Activity Recognition in a Dense Sensor Network

James Howard
jahoward@mines.edu

William Hoff
whoff@mines.edu

Colorado School of Mines
Center for Automation, Robotics, and Distributed Intelligence

Introduction

- Growing need to detect threats to security
  - Crime, terrorism
  - Accidents, disturbances
  - Fires, medical emergencies
- Local governments and private organizations
  - E.g., Minneapolis skyway
  - Need to monitor walkways covering 80 city blocks
- Federal interest in domestic security
  - E.g., Joint Terrorism Task Force (JTTF)
  - 100 field offices, operating budget of $6.4B

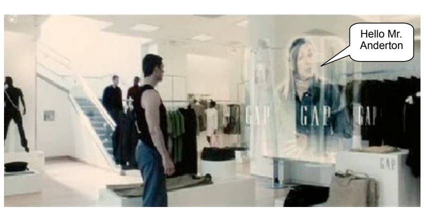
One approach – identifying people

- Can use to determine unauthorized entrance
- For private buildings, could use
  - Biometrics
  - Voluntary self position reporting (e.g., RF UWB badges)
- In public areas, could use video
  - Raises legal and privacy concerns
  - Potential for identification and localization without probable cause

Palm vein scanner

Photo traffic camera

Interactive advertising

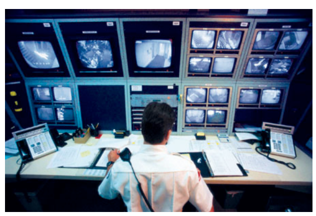


Alternative – detect unusual behavior

- Could detect unusual behavior of groups or individuals
- The key is to know when something is occurring that needs a response (e.g., to investigate further)
- Problems: large areas to monitor, lots of people, long periods
- Need to learn what is typical
- What is unusual depends on context

Video – most common approach

- Can cover area with lots of video cameras
- Use of video surveillance is growing rapidly
  - Equipment sales $8B in 2010, growing 10% per yr[1]
  - 4 million surveillance cameras in UK alone
- Problem – monitoring all that data
  - Lots of research in automatically detecting activities from video, but still in research phase
  - Current state of the art solution is to have security guards monitor



- Cost of monitoring large areas is prohibitive
  - A guard can only watch a maximum of about 15 monitors[2]
  - If a $30K/year guard[3] is responsible for 15 cameras, each camera is $2000/year

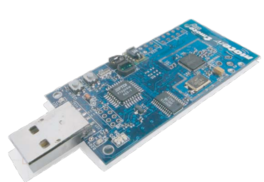
1 Simon Harris, "What's in store for VCA now and the near future," Video Content Analysis Conference, London, June 2007

2 Herman Kruegle, "CCTV Surveillance: Video Practices and Technology," Butterworth-Heinemann Press, 1995.
3 US Bureau of Labor Statistics, www.bls.gov

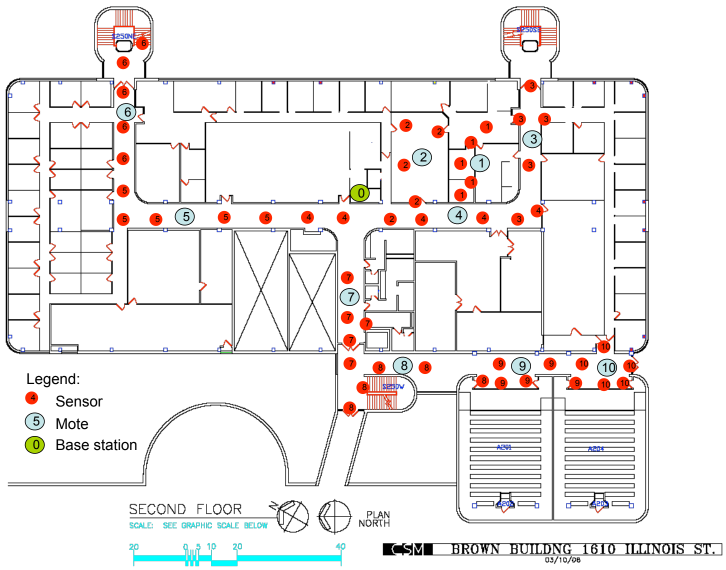
Our approach

- Deploy simple sensors in a network of small wireless computers (motes)
  - Example: infrared motion detectors which just detect a binary (yes/no) signal if a person is there
  - Gather lots of data, over a long period
- Automatically configures communications and learns layout
- Automatically learns to recognize and model typical activities
- Detect activities that are not typical
  - Only if something is sufficiently unusual is it necessary for a security guard to investigate
  - This could be as simple as directing a video camera to observe the area
- A single person could effectively monitor hundreds of locations

