



(Nie)bezpieczny trend stosowania AI do przetwarzania sygnałów radiowych

Krzysztof K. Cwalina

Politechnika Gdańska

Wydział Elektroniki, Telekomunikacji i Informatyki

Katedra Systemów i Sieci Radiokomunikacyjnych

Plan prezentacji

1. Motywacja i cel realizowanych prac
2. Charakterystyka neurotechnik
3. Zakres stosowalności neurotechnik
4. Kontradyktoryjność neurotechnik
5. Przykładowe badania efektywności neurotechnik
6. Podsumowanie, przyszłe prace i wyzwania

AI – buzzword dla komunikacji bezprzewodowej?





”

Uczenie maszynowe to rodzaj sztucznej inteligencji, która zapewnia komputerom możliwość uczenia się w warunkach ekspozycji na nowe dane, bez konieczności programowania ich wprost.

”

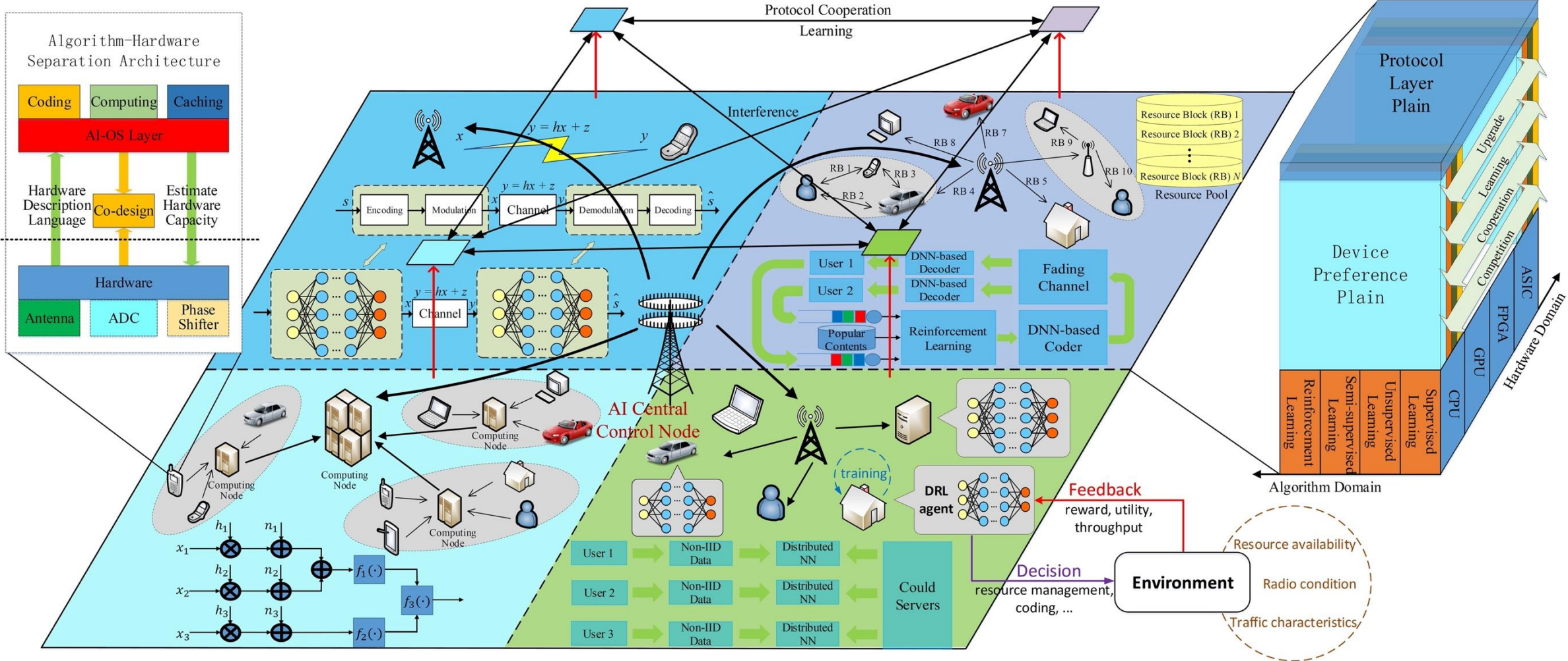
Jakub Piwowar, Google

”

Głębokie uczenie jest to sposób na ekstrakcję użytecznych cech ze zbioru danych w sposób automatyczny.

”

Lex Fridman, MIT



Z punktu widzenia rynku, dostępne są pierwsze moduły łączące AI z częścią interfejsu radiowego, więc ta technologia już jest implementowana i wprowadzana w krzemie...

W dobie wykorzystania systemów radiokomunikacyjnych bazujących na technice OFDM, niezwykle istotnym i cały czas aktualnym zagadnieniem jest rozwój warstwy fizycznej interfejsu radiowego.

