Home Heating Using GPS-Based Arrival Prediction

James Scott, John Krumm, Brian Meyers, A.J. Brush and Ashish Kapoor

Microsoft Research
{jws, jckrumm, brianme, ajbrush, akapoor}@microsoft.com

Abstract. Home heating is a major factor in worldwide energy use. We describe two experiments aimed at reducing the amount of time heating systems need to be on, without compromising occupants' comfort. The first resulted in a machine learning algorithm based on GPS data to predict when an occupant will arrive at home. The second examined how long it takes to heat homes based on temperature measurements, telling us how far in advance arrival predictions are needed. Our findings suggest that GPS-based prediction has the potential to reduce home energy consumption compared to existing methods.

1 Introduction

Home heating accounted for 47% of residential energy used and 32% of residential energy costs for U.S. homes in 2001 [3]. Since home heating uses more energy than any other residential energy expenditure (i.e. air conditioning, water heating, and appliances), increasing the efficiency of home heating is an important goal for saving money and protecting the environment. Although programmable thermostats provide the technology to solve this problem, they are underutilized. Of the 23% of U.S. homes with a programmable thermostat, only 40% use the programming feature, partly due to the perception that thermostats are hard to operate [2].

We instead envision a model where the home's heating turns on only in anticipation of an occupant's future arrival. It requires no programming on the part of the home's occupants, since it automatically learns and responds to the travel behavior of the occupants and the heating behavior of the home. This paper presents a machine learning algorithm that uses people's GPS traces to predict their future arrival times at their homes. It also examines the heating profiles of individual homes to show how far in advance we need to make arrival predictions. The time needed to heat the homes we studied (median 88 min.) was quite long relative to average commute times, suggesting arrival prediction is a valid approach to saving energy.

2 Related Work

Other researchers have explored using technology to improve home heating and conserve energy. Mozer's neural network house tried to infer patterns to anticipate the needs of inhabitants and save energy [10], while the House_n project proposed providing information to teach residents how to manage the temperature in their environment [6]. Chetty et al.'s study of households' practices for managing water, electricity and natural gas systems highlighted issues with programmable thermostats and contention caused by different temperature preferences [1].

The work most closely related to ours is from Gupta et al. at MIT [5]. They describe a GPS-controlled thermostat that automatically turns on the home's heat when a potential occupant is nearby. Their system computes how long it would take any potential occupant of the home to drive home, using the occupant's GPS-measured location. In contrast, our proposed system uses people's historical movements to predict future home occupancy, irrespective of the distance between the home and the person's current location. This focus on predicting future occupancy of a home also differs from previous work that uses GPS data such as Predestination [7], which predicts a driver's likely destination, and Liao et al.'s personal maps that try to discriminate activities and predict future transportation modes and goals [8].

3 Predicting Home Arrival Times

We decided to analyze location traces of people in an effort to predict when they would arrive at home. Our expectation was that machine learning could exploit a person's habits, based on their location, time of day, and other features to predict how long it would be before they arrive at home. We recruited six participants to log their time-stamped locations with GPS. Three participants carried a RoyalTek RBT-2300 GPS logger in their pocket, recharging its battery every night. The three other participants mounted the same type of GPS logger in their car, powered by the car's cigarette lighter. These loggers only recorded when the car was turned on. We sampled time-stamped GPS coordinates at a rate of one per minute.

Each participant lives in a detached home whose location we knew. To account for GPS noise and short excursions outside the home (e.g. mailbox), we declared the person to be "at home" if their lat/long was within 50m of their home. We segmented each person's time-stamped lat/long traces into trips consisting of contiguous periods when the person was outside the 50-meter circle. We eliminated trips shorter than 10 minutes to account for persistent GPS noise. Table 1 summarizes the data collected. Our goal is to predict when a person will arrive at home. For each lat/long point in each trip, we make a probabilistic prediction of whether or not the person will be home in a pre-specified amount of time: 30, 60, and 90 minutes.

Ppt.	Pocket/Car	Days Observed	Trips	Approx. Commute
1	pocket	51	51	35-45 min. (43km)
2	pocket	54	41	15-20 min. (10 km)
3	pocket	73	117	20 min. (1 km)
4	car	68	56	18 min. (13 km)
5	car	260	123	15 min. (3.8 km)
6	car	352	512	No commute

Table 1. Summary of GPS observations of subjects. Commute times are self reported.

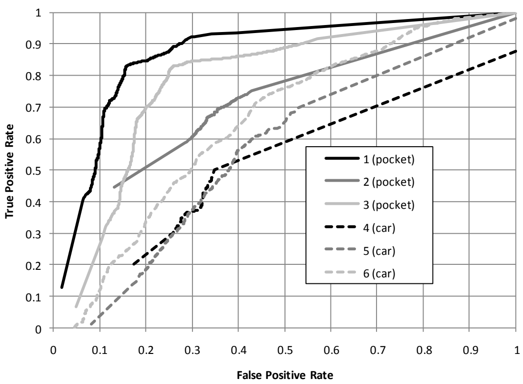

Fig. 1: ROC curve for 90-minute arrival prediction; shows the tradeoff between true and false positives

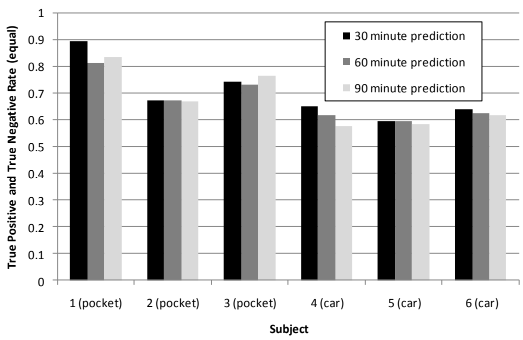

Fig. 2. Equal error rate performance; shows how well the predictors work when the false positive rate and false negative rates are equal.

