Automating Public Building Security

James Howard
William Hoff
Colorado School of Mines
{jahoward, whoff}@mines.edu

The problem of building security is a very old and extensively studied problem. Everything from complex locking mechanisms to motion sensors are meant to ensure intruders do not enter a building without cause, but what about the case when people are supposed to be there. What mechanisms are in place to ensure the safety and security of the building when it is open to the public? Due to the nature of security cameras relying on active watching during an attack, they rarely play a proactive role in security and are instead typically used for criminal forensic analysis after a crime has been committed. In cases of public spaces, how do we ensure an adequate level of security? Current implementations of security systems in public spaces are quite often inadequate.

Take for instance the Minneapolis skyway. A climate controlled pedestrian walkway connecting the buildings of downtown Minneapolis at the second or third floors. The walkway is over seven miles in length and covers nearly 80 city blocks. Security in this system is at the responsibility of each building connected to the skyway, which typically comes in the form of a security guard who shares his time behind a desk watching security cameras and patrolling the hallways.

Given its public access, high popularity and its connection to almost every major building in Minneapolis any would be attack against the city of Minneapolis may utilize this infrastructure. Criminals or terrorists would likely spy on targeted buildings determining the best avenue for attack. How is a security guard, given the limited amount of resources available to him able to discern such surreptitious spying?

To solve such a problem, we propose a system using low cost sensors and wireless motes to monitor large spaces often containing a large number of people and provide feedback about the activities within that space. Our system is not limited to buildings and could be applied to external public spaces such as sidewalks or open markets where manual analysis of threats maybe impossible due to the sheer size or timescale required for proper threat determination. Nor is our system limited by discovering actions such as spying in the form of loitering. The proposed system should be capable of discovering any anomalous activity such as riots, fires, and perhaps even muggings or heart attacks. We view the system as something used by a security guard on top of typical security cameras. We presume the probability of a guard watching a certain camera at the time of interest is low and thus we propose a system that would alert the guard to a potential problem letting the guard determine the threat and the appropriate response.

We believe the market is ripe for exactly this type of security need. Consider that since September 11, 2001 under the direction of the then Federal Bureau of Investigation Director Robert Mueller the Joint Terrorism Task Force (JTTF) has grown from 35 field offices to 100 and has an annual operating budget of 6.4 billion. The increase in size (JTTF) is primarily to increase our probability of attack prevention and not increase the efficacy of our response. Clearly domestic terrorism security is of paramount importance to the United States. Given its possibility as a preventative measure, our proposed system fills a need not currently covered by any existing security system. We look to local cities and the United States Government as potential customers.

Existing work:

Current computer vision systems operating on security camera data may be able to perform the type of activity recognition necessary to achieve our security goals, but the cost prohibitive nature of security cameras would greatly cutdown on coverage space. A typical indoor security camera costs between 40 to 60 dollars and cameras usable in out door spaces or in low light environments can begin to approach the 250 dollar range [1]. This is not including other monitoring and recording equipment. Compare this to the five to 10 dollar range of the passive infrared, light and sound sensors that are readily implementable in our system. While our system does not propose to remove security cameras, it may allow for the removal of some overlapping locations.

Consider also the legal ramifications of such a camera based monitoring system for public spaces. Due to privacy issues, any automated security system which utilizes a video first approach to security may improperly be applying privacy law and a determination could be made of the where abouts of an individual without the required probable cause (or in the case of terrorism related cases, reasonable suspicion). Take for example the case of automated traffic cameras. Based on the possibility that they may be used by law enforcement to determine the location of people in public spaces without proper legal authority, the American Civil Liberties Union has expressed public concern possibly leading to a legal case. Already it is necessary for traffic cameras to blur passengers in vehicles found in violation of traffic law. To avoid the necessity of waiting for jurisprudence considering the outcome of such a case, we believe that an anonymous sensor first approach is warranted.