KUBE TR257 motion sensor



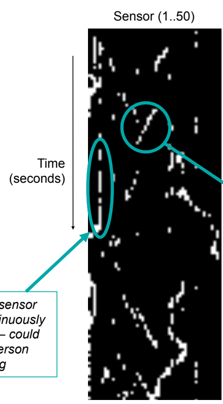
Tmote Sky mote



Sensor time plots

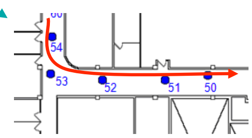
COLORADO SCHOOL OF MINES

- Sensor hits are stored in a matrix H(time,sensor#)
  - Columns are sensors
  - Rows are time intervals

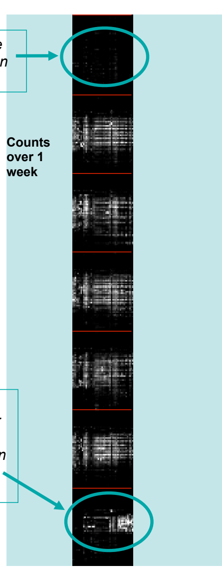


Same sensor is continuously active – could be a person loitering

- Local movement patterns appear as structures in the matrix
  - Eg., a person walking down the hall causes consecutive hits in adjacent sensors

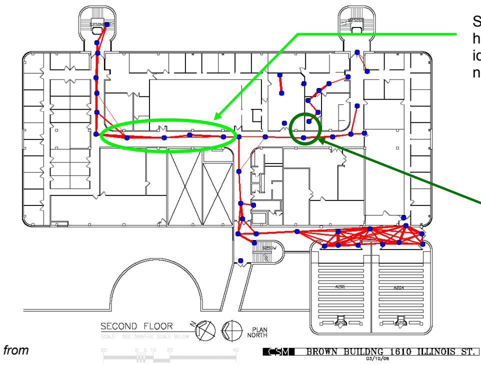


A sequence of adjacent sensor hits – could be a person walking down the hall



Neighbors of sensors

Sensors with high correlation scores comprise the neighborhood of a sensor



Sensors along hallways are identified as neighbors

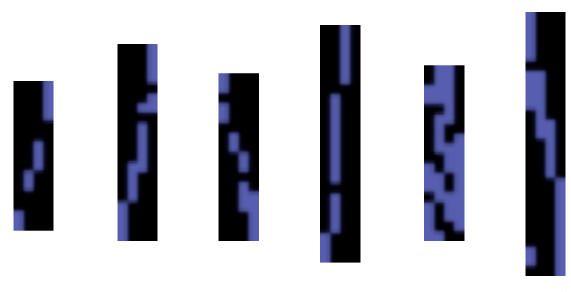
Note that these are not neighbors (no pathway between them)

Line thicknesses correspond to the value of the correlation coefficient

Based on counts from November 26 - December 1, 2007

Detecting local activity patterns

Example activities classified from trained Hidden Markov Models.



Most Probable — Least Probable

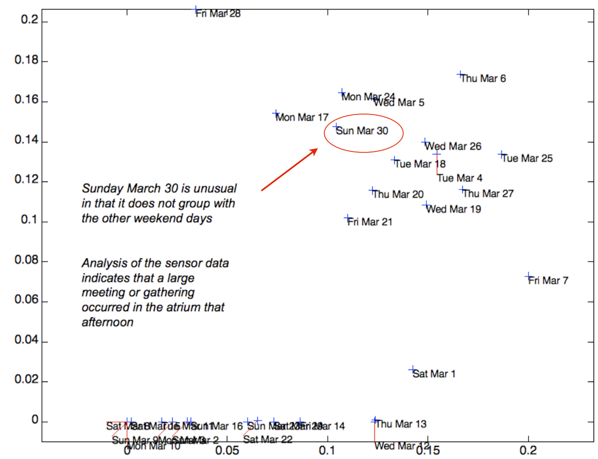
Data sampled from November 19th through December 12th from 2:50 pm to 3:00 pm

