Domain Adaptation Techniques for Fine-Grained Occupancy Estimation

Tianyu Zhang, Omid Ardakanian
University of Alberta

October, 2018

Room-level occupancy monitoring is quite challenging

- Lack of special-purpose sensors for monitoring Occupant Presence and Actions (OPA) in buildings

- Difficulty of collecting sufficient ground truth data (i.e., time series data representing the number of occupants over time)

Addressing the ground truth data collection problem

- Existing approaches
  - log the number of occupants manually, which is expensive and error-prone at scale
  - deploy high-fidelity sensors, which are perceived as privacy invasive

Addressing the ground truth data problem

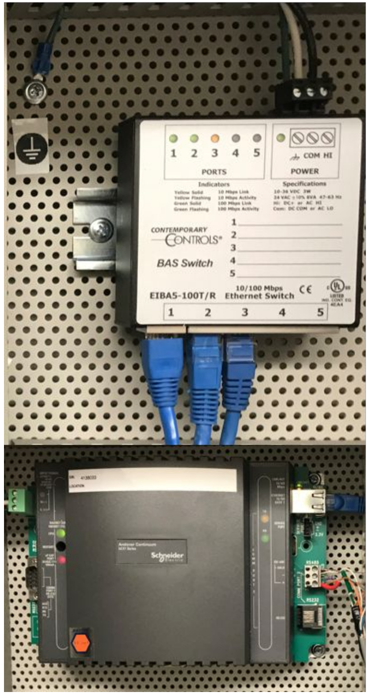
- Existing approaches
  - log the number of occupants manually, which is expensive and error-prone at scale
  - deploy high-fidelity sensors, which are perceived as privacy invasive

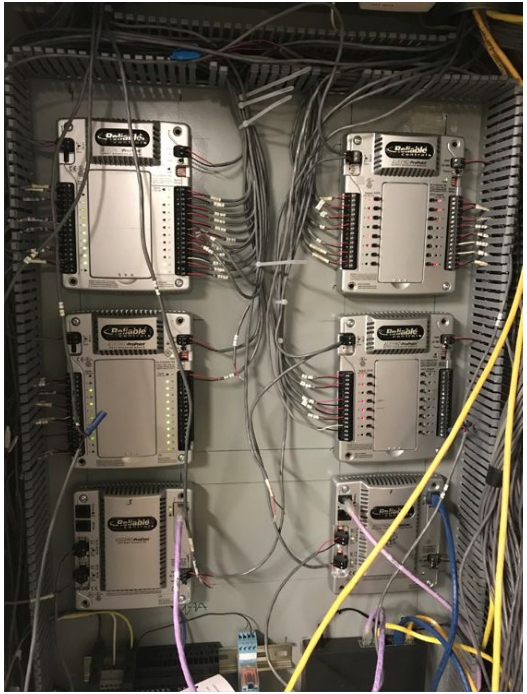
- Our approach is to adopt semi-supervised/unsupervised domain adaptation techniques to reduce the amount of ground truth data required for developing a well-suited model

Basic idea

- Estimate room-level occupancy from trend data, when possible

HVAC sensors



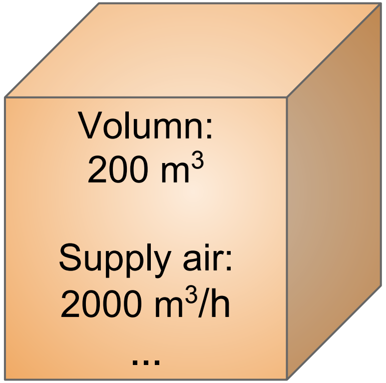


Basic idea

- Estimate room-level occupancy from trend data, when possible
- Build supervised learning models that capture short-term temporal dependencies
- Adapt and reuse occupancy models that have been trained in a controlled environment (source domain) in another `similar’ environment (target domain)!

