Transfer Learning in Wireless Channel Prediction

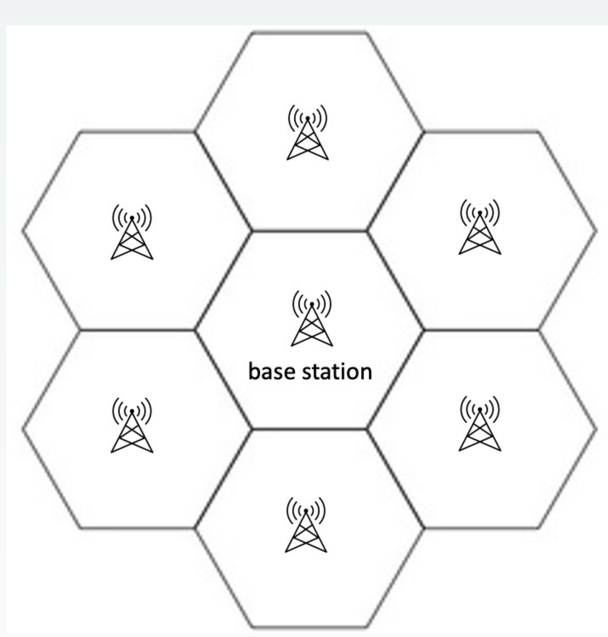
Jieying Chen, Abdalaziz Sawwan, Shuhui Yang, and Jie Wu
Department of Computer Science, Purdue University Northwest, Hammond, IN
Center for Networked Computing, Temple University, Philadelphia, PA

Introduction

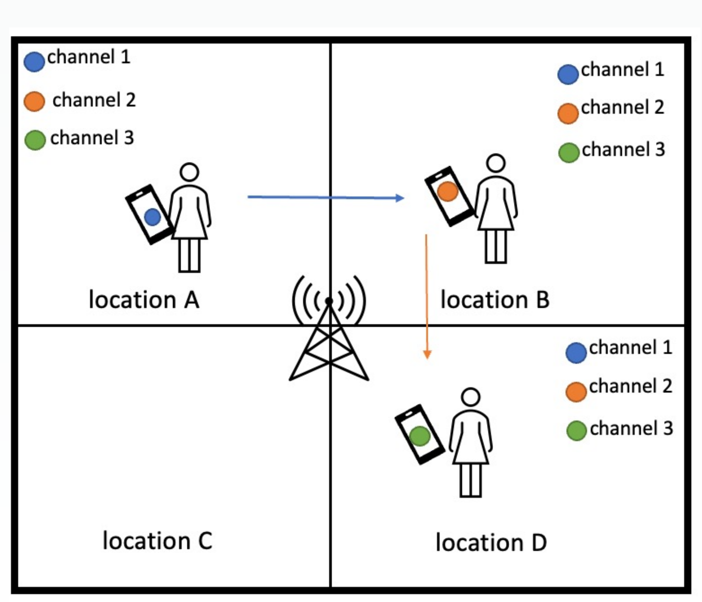
Cellular systems
- Transmit data via radiofrequency (RF) signals
- Provide RF channels for data transmission services
- Four major components
1. Cellular towers and antennas
2. Public switched telephone network (PTSN)
3. Mobile telephone switching office (MTSO)
4. Mobile subscriber units (MSU)

Channel switching
- Switch RF channels based on
1. User location
2. Signal strength
3. Availability

Transfer learning
- Transfers knowledge from one domain to another
- Subfield of machine learning
- Applications
1. Computer vision (CV)
2. Natural language processing (NLP)



Cellular systems.



RF channel switches based on user location.

TAPIA and STARS Celebration
Texas, September 2023

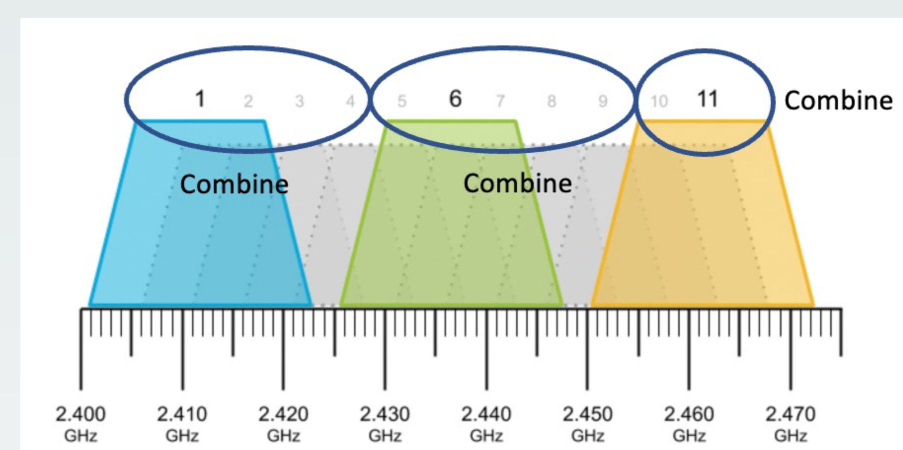
Problem Formulation

RF channel switching is
1. Time-consuming
2. Resource-intensive

Transfer learning is
1. Less time-consuming to train models
2. Less data-dependent to train models