Detecting global activity patterns

- Use a technique derived for document analysis known as Probabilistic Latent Semantic Analysis
  - Models = Words
  - Time instances = documents
- Allows us to express a document in terms of the latent classes (mixture of models) that best describe it
- Useful for the determination of context

Detecting global anomaly example

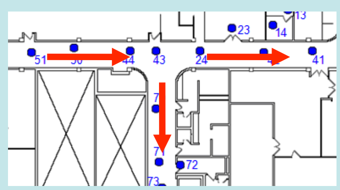
COLORADO SCHOOL OF MINES



Future work

COLORADO SCHOOL OF MINES

- Context driven local decisions
- Hierarchical decisions
- Expansion of the network to allow for different sensor types and different data configurations
- Integration with video



Time ordering of local events: Is the "walk right" event in the left hallway followed by a "walk right" in the right hallway, or by "walk down" in the lower hallway?

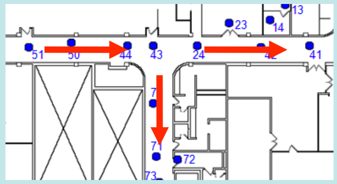
COLORADO SCHOOL OF MINES

Thank you!

We gratefully acknowledge the support of Lockheed-Martin Corp. for this research

Context-based Prediction

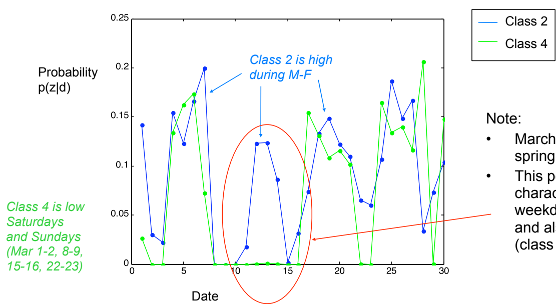
- So far we have shown how to analyze the data by looking at the histograms of local event patterns
- However, this does not take into account the time ordering of events – this might be helpful in recognizing more specific activities
- We are developing methods to automatically learn to predict events, based on recent history
  - Prediction can be used to detect anomalies
- We are using context (derived from the latent classes) to improve prediction accuracy



Time ordering of local events: Is the "walk right" event in the left hallway followed by a "walk right" in the right hallway, or by "walk down" in the lower hallway?

Results from daily analysis

- Latent class 2 seems to correlate with "weekday"
- Latent class 4 seems to correlate with "weekend" or "holiday"



Note:
- March 10-14 was spring break
- This period had characteristics of weekdays (class 2) and also weekends (class 4)

Probabilistic Latent Semantic Analysis

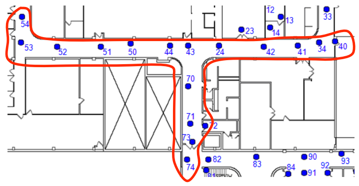
- Probabilistic Latent Semantic Analysis (PLSA) relates words *w* and documents *d* to a latent (hidden) topic space *z*
- Relies on counts of words (histograms) in documents
- It uses the iterative expectation-maximization (EM) algorithm to maximize the log likelihood function *L* of the joint probability *p(d,w)* to determine *p(z), p(w|z)*, and *p(d|z)*

$$L = \sum_{d,w} \log p(d,w) \qquad p(d,w) = \sum_{z} p(w \mid z) p(d \mid z) p(z)$$

- Our application:
  - "words" are the local activity patterns found by clustering (291 total)
  - "documents" are time intervals (such as hours or days)

Example Analyses

- Restrict analysis to 17 sensors along the hallway (for speed)
- Analyzed data for one month
- Two analyses:

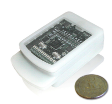


1. Daily
- Divide period into days
  - March 1-30
  - Each document is a day
  - 30 documents
- Assume 8 latent classes

2. Hourly
- Divide period into hours
  - April 1-30
  - Each document is an hour
  - 720 documents
- Assume 10 latent classes

Future Work

- Expand from 50 to 150 PIR sensors
- Replace 10 Tmote Sky motes with 30 SunSPOT motes
  - Much faster, more memory
  - Uses Java instead of TinyOS
- Add other types of sensors:
  - Sound level
  - Light intensity
  - Sonar
- Integrate with video
- Modify algorithm to work on a more global scale.



SunSPOT mote

Wireless webcam in display case