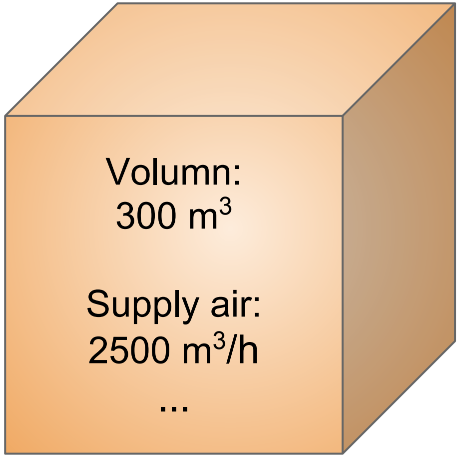
Domain adaptation

Domain A



Training data available between
January 1, 2014 and December 31, 2015

Domain B

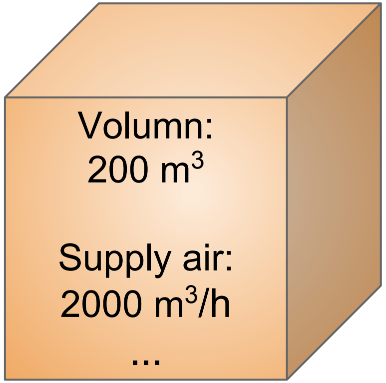


Training data available between
January 1, 2018 and January 5, 2018

Domain adaptation

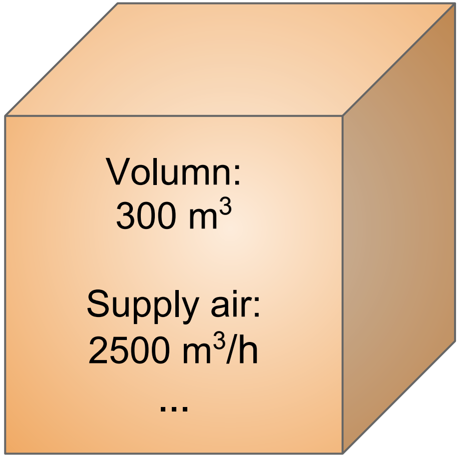
Domain A

Model A



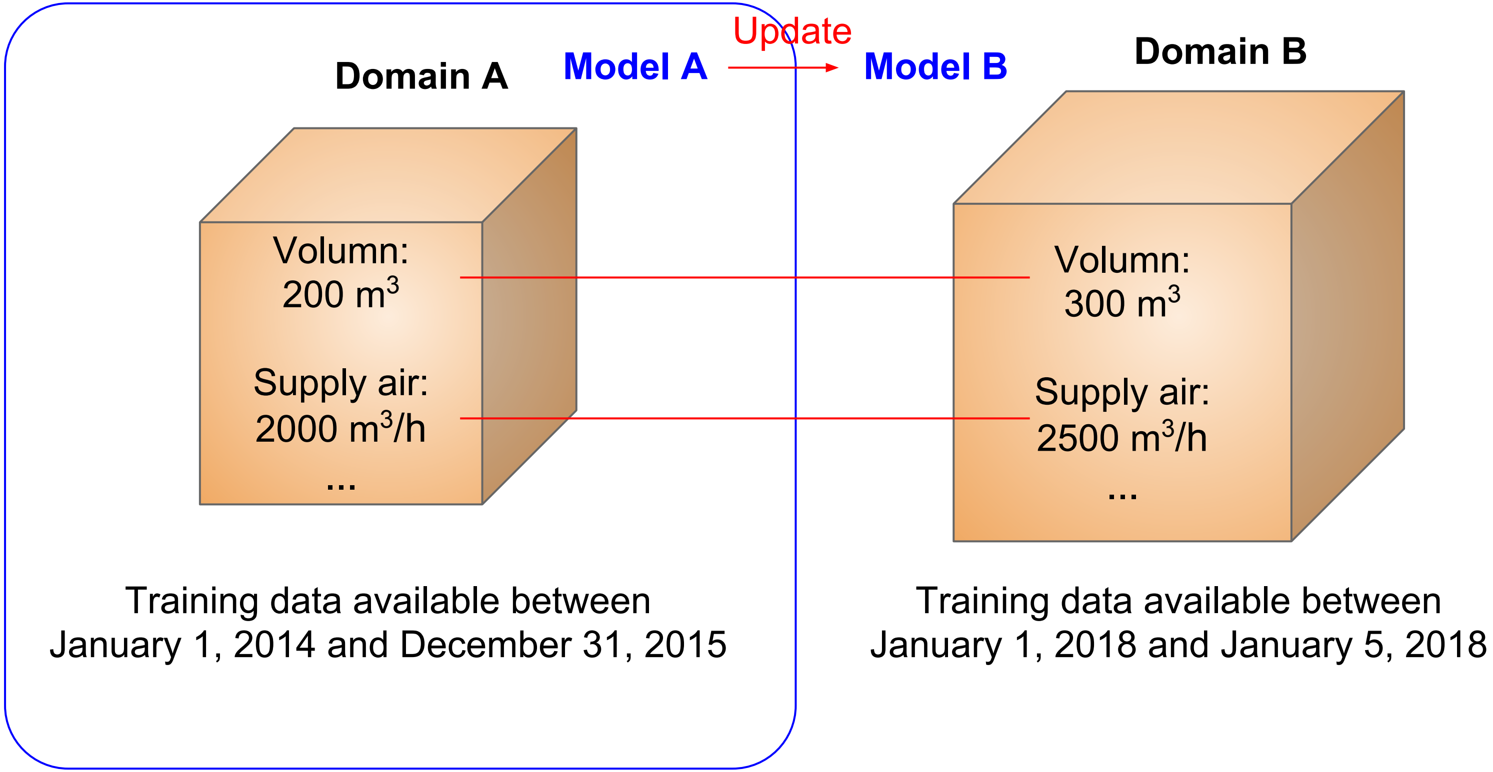
Training data available between January 1, 2014 and December 31, 2015

