide range of channels, operating conditions, hardware effects, impairments, etc and can continue to update and optimize its encoding and decoding schemes online to regain performance lost from design model-

mismatch and from changing environmental and system effects.

Q7: Beyond your own work, are there any resources that you would like to recommend, especially to those who are new in this field and want to learn more about artificial intelligence? Are there any specific resources that you recommend related to artificial Intelligence in the context of wireless and communication networks?

A7: Absolutely, IEEE has an emerging technology initiative (ETI) for machine learning in communications (MLC) which is leading a number of efforts to compile valuable resources for folks interested in diving deep in this area. https://mlc.committees.comsoc.org/
This includes datasets, code examples, curation of papers and events, videos, mailing lists, competitions, news articles, etc and is aiming to be a good starting point.
Other than this, I would encourage people to explore the software radio community, GNU Radio Conference in particular has had numerous MLC area works in recent years, and open source enabling tools for measurement and experimentation. DeepSig has provided several of our early datasets and examples for open use, & VT has also helped host datasets such as for the IEEE CTW ML Comms Workshop competitions. People interested in the area should also explore the greater machine learning, AI and deep learning research communities beyond the confines of communications systems because tons of excellent work is being done in these areas from which we can learn and apply it to our problems notably in computer vision, natural language, medical imaging, and voice processing. Attending large ML events such as ICML, NeurIPS, CVPR and others can be a quick way to be exposed to a lot of amazing work with analogues in our field, but is also a bit of a goat rodeo. I'm personally not a large advocate for social media, but I have found that following researchers in the ML and Comms communities on twitter who stick primarily to technical content propagation invaluable - and a more manageable stream of information than monitoring all of arxiv or IEEE Xplore for relevant publications.

Q8: What is your most important contribution (journal, magazine or conference article, or patents) in the topic?

A8: My most impactful contributions at this point are probably casting the sensing and communications problems as data driven deep learning problems. Specifically these were focused on treating modulation and encoding learning as an autoencoder representation problem (as described in our patent US10217047B2), and treating modulation recognition as a supervised convolutional feature learning problem on raw I/Q data. These first works had significant room for improvement, but collaborations with others really helped formalize and turn these ideas into more rigorous and thorough works and explorations, and I think the papers and open sourced tools and datasets from these works helped a lot of folks get started and interested in the area. Both of these works are best concisely described and explored further in my joint work with Jakob Hoydis in TCCN in "An Introduction to Deep Learning for the Physical Layer." Ultimately, I'm most excited by the impact of building out real world vetted systems and software leveraging these data-driven and learning based approaches to both sensing and comms systems such as what we've built and are building at DeepSig in OmniSIG and OmniPHY, which we'll be excited to continue publishing and sharing with the community.

Dr. Tim O'Shea is CTO at DeepSig and Assistant Professor at Virginia Tech in Arlington where he is focused on applied research and system development in the area of machine learning and data driven signal processing systems for wireless communications. He has led research programs including for NSF, NASA, DARPA, DOD, and industry partners, has published over 50 peer reviewed articles in the field, serves as co-chair for IEEE Machine Learning for Communications emerging technology initiative and on the editorial board for the IEEE Transactions on Wireless Communications and IEEE Transactions on Cognitive Communications and Networking. He is a core contributor to the GNU Radio project and has previously help technical positions working with CTIA, Federated Wireless, Hawkeye 360, Cisco Systems, and the DOD.

Interview with Dr. Jakob Hoydis

Nokia Bell Labs, France
Email: jakob.hoydis@nokia-bell-labs.com

Q1: Artificial intelligence has been successfully applied in many areas such as voice/video recognition and biomedical sciences. Nowadays, we are witnessing an increasingly interest in applying artificial intelligence in wireless communication problems. Do you think that artificial intelligence will experience the same success as in other areas? In other words, do you think artificial intelligence in wireless communications networks is just a hype or it will sustain its seemingly revolutionary role in the next decades? Why?

A1: As Bill Gates said, we tend to overestimate the short-term and underestimate the long-term change. The same effect is reflected in Gartner's hype cycle for emerging technologies. Especially for communications, which is a very mature field, it would be naïve to expect a disruption thanks to ML in a few years. However, I am confident that its role will continue to increase over the next decade. Communication engineers are just starting to make productive use of ML and standardization has made the first steps to enable the use of ML in 4G/5G networks. But one should not forget that we already knew how to build communications without ML. Thus, the disruption will be much smaller compared to a field such as computer vision where ML allows us to do things we could not do before.

Q2: Do you think that wireless networks will be fully controlled/designed by artificial intelligence tools with none or minimal intervention of humans in the future? If so, how far we are from that? If not, what are the limitations of artificial intelligence that prevent it from achieving that?

A2: Zero-touch configuration of networks is something colleagues at Bell Labs have been working on for quite some time. So I believe that AI-controlled networks is a realistic vision, which seems by the way far less ambitious to me than, e.g., fully autonomous cars. Some of the biggest limitations are a lack of access to real-time data, trust into technology, as well as convergence times of learning algorithms. We are also lacking good open-source simulation environments for research on reinforcement learning.

Q3: Artificial intelligence and its branch of machine learning are able to tackle many difficult problems in wireless communications. However, most of the problems in this area have been studied by researchers and engineers over the past decades using well-established techniques with strong mathematical background such as optimization, statistics and game theory. On the other hand, a common criticism is the difficulty to guarantee that machine learning solutions will always work in general scenarios or converge to the optimal solutions. Another common criticism to machine learning solutions that rely on neural networks is that they are seen as black boxes whose outputs cannot be completely explained, thus raising doubts about reliability and biases. What is your view about these aspects? Moreover, do you think that the classical solutions for wireless communications problems will be still useful in the future or they will be completely replaced by machine learning-based solutions?

A3: I believe that one should leverage as much expert knowledge as possible while designing ML solutions. So classical solutions will rather be augmented by ML than replaced. Reliability and bias are much less of an issue in our field compared to areas such as facial recognition for border control. If a packet is lost, no real harm is done (unless we operate a mission-/life-critical operation). Most algorithms in communications are validated through extensive simulations. I do not see why we cannot do the same with ML-based solutions.

Q4: Machine learning has many techniques/algorithms that can be classified in supervised learning, unsupervised learning, and reinforcement learning. The neural networks are very relevant building blocks of many machine learning solutions. In your opinion, what is(are) the most promising algorithm(s)/architecture(s)/framework(s)

from machine learning area to be applied in wireless communications problems?

A4: I think that we are closest to making productive use of supervised learning in our field although a lot of interesting research happens on self-supervised learning approaches as well as reinforcement learning. Also, meta learning seems very relevant to enable fast online training.

Q5: In your opinion, what are the most important problems to be faced by artificial intelligence in physical layer? And in a system level? What are your own short-term and long-term research plans in artificial intelligence for wireless communication?

A5: A big challenge is to enable ML algorithms that can be "configured". For example, an ML MIMO detector which works for an arbitrary number of users and modulation orders and not just for one choice of parameters. Also, implementation complexity is a big problem. If you gain 5% in performance but require 10x more compute, nobody is going to implement your solution.

Q6: Could you please briefly introduce the most recent research project(s) that you have done in this area? (Please explain the key idea(s) and interesting findings)?

A6: I have worked for several yeas on the idea of end-to-end learning of communication systems and we finally reach a point were things start to become practical and show significant gains over baselines. My biggest learning is that one should focus not only on solving an isolated problem but think about the full system implementation. We discovered many interesting ML applications through this approach. Another project has been on MIMO detection. While we struggled for a long time to develop a detector that works for any possible channel matrix, we finally discovered that it is less complex to train a very small model for every single channel realization in an online fashion than having a very big model which you train offline on a large dataset. This way of thinking is against common practice but opens up interesting possibilities.

Q7: Beyond your own work, are there any resources that you would like to recommend, especially to those who are new in this field and want to learn more about artificial intelligence? Are there any specific resources that you recommend related to artificial Intelligence in the context of wireless and communication networks?

A7: I can highly recommend the website of the IEEE Emerging Technology Initiative on Machine Learning for Communications: https://mlc.committees.comsoc.org/
Also the Best Readings on the same topic are a great starting place: https://www.comsoc.org/publications/best-readings/machine-learning-communications
Apart from that, I really enjoyed Jeremy Howard's deep learning course (https://course.fast.ai/) as well as Osvaldo Simeone's book on ML for engineers (https://arxiv.org/abs/1709.02840).

Q8: What is your most important contribution (jornal, magazine or conference article, or patents) in the topic?

A8: My work with Tim O'Shea on end-to-end learning: https://arxiv.org/abs/1702.00832

Jakob Hoydis received the diploma degree (Dipl.-Ing.) in electrical engineering and information technology from RWTH Aachen University, Germany, and the Ph.D. degree from Supelec, Gif-sur-Yvette, France, in 2008 and 2012, respectively. He is a member of technical staff at Nokia Bell Labs, France, where he is investigating applications of deep learning for the physical layer. Previous to this position he was co-founder and CTO of the social network

SPRAED and worked for Alcatel-Lucent Bell Labs in Stuttgart, Germany. His research interests are in the areas of machine learning, cloud computing, SDR, large random matrix theory, information theory, signal processing, and their applications to wireless communications. He is a co-author of the textbook "Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency" (2017). He is recipient of the 2018 Marconi Prize Paper Award, the 2015 Leonard G. Abraham Prize, the IEEE WCNC 2014 best paper award, the 2013 VDE ITG Forderpreis, and the 2012 Publication Prize of the Supelec Foundation. He has received the 2018 Nokia AI Innovation Award and has been nominated as an Exemplary Reviewer 2012 for the IEEE Communication Letters. He is currently chair of the IEEE COMSOC Emerging Technology Initiative on Machine Learning for Communications.

Interview with Prof. Mérouane Debbah

CentraleSupelec, Gif-sur-Yvette, France
Huawei France Research Center
Mathematical and Algorithmic Sciences Lab
Email: merouane.debbah@centralesupelec.fr

Q1: Artificial intelligence has been successfully applied in many areas such as voice/video recognition and biomedical sciences. Nowadays, we are witnessing an increasingly interest in applying artificial intelligence in wireless communication problems. Do you think that artificial intelligence will experience the same success as in other areas? In other words, do you think artificial intelligence in wireless communications networks is just a hype or it will sustain its seemingly revolutionary role in the next decades? Why?

A1: Artificial Intelligence is not a new topic and dates back with the famous workshop at Dartmouth College in 1956. It has gone already through several hypes and winters since that date. Typically, in the 80's, the majority of researchers were working on expert systems but years after, the AI winter came back. Today, the biggest progress in AI is not so much in the algorithmic aspects but in the computing capability as well as the huge amount of data available. And I have to admit that it is our Computer Science colleagues who are showing us the path. In terms of algorithms, the main architectures which are used date back from the 90's, with a couple of refinements. I think there is still a lot of progress to be done at the algorithmic level as the algorithms today are quite rudimentary. For the application of AI in Wireless Communications, what we see today is mostly a re-branding of Statistical Signal Processing and Optimization tools with the word AI. You can see it also by the number of courses which were called optimization and are now called Machine Learning. The content has nearly not changed but only the wording! Said that, I strongly believe that true AI will have a huge impact in the wireless Communication field with the ability to have reasoning networks (rather than just learning) and semantic capabilities (not just conveying the message without error but also the intended meaning). This will require however to build totally new mathematical foundations for AI.

Q2: Do you think that wireless networks will be fully controlled/designed by artificial intelligence tools with none or minimal intervention of humans in the future? If so, how far we are from that? If not, what are the limitations of artificial intelligence that prevent it from achieving that?

A2: I think we have to separate here what is related to automation from Artificial Intelligence. Today, many of our networks are already automated and we are progressing towards that path. There is already less and less human intervention because many of the process are designed to do things automatically. And we do not really need AI for that. Having Wireless Networks controlled by AI is really about having a wireless brain, that can take decisions with new and un-expected events. For this case, we have many issues that go from the lack of understanding on the decisions which are taken by the actual AI algorithms (can we explain things to the users and operators?) to more practical issues such as the real-time nature of the decision or the distributed nature of the data gathered, to give just a couple of examples.