Goal: Minimize channel switching

RF channels overlap in 2.4 GHz band
- Combine overlapping channels

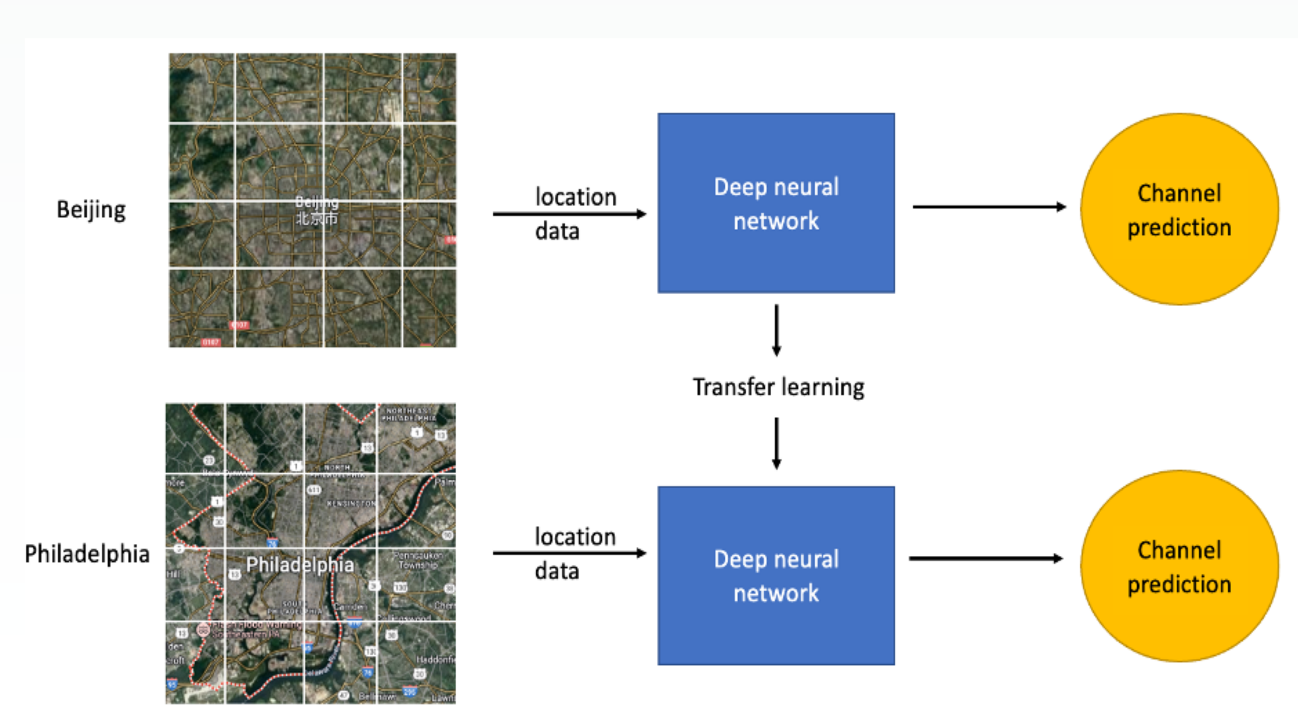


RF channel are combined in dataset.

Instance-based transfer learning

Homogeneous transfer learning
- Source and target domains have similar tasks

Instance-based transfer learning
- Corrects marginal and distribution differences
- Assign weights to the loss function of source domain
- Target domain uses the assigned weights



The knowledge gained from Beijing (City A) can transfer to Philadelphia (City B).

Weighting strategy
- $\mathbb{E}_{(x,y)\sim P^T}[\mathcal{L}(x,y;f)] = \mathbb{E}_{(x,y)\sim P^S}\left[\frac{(P^T(x,y))}{(P^S(x,y))}\mathcal{L}(x,y;f)\right]$
- $E_{(x,y)}$ is the expected risk
- x is the pattern in the
- domain
- y is the label in the domain
- $L(x,y;f)$ is the loss function that depends on the parameter f
- $P^T(x)$ is source domain, $P^S(x)$ is target
- $\frac{(P^T(x,y))}{(P^S(x,y))}$: Instances are drawn from source, and generalized to target
- Generalized instances are now the weighting parameters -> β (x, y)

- Estimate the weighting parameter
- $\min_f \frac{1}{n^s}\sum_{i=1}^{n^s}\beta_i\mathcal{L}(f(x_i^s),y_i^s)+\Omega(f)$
- β_i is the weighting parameter
- n is the number of instances
- Ω(f) is the regularizer of the re-weighted risk

Proposed Simulation Techniques

Two cities examined:
- City A (Beijing) and City B (Philadelphia)
- City A -----transfer knowledge---→ City B

Basic model for simulations
- 7-layer DNN for model training
- Latitude + Longitude ~ Channel
- Predict occupied channels of location