Wpływ wielodrogowości, przesunięcia Dopplerowskiego, czasu trwania odpowiedzi impulsowej definiują zmienny w czasie, selektywny częstotliwościowo kanał radiowy, który musi być estymowany, a jego wpływ na jakość odbioru minimalizowany (korygowany).

Dodatkowo, występujące interferencje oraz celowe zakłócenia sygnałów radiowych sprawiają, iż detekcja, identyfikacja oraz minimalizacja ich wpływu jest cały czas tematem prowadzonych prac badawczych.

- głębokie sieci oparte na... sieciach do przetwarzania obrazu np. ResNet, CENet;
- jako wejście macierze (zespolone):
 - siatki zasobów referencyjnych, siatki pilotów i danych odebranych;
 - estymaty klasycznych metod (np. Least Square);
- realizują przepróbkowanie zespolonych sygnałów (siatek zasobów czasowo-częstotliwościowych);

- szereg badań wskazuje na przetwarzanie sygnałów radiowych jak obrazów, lub... języka;
- w tym celu stosowane są duże modele językowe by rozwiązywać problemy radiokomunikacyjne;
- niestety nawet wysoko cytowane artykuły zawierają elementy upraszczające kwestie komunikacji radiowej;
 - są raczej badaniami modeli, aniżeli komunikacji radiowej – szczególnie w formie „czarnej skrzynki”;

- głębokie sieci realizujące właściwie całą warstwę fizyczną, tzw. autoenkodery
 - tak, stronę nadawczą też...
 - zakłada się, że klasyczne opisy efektywnych i kontrolowanych metod przetwarzania danych w warstwie fizycznej jest mniej efektywne niż optymalizacja całych torów nadawczo-odbiorczych w relacji punkt-punkt;
 - w praktyce: dotychczasowa wiedza? Standaryzacja? Interpretowalność?

Zalety stosowania neurotechnik:

- dla wybranych profili kanału radiowego mogą uzyskiwać zbliżone, lub lepsze wyniki estymacji kanału (rozumianej jako minimalizację MSE) względem praktycznych implementacji estymatorów MMSE;
- pod względem złożoności obliczeniowej przewyższają estymatory MMSE;
- umożliwiają zrównoleglanie obliczeń i efektywną implementację sprzętową;
 - do tego implementacja sprzętowa aktualnie może być łączona z koprocesorami neuronowymi;
- umożliwiają optymalizację działania grup modułów w warstwie fizycznej (np. połączenie estymatora kanału i korekcji jego wpływu na siatkę zasobów);

Zalety stosowania neurotechnik:

- dla wybranych profili kanału radiowego mogą uzyskiwać zbliżone, lub lepsze wyniki estymacji kanału (rozumianej jako minimalizację MSE) względem praktycznych implementacji estymatorów MMSE;
- w połączeniu z dekodowaniem (albo zastępującym dekodowanie), mogą uzyskiwać zbliżone/lepsze o 1-2 dB wartości BER dla SNR w zakresie 5 - 15 dB;
- pod względem złożoności obliczeniowej przewyższają estymatory MMSE;
- umożliwiają zrównoleglanie obliczeń i efektywną implementację sprzętową;
 - do tego implementacja sprzętowa aktualnie może być łączona z koprocesorami neuronowymi;
- umożliwiają optymalizację działania grup modułów w warstwie fizycznej (np. połączenie estymatora kanału i korekcji jego wpływu na siatkę zasobów, czy synchronizację czasowo-częstotliwościową);

Oprócz niewątpliwych zalet stosowania neurotechnik mają one wiele ograniczeń, które wpływają na ich zakres stosowalności w praktycznych rozwiązaniach:

- bardzo ograniczone pomiarowe weryfikacje;
- złożoność (oraz duża liczność) modeli i brak zasad doboru ich architektury;
- bardzo, bardzo ograniczona adaptacyjność;
 - zmienne charakterystyki odpowiedzi impulsowej kanału;
 - SNR, o SINR nie wspominając...;
 - zmienne schematy modulacyjno-kodowe, siatki pilotów, typy pilotów, alokacja zasobów, agregacja pasm;
 - ponowne używanie modeli jest czasami niemożliwe np. ze względu na struktury danych wejściowych (nawet przy używaniu metod transferu wiedzy między domenami).