Q3: Artificial intelligence and its branch of machine learning are able to tackle many difficult problems in wireless communications. However, most of the problems in this area have been studied by researchers and engineers over the past decades using well-established techniques with strong mathematical background such as optimization, statistics and game theory. On the other hand, a common criticism is the difficulty to guarantee that machine learning solutions will always work in general scenarios or converge to the optimal solutions. Another common criticism to machine learning solutions that rely on neural networks is that they are seen as black boxes whose outputs cannot be completely explained, thus raising doubts about reliability and biases. What is your view about these aspects? Moreover, do you think that the classical solutions for

wireless communications problems will be still useful in the future or they will be completely replaced by machine learning-based solutions?

A3: The best answer would be to give you my personal experience on the topic. In the years 2007-2010, I spent a lot of time working on what we called self-organized networks (SON) for Small Cells, for which by the way cognitive radio could be seen as a general case. The idea was to enable small cell networks to self-configure their parameters in order to improve their performance. I had therefore worked on Game Theoretic Techniques and had spent a lot of time implementing with my Phd students many algorithms such as Best Response Dynamics, Reinforcement Learning, Q-Learning, Multi-Arm Bandits, Trial and Error Learning, Fictitious Play, Imitation Learning and I could go like this for hours. In particular, I spent a lot of time understanding and optimizing the famous exploration versus exploitation trade-off (how much time you spend to explore the dimensions of your problem before you can actually exploit that) and it never worked! It took me years to understand the reasons behind. And so when people started to speak in 2016 about Wireless AI and the new AI based SON revolution, I was very skeptical. But something happened at the same time that totally changed my mind: since 2014, I was working with colleagues on the problem of improving Voice over long-term evolution (LTE) coverage and finding the right radio engineering planning for that. This is a quite an intricate problem as LTE is about Mobile internet and therefore, voice in an application and not a technology. In particular, even if one improves the received SNR, it does not improve the Voice over LTE performance as it depends on many other issues such as the IP protocol, the location of the servers, etc. Voice over LTE is an end-to-end metric and although we were good in modelling, it was nearly impossible for us to express things in a formula and then move forward with the optimization. The operators were quite upset as many were using their 2G and 3G network to provide voice calls (this is called circuit-switched (CS) Fall-Back) and 4G for data. With the refarming of the 2G and 3G spectrum, we had to find a solution for this issue.
The solution came from data driven approaches and is called today AI based Voice Over LTE. Our wireless Business Unit had made a lot of measurement of Voice over LTE quality for different deployment of networks. It became obvious that the solution would be an interpolation solution that could predict the new base station layout based on all the known layout deployments and previous Voice Over LTE quality. Surprisingly, the performance increase was quite incredible with factors of up to 50% and 80%.
The answer to your question is therefore a balanced answer: we will not replace all classical solutions by machine learning based solutions. However, there are strong cases where machine learning based solutions should be used and these are when models are expensive or impossible to obtain, the end-to-end objective function is not defined mathematically and often, we have a high dimensional space with many parameters that we can not capture.

Q4: Machine learning has many techniques/algorithms that can be classified in supervised learning, unsupervised learning, and reinforcement learning. The neural networks are very relevant building blocks of many machine learning solutions. In your opinion, what is(are) the most promising algorithm(s)/architecture(s)/framework(s) from machine learning area to be applied in wireless communications problems?

A4: It really depends at which level you are working on (real time transmission (RTT), radio resource management (RRM), mobile broadband (MBB)/CORE or operational support system (OSS)/self-organizing networks (SON)) and on which scenario. Typically, for problems related to peak-to-average power ratio (PAPR) non-linear compensation, LTE power control, high frequency (HF)/low frequency (LF) collaboration, failure detection or channel map reconstruction, you will be using clustering, regression or classification algorithms. If you are working on link adaptation, policy management, slice resource management or coordinated multipoint (COMP) mode selection, algorithms such as association rule mining, gaussian mixture model-hidden markov model (GMM-HMM), Dynamic Optimization or reinforcement learning would be key. Deep learning, transfer learning are suited for end-to-end performance learning,

AI base station or AI management platform for example. The response really depends on the scenario you will be tackling.
In terms of architecture, you have also to take the constraints of latency, privacy and coverage into account. You may do Cloud AI algorithms if you have no constraints to Edge or on-device AI if the constraints are extreme. Frameworks for implementing AI depend also on your problem. Typically, if you are dealing with complex numbers (in wireless, nearly everything is complex, which is not the case for images for example), then you need a framework related to Complex Neural Networks and the classical tools available such as TensorFlow may not be adequate. For example, recently, we decided to build a new framework called Mindspore to be able to federate Cloud, Edge and Device learning and provide a cooperative training/inference framework.

Q5: In your opinion, what are the most important problems to be faced by artificial intelligence in physical layer? And in a system level? What are your own short-term and long-term research plans in artificial intelligence for wireless communication?

A5: I think the most important problem today for AI and which could bring the next winter is energy efficiency. The amount of computing power used to get a single optimized parameter is not sustainable. Typically, if you use a model and you are at 80% of the performance and you use data driven approaches which consume 10,000 more times and you are at 85% of the performance, then you will be asked rapidly on the cost-performance curve. We are not yet to it but this will arrive in a couple of years and may turn out to be the next winter of AI.
I also think that we should bridge the gap between data driven and model driven approaches. Many people start AI for physical layer by throwing all the modelling that has been done for a century! AI should not replace but be incorporated in the know-how we have and communication engineers have a lot of know-how.
AI in wireless will be mostly pervasive and I think that we need to build rapidly nice frameworks for distributed learning that can cope with the actual constraints that we have in communication.

Finally, the next big AI Revolution in Wireless Networks will be about Reasoning Networks rather than learning networks. We need however to introduce a component of reasoning and this is not easy. Nowadays, there are many research groups (mine also) working on that.

Q6: Could you please briefly introduce the most recent research project(s) that you have done in this area? (Please explain the key idea(s) and interesting findings)?

A6: These last 2 years, I have been working on a couple of interesting projects, among my activity in 5G:

- One relates to an activity on large scale distributed learning where we design distributed learning strategies for multiple access based on Mean Field Games. My group has been closely collaborating with Medal Fields Pierre Louis Lions on the topic and we have a couple of interesting papers and results on that.
- The second relates to Mobile AI. Today, we have the ability to provide AI chipsets for the terminal, the edge and the cloud with different learning/inference capabilities. I have been working with my colleagues such as Prof. Mehdi Bennis on transfer learning and federated learning approaches to provide a training and inference framework that could be discussed in the standard.
- The third one is on bridging the gap between data driven and model driven approaches. I have worked on a recent paper with my colleagues A. Zappone and M. Di Renzo, “Wireless Networks Design in the Era of Deep Learning: Model-Based, AI-Based, or Both?”, IEEE Transactions on Communications, Volume: 67, Issue:10, Page(s): 7331-7376, October 2019”) which tries to tackle that and provides a neat framework.
- Finally, I am working on a project on the Mathematical Foundations of AI to go beyond learning, which is the most exciting for me at this stage.

Q7: Beyond your own work, are there any resources that you would like to recommend, especially to those who are new in this field and want to learn more about artificial

intelligence? Are there any specific resources that you recommend related to artificial Intelligence in the context of wireless and communication networks?

A7: I think the best Best Readings in Machine Learning in Communications website contains a nice overview of the papers to start with: https://www.comsoc.org/publications/best-readings/machine-learning-communications

Q8: What is your most important contribution (jornal, magazine or conference article, or patents) in the topic?

A8: I would highly recommend to read the following papers:

- Zappone, M. Di Renzo and M. Debbah, "Wireless Networks Design in the Era of Deep Learning: Model-Based, AI-Based, or Both?", IEEE Transactions on Communications, Volume: 67, Issue:10, Page(s): 7331-7376, October 2019
- J. Park, S. Samarakoon, M. Bennis and M. Debbah, "Wireless Network Intelligence at the Edge", Proceedings of the IEEE, Vol. 107, Issue: 11, Page(s): 2204-2239, November 2019
- M. Di Renzo, M. Debbah, D. Phan-Huy, A. Zappone, M. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J.Hoydis, H. Gacanin, J. de Rosny, A. Bounceu, G. Lerosey and M. Fink, "Smart Radio Environments Empowered by AI Reconfigurable Meta-Surfaces: An Idea Whose Time Has Come", accepted for publication, EURASIP Journal on Wireless Communications and Networking, 2019
- A. Destounis, D. Tsilimantos, M. Debbah and G. S. Paschos, "OLMA: Online Learning Multiple Access for URLLC Applications", INFOCOM 2019, Paris, France
- C. Bertucci, S. Vassilaras, J. Lasry, G. Paschos, M. Debbah and P. Louis-Lions, "Transmit Strategies for Massive Machine Type Communications based on Mean Field Games", ISWCS 2018, Lisbon, Portugal
- M. Chen, W. Saad, C. Yin, and M. Debbah, "Echo State Networks for Proactive Caching in Cloud-Based Radio Access Networks with Mobile Users", IEEE Transactions on Wireless Communications, Volume: 16, Issue:6, Page(s): 3520-3535,June 2017
- E. Zeydan, E. Bastug, M. Bennis, M. Abdel Kader, A. Karatepe, A. Salih Er and M. Debbah, "Big Data Caching for Networking: Moving from Cloud to Edge", IEEE Communications Magazine, vol. 54, no. 9, pp. 36 - 42, September 2016.

Mérouane Debbah (S'01–M'04–SM'08–F'15) received the M.Sc. and Ph.D. degrees from the Ecole Normale Supérieure Paris-Saclay, France. He was with Motorola Labs, Saclay, France, from 1999 to 2002, and also with the Vienna Research Center for Telecommunications, Vienna, Austria, until 2003. From 2003 to 2007, he was an Assistant Professor with the Mobile Communications Department, Institut Eurecom, Sophia Antipolis, France. From 2007 to 2014, he was the Director of the Alcatel-Lucent Chair on Flexible Radio. Since 2007, he has been a Full Professor with CentraleSupelec, Gif-sur-Yvette, France. Since 2014, he has been a Vice-President of the Huawei France Research Center and the Director of the Mathematical and Algorithmic Sciences Lab. His research interests lie in fundamental mathematics, algorithms, statistics, information, and communication sciences research. He is an IEEE Fellow, a WWRF Fellow, and a Membre émérite SEE.

Interview with Prof. Deniz Gündüz

Electrical and Electronic Engineering Department - Imperial College London, UK
Email: d.gunduz@imperial.ac.uk

Q1: Artificial intelligence has been successfully applied in many areas such as voice/video recognition and biomedical sciences. Nowadays, we are witnessing an increasingly interest in applying artificial intelligence in wireless communication problems. Do you think that artificial intelligence will experience the same success as in other areas? In other words, do you think artificial intelligence in wireless communications networks is just a hype or it will sustain its seemingly revolutionary role in the next decades? Why?

A1) I strongly believe that AI will play an essential role in the future of wireless networks. It is correct that, compared to some other application areas, wireless network design has already been a huge success, mainly built upon model-based solutions, rather than data-driven AI solutions. However, I think we are arriving at a crossroads, and the only way to satisfy the growing pressure on the wireless infrastructure to serve diverse applications with very different performance metrics and constraints, and to scale such solutions to the expected number of connected devices foreseen in the near future, is to integrate AI into the very core fabric of the network.

I think there is already significant work in this direction, including within standardisation bodies and by industry players, that shows that the current interest in AI for wireless is beyond a hype among researchers. For example, 3GPP has recently introduced the network data analytics function (NWDAF), a new network function to provide slice-level data analytics that can be used by other network functions to make decisions. Similarly, the internation telecommunication union (ITU) had formed the focus group on machine learning for future networks including 5G (FG-ML5G), which has recently released its proposal for a “Unified architecture for machine learning in 5G and future networks”. However, I should highlight that most of these efforts focus on network layer resource utilization. There are also many interesting research works exploiting ML techniques at the physical (PHY) layer. Although I believe they will also have an impact on future communication networks, their integration into practical systems might take longer.

Q2: Do you think that wireless networks will be fully controlled/designed by artificial intelligence tools with none or minimal intervention of humans in the future? If so, how far we are from that? If not, what are the limitations of artificial intelligence that prevent it from achieving that?

