From Conventional to Semantic Communications based on Deep Learning

Professor Geoffrey Ye Li
ITP Lab, Department of Electrical and Electronics Engineering
Imperial College London
London, UK
Outline

- Overview

- DL-based Conventional Communications
  - Block-Wise
  - End-to-End

- DL-based Semantic Communications

- Conclusions
Motivation

- Challenges in current/conventional communication systems
  - Mathematical models versus practical imperfection
  - Block structures versus global optimality
  - Complexity and performance of optimization
  - Spectrum efficiency limited by Shannon capacity

- Why deep learning?
  - No need for models for data-driven method
  - End-to-end loss optimization for global optimality
  - Deep learning enabled end-to-end and semantic communications
Block Structure or End-to-End for Conventional Communications

Conventional Communications: Transmit symbols or bits, following Shannon Limit
DL in Physical Layer Conventional Communications

From Symbol to Semantic Transmission

- Three Levels of Communications: Shannon and Weaver
  - Transmission of symbols (Shannon Paradigm)
    following Shannon limit & well-developed near limit
  - Semantic exchange of source information
    semantic communications (transmission of intelligence)
  - Effects of semantic information exchange

- Semantic Communications: Significantly improved efficiency!

Conventional Communications

- Only consider the data recovery accurately
- Information redundancy are removed in entropy-domain
- All information (including useless and irrelevant) is transmitted to the receiver, part is useless for the target network
Semantic Communications

- **Feature** networks and **action** networks considered
- Information redundancy removed in **semantic domain**
- Only **useful and relevant** information transmitted to the receiver
- The features can serve **different** action networks

![Image Diagram](Image Diagram Description)
Outline

● Overview

● DL-based Conventional Communications
  * Block-Wise
  * End-to-End

● DL-based Semantic Communications

● Conclusions
Channel Estimation (CE) and Signal Detection (SD)

- **Related works:**
  - MMSE for channel estimation
  - Neural networks and DL in equalization and decoding

- **Challenges:**
  - Nonlinear distortion and interference

- **Innovations:**
  - DL for joint channel estimation and symbol detection
  - DL-based method: robust and insensitive to nonlinear distortion and interference
Traditional CE and SD

Robust Channel Estimation for OFDM Systems with Rapid Dispersive Fading Channels

Ye (Geoffrey) Li, Senior Member, IEEE, Leonard J. Cimini, Jr., Senior Member, IEEE, and Philip Borjesson, Fellow, IEEE

Abstract—Channel estimation for OFDM (MMSE) channel estimation has been studied. The bit rates in OFDM systems by using column diagonalization are-error robust estimator, that is, an estimator that is not sensitive to the channel statistics. Computer simulation demonstrates that the performance of OFDM systems using coherent demodulation based on our channel estimator can be significantly improved.

Index Terms—channel estimation, channel estimation, frequency-selective fading channels, OFDM systems.
DL-based CE and SD

- **Input:** received pilot OFDM block + received data OFDM block
- **Output:** recovered data

Model-Driven DL

- Relying on relatively accurate model
- Exploiting rich domain/expert knowledge
- Easy to train with a small amount of data
- Explainable and predictable neural networks
- Deep unfolding: a popular model-driven approach

Example: MIMO Detection

- **MIMO System:**
  \[ y = Hx + n \]

- **Goal:** estimating \( x \) from received signal \( y \) and channel matrix \( H \)

- **Conventional Detectors:**
  - Optimal detector: **ML** detector, high complexity
  - Linear detectors: **ZF, LMMSE**, low complexity but poor performance
  - Iterative detectors: **AMP**-based detector, **EP**-based detector, excellent performance, moderate complexity, performance **degradation** with ill-conditioned channel matrix

- **Motivation:** **deep learning** to perform iterative detection

Outline

- Overview
- DL-based Conventional Communications
  - Block-Wise
  - End-to-End
- DL-based Semantic Communications
- Conclusions
Why End-to-End Learning?

- **Architecture:**
  - Representing both transmitter and receiver by DNNs
  - Leaning to encode transmit symbols at transmitter
  - Learning to recover transmit symbols at receiver

- **Merits:**
  - Achieving global optimum
  - Universal solution to different channels
  - Beating current state-of-arts
E2E based on Reinforcement Learning

- **Reinforcement Learning Formation:**
  - Agent: transmitter
  - Environment: channel + receiver
  - States: source data
  - Actions: transmit signals

- **Advantage and Disadvantage:**
  - Unnecessary for channel modeling
  - Hard for continuous action in reinforcement learning

E2E based on Conditional GAN

- Using CNN to address curse of dimentionality
- Conditional GAN: modelling the channel output distribution
- Surrogate of real channel when training the transmitter
- Received pilots as a part of conditioning for unknown channel