There are a limited number of research laboratories [2, 3, 4] working on similar activity recognition problems using a sensor first approach, but they appear not to focus on the same large area data sets that we do. Much of the existing research in activity classification either deals with a small number of people, a limited area, short tracking duration or uses video for its content. Also much of the existing work requires a large amount of human interaction and intuition about the environment to perform the proper modeling.

Our Approach:

Due to the potential for human error and to reduce setup time, we look to implement a system that dynamically determines the topology of our environment. We then use the topological information to discover local spatial neighborhoods around each sensor. These spatial neighborhoods are then used to create local patterns of activity which will be used for the determination of anomalies. We finally analyze the distribution of these local patterns to determine changes in activity patterns and predict the outcome of the activity.

Experimental Testbed

Using the Colorado School of Mines Brown Building as our testbed, we have placed 10 wireless Tmote sky's and 50 passive infrared sensors though out the second floor (Figure 1). This location provides a good experimental test bed as it contains offices, class rooms and research laboratories. Data is sampled every second across the network and transmitted to a central database. Because data is so simple compared with traditional video, years worth of data maybe recorded and stored on a relatively inexpensive computer compared to the expensive storage methods required for video surveillance.

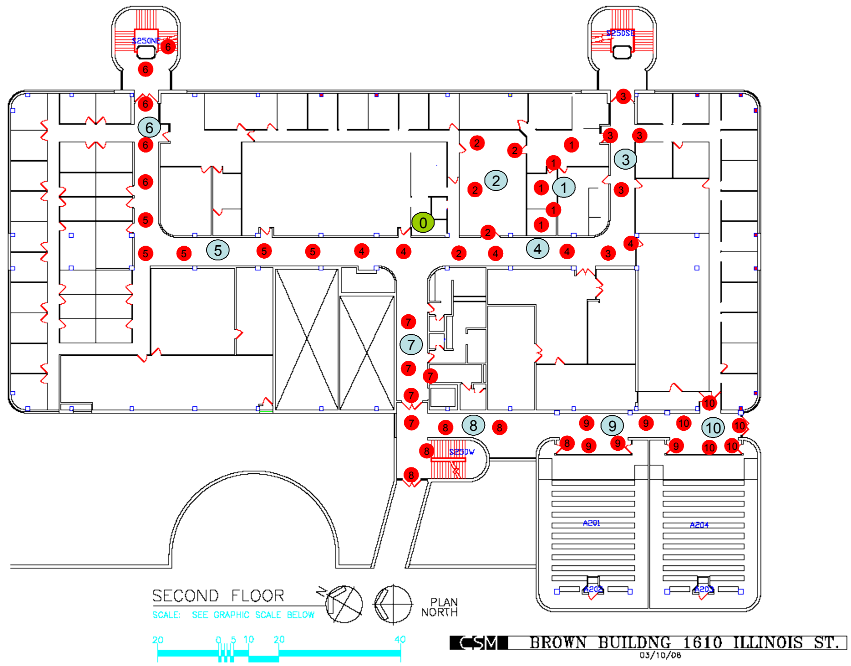

Figure 1: Mote and sensor locations

As the final piece of our Lockheed Martin funding (which ended in January 2009) 10 Sun SPOT motes were purchased with the goal to expand the network to the third floor. The extra data collected from this expansion is expected to give us a better idea of how different environments affect motion.

Determining Local Neighborhoods:

As stated before, because we are looking to keep our system with as little human setup cost as possible, we look to build a global sensor topology automatically. To do this we look to Gabriel Graphs [5] based on the inverse cross correlation score derived from the sensor activations (Figure 2). Not only does this approach yield a graph which allows for the creation of local neighborhoods, but it can lead to insights into the sensed environment that are not readily apparent and may be over looked by human creation of sensor neighborhoods.

Consider figure two, which shows a sensor topology derived by a single day of readings in February 2007. Notice how most of the spatially close sensors are also neighbors in this activation space. There are a few notable exceptions however. The link from sensor number 44 to sensor number 84 for example does not seem intuitive. Indeed it may not have an obvious explanation, but it is plausible that this link exists due to both sensors being about the same relative location from a major classroom. Both rooms may have been on similar class room schedules thus leading to this strong correlation.