A2: I think there is already little intervention from humans on the operation of the wireless networks. Human impact is more on the design level, and many current protocols and algorithms are designed based on human expertise and intuition. In this regard, I would like to point to the body of work (by Mung Chiang and others) that have reverse engineered many existing networking protocols, such as TCP/IP, which have originally been designed based on engineering heuristics, and shown that they inherently solve some network utility optimization problem. While I find this very interesting, I also believe that as the complexity of the networks grow with the need to serve users with diverse needs and constraints, we will reach the limit of our intuitions, and need a more automatic way of solving these optimization problems. I believe AI/ML is the right approach to solve these complex distributed optimization problems, and this will reduce the human intervention on wireless networks even at the design level as we will have more and more self-organised architectures based on modern ML tools.

Q3: Artificial intelligence and its branch of machine learning are able to tackle many difficult problems in wireless communications. However, most of the problems in this area have been studied by researchers and engineers over the past decades using well-established techniques with strong mathematical background such as optimization, statistics and game theory. On the other hand, a common criticism is the difficulty to guarantee that machine learning

solutions will always work in general scenarios or converge to the optimal solutions. Another criticism to machine learning solutions that rely on neural networks is that they are seen as black boxes whose outputs cannot be completely explained, thus raising doubts about reliability and biases. What is your view about these aspects? Moreover, do you think that the classical solutions for wireless communications problems will be still useful in the future or they will be completely replaced by machine learning-based solutions?

A3: First of all, I do not agree with the often-cited criticism of ML-based solutions not providing performance guarantees, as opposed to current model-based solutions. Model-based solutions provide performance guarantees only under the model assumptions. For example, we can prove that a certain code is near capacity-achieving for a Gaussian channel, but in practice we may not have a Gaussian channel, and we do not have any theoretical performance guarantees in most cases, and we are limited by experimental evaluations. Some ML solutions can even be considered to be more robust as they can adapt to the environment in an online manner, and experimental results often show improved performance. Moreover, some ML algorithms do come with convergence guarantees under certain assumptions (on the model or the objective function), just like model-based solutions.

I agree that the interpretability is a serious concern, particularly for the robustness of the system. Engineers, especially for critical systems, need to understand potential implications of their design in response to changes in the environment or system parameters. This is a concern for general ML research, but there are many interesting results in this direction, and I believe as wireless researchers we should also work towards a better understanding of the implications of ML solutions on both the user and the network level. However, I personally see this as a research challenge rather than a roadblock for the adoption of ML techniques in wireless networks. As engineers we know very well that if something works, we need to understand and use it, rather than dismiss it as magic.

Of course, we do have many model-based solutions that work incredibly well, such as the near optimal channel codes, so there is less to expect from ML for those problems. However, in wireless communications more often than not we do not have (near) optimal model-based solutions, either because the system is very hard to model, or even with an accurate model, they lead to very difficult (e.g., NP-hard) problems. For example, for channels such as underwater acoustic, or optical communication channels, which are very difficult to model, we do not have good structured codes, and ML techniques have already provided promising results. Similarly, many distributed resource allocation problems lead to NP-hard optimization problems, which are typically solved through relaxation leading to suboptimal solutions. Recent results show that neural networks with stochastic gradient descent can learn to achieve better results.

I believe that the future (at least the near future) lies somewhere in between: we will continue to use many of our current model-based designs, but they will be combined and enhanced with ML-based solutions. A good example for such interaction is from the image compression domain. Current codecs apply some transform coding (DCT, wavelet, etc.), followed by quantization and entropy coding. We know that entropy coding can approach the fundamental theoretical limits; however, it requires an accurate model of the underlying source distribution. Recent state-of-the-art results apply ML tools to learn the distribution of quantized latent variables. Combined with a convolutional neural network replacing the transform coder, this has led to the first ever neural network based image compressor that beats the best-known image compression codec (BPG). I think this is a very good example of a great combination of modern ML tools with optimal structured code design, and I expect to see many more such applications.

Q4: Machine learning has many techniques/algorithms that can be classified in supervised learning, unsupervised learning, and reinforcement learning. The neural networks are very relevant building blocks of many machine learning solutions. In your opinion, what is(are) the most promising algorithm(s)/architecture(s)/framework(s) from machine learning area to be applied in wireless communications problems?

A4: I think each class of ML algorithms you mentioned has their application areas in wireless communication problems. Unsupervised learning can be used for anomaly detection, or for dimensionality reduction in source and channel coding. Supervised learning has many applications from detection and channel decoding to distributed power allocation. Reinforcement learning can be particularly relevant for distributed scenarios, or when an optimal action has to be identified in an online fashion through interactions with the environment, e.g., from content caching to power allocation and spectrum sensing.

In terms of solution tools, I think nobody can deny the prominence of deep learning in all these frameworks. While I believe deep learning has a lot of potential in solving wireless communication problems, I also would like to caution against throwing a deep network at every problem. I often see papers proposing deep learning solutions for problems that we have optimal solutions with much simpler methods. I think as a community we are still at the learning stage. Hopefully we will have a better picture of the efficacy and appropriateness of different methods for different problems as the current deep learning frenzy settles down.

Q5: In your opinion, what are the most important problems to be faced by artificial intelligence in physical layer? And in a system level? What are your own short-term and long-term research plans in artificial intelligence for wireless communication?

A5: As I have mentioned earlier, we have many problems in the physical layer for which we do not have good solutions. I think we should first target those problems, which hold more potential for impact. Joint source-channel coding, channel estimation, channel state information feedback, communication with feedback, or resource allocation for various distributed or limited-communication scenarios are some examples that we work on in my group. One of the challenges that we are currently tackling is over-the-air training, particularly for distributed settings. When we train neural networks to identify the optimal actions of multiple nodes, e.g., transmitter and receiver in the case of autoencoder-based joint source-channel code design, we need to backpropagate gradients from the receiver to the transmitter. While this is easy to do offline, it is not clear how it can be done if we want to train the networks at the time of implementation. In the system level, I believe reinforcement learning can have a huge impact on resource management on all layers of the network architecture; however, again a critical challenge is to factor in the complexity of these solutions and the time scales for convergence with respect to the network dynamics.

In the long term my group will continue to explore ML applications in communications and networking domains. In parallel, we are also exploring the other side of the same coin: how we need to (re)-design communication systems to enable ML applications at the edge. Today we have more and more edge devices collecting a lot of data, and learning from this data holds many potentials, but also brings new challenges as edge devices are typically limited in bandwidth and power resources. There is also growing privacy concerns against offloading all the data to a cloud server for processing. Therefore, we are looking at distributed/ federated learning algorithms at the wireless edge. Considering the growing demand for ML -- from mobile phones to autonomous vehicles, drone networks and IoT devices, I believe adapting our communication networks to the needs and constraints of ML algorithms will be a pressing challenge in the medium to long term.

Q6: Could you please briefly introduce the most recent research project(s) that you have done in this area? (Please explain the key idea(s) and interesting findings)?

A6: One of the exciting projects we are working on these days is ML-based joint source-channel coding. Today almost all communication systems are digital and designed based on the separation principle. For example, images are first compressed to get rid of redundancy and then channel coded against noise and interference. We have highly advanced codes for both compression (JPEG/ JPEG2000/ BPG) and channel coding (Turbo/ LDPC/ polar codes), fruits of decades-long research in both domains. However, we know that separate source and channel coding is inherently suboptimal, even if we employ optimal codes for each of the component, and in general we do not have good practical joint source-channel coding schemes.

We have replaced both the encoder and decoder with deep neural networks, and trained them jointly, which can be considered as an autoencoder with a nontrainable channel layer in between. This surprisingly achieves a better performance compared to state-of-the-art source and channel coding systems, even assuming an ideal Gaussian noise channel – for which we have near-capacity achieving channel codes. More interestingly the neural network learns to communicate more like an analog communication system, achieving graceful degradation with channel quality, as opposed to digital systems, which suffer from the cliff effect; that is, the reconstruction quality falls sharply if the channel quality goes below a certain threshold, and saturates at the level dictated by the compression scheme no matter how good the channel is. This property is especially attractive when broadcasting to many receivers, or when transmitting over a time-varying channel.

In another project, we study automatic modulation detection, a very popular application of ML in wireless communications these days. It has been shown that using only a limited number of time samples from a transmitted signal, it is possible to detect its modulation scheme with very high accuracy. This can potentially be used maliciously, as modulation detection is the first step in many attacks. We have studied how we can avert such attacks without impacting the performance of our communication system. This requires shaping the modulation constellation at the transmitter in an intelligent manner. We have used tools developed for adversarial attacks against neural networks, that have recently gained popularity by showing the vulnerability of deep learning based classifiers: a single pixel change in an image can fool them. While the goal there is to distort the image without being noticed by a human observer, while still fooling the classifier, in our case the goal is to fool the classifier of the attacker without damaging the legitimate receiver's accuracy. Also note that all the changes applied by the transmitter goes through a noisy channel in our problem, as opposed to directly modifying the data samples. We have shown that it is indeed possible to communicate reliably without being intercepted by a malicious attacker.

We are also doing a lot of exciting work on distributed computation as well as distributed/ federated edge learning, which as I said, looks at how we can adapt our communication systems in order to increase the speed and accuracy of ML algorithms among wireless agents.

Q7: Beyond your own work, are there any resources that you would like to recommend, especially to those who are new in this field and want to learn more about artificial intelligence? Are there any specific resources that you recommend related to artificial Intelligence in the context of wireless and communication networks?

A7: In terms of AI/ML, resources are almost unlimited. I would recommend a fresh starter to follow one of the popular online courses. They provide a structured introduction to the basic ideas and tools. There are very good courses on general ML, and on deep learning or reinforcement learning. For more specialised topics, say federated learning, or applications of ML in communication systems, my suggestion is to read the fundamental papers, and there are also plenty of talks available online. One of the challenges of this research area is the speed of development. It is quite challenging to catch up with new ideas in ML, or their applications to wireless as the number of papers has exploded. I suggest following some of the main conferences and journals to be aware of some of the most important developments and research trends in general, and keeping an eye on arXiv preprints as these days most people post their results on arXiv even before submitting to a conference.

Q8: What is your most important contribution (journal, magazine or conference article, or patents) in the topic?

A8: For a general introduction to the field, I would recommend the tutorial paper we have written for the Special Issue on Machine Learning in Wireless Communication we have edited recently:

D. Gunduz, P. de Kerret, N. Sidiroupoulos, D. Gesbert, C. Murthy, M. van der Schaar, "Machine learning in the air", IEEE Journal on Selected Areas in Communications, vol. 37, no. 10, pp. 2184-2199, Oct. 2019.

The work on deep joint source-channel coding is explained in the following paper:
E. Bourtsoulatze, D. Burth Kurka and D. Gunduz, “Deep joint source-channel coding for wireless image transmission”, *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 567 - 579, Sep. 2019.
The following is an application of reinforcement learning to wireless content delivery, but the tools and results can be relevant for other applications in wireless communications:
S. O. Somuyiwa, A. Gyorgy and D. Gunduz, “A reinforcement-learning approach to proactive caching in wireless networks”, *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 6, pp. 1331 - 1344, Jun. 2018.

Deniz Gündüz received his M.S. and Ph.D. degrees from NYU Tandon School of Engineering (formerly Polytechnic University) in 2004 and 2007, respectively. After postdoctoral positions at Princeton and Stanford Universities, se served as a research associate at CTTC in Spain for three years. In 2012 he joined the Electrical and Electronic Engineering Department of Imperial College London, UK, where he is currently a Reader (Associate Professor) in information theory and communications, serves as the deputy head of the Intelligent Systems and Networks Group, and leads the Information Processing and Communications Laboratory (IPC-Lab). His research interests lie in the areas of communications, information theory, machine learning, and privacy. Dr. Gündüz is an Editor of the IEEE Transactions on Wireless Communications and IEEE Transactions on Green Communications and Networking. He served as a Guest Editor of the IEEE JSAC Special Issue on Machine Learning in Wireless Communication (2019). He is a Distinguished Lecturer for the IEEE Information Theory Society (2020-21). He is the recipient of the IEEE Communications Society - Communication Theory Technical Committee (CTTC) Early Achievement Award in 2017, a Starting Grant of the European Research Council (ERC) in 2016, and best paper awards at several conferences.

Feature Topic: Terahertz Communications
Editor: Hadi Sarieddeen

King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
Email: hadi.sarieddeen@kaust.edu.sa

Terahertz (THz)-band communications are expected to play a pivotal role in the upcoming sixth-generation (6G) of wireless mobile communications, enabling ultra-high bandwidth and ultra-low latency communication paradigms. Towards this end, high capacity THz links have been advocated to replace wired backbone connectivity in network backhauls and data centers. The holy grail of THz communications, however, is to enable indoor and outdoor mid-range mobile wireless communications, in the context of device-to-device, vehicular, and personal communications.