Performance for WINNER II Channels

- Similar BER at low SNR
- Better at high SNR
Outline

- Overview
- DL-based Conventional Communications
  * Block-Wise
  * End-to-End
- DL-based Semantic Communications
- Conclusions
Information Content of English & Semantic Encoding

• Encoding English Words Letter-by-Letter
  o In English, on average there are 4.5 letters per word
  o 5.5 characters per word if including space
  o 5 bits to encode each letter (26 letters)
  o 27.5 bits/word (5*5.5=27.5) Need a codebook of 26 letters

• Encoding English Words Word-by-Word
  o 171,476 English words (from Google)
  o 18 bits/word (2^{17} < 171,476 < 2^{18}) Need a codebook of 171,476 words

• Encoding English Semantically
  o i.e., only 1 bits if answering YES or NO a question
  o …. More Efficient! Need an extremely huge codebook
Semantic Transceiver

- **Transceiver**
  - **Transmitter**
    \[ X = C_\alpha (S_\beta (s)) , \]
  - **Receiver**
    \[ Y = HX + N, \]
    \[ \hat{s} = S_X^{-1} (C_\delta^{-1} (Y)) \]

- **Channels**
  - Physical channel noise is caused by the physical channel impairment
    - AWGN, fading channels...
  - Semantic channel noise refers to misunderstanding
    - Caused by interpretation error and disturbance in estimated information.

Transformer Structure

Transformer based semantic communication

- Merge the traditional communication and semantic into DNNs
- Transformer can learn the semantic in text
  - e.g., “it” completes pronoun reference “the animal”

Loss Function

- Loss function used to train the transceiver

\[ \mathcal{L}_{\text{total}} = \mathcal{L}_{\text{CE}}(s, \hat{s}; \alpha, \beta, \chi, \delta) - \lambda \mathcal{L}_{\text{MI}}(x, y; T, \alpha, \beta) \]

- **Cross-Entropy**: Through reducing the loss value of channel encoder, the network can learn the syntax, phrase, the meaning of words

\[ \mathcal{L}_{\text{CE}}(s, \hat{s}; \alpha, \beta, \chi, \delta) = -\sum_{i=1} q(w_i) \log(p(w_i)) + (1 - q(w_i)) \log(1 - p(w_i)) \]

- **Mutual Information**: maximizing achieved data rate

\[ \mathcal{L}_{\text{MI}}(X, Y; T) = \mathbb{E}_{p(x,y)}[f_T] - \log(\mathbb{E}_{p(x)p(y)}[e^{f_T}]) \]
Two-Step Training

- Maximizing mutual information
- Train the whole model
Performance Metrics

● BLEU score
  ➢ Compare the difference between words in two sentences

\[
\log \text{BLEU} = \min \left( 1 - \frac{l_\hat{s}}{l_s}, 0 \right) + \sum_{n=1}^{N} u_n \log p_n.
\]

  – \( l_s \) is the length of sentence \( s \), \( l_\hat{s} \) is the length of sentence \( \hat{s} \)
  – \( p_n \) is the n-grams score, \( u_n \) is the weights of n-grams

● Sentence Similarity
  ➢ Use siamese network to compute the semantic similarity

\[
\text{match} (\hat{s}, s) = \frac{B_\Phi(s) \cdot B_\Phi(\hat{s})^T}{\|B_\Phi(s)\| \|B_\Phi(\hat{s})\|}
\]

  – \( B_\Phi(\cdot) \) is the BERT model
  ➢ Sentence, \( s \), will be mapped into semantic vector space, \( B_\Phi(s) \), by BERT model
  ➢ Similarity is computed by measuring distance between \( B_\Phi(s) \) and \( B_\Phi(\hat{s}) \)
Simulation Setting

● **Dataset**
  ➢ The proceedings of the European Parliament

● **Proposed network architecture**
  ➢ Transmitter:
    – 3 layers of Transformer encoder and 2 dense layers
  ➢ Receiver:
    – 2 dense layers and 3 layers of Transformer decoder

● **Benchmark**
  ➢ Deep Learning based joint source-channel coding (DL based JSC coding)
  ➢ Traditional methods
    – Source coding: Huffman coding
    – Channel coding: Turbo code
    – Modulation: 64-QAM
All deep learning approaches are more competitive in the low SNR regime.

The tendency in sentence similarity is much closer to human judgment.

- In SNR = 12 dB, 20% BLEU score = approximate 0 sentence similarity
- People are usually unable to understand the meaning of texts full of errors
Outline

- Overview
- DL-based Conventional Communications
  * Block-Wise
  * End-to-End
- DL-based Semantic Communications
- Conclusions
Conclusions

- For Conventional Communications
  - Robust to nonlinear distortion, interference, & frequency selectivity
  - Improving performance of iterative detectors and adapt to complicated channels

- End-to-end Communication Architecture
  - Enabling global optimization of transceiver
  - Potentially reducing the complexity

- Semantic Communications
  - Significantly improving transmission efficiency
  - Future of wireless communications