Module Overview

EN5730 Machine Learning for Communications

Samiru Gayan

Department of Electronic and Telecommunication Engineering
University of Moratuwa

2024

Developments of Wireless Communications

Wireless Communications

- The success of wireless communications stands as one of the most remarkable achievements in the history of science and technology, revolutionizing interactions between people and machines.
- In 1948, Claude Shannon introduced a groundbreaking mathematical framework that defined how information is transmitted through a communication channel and established the fundamental limits of such systems.
- Shannon's information theory has been regarded as the foundation of modern communication technology.
- Since then, wireless communications have experienced significant technical revolutions and have evolved through five generations.

Wireless Communications

Generation	Remarks
1G	• Introduced in the 1980s • Analog cellular networks with basic voice communication • Frequency-division multiple access (FDMA)
2G	• Introduced in the 1990s • Digital cellular networks with both voice and text communications • Time-division multiple access (TDMA)
3G	• Introduced in the 2000s • Higher data rates to support various services, including Internet access and global positioning system (GPS) • Code division multiple access (CDMA) and wideband CDMA (WCDMA)

Wireless Communications

Generation	Remarks
4G	• Introduced in the 2010s • multiple-input multiple-output (MIMO) • Orthogonal frequency-division multiplexing (OFDM) • Much faster data speeds and low transmission delay • High-quality video transmission and mobile Internet access
5G	• Introduced in the 2020s • Massive MIMO, millimeter-wave (mmWave) • Supports ultrareliable and low-latency communications (URLLC) • Supports massive machine-type communications (mMTC) • Supports enhanced mobile broadband (eMBB) communications

5G beyond and 6G

Key Performance Indicator (KPIs)	4G-LTE	5G	6G
Deployment year	2000	2020	Yet to be implemented
Core architecture	Internet	Internet	Internet
Core networking	Internet	Internet	Internet
Multiplexing bandwidth	OFDMA/SC-FDMA (1.4 Mhz–20 Mhz)	BDMA/FBMC (60 GHz)	OMA/NOMA (up to 3 THz)
Per device peak data rate	1 Gbps	10 Gbps	1 Tbps
Switching	Packet switching	Packet switching	Packet switching
Forward error correction	Turbo codes	LDPC codes	LDPC codes
E2E latency	100 ms	10 ms	1 ms
Maximum spectral efficiency	15 bps/Hz	30 bps/Hz	100 bps/Hz
Mobility support	Up to 350 km/h	Up to 500 km/h	Up to 1000 km/h
Satellite integration	No	No	Fully supported
AI supported	No	Partially supported	Fully supported
Autonomous vehicle supported	No	Partially supported	Fully supported
XR supported	No	Partially supported	Fully supported
Haptic communication	No	Partially supported	Fully supported
Visible light communication (VLC)	No	No	Yes
Maximum frequency	6 GHz	90 GHz	10 THz
Architecture	MIMO	Massive MIMO	Intelligent surface
Service Level	Video	AR,VR	Tactile
Connectivity density	10^5 Devices/km^2	10^6 Devices/km^2	10^7 Devices/km^2
Area traffic capacity	0.1 Mb/s/m^2	10 Mb/s/m^2	1 Gb/s/m^2
Network energy efficiency	1$\times$	10–100$\times$ of 4G	10–100$\times$ of 5G
Spectrum efficiency	1$\times$	3$\times$ of 4G	5–10$\times$ of 5G
Reliability	99.99	99.999	99.99999

5G beyond and 6G

- Many emerging techniques have been proposed and investigated.

