MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Activity Mining in Sensor Networks

Christopher R. Wren, David C. Minnen

TR2004-135 December 2004

Abstract

We present results from the exploration of activity discovery based on impoverished sensors. Networks of low-cost, low-power, low-bandwidth sensors are a practical way of gathering context awareness in buildings. They are more widely applicable than dense networks of cameras because of their low component cost, low installation cost, and low privacy cost. However impoverished sensors pose a significant challenge for activity monitoring due their low capability. We build on our behavior understand work with impoverished sensors to show results relating to behavior discovery and novel event detection.

Presented as an invited poster at NIPS:WARD 2004

Decemer 17, 2004: Invited poster presentation at the 2004 Neural Information Processing Systems Workshop on Activity Recognition and Discovery

1 Sensors

17 1-bit motion detectors, each with 3x4m field of view.

We capture the data from 1394 board cameras and then compute the bit stream from a hypothetical motion detector network. This allows flexible experimental design and the collection of validation data,

2 Model

We employ a composite Hidden Markov Model structure that is well-suited to representing a multitude of activities in a unified framework.

The model parameters, λ, are estimated from the data with a fast, hierarchical method[2].

The model presented here has 124 individual paths learned from three weeks of data consisting of 2.7 million non-zero motion vectors: 17D binary vectors, one bit per camera at 7.5Hz.

3 Entropy

Each symbol in the database was compared to each of the 124 HMMs. The likelihoods $p(O|\lambda)$ were employed two ways:

1. The maximum likelihood model labels the data point as potentially being a sample from that model.

2. The entropy of the sample is computed as

$$\sum -p(O|\lambda) \log \left(p(O|\lambda)\right)$$

4 Novelty

This entropy is a measure of how well the observation is being modeled. The canonical example of a particular gesture should have a high likelihood for the given model and low for the others, and so have a low entropy. For a sample that is not well explained by any particular model, the entropy will be high. These are taken to be atypical, or unusual events[1].

5 Results

We captured the observations of 17 motion detectors for 3 weeks in the $175m^2$ experimental space. That dataset consists of over 13 million observations, of which 2.7 million contain some motion information. We clustered that data into 124 Hidden Markov Models. The goal of the work was to recognizes elevator approaches, so half of the models were restricted to learning a subset of the

1800 labeled elevator-related events. The other 62 models were clustered out of a subset of the 2.7 million non-trivial observation sequences.

In this section we show samples of just a tiny fraction of those 2.7 million sequences. First we will show real example traces that are the best match for individual cluster models. they can be thought of as the exemplars of their clusters. Specifically, they have a high likelihood of being generated by their given model, and a low entropy, meaning that they have uniformly low likelihood of being generated by the other models.

Then we will pull out the most atypical sequences: those that have a high entropy: meaning that there is no clear, single model that explains these sequences well.

It is important to remember that while these plots show real validation data plotted on a map, the algorithms only get motion data from $3x4$ meter receptive areas, and have no model whatsoever of the geometry of the space. These plots are presented for the reader, but have very little do do with the bit streams that were used to generate the models, or the probability and entropy measures that helped select these sequences for review.

It is also important to remember that these sequences were ranked by the algorithm in to most typical to least typical. A human has sorted through the top 200 sequences at each end of the spectrum (out of 2.7 million, remember), and sorted them into meaningful categories by hand. This was done to help the reader (and the author) successfully understand significance of the data.

Figure 1 illustrates that a significant portion the recovered clusters explain transit motions through the space: walking from one place to another. These plots (as well as those below) are color coded: red from the beginning to blue at the end. They suggest that perhaps the clustering window was too short: that we should have possibly allowed for longer sequences to be considered to catch whole transits in their entirety. On the other hand, these clusters represent the most highly recurrent parts of the larger transits: and so make up an alphabet of transits through the space. Exploration of this trade-off between holistic gestures and atomic components is left as future work.

