Forecasting Building Occupancy Using Sensor Network Data

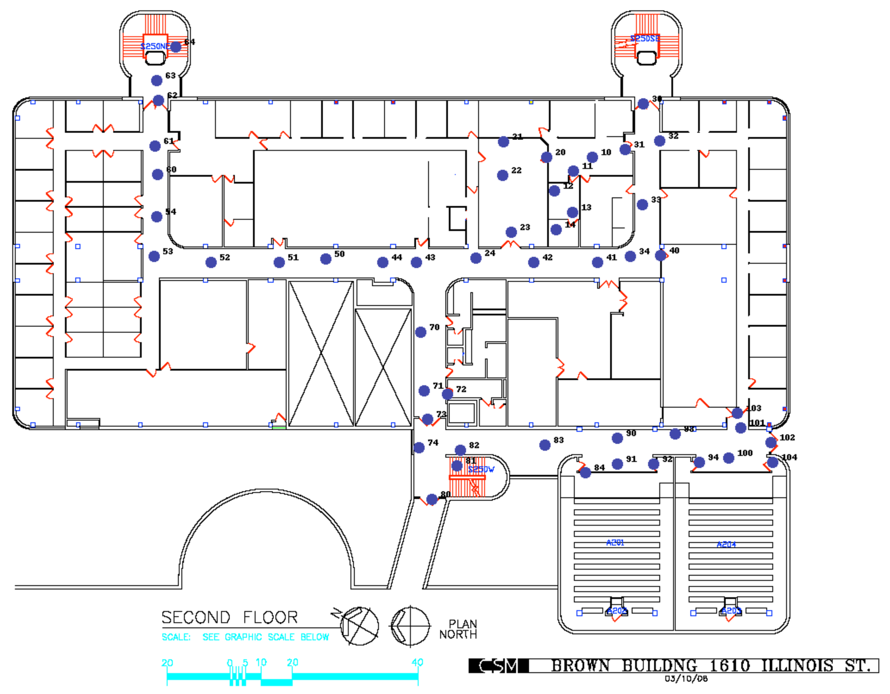
James Howard, Dr. William Hoff
Center for Automation, Robotics, and Distributed Intelligence
Colorado School of Mines

Motivation

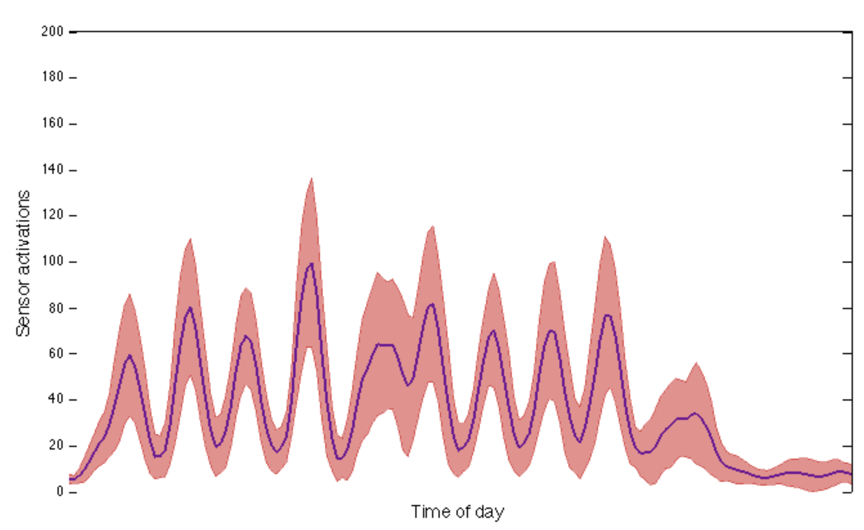
According to the U.S. Department of Energy, energy used for heating and cooling accounts for approximately 35 - 45% of the total expenditure of energy within a building [1]. Accurate estimates of building occupancy allows modern heating and cooling control systems to dynamically forecast building demands and respond efficiently to those demands thus reducing energy consumption.

Data

The data measures counts of people over time at various positions within a building. The time resolution of occupancy density is an important factor in our work. Results shown here are based on total counts per 10 minutes.