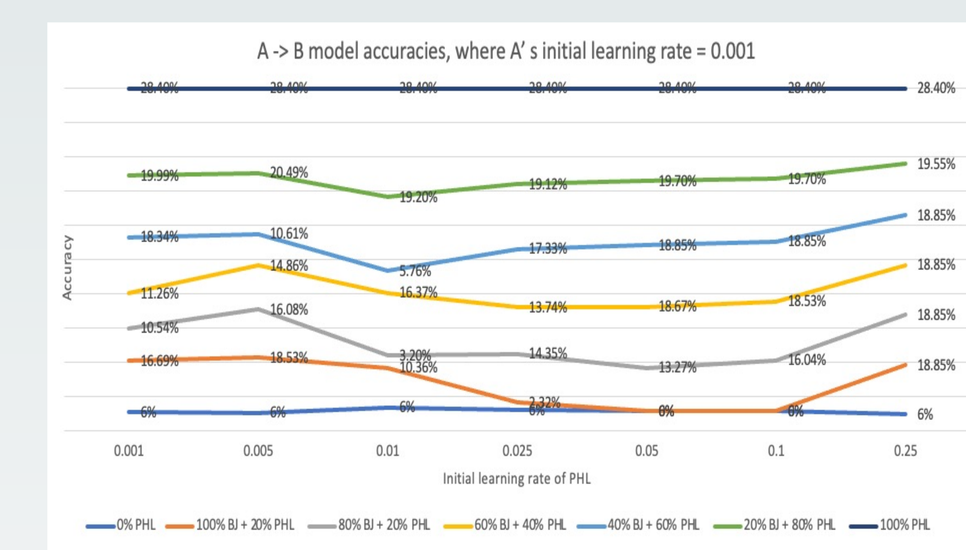
Baseline and Upper bound
- Baseline: averaging 100 accuracy of City A –transfer → City B results
- Upper bound: result of training City B model

Varying training data ratio
1. Train certain % of A data
2. Use the same model, train certain % of B data
3. Different combinations of % A and B data

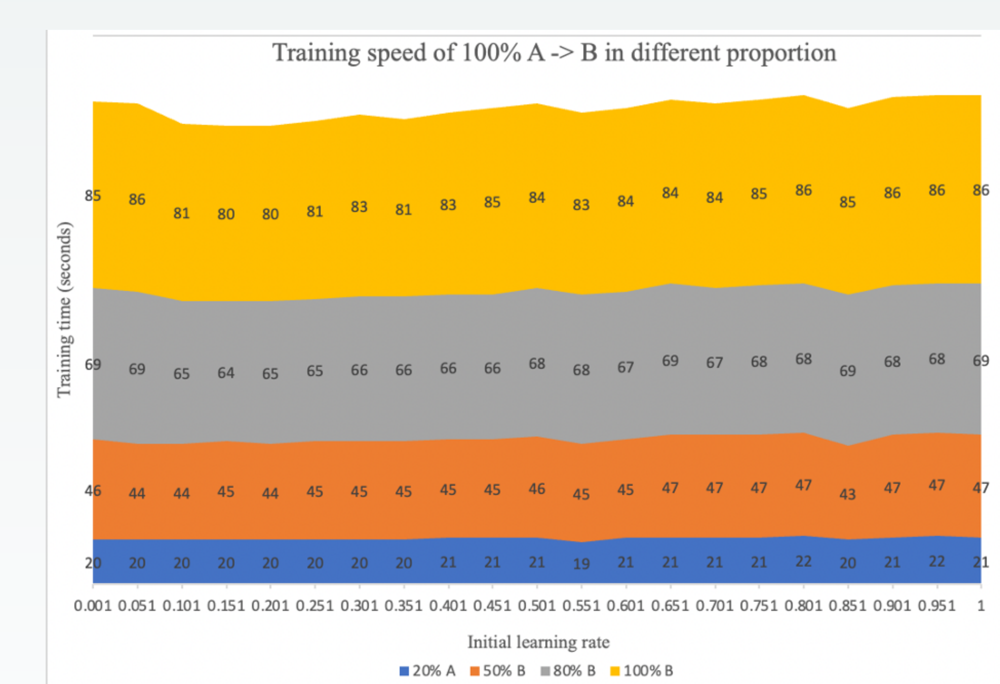
Fine-tune initial learning rates
1. Assign a set of learning rate (LR) for City A
2. For each LR of City A, a set of LR assigns to City B
3. For each pair of LR (LR_A, LR_B), train model with different combinations of Cities A and B data

Training time vs. accuracy
1. Assign a set of LR for training
2. Cities A and B use the same LR
3. Train City A --> B in different %
4. Record training time and accuracy

Evaluation



- As % B increases, accuracy increases.



- Training speed of A → B.
- Uses less training time while maintaining some level of accuracy.

Conclusion

- Transfer learning is feasible in predicting the used channels in each location
  - Propose simulation techniques
  - Fine tune DNN model and adjust data ratio to generate test results
  - Experiments verify the performance
  - Useful when time and data are scarce