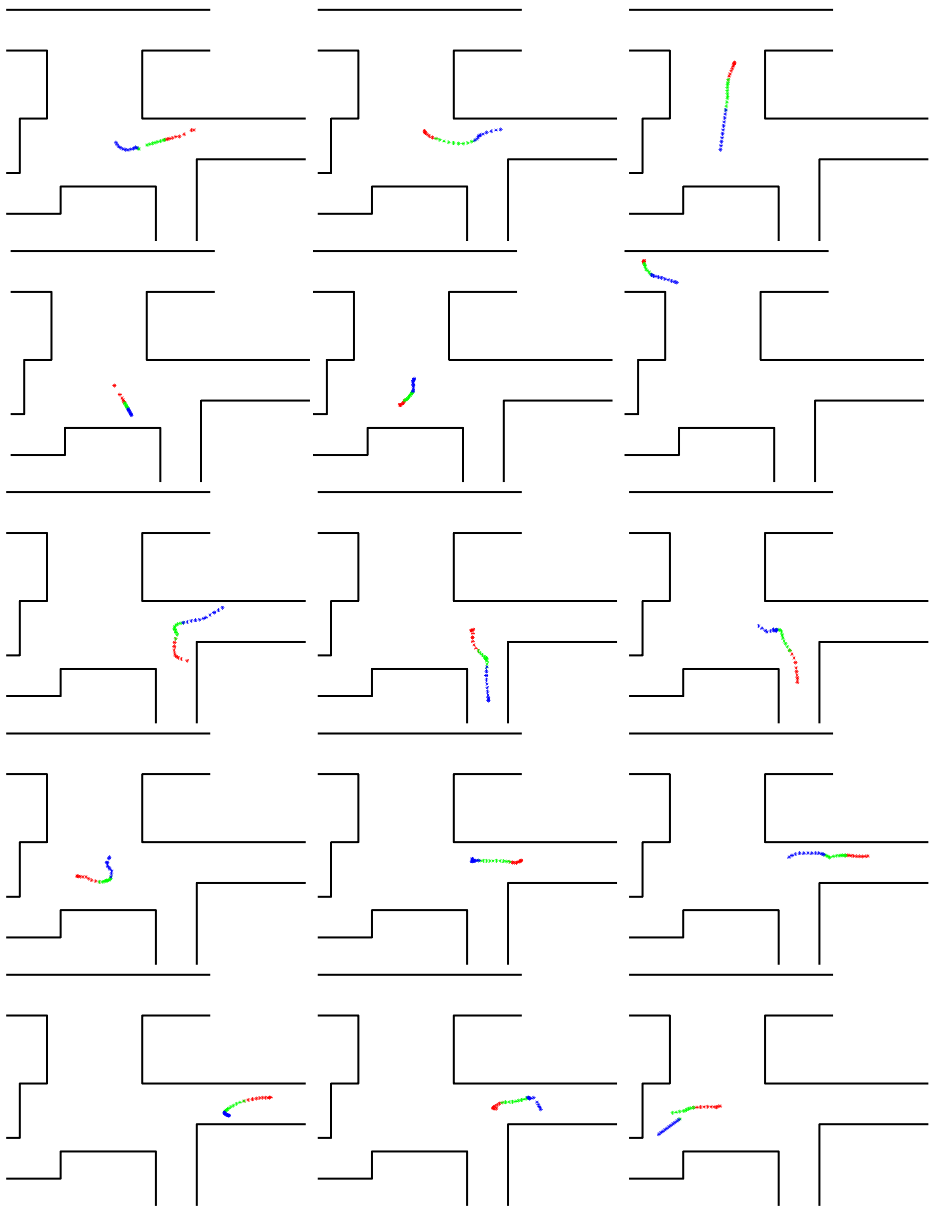

Figure 1: Typical behaviors that seem to capture transitions through the space.