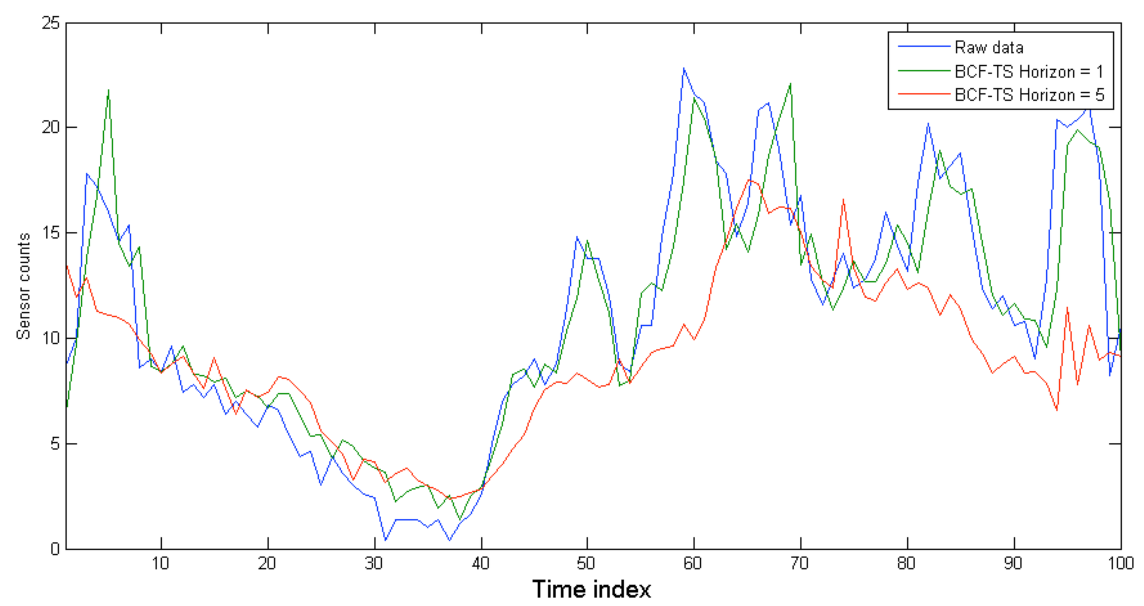
Sensor locations for classroom building.



Counts of people with one standard deviation on Wednesdays.

Forecasting Models

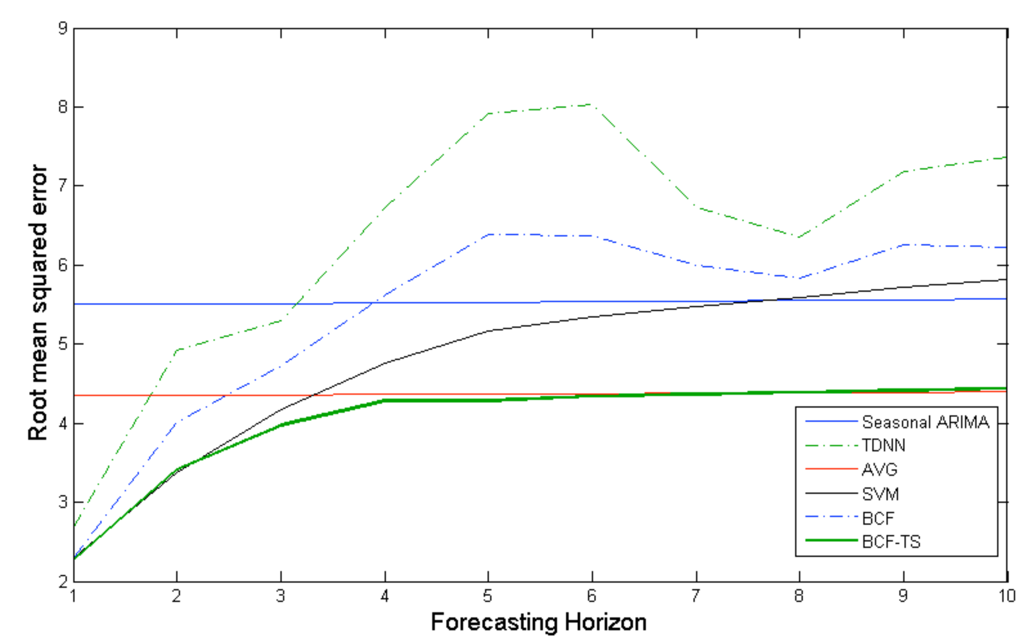
Common models used to forecast time series include: Seasonal ARIMA, Time Delayed Neural Network, Support Vector Machine, and Historic Averages. However, when models are used individually they have errors that are unique to the model.