We chose these times based on the home heating times we found, detailed in the next section. These choices led to a separate binary classifier for each participant and each time interval, where the classifier categorized each point as either negative (will not be home within pre-specified time) or positive (will be home within pre-specified time). We used a nearest neighbor approach with a learned distance metric based on five features computed at every point in a trip: hour of day (integer), day of week (integer), latitude (real), longitude (real), and hours since last home (real). For each classifier, we transformed the feature space with a matrix designed to better separate the two classes, as explained in [4].

For testing, we left out one trip at a time and compared each point in the test trip with all the remaining trips for a given participant. Figure 1 shows a ROC curve for 90-minute prediction for each of our participants. Each point on each ROC curve represents an operating point. In general, the price of more true positives (i.e. heat turning on in time) is more false positives (i.e. heat turning on prematurely). A more succinct summary of prediction performance is the equal error rate, where the false positive rate is equal to the false negative rate. Figure 2 shows the true positive rate (and true negative rate) at the equal error rate setting. This rate varies from a low of 0.58 to a high of 0.83 for 90-minute predictions.

4 Measuring Home Heating Times

In order to realize a home heating system based on location prediction, we also need to be able to predict how far ahead of time we would need to start the heating. To explore the feasibility of this, we conducted a home heating/cooling measurement study. We deployed Thermochron DS1922L iButton temperature loggers in 13 homes over two weeks. The homes belonged to people at our company and were located in and around Redmond, USA. All were wood frame constructed, the oldest in 1965 and the newest in 2008, and used gas furnaces with forced air heating. A raw trace covering a typical day is shown in Figure 3.

Each home received 3 iButtons, one placed on the central heating unit ("main heater"), one outside and out of direct sunlight ("outside"), and one placed near the home's thermostat ("inside"). The iButton sensor can measure from -40.0F to 185.0F with a resolution of 0.11F. We recorded samples every 5 minutes for 14 days. We chose the iButton sensors for their small size and robustness (they look like a small coin battery), thus facilitating deployment by home owners themselves.

We noted the set-point temperature of each house by finding the warmest steady-state temperature. We identified "heating periods" by examining the "main heater" trace for times that the house was below the set-point and the furnace stayed on until the set-point temperature was reached. Four of the homes had the thermostat set to a constant temperature so we were unable to extract useful, extended heating periods from them. However, they represent homes where using automatic heating control could have a large impact.

We observed that each heating period was well-modeled by a linear increase in temperature in the house over that time. Figure 4

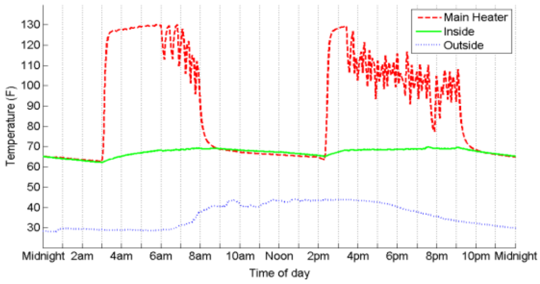

Fig. 3. Temperature record for a single house and day

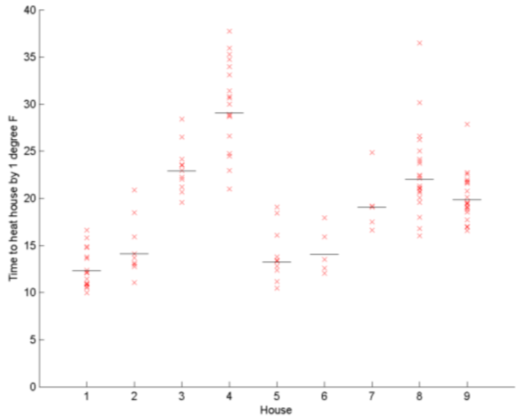

Fig. 4: Heating rates measured in the study for each heating period in each house. Black lines are medians.