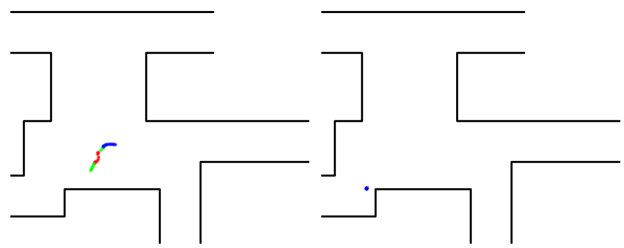

Figure 2: Typical interactions with elevator buttons and locked doors.

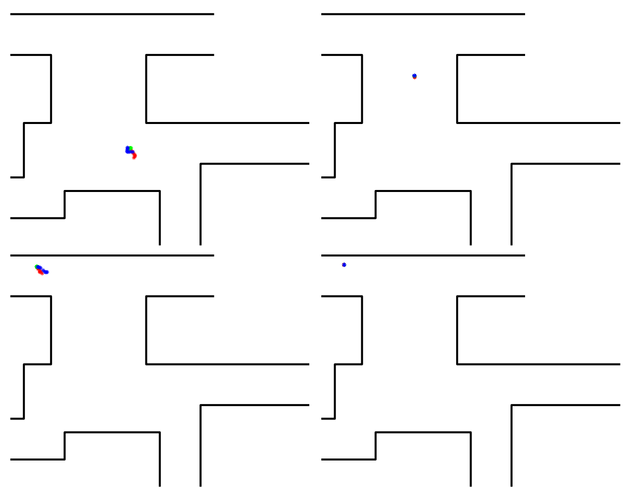

Figure 3: Typical styles and locations for loitering.

Figure 2 shows a pair of clusters that seem to be explaining the common behavior of interacting with buttons and security devices. The left sequence shows someone walking forward, pressing the call button and stepping back to wait for the elevator. opening doors that are typically locked. The right sequence shows someone interacting with the security device on a door that is always locked. See Figure 6 for sample traces that involve interacting with locked doors that are typically unlocked.

Figure 3 shows loitering behavior. There are some parts of the space where loitering is expected: in front of the elevators, in the reception area, and near the kitchen area.

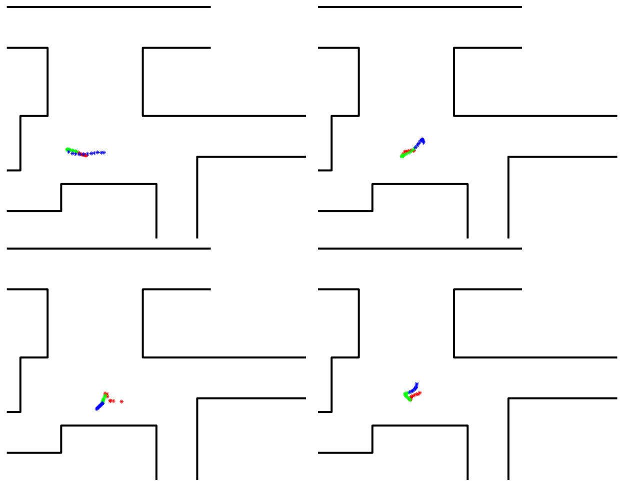

Figure 4: Typical pacing in front of the elevator.

Figure 4 shows an alternate way to wait for the elevators: pacing. Pacing is common enough in the experimental population that the system assigned clusters to the behavior. It is interesting to note that several examples of significantly more energetic pacing ended up in the top 200 atypical samples list despite the existence of these cluster models. See those atypical samples in Figure 7.