- Reconfigurable intelligent surface (RIS)
- Integrated sensing and communications (ISAC)
- Semantic communication

AI in Communications

Machine Learning (ML)

- ML models can be generally divided into three types:
  - supervised learning
  - unsupervised learning
  - reinforcement learning
- The neural network is a novel structure that involves layers of interconnected neurons (nodes) to model complex relationships and learn from diverse types of data.
- The neural network with multiple layers is called the deep neural network (DNN) which is enabled by deep learning (DL).
- The DNN-based architecture includes a variety of well-known models, such as
  - recurrent neural network (RNN)
  - convolutional neural network (CNN)
  - graph neural network (GNN)
  - long-short-term memory (LSTM) network

Motivation

- Since the data traffic load for wireless communication will increase dramatically in the future, it is difficult for existing wireless communication systems to meet the ever-growing requirements of various intelligent wireless applications.
- AI/ML technology has been regarded as a powerful approach and brought many benefits in the design of the next-generation wireless communication systems.

Motivation: Signal Processing at the Physical Layer

- Traditional networks rely on accurate mathematical models for channel parameter estimation, which often fails to find an accurate model when the wireless propagation environment is time-varying and complex.
- ML is an alternative method to facilitate adaptive channel modeling and estimation by learning from the massive recorded data, relaxing the constraint for accurate mathematical models.

Motivation: Superior Data Processing

- The explosion in the number of wireless devices challenges the traditional ways of data storage and processing.
- ML algorithms can be used to identify patterns in raw data and remove redundant information to optimize the storage and processing spaces of data.
- With the help of NLP and DL, AI technology also allows us to rethink the traditional information theory and investigate post-Shannon wireless communications at the semantic level.

Motivation: Resource Allocation and Network Optimization

- Integrating AI/ML into wireless communications also benefits resource allocation and network optimization, ensuring the efficiency and scalability of wireless networks.
- Resource allocation problems can be generalized as optimization problems under some constraints (e.g., energy efficiency, transmission delay, and throughput).
- Traditional optimization tools can solve the problem efficiently when the objective function satisfies some assumptions, such as convex property, continuous property, and differential property.
- However, the complexity of the traditional optimization algorithms grows exponentially with the network scale, which is not acceptable for large-scale real-time applications.
- In this regard, AI/ML is seen as an effective tool for solving the challenging optimization and resource management problems in wireless communication systems.

Learning based Physical Layer Processing

Channel Modeling and Estimation

- Conventional methods for channel modeling can be categorized into deterministic approaches, e.g., ray tracing, and stochastic approaches, which usually require extensive calculations and exact descriptions of the environment.
- In contrast to these conventional approaches, generative models can adopted to capture the wireless channel effects and produce channel parameters due to their capacity to extract underlying properties from observed data.
- Furthermore, generative model-based channel modeling methods can deal with more complex communication environments and offer higher modeling accuracy.

Channel Modeling and Estimation

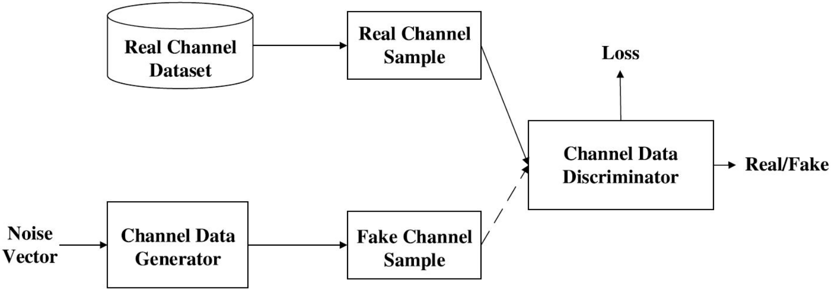

Figure: Framework of GAN-based wireless channel modeling.

- The training stops when the discriminator cannot distinguish the real and fake samples.
- At this point, the generator has learned the distribution of the real channel samples and it can be directly extracted as the target channel model.

Signal Detection and Decoding

- Consider a MIMO system:
$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n},$$
where $\mathbf{y} \in \mathbb{C}^{M}$ denotes the received signal vector, $\mathbf{H} \in \mathbb{C}^{M \times K}$ denotes the channel matrix and $\mathbf{n} \in \mathbb{C}^{M}$ denotes the Gaussian noise vector.
- The goal of MIMO detection is to estimate the symbol vector $\mathbf{x}$ based on the observation $\mathbf{y}$ and the knowledge of $\mathbf{H}$.
- Maximum likelihood detection:
$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathcal{C}} ||\mathbf{y} - \mathbf{H}\mathbf{x}||^2 \tag{1}$$
- Due to the discrete nature of $\mathbf{x}$, optimal MIMO detection belongs to integer LS problems, known to be NP-hard.