Due to the lack of compact and efficient THz devices (the so-called *THz gap*), THz-band applications have been traditionally restricted to the areas of imaging and sensing. However, following recent advancements in THz signal generation, modulation, and radiation, the THz band is opening up for everyday applications. THz transceiver designs are mainly electronic and photonic. While photonic technologies have a data rate advantage, electronic platforms can generate higher power. Nevertheless, since satisfying emerging system-level properties requires designing efficient and programmable devices, as opposed to perfect devices, integrated hybrid electronic-photonic THz systems are emerging, as well as compact graphene-based plasmonic solutions.

Like any new technology, THz communication is attracting both praise and criticism. Is pushing microwave communications beyond the well-established millimeter-wave (mmWave) band worth the effort? And why should we settle for THz communications if high data rates can be supported by the more mature visible light communications (VLC)? In fact, it is not yet clear how future THz communication systems can combat the inherent limitations at high frequencies. For instance, high propagation losses and power limitations result in very short communication distances and frequency-dependent molecular absorptions result in band-splitting and bandwidth reduction. Skeptics even go beyond these technical issues and raise health (due exposure to THz radiation) and privacy (due to high resolution sensing) concerns.

This feature topic highlights some facts and debunks some myths surrounding this emerging technology. The contributors summarize the latest advancements in THz technology and discuss how THz communications can reap the benefits of both mmWave and VLC communications. In the following sections, we present one position paper and three interviews with leading experts in the field. The paper is written by Prof. Cyril C. Renaud, who is an expert in THz photonics. The interviews are carried out with Prof. Daniel Mittleman (a pioneer of THz technology), Prof. Josep M. Jornet (active in the field for more than 10 years), and Dr. Onur Sahin (involved in THz research through industry). I take this opportunity to thank them all for taking the time to share with us their valuable insights.

Hadi Sarieddeen (S'13-M'18) received the B.E. degree (summa cum laude) in computer and communications engineering from Notre Dame University-Louaize (NDU), Zouk Mosbeh, Lebanon, in 2013, and the Ph.D. degree in electrical and computer engineering from the American University of Beirut (AUB), Beirut, Lebanon, in 2018. He is currently a postdoctoral research fellow in the Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division at King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia. His research interests are in the areas of communication theory and signal processing for wireless communications, with emphasis on large, massive, and ultra-massive MIMO systems and terahertz communications.

Position Paper: The future of broadband wireless communication: Is THz photonics the answer?

Prof. Cyril C. Renaud
Department of Electronic and Electrical Engineering, University College London, London, UK

I. Abstract

In an era where we have increasing demands for high data rate both in wired and wireless communication, new solutions are required to meet it. For the wireless channel, a number of solutions have now emerged from massive MIMO radio to optical wireless. In this paper, we will discuss the advantages and disadvantages of using THz carrier frequencies for the wireless signal and why they might offer the best compromise between available bandwidth and ease of alignment. We will also discuss the issue of integrating seamlessly the wireless and wired channel and the advantages it could bring in future network architectures. Within such a potential development we will discuss the advances in demonstrations of wireless bridges at THz carrier frequencies over a fibre network as an argument to the advantages of using photonic solutions compared to electronic ones. Finally, we will look into the potential for photonic integration to create a viable THz photonics wireless technology.

II. Introduction

The increase on data traffic, in particular with Internet protocol has been exponential for a number of years and is now reaching several hundred Exabytes per month [1]. While the strain is felt in all part of the network, it is particularly noticeable in the wireless channels going from mobile to backhaul. This means that the expected data rate in the wireless channel is expected to exceed 100 Gbit/s within the next 10 years [2]. to reach such data rate in the number of wireless channel scenarios, a number of solutions have emerged, either through increased complexity through massive multiple input multiple output architecture combined with higher level modulation formats using standard microwave carriers to wireless optical signals. All these solutions offer different benefits and have the potential to reach the required data rate for future wireless channels.

The use of standard radio channel at microwave frequency would enable the use of existing technologies in smaller wireless cells, however this comes at the cost of still using a highly congested part of the spectrum. It then becomes clear that it might be worth to investigate higher carrier frequencies. Indeed,

a 2017 technology review by Ericsson highlights the importance of higher carrier frequencies for future wireless link [3]. While the review focuses on the W-band (75-110 GHz) and D-band (110-170 GHz) is notes that even higher frequencies will be required. As seen in Figure 1, this comes from the amount of unallocated bandwidth above 300 GHz, the THz region. For example, around 300 GHz we can find two bands with respectively 68 and 46 GHz of bandwidth, which should enable 100 Gbit/s transmission with relatively simple modulation formats. However, the free space path losses (FSPL) at these frequencies become a dominating factor and beam collimation (high gain antenna) will be required for an operational link, while steering will be necessary for mobile applications.

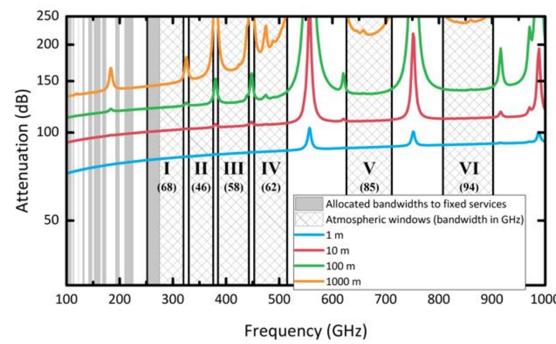

Fig. 1. Atmospheric attenuation at frequencies from 100 GHz to 1 THz.

In a similar idea of increasing carrier frequencies, one could go further and use optical frequencies. There the bandwidth available is counted in THz, and modulation technologies in the fibre network have already achieved data rate beyond 100 Gbit/s. However, in such a link FSPL is even more a limitation, therefore collimation, and as a consequence alignment, are harder to perform than at lower frequencies.

In this paper, we will discuss in our first part, why THz communication should offer the best compromise between the different solutions for the future wireless channel technology. We will then move onto discussion photonics technologies and their advantages. We will finish with a short discussion on what still needs to be developed to have a viable photonic THz technology for future wireless networks.

III. THz communication

As already discussed, THz frequencies (above 300 GHz) give access to sufficient bandwidth to meet the future network requirements, however technologically the THz range has been hard to access due to a lack of sources and was confined in niche applications such as space technologies, with the use of lower frequencies oscillators and multiplier chains. In the last 20 years advances in semiconductor technology has opened up the lower part of the THz range and electronic and photonic components are now available to operate at frequencies above 300 GHz. For example, using electronic components and in particular III-V high electron mobility transistor, a link at 240 GHz was demonstrated with over 100 Gbit/s data rate [4]. Further, development in SiGe transistor is going at pace and frequencies in the W-band are already achievable, which could lead to large production of silicon-based electronic systems for THz communication.

Once sources are available the key is to use the available bandwidth in the channel effectively. First, we note that at 300 GHz for a backhaul 1 km link would be facing 140 dB of FSPL and 1 dB of atmospheric absorption loss. So, for a typical source power of -10 dBm (or about 10 dBm with amplification) and receiver sensitivity of the order of -30 dBm, we would require antenna gain over 50 dBi at both transmitter and receiver to compensate for the FSPL (see table 1 for an example of a link budget).

Table 1: Example of a 100 Gbit/s, 1km link budget at 300 GHz carrier

Data Rate	100	Gbit/s	
Source power	0	dBm	
Tx antenna gain	50	dBi	
Transmission Loss	143	dB	1km, 1dB absorption
Rx antenna gain	50	dBi	
Received power	**-43**	**dBm**	
IF power	-53	dBm	10dB CL
IF equivalent noise	-170	dBm/Hz	
E_b/N_0	**9**	**dB**	
System Margin	2	dB	QPSK at BER=10^{-3}

Table 2: Example of THz wireless link demonstrations, extracted from [5].

Data rate (Gbit/s)	Distance (m)	Carrier (GHz)	Technology	Ref.
100	20	237.5	Ph./Electronic	[4]
100	850	240	Electronic	[8]
64	1	300	Electronic	[9]
40	10	300	Photonic	[10]
48	0.5	300	Photonic	[11]
3	50	340	Electronic	[12]
32	0.5	385	Photonic	[13]
46	2	400	Photonic	[14]
50	0.5	330	Photonic	[15]
60	0.5	400	Photonic	[16]
2.5	3	625	Electronic	[17]

It is then clear that with current sources and receiver technologies THz links at 100 Gb/s are achievable, when one is using appropriate antenna gain. Further, while the example used is for a backhaul link of 1 km, the vast majority of envisaged applications are expected to be with shorter links (below 100 m) [5]. This includes for example rack to rack communication in data centres [6], kiosk to mobile device high data rate link, indoor wireless communication etc…

With that in mind, one can see as shown in [5] (table 2) that THz links have now been demonstrated with data rate beyond 100 Gbit/s over distances fully relevant for the envisaged applications. The bandwidth on offer, the less stringent alignment requirements and the 2 order of magnitude lower losses in fog compared to optical wireless [7] make, therefore, THz wireless technology a clear compromise solution for future wireless channels.

IV. THz photonics and Wireless Bridges

As seen in table 2, a lot of the early demonstrations with some of the best performances have been done using photonic techniques [5]. All these techniques rely on the use of standard optical communication systems, or radio over fibre techniques to generate the data signal, and heterodyning with a second laser oscillator in a photomixer to generate the modulated THz signal, while the receiver is typically a Schottky barrier diode (SBD) used as a sub-harmonic mixer. The advantage of such an approach is that the generation of high data rate signals on an optical carriers is fully developed while the advance in the development of uni-travelling carrier photodiodes (UTC-PD) as efficient photomixers [18] has enabled photonic techniques to be used to emit power close to 0 dBm at 300 GHz. There is a further obvious advantage of using such techniques as development of future communication networks are almost exclusively seen as putting fibre links as close to the access point as possible. In that case, a technique that does not require detection, decoding and remodulation to transfer to the wireless channel is clearly advantageous, should save some energy and be transparent to the overall wired network.

For that last scenario, where the wireless part of the network will integrate transparently with a fibre network, one would require a sub-THz communications wireless bridges which refer to wireless links connecting two portions of a fibered network. In that case, the received signal should be up-converted to the optical domain and then transmitted to an optical receiver through an extra portion of fibre. THz would be the only part of the spectrum, apart from free space optics, that offer sufficient bandwidth to match the data rate available per channels in a fibre network.

In Figure 2a, we represent the basic schematic of a THz wireless bridge, where the optical signal is generated at a central office (CO) and transmitted through fibre to a remote antenna unit (RAU), where photomixing occurs (O/THz). The signal is then received by another remote antenna unit where a THz to optical (THz/O) conversion system will modulate another laser. That signal is then sent through fibre to an optical network unit (ONU) that could either detect the signal directly (DD) or coherently.

The key to that system is the THz/O converter. In figure 2b, we show a set of solution for such a converter, which could be either using a high-speed modulator such as the one found in [19]. However, in current available technology these high-speed modulators still need development. Alternatively, as seen in figure 2b, one can down convert the THz signal either to baseband or to an intermediate frequency (IF) and use a standard optical modulator combined with optical filtering to generate the new optical channel. This would offer the opportunity to pick higher performances components to create the THz/O conversion and reduce the losses incurred in that process at the cost of increased complexity.

One example of such a system is using a THz mixer to go to an IF. As seen in figure 3, once could envisage sending 5x20 Gbit/s single side band optical channels for a total throughput of 100 Gbit/s. For this experiment, this was transmitted through 10 km of single mode fibre (SMF), photomixed with a second laser in a UTC-PD to generate a 250 GHz carrier signal. This was then detected by a sub-harmonic mixer to generate an IF. This amplified IF was then used to modulate an optical oscillator and transmitted through 40 km of fibre to be coherently detected at the ONU using standard optical communication digital signal processing. As seen in in Figure 3, despite the added noise figure to the link due to O/THz and THz/O conversion, all channels were successfully transmitted, enabling a 100 Gbit/s wireless bridge.