Domain B

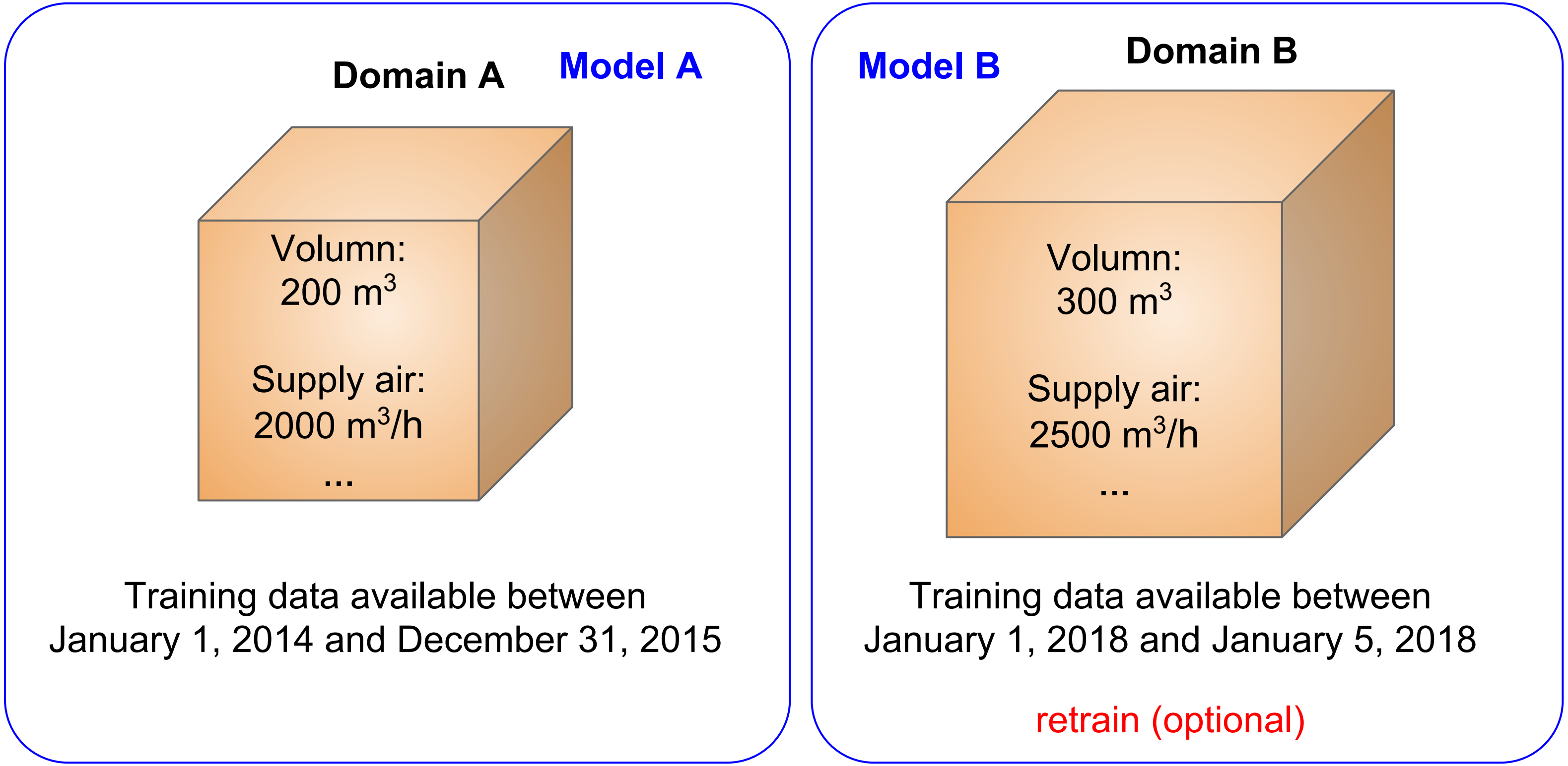


Training data available between January 1, 2018 and January 5, 2018

Domain adaptation



Domain adaptation



Data-driven modelling techniques

A well-suited model should describe the complex nonlinear relationship between historical sensor data and the number of occupants

- standard supervised learning techniques, e.g., Logistic Regression (LR) and Support Vector Regression (SVR)

Data-driven modelling techniques

A well-suited model should describe the complex nonlinear relationship between historical sensor data and the number of occupants

- standard supervised learning techniques, e.g., Logistic Regression (LR) and Support Vector Regression (SVR)

Since the number of occupants varies over time, the model should capture this temporal dependency