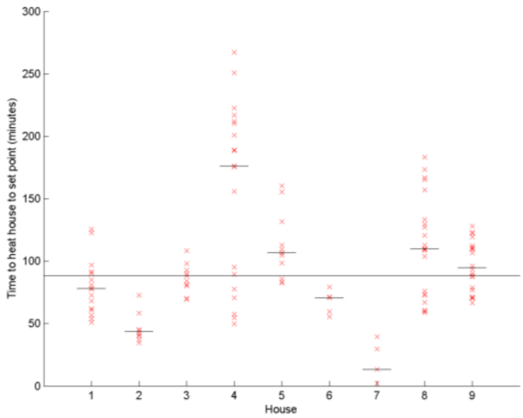

Fig. 5. Heating period durations. Black lines are medians; long black line (overall median) at 88 mins.

shows the average time each furnace required to raise the temperature of the house by 1F for each individual heating period. The variation in this time, as much as a factor of two for individual houses, was due in part to the outside temperature (which varied from 28F to 58F over the course of the logging) and in part to other factors which we had not measured, such as wind speed, solar heating effects, the opening or closing of ventilation, etc. Given the large observed range (a factor of 3.8 between houses) during the same local weather conditions, it appears necessary to model each house's heating capabilities separately.

However, the empirical data we collected can help give a measure of how far in advance a deployed GPS prediction system would need to predict. Figure 5 shows the duration of all heating periods in our measurements – i.e. the time taken by the furnace to heat from whatever temperature it started at up to the set-point. This measure implicitly captures the other factors affecting heating times, as well as the cooling behavior of houses and the normal occupancy pattern, since the heat duration depends on how cool the house became while unoccupied. The overall median duration of a heating period was 88 minutes.

5 Discussion

In combining these two analyses we find that during our study period the time required to heat a home (median 88 minutes) was significantly longer than the commute times we found and the U.S. national average commute time of 24.3 minutes [11]. This delay is also certainly long enough to discourage participants from relying on manual adjustment of the thermostat when they returned to the home. In addition, we have shown that GPS data can be used for arrival prediction for homes at a long enough timescale to be useful for controlling home heating in many cases.

While we have not quantified the level of energy savings that this approach will provide, we refer the reader to Gupta et al. [5] who show that location-based techniques can outperform standard manual and programmable thermostat solutions. Gupta et al. showed improvements in heating efficiency when the drive-home time was 90 minutes or greater, so our work relaxes a key constraint from that work.

The uncertainties inherent in future location prediction and in predicting the heating time required for a home lead to an interesting tradeoff between a "conservative" system that heats if there is even a slight chance that is necessary, and a more "optimistic" system that heats only if there is a greater expectation of need. By this metric the Gupta et al. system is at the conservative end. However, moving towards the optimistic end directly translates into energy saving at the cost of sometimes experiencing a colder house than desired.

In addition to heating, we have looked briefly at how homes cool after being heated. We found that the exponential cooling models prevalent in the literature (e.g. [9]) do not fit our measurements well. Instead, many of the houses in our study showed an initially quicker drop that flattened out to a long-term exponential trend. Despite this drop, many houses stay relatively warm for useful periods – the median time for a house cooling by 1F was 35 minutes. So, if a drop of 1F near departure is regarded as acceptable by occupants, departure prediction and preemptively turning the heat off could result in additional energy cost and environmental savings.

6 Concluding Remarks

Motivated by the potential to save energy in heating homes using GPS-based arrival prediction, we have conducted two preliminary feasibility studies. We have developed and evaluated a machine learning based home arrival prediction algorithm using GPS traces. We have also conducted measurement studies showing that, for our participants, their commute times and the time that their homes require to re-heat indicate that GPS-based arrival prediction has the potential to save significant energy.

In future work we intend to implement GPS-based arrival and departure prediction "in the wild". This entails performing real-time temperature sensing and heater control, building and training an empirical model of heating and cooling times, addressing multiple occupancy issues, and providing a user interface for occupants to obtain status information and exert manual control.

7 References

1. Chetty, M., Tran, D., Grinter, R.E. Getting to Green: Understanding Resource Consumption in the Home. Proc. Ubicomp 2008, ACM Press (2008), 242-251.
2. Emerson Climate Technologies Survey http://www.emersonclimate.com/NEWS/News/CC_002199.htm.
3. Energy Consumption by Sector http://www.eia.doe.gov/emeu/aer/consump.html.
4. Goldberger, J., et al., Neighbourhood Components Analysis, Proc. NIPS 2005, 513-520.
5. Gupta, M., S.S. Intille, and K. Larson, Adding GPS-Control to Traditional Thermostats: An Exploration of Potential Energy Savings and Design Challenges. Proc. Pervasive 2009. Springer (2009).
6. Intille, S., Designing a home of the future. IEEE Pervasive Computing, vol. April-June, pp. 80-86, 2002.
7. Krumm J. & Horvitz, E., Predestination: Inferring Destinations from Partial Trajectories, Proc. UbiComp 2006, Springer, 243-260.
8. Liao, L., Patterson, D. J., Fox, D., and Kautz, H., Building Personal Maps from GPS Data. Proc. MOO 2005.
9. Meredith, D. Intro to Building Environmental Systems http://www2.fe.psu.edu/~dxm15/aet121/frames.htm
10. Mozer, M. C., The neural network house: An environment that adapts to its inhabitants. Proc. of the AAAI Spring Symposium on Intelligent Environments, AAAI Press (1998), 110-114.
11. U.S. Census American Community Survey http://www.census.gov/acs/www/Products/Ranking/2003/R04T040.htm