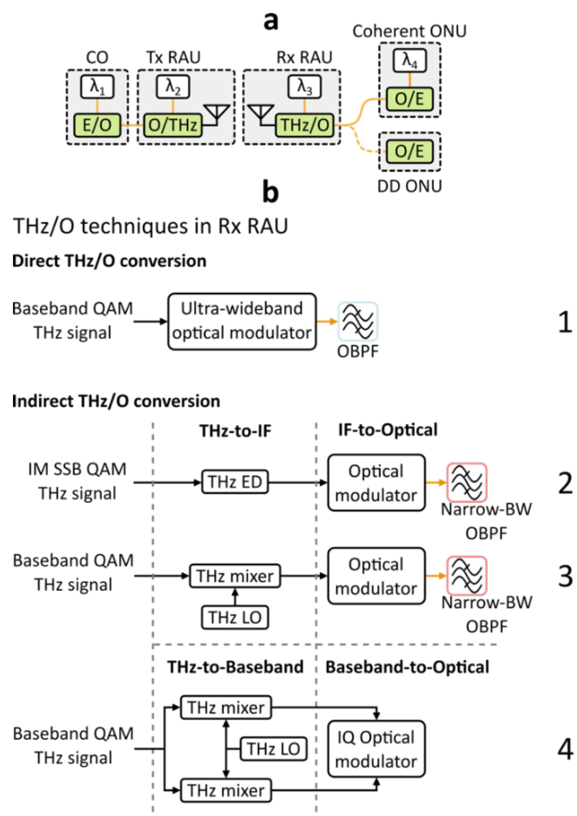

Fig. 2. (a) schematic representation of a wireless bridge based on photonic THz generation and (b) methods for THz-to-optical conversion in the Rx RAU (only schemes supporting higher-order modulation are considered). Note that only schemes 2 and 3 are compatible with a direct detection (DD) optical network unit (ONU). For schemes 1 and 4, a coherent optical receiver must be employed to recover the signal. CO: central office; OBPF: optical band-pass filter; IM: intensity modulated; SSB: single sideband; QAM: quadrature amplitude modulation; ED: envelope detector.

This clearly demonstrates that within a fully wired-wireless hybrid network where the wired network is based on high data rate fibre techniques, photonic THz wireless solution are extremely attractive. They would offer the required throughput within a frequency range that is a good compromise between microwaves and optical wireless while remaining transparent to the fibre network.

V. What is next?

While, to date, there has been a number of demonstrations of photonic THz wireless links and wireless bridges with data rate beyond 100 Gbit/s, most of them have been done as a laboratory benchtop demonstration. The technology still needs to progress to be fully be implemented in systems, in particular in having fully integrated RAUs.

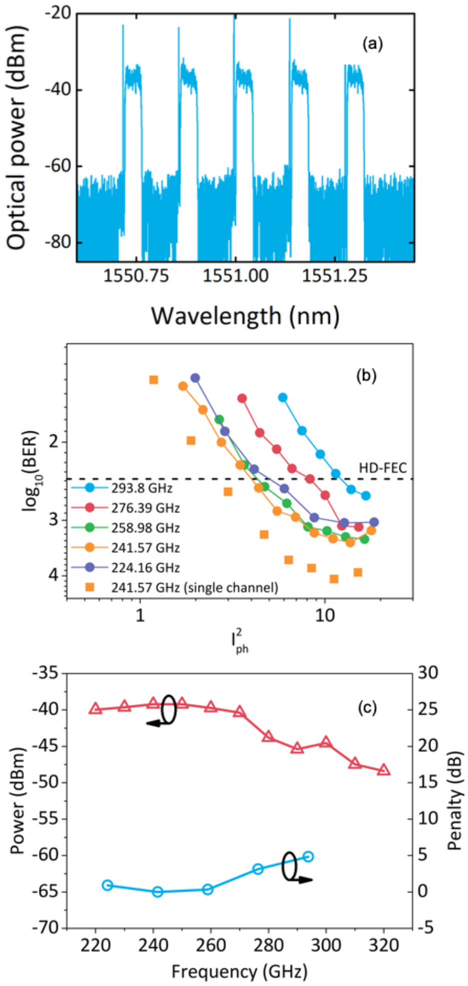

Fig. 3. (a) 5 single side band channels with 20 Gbit/s and 18 GHz channel spacing. (b) Received BER at the ONU for each channel showing successful transmission beyond the HD-FEC limit for all channels. (c) Power received at different carrier frequency and penalty to single channels transmission for each channel.

There is a lot of promise in the different works done currently, both with the development of high frequency electronic components to offer amplification solutions (higher transmitted power), to better high-speed modulators for the THz/O conversion. However, these developments are still done at single device level which is not competitive to silicon CMOS based technology in term of integration. While, as mentioned, silicon technology operation frequency is ever increasing, it is clear that to be competitive THz photonic technology will need to reach a higher level of integration.

For the purpose of the argument, we can have a look at a photonic integrated circuit (figure 4) [20] that was developed for the W band. That circuit could be used as a transceiver and would only be missing the modulation and amplification to be used as a full wireless bridge RAU. The actual key components used for the two-way conversion is the UTC-PD that is here integrated with lasers, optical modulators and amplifier. This component has been demonstrated as a transmitter for frequencies above 2THz while it was also demonstrated as a receiver up to 600 GHz [21]. It then becomes clear that there would be a path to create fully integrated transceivers as part of the remote antenna unit and that THz photonic technology is a clear and strong contender for future wireless channels for data rate above 100 Gbit/s.

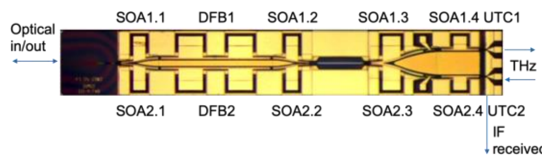

Fig. 4. Photonic integrated chip including lasers and UTC_PDs to be used as a wireless transceiver.

References

[1] Cisco; and/or its Affiliates, The Zettabyte Era: Trends and Analysis. Cisco Vis. Netw. Index, 1-24 (2014).

[2] T. Kürner, S. Priebe, "Towards THz communications – status in research standardization and regulation," *Journal of Infrared, Millimeter and Terahertz Waves*, Vol. 35, pp. 53-62, 2014.

[3] J. Edstam, J. Hansryd, S. Carpenter, T. Emanuelsson, Y. Li, and H. Zirath, "Microwave backhaul evolution-reaching beyond 100Ghz," 2017.

[4] S. Koenig, D. Lopez-Diaz, J. Antes, F. Boes, R. Henneberger, A. Leuther, A. Tessmann, R. Schmogrow, D. Hillerkuss, R. Palmer, T. Zwick, C. Koos, W. Freude, O. Ambacher, J. Leuthold, and I. Kallfass, "Wireless sub-THz communication system with high data rate," *Nat. Photonics*, vol. 7, no. 12, pp. 977–981, 2013.

[5] T. Nagatsuma, G. Ducournau, and C. C. Renaud, "Advances in terahertz communications accelerated by photonics," *Nat. Photonics (Review)*, vol. 10, no. 6, pp. 371–379, 2016.

[6] Alan Davy, Luis Pessoa, Cyril C. Renaud, Edward Wasige, Mira Naftaly, Thomas Kürner, Glenn George, Oleg Cojocari, Niamh O'Mahony, Marco AG Porcel, “Building an end user focused THz based ultra high bandwidth wireless access network: The TERAPOD approach,” Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2017 9th International Congress on, Munich, 2017.
[7] ITU-R P.840-3, "Attenuation due to clouds and fog," 1999.
[8] I. Kallfass et al., “64 Gbit/s transmission over 850 m fixed wireless link at 240 GHz carrier frequency,” *Journal of Infrared Milli. Terahz Waves*, Vol. 36, pp. 221-233, 2015.
[9] I. Kallfass et al., “Towards MMIC-based 300 GHz indoor wireless communication systems,” *IEICE Trans. Electronics*, Vol. E98, pp. 1081-1088, 2015.
[10] A. Kanno et al., “Coherent terahertz wireless signal transmission using advanced optical fiber communication technology,” *Journal of Infrared, Millimeter, and Terahertz Waves*, Vol. 36, pp. 180-197, 2015.
[11] T. Nagatsuma, et al., “Terahertz wireless communications based on photonics technologies,” *Optics Express*, Vol. 21, pp. 477-487, 2013.
[12] C. Wang et al., “0.34-THz wireless link based on high-order modulation for future wireless local area network applications,” *IEEE Trans. Terahertz Science and Technology*, Vol. 4, pp. 75-85, 2014.
[13] G. Ducournau et al., “32 Gbit/s QPSK transmission at 385 GHz using coherent fibreoptic technologies and THz double heterodyne detection,” *Electronics Lett.*, Vol. 12, pp. 915-917, 2015.
[14] G. Ducournau et al., “Ultrawide bandwidth single channel 0.4 THz wireless link combining broadband quasi-optic photomixer and coherent detection,” *IEEE Trans. Terahertz Science and Technologies*, Vol 4, pp. 328-337, 2014.
[15] T. Nagatsuma & G. Carpintero, “Recent progress and future prospect of photonics enabled terahertz communications research,” *IEICE Trans. Electronics*, Vol. E98-C, pp. 1060-1070, 2015.
[16] X. Yu et al., “60 Gbit/s 400 GHz wireless transmission. presented at Intern. Conference on Photonics in Switching,” (PS2015), Florence Italy, 2015.
[17] L. Moeller, J. Federici, K. Su, “2.5Gbit/s duobinary signalling with narrow bandwidth 0.625 terahertz source,” *Electronics Lett.*, Vol. 47, pp. 856-858, 2011.
[18] E. Rouvalis, C. C. Renaud, D. Moodie, M. J. Robertson, A. J. Seeds, “Continuous Wave Terahertz Generation from Ultra-Fast InP-based Photodiodes,” *IEEE transactions on Microwave Theory and Techniques*, Vol. 60, pp. 509-517, 2012
[19] Y. Salamin, B. Baeuerle, W. Heni, F. C. Abrecht, A. Josten, Y. Fedoryshyn, C. Haffner, R. Bonjour, T. Watanabe, M. Burla, D. L. Elder, L. R. Dalton, and J. Leuthold, “Microwave plasmonic mixer in a transparent fibre–wireless link,” *Nat. Photonics*, no. Mim, 2018.
[20] F. Van Dijk, G. Kervella, M. Lamponi, M. Chtioui, F. Lelarge, E. Vinet, Y. Robert, M. J. Fice, C. C. Renaud, A. Jimenez, G. Carpintero, “Fully Monolithically integrated InP Heterodyne millimeter wave source for high data rate transmission,” *IEEE Photonics Technology Letters*, Vol. 26, PP.965-968, 2014
[21] C. C. Renaud, M. J. Fice, L. Ponnampalam, M. Natrella, C. Graham, and A. J. Seeds , “Uni-Travelling Carrier photodetectors as THz detectors and emitters,” Photonics West 2015, San Francisco, 2015.

Professor Cyril C. Renaud (SMIEEE) received the degree of engineering from the Ecole Supérieure d’Optique, Orsay, France, and the Diplôme d’Etudes Approfondies (D.E.A.) in Optics and Photonics from the University Paris XI, Orsay, France, in 1996. He spent one year as a project engineer with Sfim-ODS, working on the development of microchips lasers and portable range finders. He, then, joined the Optoelectronics Research Centre, University of Southampton, Southampton UK, in 1998, to work on diode pumped high-power ytterbium-doped fibre-lasers, with particular interest on Q-switched systems and 980-nm generation. This work led to the award of a PhD in 2001. He is currently a Professor of Photonics at University College London, Director of Graduate Research and the programme director for the

UCL/Cambridge Doctoral Training Centre in Integrated Photonic and Electronic Systems. His work on photodiodes, integrated photonic and THz photonic has led to over 180 publications in peer reviewed journals and international conferences, attracting over 3200 citations, and 3 patents.

Interview with Prof. Daniel Mittleman

Brown University, Rhode Island
Email: daniel_mittleman@brown.edu

Q1: What is, in your opinion, the most appropriate way to define THz communications? What frequency ranges constitute the THz band and what are the key system performance requirements?