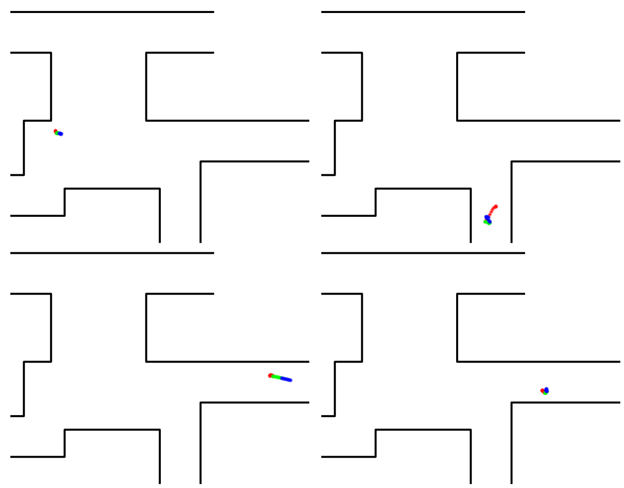

Figure 5: Atypical places to loiter.

Figure 5 illustrated the first of the atypical samples traces. Not that these are not exemplars of any particular cluster: they are samples that were poorly explained by all the clusters as a whole. These particular traces show people loitering in places where loitering is not expected by the models.

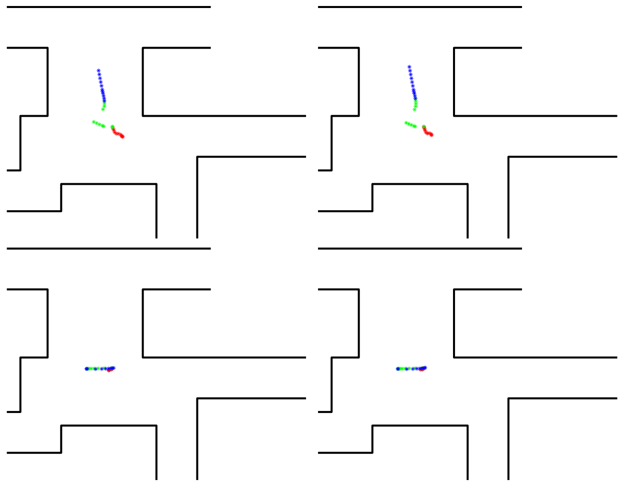

Figure 6: Dealing with a door that is typically open.

Figure 6 illustrates the case of people having to negotiate the front door between the elevator lobby and the reception area when it is locked. This door is usually unlocked and propped open during business hours. It is closed and locked at night. The plots show both people unlocking and opening the doors and well as people pacing outside the door, waiting to be let in.

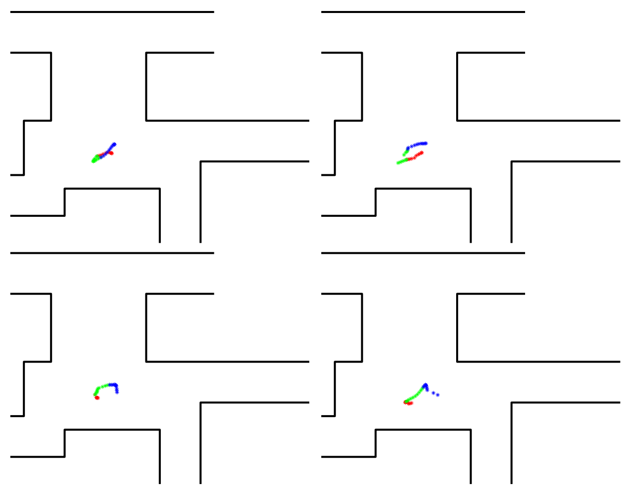

Figure 7: Atypical (energetic?) pacing in front of the elevator.

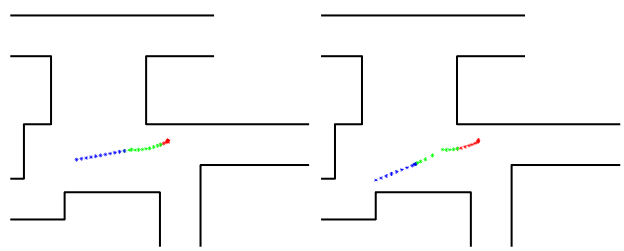

Figure 8: Atypically fast transits.