Of course this relationship between 44 and 84 will only be valid in instances where classrooms are letting out. For a given time of day the network topology will certainly change. By allowing this change, we increase the networks capability to handle local variances in the environment.

Determining Local Activity:

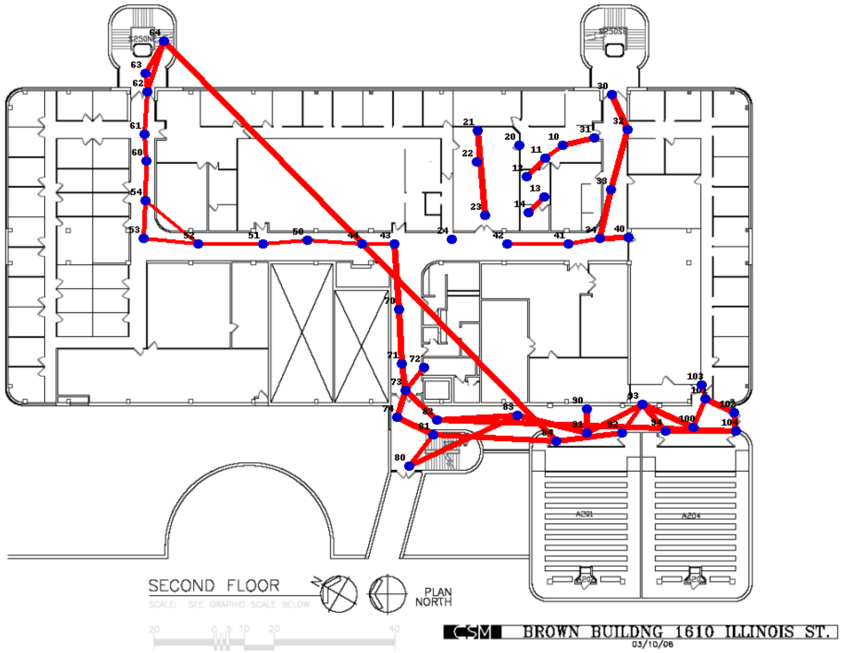

Figure 2: Gabriel Graph using cross-correlation scores

Given our sensor topology, we can then begin to define local activities around individual sensors. To calculate these local activities we implement a K-means clustering algorithm to determine the the most likely set of activities for an individual sensor. Activities are represented by a matrix as shown in figure 3. The x axis of the matrix represents the neighborhood of sensors and the y axis time. The color of each pixel within the matrix is representative of activity at that sensor at that time. Black represents no activity while white represents strong activity. Work is proposed utilizing the principal of minimum description length to determine the optimal number of clusters for each sensor. We are also looking to representing local activity as a Hidden Markov Model.

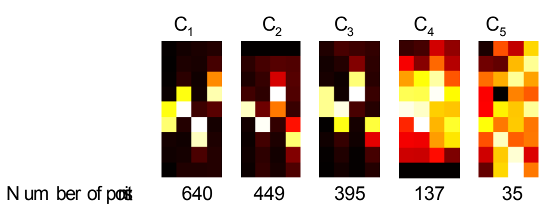

Figure 3: Representation of five actions for a given sensor

Determining Context:

If we return to the Minneapolis skyway scenario clearly there are going to be spatial variations in activity. These variations should be captured by our per sensor local activity representation. For example, it makes sense that people will loiter longer in food court locations than in front of an office building. Extending this to the temporal domain, it should also come as no surprise that people will loiter longer during certain times of the day than others. Behavior and environment are not independent variables and within an environment time and space are important when determining what is anomalous and what isn't.