Forecasting example on raw data for horizon of 10 minutes and 50 minutes.

Combined forecasting

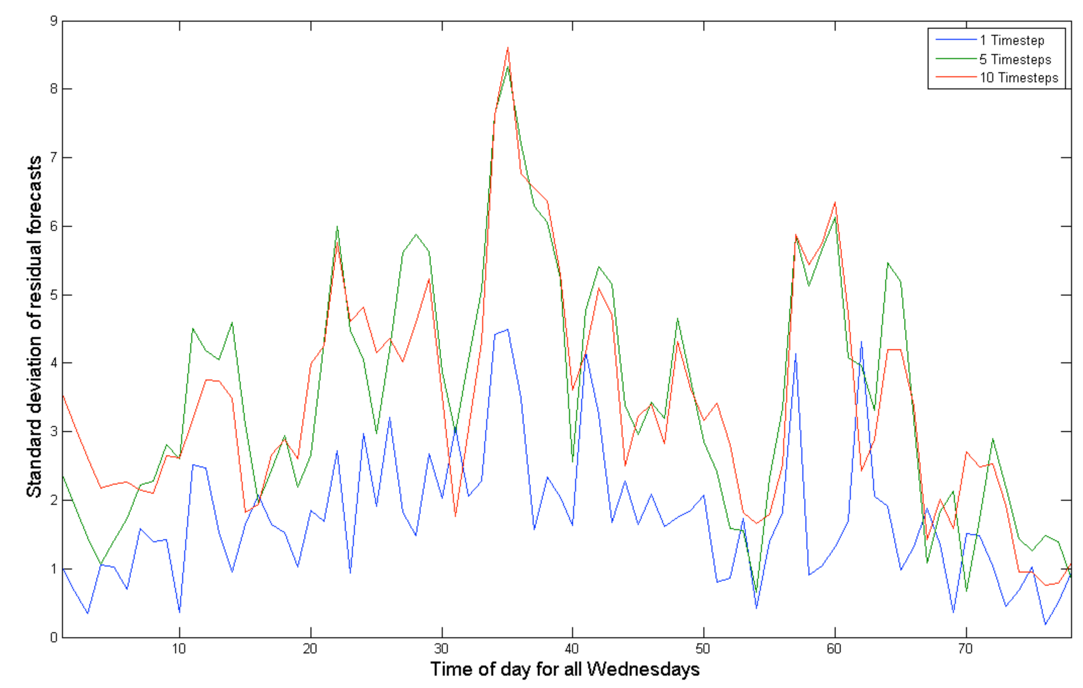
By combining multiple forecasts we can produce a more accurate forecaster. We use Bayesian combined forecasting (BCF) [2] which has historically been shown to be an accurate multiple model forecasting approach for time series.



Forecasting accuracy for various models.

Improving BCF (example)

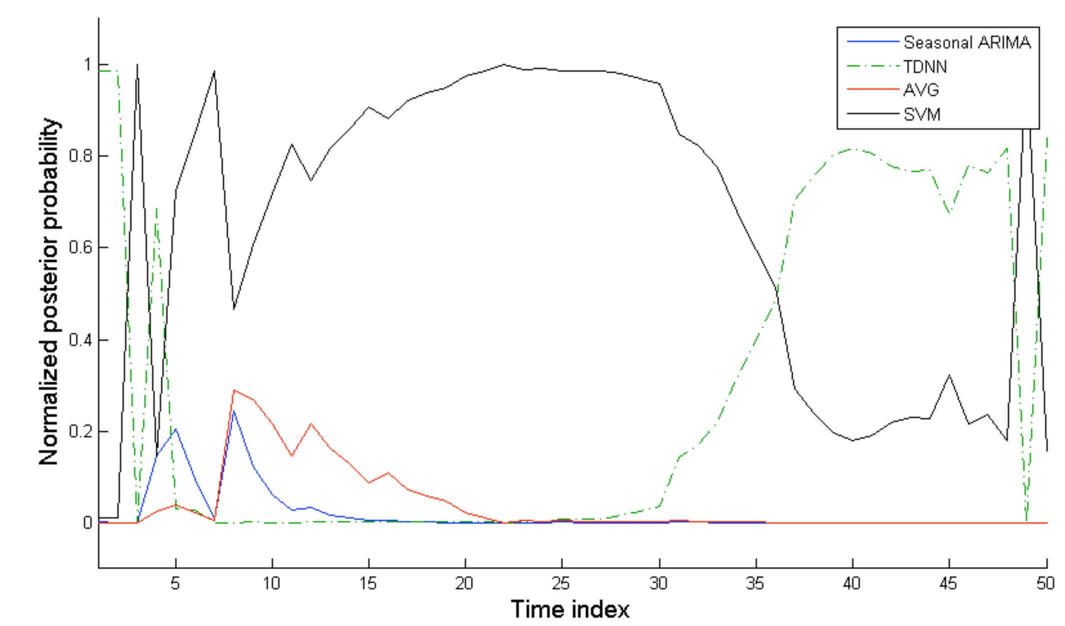
As the forecasting horizon grows, traditional Bayesian combined forecasting is no longer accurate. One of our improvements is to more accurately model forecasting accuracy. To do this we model the distribution of forecasting noise as a mixture of Gaussians instead of a white noise Gaussian.