Signal Detection and Decoding

- To address the challenge, DL-based MIMO detectors have been developed by unfolding the projected gradient descent method and incorporating learnable weights.

  Eg. DetNet, OAMPNet
- GNN-aided expectation propagation (EP) algorithms - GEPNet
- Online learning based methods - MMNet
- Stochastic sampling-based methods - Markov Chain Monte Carlo (MCMC)

End-to-End Communications

- End-to-end communication transforms the traditional communication system into a data-driven framework.
- In this novel end-to-end paradigm, transmitters and receivers are jointly trained based on an end-to-end loss function rather than separate channel coding/decoding and modulation/demodulation.
- The end-to-end communication architecture can be viewed as an autoencoder, where the transmitter functions as the encoder network, and the receiver serves as the decoder network.

End-to-End Communications

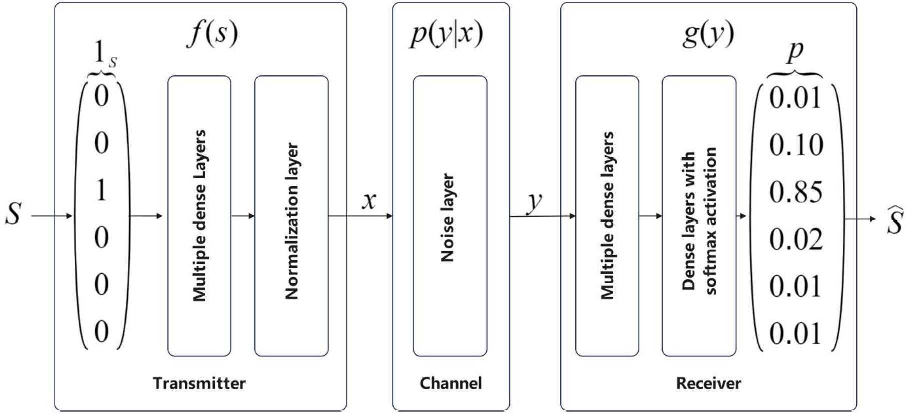

Figure: Communications system over an AWGN channel represented as an autoencoder.

In this Course

Intended Learning Outcomes

- Understand the fundamentals of machine learning (ML) and its applications in communication systems.
- Identify different machine learning approaches relevant to the physical layer.
- Design and implement machine learning models to solve physical layer problems.
- Discuss 3GPP standardization on machine learning in communication technologies.

Course Structure

- A three hour lecture every week - Sunday 1 pm to 4 pm
- Project Work

The Road Ahead

- Topic 0 - Overview of the Course
- Topic 1 - Introduction and ML Basics
- Topic 2 - Review Fundamentals of Wireless Communications
- Topic 3 - Applications of ML for Communications
- Topic 4 - 3GPP Standardization on Machine Learning Activities, Open Problems and Challenges

Useful Text Books

- Yonina C. Eldar, Andrea Goldsmith, Deniz Gündüz, and H. Vincent Poor, "Machine Learning and Wireless Communications", Cambridge University Press, 2022.
- D. Tse and P. Viswanath, "Fundamentals of Wireless Communication" Cambridge University Press, 2012
- Andrea Goldsmith, "Wireless Communications" Cambridge University Press, 2005

Assumed Knowledge

- Basic Python programming
- Basics of linear algebra and calculus
- Fundamentals of telecommunications

Getting Help

- Outside the lectures
  - Email: kasunh@uom.lk or samirug@uom.lk

References

1. Z. Qin et al., "AI Empowered Wireless Communications: From Bits to Semantics," in Proceedings of the IEEE, vol. 112, no. 7, pp. 621-652, July 2024.

- All the images have been taken from [1].

The End