A1: The IEEE Transactions on Terahertz Science and Technology defines the terahertz range as running from 0.3 THz to 10 THz. However, that journal has published quite a few articles (including their most highly cited paper ever) which mostly (or entirely) focus on frequencies below that range – specifically, starting from 100 GHz, rather than 300. From my point of view, one can draw a fairly natural distinction between millimeter-wave systems that operate below 100 GHz (including, for example, the millimeter-wave bands of 5G, existing automotive radar near 77 GHz, active denial crowd-control systems in the 95 GHz range, and next-generation imaging systems at 94 GHz) and those which operate above 100 GHz (of which there are far fewer familiar examples). So I'm pretty comfortable with a definition that starts at 100 GHz and goes up from there. In my lexicon, any system designed for wireless communications at 100 GHz or above should be considered a "THz communications" system. It seems unlikely that any such system will ever be useful at frequencies much above 1 THz, simply because the atmospheric water vapor absorption spectrum becomes decidedly congested at these higher frequencies. So, there will always be a natural gap for wireless communications between the high end of the THz range and the realm of free-space optics.

Addressing the question of key system performance requirements would require me to write a book, not a paragraph.

Q2: THz communications have been a subject of both praise and criticism. Since communication technologies are already mature at the neighboring bands, the mere necessity for exploiting the THz-band is questionable. Is pushing microwave communications beyond the well-established millimeter-wave band worth the effort? And why should we settle for THz communications when cheap off-the-shelf light-emitting diodes can support much higher data rates in visible light communications? Is this just a THz-hype or will THz communications sustain their seemingly revolutionary role in future communication systems?

A2: The ever-increasing demand for bandwidth is one obvious reason why we will need to move to higher frequencies than those which are currently used in wireless systems. It seems to have been worth the effort to include millimeter-wave bands in the 5G standard, so I suppose it will continue to be worth the effort to push to even higher frequencies where even higher data rates can be supported. It may be worth pointing out that it is not possible to send uncompressed 8K video via a wireless link, using any 4G system, or even any envisioned 5G system. The bandwidth simply cannot support it. So, we are already building devices which suggest the need for more bandwidth than even 5G will be able to provide. This argues that the use of higher frequencies will be mandatory, at some point.

You are quite correct in pointing out that free-space optics (e.g., at 1.5 microns) is a competing technology. In my view, the two ideas (FSO vs. THz) both have merit, and both have problems. I do not agree with your statement that FSO can support much higher data rates – have they reached a terabit per second yet? Anyway, one could write a very long article on the relative trade-offs between the two. Just to give one example: FSO signals are much more susceptible to disruption by, e.g., atmospheric turbulence (scintillation effects) or fog, whereas THz beams may be more susceptible to snow (although maybe not rain – it is a common misconception that rain would kill THz propagation). In the end, it is impossible at this moment to say which of these two very different technology platforms is superior, in part because the answer depends to a great extent on the details of the scenario. Probably, both will be useful, each in different situations.

Q3: Following recent advancements in electronic, photonic, and plasmonic technologies for THz transceiver design, the so-called THz-gap is closing. Is there a race/competition between these three technologies (Please comment on the strong and weak aspects of each)? Which technology supports the best range of reconfigurability for adaptive cognitive applications (the interest of our readers)? Is there a clear winner or will we settle for hybrid solutions?

A3: I would not characterize this as a 'race'. As above, different technologies will be valuable in different situations. As it stands right now, photonic technologies have the clear advantage in data rate (e.g., already demonstrated hundreds of gigabits per second), where electronic platforms remain superior in their ability to generate higher power. It is too early to say what will end up being the 'transceiver of choice'. Although, as usual, one should never bet against silicon. I have recently co-authored an article on the convergence of electronic and photonic technologies in the terahertz range, with an eye towards the impact of this convergence on future systems for communications and sensing. See here: https://www.brown.edu/research/labs/mittleman/sites/brown.edu.research.labs.mittleman/files/uploads/Sengupta_Review.pdf

I'm afraid that I do not know what an 'adaptive cognitive application' is, so I can't really comment on which technology will support that best.

Q4: What are, in your opinion, the most disruptive THz breakthroughs that have emerged in the past few years? What do you think are the most important remaining technical challenges or open problems in the field?

A4: Some of the most important recent THz breakthroughs are in areas of fundamental science: specifically, the ability to generate extremely high THz peak fields for the purpose of driving nonlinear responses in materials, and the ability to perform imaging and spectroscopy at length scales well below the wavelength. These two have both been truly transformative.

In the realm of engineering, I would point to the very dramatic advances in silicon and SiGe integrated circuit technologies, which have accomplished incredible things in the last 5 years. These advances will enable many commercial systems that would otherwise be unrealistic due to cost or form factor considerations, including in particular networks.

I have recently written a fairly extensive review article on these topics, to which you could point your readers: "Invited Perspective: Terahertz Science and Technology," D. M. Mittleman, Journal of Applied Physics, 122, 230901 (2017).

Q5: It is argued that the breakthrough that this field will introduce is not solely driven by the high achievable data rates, but more profoundly by the combination of THz communications, THz sensing and imaging (traditional THz applications), and high-accuracy localization applications. Do you see real potential in such application merges? Can you envision a role for machine learning and artificial intelligence in this regard?

A5: I do not entirely agree with the statement. I think the high achievable data rates alone will be transformative. However, I also see potential in hybrid systems which accomplish not only comm, but also those other things that you mentioned. I think it's clear that multi-functional systems are going to be valuable in countless ways. Of course, artificial intelligence could play an important role in the operation of these systems, if for no other reason than that the systems will be capable of generating a LOT of data, very quickly. But it is far too early to be specific about the details. Nobody has yet even built a terahertz network, let alone a multi-functional one.

Q6: How do you describe the interest/involvement of industry in THz communications? When do you think we will start to see commercially available solutions?

A6: Today, most industrial players are focused on 5G, and therefore not thinking very hard about things beyond 5G, yet (with a few small exceptions). There are already quite a few commercial deployments of terahertz systems, but mostly in the realm of non-destructive evaluation, in areas like automotive, pharmaceutical, and manufacturing. In the realm of networks for communications, it is premature to think about commercial deployment. It's still a research topic. I would love to see more interest in this area from the telecom giants, but I'm not holding my breath.

Q7: Could you please briefly introduce the most recent research project(s) that you have done in this area (Please explain the key idea(s) and interesting findings)? What are your own short-term and long-term plans?

A7: One interesting recent project in my group involved a study of the security of directional THz wireless links. Directionality offers a higher level of security, but vulnerabilities to eavesdropping and jamming still exist. This enhanced (but not perfect) security is yet another reason why one might consider moving to higher frequencies, beyond 5G. Our 2018 article in Nature (vol. 563, pp. 89-93) was the first to consider the question of vulnerability to eavesdropping in the terahertz range:

https://www.brown.edu/research/labs/mittleman/sites/brown.edu.research.labs.mittleman/files/uploads/Ma_eavesdropping_0.pdf

We are also thinking about various aspects of how a multi-user network would operate. For example, how does the network know where to direct the signal for a given client? How does the network allocate resources and steer the signal for mobile clients? How many of these directional links can be packed into a single LAN at a given frequency without interfering? How well do NLOS links work? The answers to all of these questions are unique for terahertz signals – they are not merely extensions of the similar considerations at lower frequencies.

Q8: As major contributions to THz technology are still at the level of transceiver design, what advice do you give to researchers who are approaching this field from a communication system, signal processing, and networking perspective? What is the take-away message that you prefer to leave our readers with?

A8: The main take-away is that the field needs more people like you! What we have found recently is that the signal processing and networking considerations are really very different in the THz realm (as compared to networks at lower frequencies) – and not only that, the ideas are really closely linked to the transceiver or device architectures that one chooses to employ. We have been developing device concepts that (a) cannot operate well at lower frequencies, so they are unique to the THz bands, and (b) enable new MAC protocols that cannot be even considered for networks at lower (or higher) frequencies. Optimizing the protocols requires an understanding of the devices, and optimizing the devices requires knowledge of what these protocols are trying to accomplish and under what constraints. So one really requires a close collaboration between signal processing/networking/communications people and hardware/device/physics people. Three years ago, there were no such collaborations on earth. Now, I am aware of a few, but still not too many. The problem is too big for just a few people to be tackling it.

Dr. Mittleman received his B.S. in physics from the Massachusetts Institute of Technology in 1988, and his M.S. in 1990 and Ph.D. in 1994, both in physics from the University of California, Berkeley , under the direction of Dr. Charles Shank. He then joined AT&T Bell Laboratories as a post-doctoral member of the technical staff, working first for Dr. Richard Freeman on a terawatt laser system, and then for Dr. Martin Nuss on terahertz spectroscopy and imaging. Dr. Mittleman joined the ECE Department at Rice University in September 1996. In 2015, he moved to the School of Engineering at Brown University . His research interests involve the science and technology of terahertz radiation. He is a Fellow of the OSA, the APS, and the IEEE, and is a 2018 recipient of the Humboldt Research Award. He currently serving a three-year term as Chair of the International Society for Infrared Millimeter and Terahertz Waves.

Interview with Prof. Josep Miquel Jornet
Northeastern University, Boston
Email: j.jornet@northeastern.edu

Q1: What is, in your opinion, the most appropriate way to define THz communications? What frequency ranges constitute the THz band and what are the key system performance requirements?

A1: There have been historically different definitions for the THz band. Traditionally, on the one hand, for RF engineers, "anything above 100 GHz" could be considered THz band. On the other hand, for optical engineers, any frequency below 10 THz (the far infrared) was already THz band. From these, one can define the THz band as the frequency spectrum between 100 GHz and 10 THz. Nevertheless, according to the ITU-R, the THz band can be closely mapped to the Tremendously High Frequency (THF) band, between 300 GHz (right after the millimeter-wave spectrum) and 3 THz (not yet infrared).

Q2: THz communications have been a subject of both praise and criticism. Since communication technologies are already mature at the neighboring bands, the mere necessity for exploiting the THz-band is questionable. Is pushing microwave communications beyond the well-established millimeter-wave band worth the effort? And why should we settle for THz communications when cheap off-the-shelf light-emitting diodes can support much higher data rates in visible light communications? Is this just a THz-hype or will THz communications sustain their seemingly revolutionary role in future communication systems?

A2: There have been and still are many skeptics about the potential of the THz band for communications. This is the result of decades of discouraging results mainly due to the lack of capable device technologies and the required communication expertise to support THz communications. The THz band is between the realm of micro/millimeter-waves and the realm of optics. In the micro/millimeter-wave realm, we think of electromagnetic radiation as waves and we generally generate the signals using electronic devices. In the optical realm, we prefer to model electromagnetic radiation in the form of particles (i.e., photons) and, thus, we deal with photonic devices. Whether with waves or particles, we are still talking about electromagnetic energy but, unfortunately, people tend to not connect the two (this is the result of traditionally teaching electromagnetics focusing on wave theory only, and then teaching optics in totally separate courses, without linking the two approaches). The THz band lies in between the two realms and, as such, is outside the comfort zone of microwave engineers and optical engineers.

With this background information in mind, let me answer your question. The THz band offers unique compromises between micro/millimeter-waves and optical wireless communications. At THz frequencies, we already have tens to hundreds of consecutive GHz of bandwidth, much more than at millimeter-wave frequencies (only a few GHz) and comparable to that of optical wireless systems. Such bandwidth comes indeed with a more challenging wireless propagation channel, but this is still much better than the optical wireless channel (after all, at THz frequencies the wavelength is much larger than that of optical frequencies). So, are there opportunities for THz communications? Plenty. Is there hype on THz communications? That is probably also true. I have been working on this field for over ten years, first as a PhD student and then as an independent faculty. Ten years ago, people were still discussing whether millimeter-wave communications made sense. Only a few visionaries, such as Professor Akyildiz, decidedly invested in THz communications. Now, everyone seems to jump on this. THz communications will happen and are here to stay, I have no doubt.

Q3: Following recent advancements in electronic, photonic, and plasmonic technologies for THz transceiver design, the so-called THz-gap is closing. Is there a race/competition between these three technologies (Please comment on the strong and weak aspects of each)? Which technology supports the best range of reconfigurability for adaptive cognitive applications (the interest of our readers)? Is there a clear winner or will we settle for hybrid solutions?