Our system implements a probabilistic latent semantic analysis (6) of the distribution of local activities to determine when these temporal changes occur. We then use this information to create

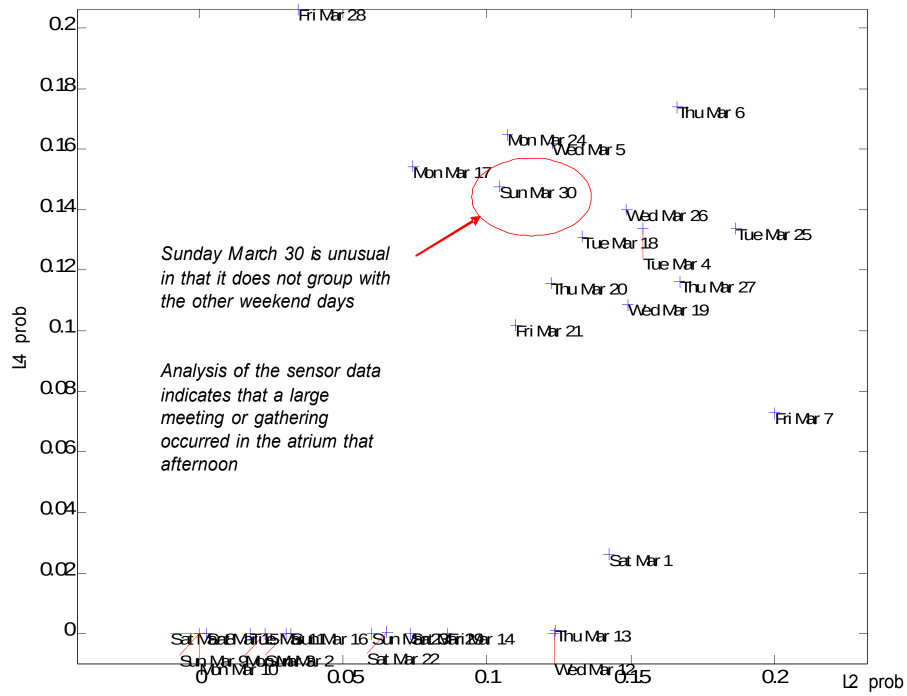

Figure 4: Latent class 2 vs Latent class 4. Shows an example of anomaly.

a set of latent classes which each become experts at a description of a different part of our data. Looking at figure 4 we can see that for the problem of determining action for a given day of the week, latent class two and latent class four seem to form a nice set of clusters for determining the day of the week. Points in the upper right hand side of the graph tend to represent weekdays while points in the lower left represent weekends.

Notice the red circled point is a Sunday, but has happened to be clustered with other weekday points. We reference this point to be anomalous and flag it for further analysis. Looking at the raw data for this day shows it to have a large gathering in the atrium that afternoon. This is a short example of how our system has the capability to discover anomalies.

Although the system is still in flux and many of the automated features we envision are not yet implemented, we believe the system with additional work could assist the United States in becoming a more secure nation will still maintaining the absolute privacy of every citizen.

References:

1. Super Circuits Security Cameras, http://www.supercircuits.com/Security-Cameras/
2. S. S. Intille and A.F. Bobick, "Recognizing Planned Multiperson Action", *Computer Vision and Image Understanding*, pp. 414 – 445, 2001.
3. S. Gong and T. Xiang, "Recognition of Group Activities using Dynamic Probabilistic Networks", *Proc. of the Ninth IEEE International Conference on Computer Vision (ICCV'03),* 2003.
4. Wren, C.R. and Tapia, E.M., "Toward scalable activity recognition for sensor networks," *Proc. of Second International Workshop on Location and Context Awareness* (Springer-Verlag Lecture Notes in Computer Science Vol.3987), pp. 168-85.
5. Gabriel, K.R. And Skokal, "A new statistical approach to geographic variation analysis", *Systematic Zoology*, 1969.
6. Hofmann, T., "Probabilistic Latent Semantic Indexing", *Proc. of the Twenty-Second Annual International SIGIR Conference on Research and Development in Information Retrieval*, 1999.

None of the above work is bound by any intellectual property protection.