- neural network models that have “memory”, e.g., Long Short Term Memory (LSTM) and Nonlinear AutoRegressive network with eXogenous input (NARX)

Data set

- Four rooms in a campus building
  - Two study zones, 125m^2, seating capacity: 36, supply air pressure: 3000m^3/h
  - Two classrooms, 139m^2, seating capacity: 85, supply air pressure: 4800m^3/h
- Two sensors were installed in each room, reporting the damper position and CO_2 level to the BMS in one-minute intervals

Comparing different domains

Similarities:

- Sensing modalities
- Surrounding environment
- Building envelop

Differences:

- Floor area
- Supply air pressure
- Occupancy pattern

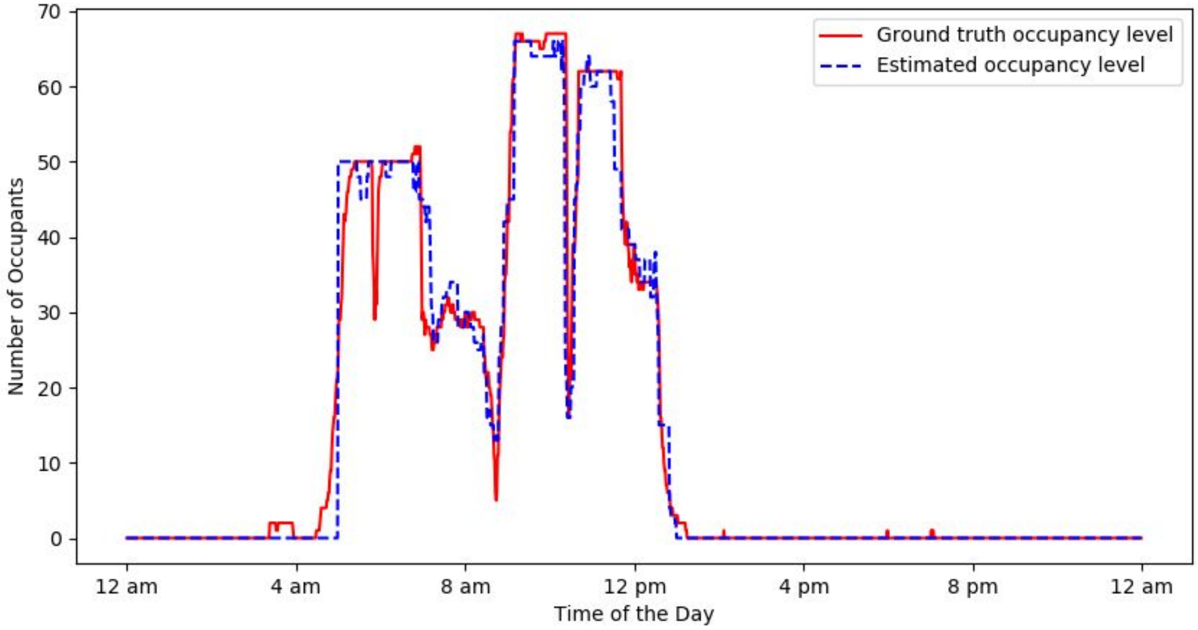
Adapting the model to the new environment