A3: As a communications engineer, I can only get excited when I see the progress in all the possible device technologies. In a summarized way, currently, the highest power THz transceivers have been developed in the electronics approach and, more specifically, through Schottky-diode-based frequency multiplying chains. For example, the NASA Jet Propulsion Laboratory (JPL) has demonstrated THz up and down converters from 100 GHz all the way up to 4 THz with transmission powers ranging from 200 mW to a few mWs, respectively – this is at least two orders of magnitude higher than any other technology. Other electronic approaches, like silicon CMOS and silicon-germanium BiCMOS, have other advantages, such as compatibility with existing fabrication processes and compactness, but do not have the power. If instead of power, the main design driver is data-rate, photonic-based systems are leading. The photonics approaches mainly consist in down-converting modulated optical signals to the THz band utilizing photomixing and photoconverting processes. Their power is much lower, but high-speed optical modulation works and converting that at THz frequencies is a good strategy.

However, whether electronic or photonic, there is a fundamental challenge: in both cases, we are trying to generate THz signals starting either from microwave/millimeter wave signals to be up-converted or from optical signals to be down-converted. Every time there is a conversion, we lose energy at the very least in the generation of harmonics, which further affect the overall efficiency of the system.

Instead, by leveraging new plasmonic physics in new nanomaterials and nanostructures, we can create new transceivers and antennas that intrinsically operate at THz frequencies. Among others, graphene, a two-dimensional nanomaterial with unique electrical, optical and mechanical properties, can be utilized to develop direct THz signal sources (from DC to THz), direct THz signal modulators (able to manipulate amplitude, frequency and phase of THz signals), and on-chip THz antennas and antenna arrays. Of course, compared to the electronic and photonic approaches, which have been refined over decades, graphene-based plasmonic technology is much less mature (among others, this material was first obtained experimentally in 2004). Some people will see this as a challenge, but I personally see this as an opportunity: instead of adapting the communication system to already designed devices (constrained optimization) we can jointly design the devices and the communication solutions for the greater goal: data-rates, latency, connectivity.

Q4: What are, in your opinion, the most disruptive THz breakthroughs that have emerged in the past few years? What do you think are the most important remaining technical challenges or open problems in the field?

A4: Many things have happened in the last ten years. From the device perspective, NASA JPL has demonstrated THz transmitters with more than 100 mW at 300 GHz or few mWs at 1 THz, while Northrop Grumman has demonstrated the first electronic power amplifier operating at 1 THz. Moreover, new materials and structures have entered the game and demonstrating transforming approaches to the generation, modulation and radiation of THz signals. From the communication perspective, in 2008, the first channel measurements and data transmissions at 300 GHz were reported. This year, we have experimentally demonstrated error-free multi-Gigabit-per-second links at 1 THz, all while studying the true THz wireless channel. Moreover, it has been experimentally demonstrated that THz links can be established in non-line-of-sight conditions through first-order reflections. This has further motivated the development of mechanisms to create and leverage spatial diversity, including ultra-massive MIMO schemes in transmission, reception and, more recently, reflection.

Moving forward, it is time to step up the game and go beyond channel modeling (there are many works since 2010 which “newcomers” to the field should not miss) and start addressing theoretically and experimentally real problems, including synchronization of ultra-broadband THz signals at the physical and logical levels, real-time channel estimation and equalization of ultra-broadband channels, spectrum access and sharing policies for ultra-fast networks, neighbor discovery with ultra-directional systems at the transmitter and the receiver, or connectivity in mobile THz networks, to name a few.

In parallel to all the technical work, spectrum policies need to accompany the development of THz communications. It was not until earlier this year that the US Federal Communications Commission made the first attempt at regulating the spectrum above 95 GHz. While this has been very exciting, currently, only “a few GHz here and there” have been allocated for communications. This is far from the tens of GHz of consecutive bandwidth that motivate the use of the THz band. Therefore, there is still plenty of work, but this will happen.

Q5: It is argued that the breakthrough that this field will introduce is not solely driven by the high achievable data rates, but more profoundly by the combination of THz communications, THz sensing and imaging (traditional THz applications), and high-accuracy localization applications. Do you see real potential in such application merges? Can you envision a role for machine learning and artificial intelligence in this regard?

A5: There is indeed potential for meaningful joint communications and sensing at THz frequencies. The reason is the following. As you might recall from your physics or electromagnetics class, the energy of a photon is related to the Planck constant and the frequency of the signal. The higher the frequency, the higher the energy. At low frequencies, we cannot distinguish individual photons because they have low energies and, therefore, we talk about electromagnetic waves. At optical frequencies, the very high frequencies lead to very high energies and, thus, we generally talk about photons. The higher the energy of a photon, the more likely it interacts with other particles and matter. This is why there is an entire field called "light-matter interactions", which studies how light interacts with particles, materials and objects both "geometrically" (e.g., reflections, diffraction) as well as "physically" (e.g., absorption). At THz frequencies, photons start having meaningful energies and, thus, these can be used to extract the materials properties in the form of unique electromagnetic signatures (e.g., through THz spectroscopy).

Of course, in parallel to all these, THz waves can be used in radar-type applications for localization. In radar, the resolution is determined by the wavelength of your signal, among many others. At THz frequencies, the wavelength is under one millimeter, which leads to very precise localization. If now you combine communications, sensing/imaging and localization, you get a very complete and complex system, only reproducible at optical frequencies (not possible in the microwave/millimeter-wave realm). So, yes, I see potential in such, as these applications come from the physics, not from the hype.

When it comes to whether machine learning can help here, let me just say the following. A machine, as of today, cannot learn if there is no teacher. In this case, the teacher is a well-defined labelled dataset. The good news is that THz sensing and imaging is the oldest application of THz technologies (decades old) and there are extensive datasets available describing frequency-dependent absorption of a myriad of materials. Similarly, while slowly, THz channel measurements are coming up or, at least, the platforms to collect such datasets are becoming more available and more affordable. Models to jointly describe communications and sensing might look too complex or, if simplified, might not be accurate. Only in that case, it is reasonable to adopt data-based approaches including machine learning.

Q6: How do you describe the interest/involvement of industry in THz communications? When do you think we will start to see commercially available solutions?

A6: The industry involvement in THz communications has been rather anecdotic. As of today, there are very few options if you want to acquire a THz communication system. One of the main companies (behind many of the THz communications testbeds and many times in partnership with major equipment vendors) is Virginia Diodes, Inc. They commercialize Schottky-diode-based frequency up & down converters at frequencies of up to 1 THz. Besides them, expect to see many key wireless industry players showing 120-140 GHz systems in the very near future. The D-band (from 110 to 170 GHz) is for many considered "the next 60 GHz band" and, while it can be discussed whether this is sub-THz or high millimeter-wave, it is a step in the right direction.

For the low THz frequencies, I believe the fundamental research is done. As of today, we know very well how to create a point-to-point multi-Gigabit-per-second link above 100 GHz (there is even a standard for that!) and, therefore, it is just logical that industry enters the game. If you think how millimeter-waves started, after the first WiGig standard ten years ago, the interest in 60 GHz raised quickly. While much work needs to be done to move from point-to-point links to actual mobile networks (that's still happening at millimeter waves frequencies!), I expect more key wireless industry players entering the game.

Q7: Could you please briefly introduce the most recent research project(s) that you have done in this area (Please explain the key idea(s) and interesting findings)? What are your own short-term and long-term plans?

A7: We have learned many things in THz communications since our first papers in 2009. After many years of channel modeling and physical layer design, our latest projects are focused on two things. On the one hand, we are working towards developing experimental testbeds to validate our analytical models. In this direction, thanks to the US National Science Foundation, we currently have the only testbed in the world able to communicate at true THz frequencies, i.e., in the first absorption-defined window above 1 THz. More specifically, we are able to transmit and receive any user-defined data frame structure (0s and 1s), with single or multi-carrier amplitude, frequency and phase modulations with over 30 GHz of modulation bandwidth. We are using this platform for many things, including channel modeling (both in time and frequency domains), ultra-broadband channel estimation and equalization, waveform design and modulation, and testing of time, frequency and phase synchronization algorithms. Our next goals involve expand the platform to support and study multi-band simultaneous transmissions across 120 GHz, 240 GHz and 1 THz. Of course, part of our work involves sharing all the collected experimental data with the wireless research community. On the other hand, we are working on the development of networking protocols (link and network layers) tailored to the peculiarities of the THz-band channel and the capabilities of THz devices. Such protocols need to support mobile indoor and outdoor THz networks in different scenarios from inside an office to outdoor in a city, between planes at 30,000 feet and across satellites above the atmosphere. For these, obviously, we will have to enhance our testbed to support real-time protocol testing. In parallel to all these, our work on fundamentally new types of devices for THz communications, that build upon our knowledge of graphene-based plasmonic THz devices, keeps evolving and, as the technology matures, we hope to have working prototypes within the next 2-3 years.

Q8: As major contributions to THz technology are still at the level of transceiver design, what advice do you give to researchers who are approaching this field from a communication system, signal processing, and networking perspective? What is the take-away message that you prefer to leave our readers with?

A8: Terahertz communications are going to happen or, in fact, are already happening. Be ready to face many "naysayers", who have helped to propagate (no pun intended) some myths about THz communications. The only rules that cannot be changed are the rules of physics, and for many of the applications that we have mentioned here, physics are on our side (of course, not for everything). It might take some time before we have a good technology on the table, but it will happen.

Having said this, even when THz devices become more available, their cost might be prohibitive for many. This should not stop researchers to enter the field. The way in which research is evolving is that not every institution needs to have a testbed for every possible technology. Collaboration across research labs makes more sense than ever and, despite sometimes there might be political interference, the beauty of academia is that we can all collaborate, exchange ideas and work together towards the bigger goal. Take this as an open invitation to use our testbed.

Josep M. Jornet (M'13) is an Associate Professor in the Department of Electrical and Computer Engineering at Northeastern University, in Boston, MA. He received the B.S. in Telecommunication Engineering and the M.Sc. in Information and Communication Technologies from the Universitat Politecnica de Catalunya, Barcelona, Spain, in 2008. He received the Ph.D. degree in Electrical and Computer Engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, GA, in 2013. From August 2013 and August 2019, he was an Assistant Professor with the Department of Electrical Engineering at the University at Buffalo, The State University of New York. He was the recipient of the Oscar P. Cleaver Award for outstanding graduate students in the School of Electrical and Computer Engineering, at Georgia Tech in 2009. He also received the Broadband Wireless Networking Lab Researcher of the Year Award in 2010. In 2016, 2017 and 2018, he received the Distinguished TPC Member Award at the IEEE International Conference on Computer Communications (INFOCOM). In 2017, he received the IEEE Communications Society Young Professional Best Innovation Award, the ACM NanoCom Outstanding Milestone Award and the UB SEAS Early Career Researcher of the Year Award. In 2018, he received the UB Exceptional Scholar Award, Young Investigator Award, and the UB SEAS Early Career Teacher Award. In 2019, he received the NSF CAREER Award. His current research interests are in Terahertz-band communication networks, Wireless Nano-bio-sensing Networks, and the Internet of Nano-Things. In these areas, he has co-authored more than 140 peer-reviewed scientific publications, 1 book, and has also been granted 3 US patents. These works have been cited over 6,800 times (h-index of 37). He is the Editor-in-Chief of Elsevier's Nano Communication Networks Journal, and has organized multiple special issues on THz communications in several IEEE magazines and journals, in addition to serving on the TPC of the main conferences in the field.

Interview with Dr. Onur Sahin

InterDigital Inc., London
Email: onur.sahin@interdigital.com

Q1: What is, in your opinion, the most appropriate way to define THz communications? What frequency ranges constitute the THz band and what are the key system performance requirements?

A1: The definition of THz communications has been historically a factor of the frequency ranges the THz signals span, albeit without a common definition yet. Academic publications, particularly in applied physics and devices domain, have broadly considered 100GHz-10THz bands. However, over the last decade, the communications society seems to have narrowed it down to the bands between 100GHz through 3THz, supported with some encouraging implementation results in the RF and devices up to 1THz. From regulatory bodies' perspective, ETSI and ITU-R consider bands between 300GHz-3THz (corresponding to signal wavelengths between 1mm-0.1mm) as THz spectrum, while other definitions include lower end of 100GHz and higher end of 10THz.

In my opinion, at least in the upcoming decade or so, the commonly accepted definition of "THz communications technology" and underlying frequency bands will be 100GHz-1THz which might allow commercial grade demonstrations and implementations within this time frame.