Figure 7 seems to show people pacing while waiting for the elevator. It is unclear exactly why these samples are not well explained by the elevator pacing models from Figure 4, except that they seem to be more energetic.

The dynamic time warping that is part of the Viterbi decoding process will usually ignore differences in timing, however large departures from typical timings can still be detected. Figure 8 seems to capture samples of people transiting the space rapidly.

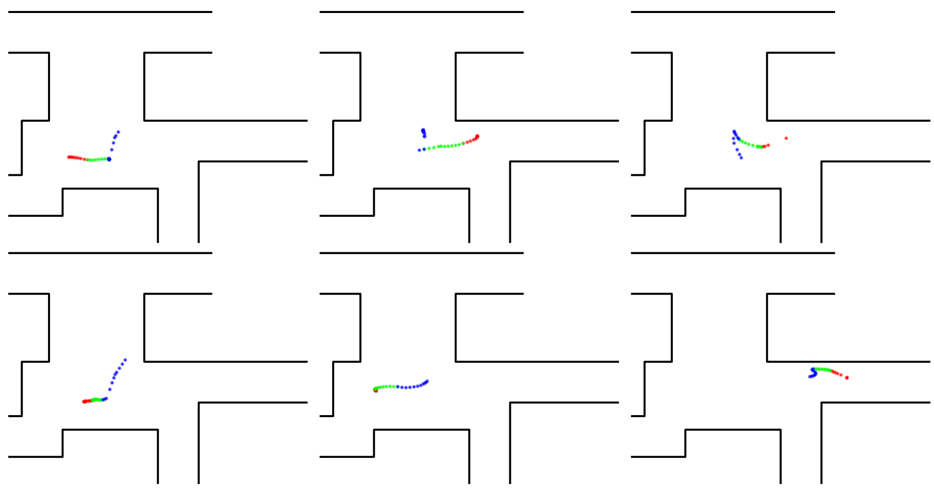

Figure 9: The system seems suspicious of indecidion.

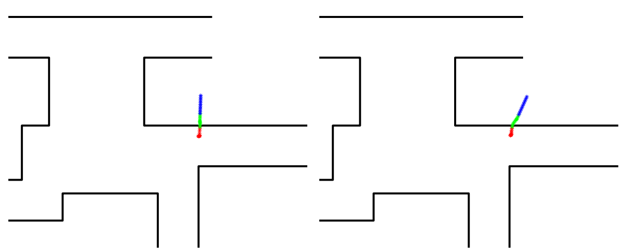

Figure 10: Some atypical gestures are actually illusions caused by sensor failures. These examples were generated by system initialization transients.

Figure 9 shows sample traces of people changing their minds: either leaving an area where they were loitering or changing direction in mid-course. Presumably people loitering near the elevators and then entering the elevators is typical, but loitering and then leaving some other way seems to be atypical.

Figure 10 shows some illusory sample traces that are caused by transients in the system during initialization. Obviously a real system would automatically drop these kinds of samples as part of the initialization sequence. The research system had temporary failures throughout the experimental run that injected this kind of noise in to the data stream.

6 Conclusion

These results offer compelling evidence in support of the notion that subtle models of building-scale behaviors can be captured by networks of impoverished senors. Much work remains to be done to find the optimal sensor node configuration, sensor modalities, and modeling methods. It is also important that these results be validated with data from larger areas and from different use categories.

References

[1] Christopher M. Bishop. *Neural Netowrks for Pattern Recognition.* Oxford University Press, 1995.

[2] David C. Minnen and Christopher R. Wren. Finding temporal patterns by data decomposition. In *Sixth International Conference on Automatic Face and Gesture Recognition.* IEEE, May 2004. also MERL Technical Report TR2004-054.