Since we have the same features (sensing modalities, meteorological factors) in both domains, we do not need to change the architecture of our neural network models

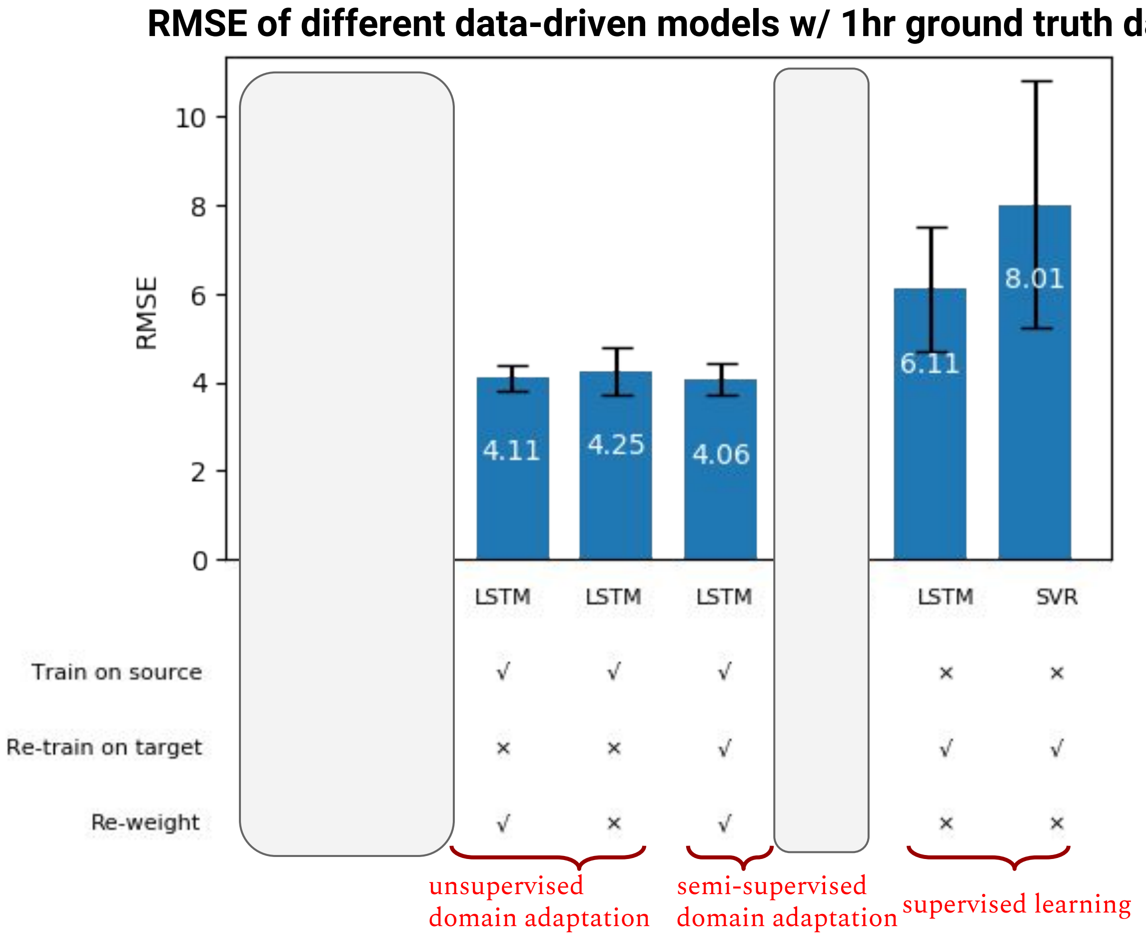
Reweighting: updating the weights in the trained model to correct for apparent differences between the two domains