Standard deviation of forecast residuals per ten minutes for all Wednesday data.

Likelihood Computation

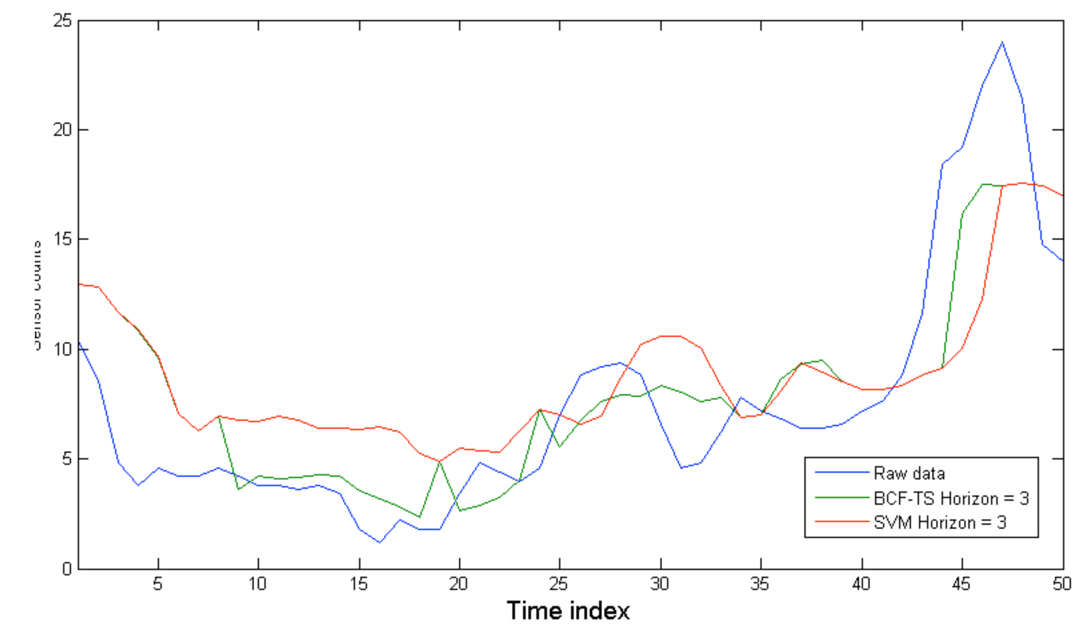
Based on prior trained noise distributions, likelihoods can be computed for each component forecasting model. The likelihoods are then combined to achieve a final forecast.



Likelihood of all forecasting models for a given chunk of data.

Results

For forecasts up to six time steps (60 minutes) into the future our improved BCF shows considerable improvement at forecasting building occupancy data over any of the component models. Beyond six time steps a there does not appear to be a strong correlation to future occupancy. For forecasts beyond six time steps into the future a standard historic average appears to be best. Similar results were achieved on both classroom [3] and office building datasets [4].



Example forecast of BCF (green) vs SVM (red). BCF is considerably more accurate in many areas.

References

(1) US DOE. Building Energy Databook, 2008

(2) V Petridis, et. al. A Bayesian multiple models combination method for time series prediction. Journal of intelligent and robotic systems, 31(1):69–89, 2001

(3) W. Hoff, J. Howard, Activity recognition in a dense sensor network. International Conference on Sensor Networks, 2009

(4) C. Wren, et. al. The MERL motion detector dataset: 2007 workshop on massive datasets. ICMI Workshop on Massive Datasets, 2007

Acknowledgements

National Science Foundation (grant CNS-0931748), The Lockheed Martin Corp. and Northrop Grumman Corp.