Oprócz niewątpliwych zalet stosowania neurotechnik mają one wiele ograniczeń, które wpływają na ich zakres stosowalności w praktycznych rozwiązaniach:

- problematyka implementacji sprzętowej;
 - wymuszenie niejako przeprojektowywania istniejących warstw sprzętowych!
- mało (kilkanaście/kilkadziesiąt) jest prac dotyczących (w kontekście radiokomunikacji) podatności sieci na wszelkiego typu zakłócenia;
 - klasyczne sygnały zakłócające wąskopasmowe/szerokopasmowe/CW;
 - ustrukturyzowane sygnały kontradyktoryjne;

Oprócz niewątpliwych zalet stosowania neurotechnik mają one wiele ograniczeń, które wpływają na ich zakres stosowalności w praktycznych rozwiązaniach:

- truizm, ale... potrzebne są duże liczności danych (big data) oraz czas wyuczenia sieci...
 - a gotowe i wyuczone modele są w większości niedostępne...
 - powtarzalność badań naukowych czasami niemożliwa do realizacji?
 - problem ze zbieżnością modeli;
- podejście do problemów w stylu „czarnej skrzynki”;
 - bardzo mocno ograniczona interpretowalność modeli;
 - brak specjalistycznej wiedzy prowadzi do problemów np. z wrażliwością sieci;
 - dekompozycja i analiza modeli (interpretowalność).

A to tylko wybrane problemy...

Dotychczas przyjmowaliśmy założenie, że mamy pewną wielowymiarową, nieliniową funkcję $\mathbf{F}$ modelowaną przez sieć:

$$\mathbf{F} : \mathbf{X} \rightarrow \mathbf{Y}$$

gdzie $\mathbf{X}$ to macierz danych wejściowych, a $\mathbf{Y}$ to macierz danych wyjściowych, przy czym $\mathbf{F}$ musi być różniczkowalna. W takim przypadku możemy proces uczenia w ogólności opisać jako:

$$\theta_{n+1} = \theta_n - \alpha \cdot \nabla \mathcal{L}(\mathbf{F}(\theta_n, \mathbf{X}))$$

gdzie α to współczynnik uczenia, $\mathcal{L}$ to funkcja kosztu, θ to parametry modelu.

A co w przypadku gdy kolejny wektor danych wejściowych $\mathbf{X}^* \neq \mathbf{X}$?

Zaburzenie (celowe lub przypadkowe) danych wejściowych wpływa na efektywność działania sieci neuronowych - w tym także neurotechnik odbioru radiowego.

W takim przypadku zasadnym jest rozpatrywanie celowych ataków i wprowadzanie celowych interferencji wnikających i propagujących się wewnątrz modeli neuronowych w odbiorniku spełniających warunek

$$\operatorname{argmax}_{X^*} \mathcal{L}(F(\theta_n, X^*)),$$

czyli tzw. **sygnały kontradyktoryjne (ang. adversarial)**. Co wraz z wprowadzaniem neurotechnik otwiera nową kategorię interferencji radiowych.

Zagadnienie atakowania sieci neuronowych jest przedmiotem aktualnie prowadzonych badań w literaturze (oczywiście głównie w informatyce) i dzielone są na ataki:

- z pełną wiedzą o modelu (white box);
- z brakiem wiedzy o modelu (black box);

których skuteczność silnie zależy od zakresu wiedzy jaki posiadamy o atakowanym modelu, jego odpowiedzi lub innych parametrów np. rozkładu zbioru uczącego.

Póki co, więcej pytań niż odpowiedzi:

- jak to się ma do radiokomunikacji?
- czy posiadamy pełną, częściową czy nie mamy wiedzy modelach?
- czy da się to zrealizować w odbiorniku?

Skupiając się w pierwszej kolejności na atakach z pełną wiedzą o modelu, np.:

- FGSM (Fast Gradient Sign Method Attack)

$$X_{adv} = X + \epsilon \cdot \text{sign}(\nabla \mathcal{L}(F(X, Y)))$$

- BIM (Basic Iterative Method Attack)

$$X_{adv} = X + \alpha \cdot \text{sign}(\nabla \mathcal{L}(F(X_{adv}, Y)))$$

- wygenerowana interferencja nie jest ustrukturyzowana względem znanych i klasycznych sygnałów interferujących;
- czy atak zadziała na modelu bez pełnej wiedzy o jego architekturze, modelu i danych uczących?
 - możemy stworzyć model bliźniaczy... a właściwie już go mamy;
- neurotechniki bazujące na analizie czasowo-częstotliwościowej (np. spektrogramowej) mogą być wrażliwe na emitowane sygnały nie bezpośrednio w paśmie pracy;
- jak uodpornić się na tego typu ataki?
 - póki co... po prostu poszerzać modele o zbiory danych kontradyktoryjnych;
 - brzmi jak utopia Chata GPT: masz problem -> więcej danych;
 - istnieją metody np. bazujące na transformacji wyuczonych modeli;

The background is a light cream color with faint, thin-lined geometric patterns of interconnected dots and lines in the upper corners. The bottom third of the image is a dark blue band featuring glowing, wavy lines and a field of small, out-of-focus light points.

Dziękuję za uwagę