Retraining: training the model with limited data after reweighting can significantly reduce the training and test errors

Semi-supervised domain adaptation using one day ground truth data

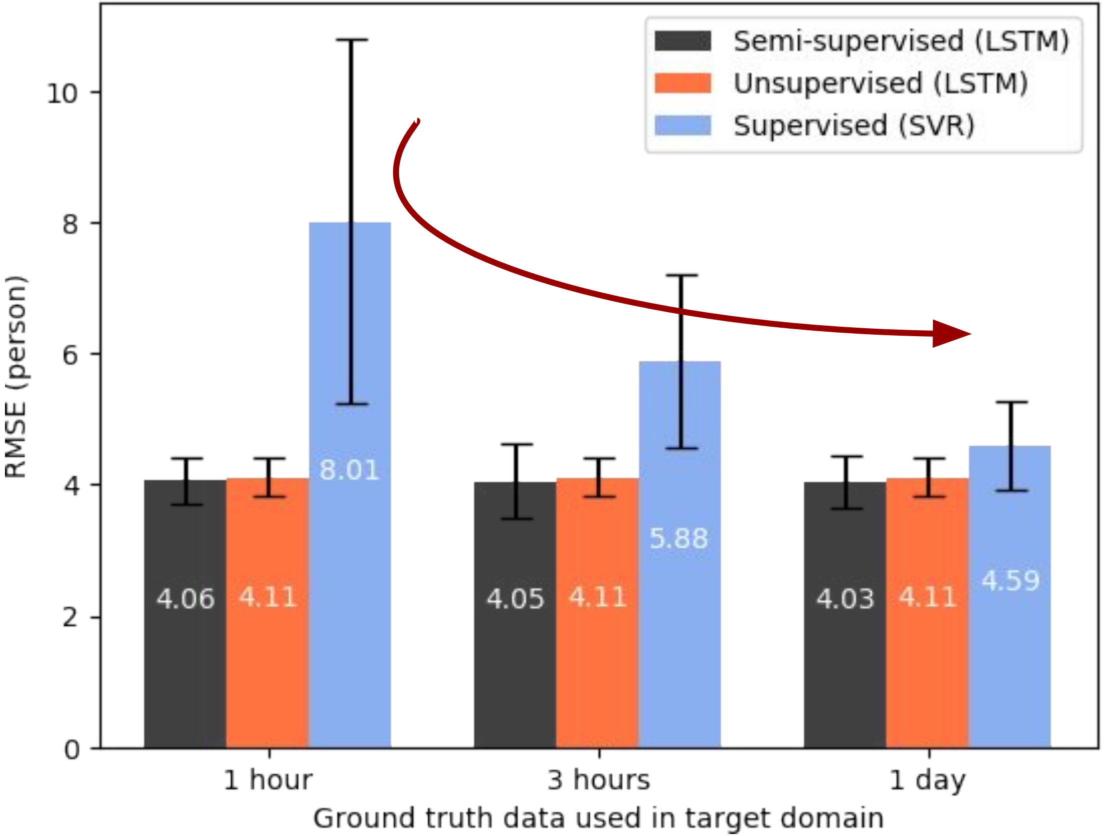


Evaluation



RMSE of different data-driven models w/ 1hr ground truth data
RMSE
0
2
4
6
8
10
4.11
4.25
4.06
6.11
8.01
LSTM
LSTM
LSTM
LSTM
SVR
Train on source
Re-train on target
Re-weight
√ √ √ × ×
× × √ √ √
√ × √ × ×
unsupervised domain adaptation
semi-supervised domain adaptation
supervised learning

How does more ground truth data improve the results?



Future work

- Inter-building analysis with domain adaptation
  - train the model on one building and apply to another one

- Sensors calibration using domain adaptation

Future work

- Inter-building analysis with domain adaptation
    - train the model on one building and apply to another one

- Sensors calibration using domain adaptation

- Leverage a collection of models developed on multiple source domains, each containing a subset of features

ardakanian@ualberta.ca

https://webdocs.cs.ualberta.ca/~oardakan/