For the key system performance requirements, it is essential to realize that "THz technology" corresponds to an umbrella term that contains multiple sub-systems under its definition. These range from nano-networks, e.g. internet of nano-things including sensors with nm form factors, to micro and macro scale deployments for ultra-high throughput (>1Tbps) cellular and mobile use-cases, as well as highly precise sensing and positioning solutions offered by THz bands among many others. Naturally, each sub-system will have its own detailed requirements. However, as in all commercially successful technology solutions, each of these sub-systems will need to deliver feasible size, weight, power, and cost (SWaP-C) KPIs, which ideally satisfy the demand of the corresponding use-cases. [In fact the lack of feasible SWaP-C KPIs that could enable most of the promising THz use-cases is one of the primary reasons that undermine commercial success and proliferation of the technology so far.]

Q2: THz communications have been a subject of both praise and criticism. Since communication technologies are already mature at the neighboring bands, the mere necessity for exploiting the THz-band is questionable. Is pushing microwave communications beyond the well-established millimeter-wave band worth the effort? And why should we settle for THz communications when cheap off-the-shelf light-emitting diodes can support much higher data rates in visible light communications? Is this just a THz-hype or will THz communications sustain their seemingly revolutionary role in future communication systems?

A2: I think all these three technologies have distinctive capabilities that will allow them to be the de-facto solutions for different use-cases. More particularly, the THz technology has a unique potential, since it amalgamates the two most important features of mmW and VLC technologies. These are, very large bandwidths enabling ultra-high throughput data modulation above 100Gbps (compared with mmW) and favorable NLOS/scattering, relatively good obstacle penetration/loss characteristics, and immunity to ambient light based interference (compared with VLC). For these reasons, THz can be seen as an attractive solution for many of the critical and widely deployed use-cases in beyond-5G. Consider, for instance, next generation mobile broadband communications, supporting 100Gbps and above wireless links in relatively high mobility scenarios and typical environments with obstacles. This will be one of the most important use-cases in beyond 5G systems as has been in all generations so far. The THz technology is possibly the only option to offer technically and commercially feasible solution for this, along with many other mobile broadband use-cases. Additionally, high-resolution sensing in foggy or rainy weather conditions, and nano-device networks will clearly rely on THz technology-based solutions instead of mmW or VLC.

Q3: Following recent advancements in electronic, photonic, and plasmonic technologies for THz transceiver design, the so-called THz-gap is closing. Is there a race/competition between these three technologies (Please comment on the strong and weak aspects of each)? Which technology supports the best range of reconfigurability for adaptive cognitive applications (the interest of our readers)?

Is there a clear winner or will we settle for hybrid solutions?

A3: Each of these approaches has its own merits and can better adapt to different requirements of the THz technology, such as high frequency capability of photonics, and higher power output of CMOS based solutions. From a practical implementation and commercialization perspective, we see clear advantage of electronics-based transceiver design, particularly leveraging CMOS technology in the transceiver baseband unit. In the RF front-end component of the transceiver though, CMOS is known to have bottlenecks for the bands above 300GHz-400GHz and scaling unfortunately does not seem to improve performance as in baseband. For these bands, monolithic microwave integrated circuits (MMIC) based devices using HBT and HEMT processes appear as very capable options.

Regarding photonics-based RF solutions, there has been substantial progress over the last decade using uni-traveling carrier photodiode (UTC-PD) and quantum cascade laser (QCL) components. Despite the potential of the photonics-only based solutions however, a practical transceiver based on photonics devices still seems to be challenging. This approach currently lacks practical implementation either resulting in very large form factors, cryogenic cooling requirements, or very low output power.

For higher bandwidth and highly tunable operations, hybrid photonic-electronic are currently a focus of interest, and demonstrate admirable output powers along with high flexibility and fast data modulation, which are critical in any wireless technology and beyond-5G systems. On the flip side, the hybrid solutions seem to provide lesser performance benefits at the receivers and require highly precise synchronization between the transmitter and receiver. Going forward, I expect all three options to continue their progress towards mature and desirable solutions for specific capabilities to be offered in THz technology. However, CMOS based solutions operating at the 400GHz and lower bands will possibly the first widely deployed products among others. The adaptability and low-cost advantages of electronic, CMOS based technologies similarly make them the most suitable option for the adaptive cognitive applications in my opinion.

Q4: What are, in your opinion, the most disruptive THz breakthroughs that have emerged in the past few years? What do you think are the most important remaining technical challenges or open problems in the field?

A4: Over the last 5 years, we have seen substantial progress in the THz transceiver technology focusing on the design of very challenging THz signal generating and detection modules. In the CMOS camp, single-chip transceivers operating in the 250GHz bands, achieving up-to 80Gbps with practical form factor and output power is demonstrated. A hybrid electronic-photonic solution, using UTC-PD at the transmitter with photomixer and MMIC based receiver is shown to demonstrate 100Gbps data rates in 273.5 GHz band up-to 40m distances. Furthermore, photonics based solutions are also capable of operating at 1 THz band. All of these are some examples in the THz device technology state-of-the-art, demonstrating the scale of development that has been made fairly recently.

Additionally, the THz technology also requires very directional antennas with ideally steerable features with beamforming capabilities. Recently, we have seen advances in graphene based plasmonic antennas compatible in nano scale along with plasmonic patch antennas operating around 700GHz bands. It is instructive to note that the Graphene based patch antenna array in Yagi-Uda MIMO configuration with beamsteering capabilities is also reported.

For the baseband design, we see the first examples of practical Tb/s receivers operating in the mobile terminal power budgets and size constraints (e.g. around 1pj/bit and 10mm2 chip area). The system-level discussions to enable multi-user THz network have already been initiated and some compelling initial solutions in terms of neighbor discovery, synchronization, directional channel access mechanisms are provided under the IEEE 802.15.3d standardization group.

In short, we have been observing and will continue to see advancements in all building blocks of THz technology. The key challenge, which is possibly the most important and challenging obstacle remaining in this domain, is to design and develop an integrated THz communication unit that is composed of transceiver, antennas, and adaptable to higher layer, e.g. medium access control procedures, all within practical and commercially feasible SWaP-C constraints.

Q5: It is argued that the breakthrough that this field will introduce is not solely driven by the high achievable data rates, but more profoundly by the combination of THz communications, THz sensing and imaging (traditional THz applications), and high-accuracy localization applications. Do you see

real potential in such application merges? Can you envision a role for machine learning and artificial intelligence in this regard?

A5: All of these functionalities, e.g. THz sensing and imaging, and high accuracy localization applications are some of the most compelling technologies that THz is able to offer. I certainly believe that these applications have substantial potential to be commercially successful. In fact, we see a similar trend in 60 GHz technology and WiGIG chipset solutions. The WiGIG chipset market, offering multi-Gbps mainly for personal-area networks, has experienced a relatively slow uptake so far. However, we have already started seeing many interesting and complementary applications of mmW in sensing and imaging, by providing solutions for diverse use cases such as object recognition, elderly care, health-care, new user interfaces, etc, with products by Google, Qualcomm, and others. Considering substantially higher resolution sensing, and better localization capabilities, I believe that these features will be in-built features of the THz communications links, which will also be used to optimize the underlying communication networks.

For these types of applications, e.g. object recognition, tracking, localization, etc., machine-learning provides sophisticated set of tools and solution frameworks. This is due to the fact that these problems lack a well-defined underlying mathematical abstraction or model of the problem, or need highly complex techniques in feature engineering with limited number of well-known algorithms. Moreover, both via simulation and emulation platforms, as well as real-time demos, the data available in these experiments will be abundant thanks to very low-cost (almost free) nature of wireless data collection. Therefore, machine learning and in particular deep learning will offer very useful framework in developing very interesting radio based use-cases which we currently lack.

Q6: How do you describe the interest/involvement of industry in THz communications? When do you think we will start to see commercially available solutions?

A6: The wireless industry has so far shown limited interest and involvement in the design and development of THz communications as the market demand and commercial potential of the technology have not been proven to be significant yet. This has surely to do with the technical challenges observed in the transceiver component, which further requires increased investment. Therefore, the field has been driven mostly by academic contributions. However, the development of first wireless technology standard operating in the THz spectrum, IEEE 802.15.3d in 2017, and new spectrum allocations by regulatory bodies are clearly solid indicators of the industries mid-to-long terms visions and interests in these bands. Also, FCC’s recent allocation of a total of 21.2GHz of spectrum between 116GHz and 246GHz bands for unlicensed usage is a very important step forward that will surely attract a more dedicated focus from industry in the near future.

In my opinion, we will start seeing integrated and commercially viable CMOS or MMIC based solutions for the 100GHz-400GHz bands at least, around or just after 2025 timeframe. Initial products and solutions will highly likely be in the infrastructure deployments, e.g. backhaul/fronthaul and data-centers connectivity where form factors and power budgets are not as stringent as mobile terminal use-cases. The proliferation of the technology and related products in the mobile broadband and/or THz-enabled sensing and monitoring applications will be contingent on the mobile data-rate demands as well as commercial opportunity these applications might bring.

Q7: Could you please briefly introduce the most recent research project(s) that you have done in this area (Please explain the key idea(s) and interesting findings)? What are your own short-term and long-term plans?

A7: Over the last years, I have been involved with the development of Tb/s baseband solutions for THz systems targeting practical power budget and form-factor constraints for mobile terminals. This research is carried out under Pan-European collaborative project named EPIC, which is funded under EC H2020 Beyond-5G program. The project takes a bottom-up approach by targeting the design of a major building block in ultra-high-throughput wireless transceivers and focuses on the forward-error-correction (FEC) module which is computationally the most complex unit in baseband chain. Our analysis shows that the silicon node scaling will provide limited improvements in terms of the baseband computations and the power density on the silicon chip. This is because of the diminishing effect observed in Moore’s Law in the future silicon generations, which will be a major bottleneck. Therefore a holistic approach that incorporates ASIC architectures with baseband algorithms in a unified design framework is the only viable option in achieving the Tb/s bottleneck in mobile terminal constraints.
We have made substantial progress and have designed Polar and LDPC based ASIC decoders achieving Tb/s

data rates within the practical energy efficiency (~1pj/bit) and power density budgets (~0.1W/mm2).

Q8: As major contributions to THz technology are still at the level of transceiver design, what advice do you give to researchers who are approaching this field from a communication system, signal processing, and networking perspective? What is the take-away message that you prefer to leave our readers with?

A8: In my opinion, the success of a commercially viable THz technology is highly aligned with the feasibility of holistic and integrated end-to-end IP and chipset implementation solutions. This brings very interesting innovation opportunities at the boundaries and interfaces of key building block of the overall solution, as well as at the intersection of device technology and signal processing algorithms. For instance, novel approaches at the functional separation of RF, antenna, and baseband elements, leveraged hybrid analog and digital architectures and algorithms for a feasible THz technology solution will be critical. Furthermore, since the computational requirements of THz systems are already pushing the boundaries of the state-of-the-art in the device technologies, particularly in silicon node generations, the corresponding technical limitations have to be factored in the overall design framework of the technology. I believe most of the building blocks in THz technology will need a holistic design of the ASIC architecture and underlying signal processing algorithms.

Hence, researchers in the THz field will greatly benefit from having a system's view in approaching the design challenges in the THz technology. Incorporating a design space exploration in the target design of the constraints of the major building blocks, e.g. baseband algorithms, RF design, device architectures, networking, etc,. This will surely require a level of understanding in each of the fields complementing their specific focus areas. Therefore, a true multi-disciplinary approach is indispensable to achieve a widely deployed and commercially successful THz technology.

Dr. Onur Sahin received his B.S. degree in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, in 2003 and Ph.D. degree in electrical engineering from the Polytechnic Institute of New York University, USA in 2009. He is currently a Senior Staff Engineer at Innovation Labs, InterDigital Europe. His primary research and development interests are on the next generation telecommunication and wireless systems (including 5G and beyond). Dr. Sahin has held technical lead positions at multiple projects on next generation cellular and Wi-Fi systems including 5G NR, LTE-A and IEEE 802.11 standards. He currently leads Beyond-5G ultra-high throughput (Tbps/THz) wireless technology design and development at InterDigital. Dr. Sahin is the co-author of over 50 peer-reviewed scientific articles and co-inventor of 25 patents and patent applications. He is co-recipient of the 2018 IEEE Signal Processing Society Best Paper Award, 2016 Journal of Communication Networks Best Paper Award, and InterDigital Innovation Awards in 2012 and 2015.