Proc. of 10th International Conference on Information Fusion, Quebec City, QC, Canada, 2007, Jul. 9-12.

Recognizing Wide-Area and Process-Type Activities

Raymond D. Rimey
Radical Innovation Technology Center
Lockheed Martin IS&S
Denver, Colorado, USA
raymond.d.rimey@lmco.com

William Hoff
Engineering Division
Colorado School of Mines
Golden, Colorado, USA
whoff@mines.edu

Jae Young Lee
Dept. of Mathematical and Computer Sciences
Colorado School of Mines
Golden, Colorado, USA
jaelee@mines.edu

Abstract – *New methods are presented to model, visualize and automatically recognize wide-area activities, which essentially are activities that span large areas (such as a facility or urban neighborhood) and that usually span long time intervals (such as hours and weeks). We introduce the no-go topology method and the chokepoint-observation interaction method, and then show how new algorithms can be built on them to recognize a category of wide-area activity, called process-type activities. Experimental results are presented for recognizing a manufacturing process observed using persistent GMTI sensor data. Then we present experimental results illustrating how an interesting activity can be detected as a deviation from a learned wide-area normalcy model, and how new wide-area activity patterns can be discovered using simple visualizations of the results.*

One objective of this paper is to demonstrate that it is theoretically possible to recognize wide-area and process-type activities in built-up environments using GMTI data. The results presented here use somewhat ideal sensor data (small positional error ellipses, continuous GMTI observations, repetitive activities) and our approach is to move toward realistic parameters in operational situations (larger error ellipses, fewer observations, figuring out how to exploit additional kinds of activities).

Keywords: Activity recognition, Built-up area, GMTI, Motion pattern analysis, Persistent surveillance, Process-type activity, Urban operations, Wide-area activity.

1. Introduction

Recognizing activities using various kinds of sensor observations is a problem of increasing interest to the military and the intelligence community. Intelligence analysis is expanding from looking mostly at isolated snapshots in time to looking at things that happen continuously over time. In the past, many intelligence tasks involved determining "what is where", detecting a relatively isolated object within a relatively large spatial area. Some tasks have involved detecting a single large-scale event, for example, the vehicles have begun an attack, or the facility has initiated a test. Often such events are inferred by the analyst from one or two snapshots in time. Sometimes sensors are used to continuously track vehicles over time, but typically this provides only a real-time "what is where" answer. More sophisticated time-based analysis was theoretically possible, but historically there was not a strong driving need for it.

Many of the new problems faced by intelligence analysts today involve time-based patterns and, in particular, activities. The area of operations today is likely to be an urban environment, such as a city or town, or other built-up environments, such as a multi-building facility. Typically a large number of people and vehicles inhabit these areas. There are many vehicles that can be detected, they are not isolated, and they all look somewhat similar. Instead of detecting a relatively isolated vehicle, the problem involves detecting an activity that is mixed in with many other activities within the same space-time volume. The activity involves the movement and interaction of people, vehicles, equipment and materials. In crude terms the problem is to determine "what are they doing" instead of "what is where". For example, a tactically-oriented analyst may need to distinguish the movement pattern around an insurgent safehouse from the movement pattern around a laundry service, and a strategically-oriented analyst might detect preparations leading up to a test from the associated activity movement patterns within a facility.

This paper addresses *wide-area activities*, which means an activity that spans a relatively large area, such as an urban neighborhood or a large facility. Wide-area activities typically involve a relatively large number of agents (vehicles or people), for example tens of vehicles. And the activity typically requires a large amount of time to perform, partly because all the agents need time to traverse the large area. This is admittedly a loose definition for a wide-area activity, however it encompasses many problems of interest to the military, problems which have also not been addressed to date.

Recognizing activities requires observations over time. The recognition problem can be framed using probabilistic models of activities and evidence. Recognition performance is a function of the discrimination power of the set of observational evidence relative to the structure of a specific activity to be recognized. Observational evidence will generally come from multiple sources, meaning multiple types and instances of sensors. Continuous motion tracking of all moving agents and objects is the best possible type of observation data for automatically recognizing time-based movement patterns, but it is rarely available in built-up areas. Less frequent (or lower quality, or lower information) observations are more realistic but in fact can be sufficient in some cases for good recognition performance.

Happily, sensors are proliferating. The term *persistent surveillance* refers to the ability, of a single sensor system or a combination of them, to observe an area of interest continuously. A persistent surveillance sensor system provides the type of observation data needed to recognize activities. Persistent surveillance capabilities currently exist (though they are not yet common) and will become increasingly available over time. Examples utilize GMTI, UAV video, ground-based sensor networks, frequent image collections, and AIS ship tracks. Sensors also include the soldier-as-sensor concept, for example the set of observations made by all patrols in an urban neighborhood over the last month.

Urban and other kinds of built-up environments provide challenges to persistent surveillance sensor systems, often resulting in low quality observations. For example, multi-path reflections effect the quality of GMTI, occlusions effect all sensors, and the density of moving agents makes it difficult to track an individual. Urban and built-up environments contain many activities that all overlap in space and time, which presents an additional challenge. Activity patterns of interest must be detected within massive clutter, where the clutter consists of other activities being performed by other vehicles in the same space-time volume as the activity of interest.

This paper makes the following contributions:

- A new category of wide-area activity is identified, *process type activities*, which involve structured repetitive material transfer movements. Many problems of interest to the military and intelligence community include processes.
- The *no-go topology* method is introduced for interpreting (constraining) noisy position observations (e.g., GMTI) in large built-up areas.
- The *chokepoint-observation interaction* method is introduced for detecting events from noisy position observations near a small local area.
- Methods are introduced for estimating the parameters of a process-type activity model and for identifying the most likely pre-existing activity model for an observed dataset. We demonstrate that wide-area activities can theoretically be identified in built-up environments using GMTI data.
- Experimental results show how much positional error can be tolerated in GMTI observations while still obtaining good discrimination of activities.
- A combination of clustering methods and Markov models are used to represent normal patterns of activity in a wide area. Experimental results illustrate how an interesting activity can be detected as a deviation from the normalcy model, and how new activity patterns can be discovered using simple visualizations of the results.

This paper is organized as follows. Section 2 summarizes related work. Section 3 introduces the no-go topology method. Section 4 introduces the chokepoint-observation interaction method. Section 5 introduces our model for a process-type activity, associated estimation and classification methods, and experimental results. Section 6 presents our work on wide-area normalcy deviations and visual pattern analysis. Section 7 summarizes the paper.

2. Problem Dimensions and Related Work

An *activity* involves multiple agents, manipulating resources, interacting with each other, interacting with the world, and working over time to achieve an objective. Following are some ways to characterize the space of all activity types.

<u>The number of agents (people or vehicles) performing the activity:</u> Most existing work (video) addresses 2-5 agents. A small amount of work (GMTI, instrumentation for urban planning) involves 1000's of vehicles treated as flows or statistics. Our work addresses many 10's of vehicles.

<u>The extent of the area where the activity is performed:</u> Some familiar examples of the spatial extent are a room, a street intersection, a sports field, a city block, a facility, a city neighborhood, a city, a county. Most existing work (video) addresses extents as big as a street intersection. Most existing GMTI work addresses convoys moving across road networks in a country-sized area. Transportation analysis involves country-sized areas that span airplane, shipping, train or road networks. Our work addresses facility-sized and neighborhood-sized spatial extents.

<u>The duration of the activity:</u> Some example durations are 1 minute, 5 minutes, 20 minutes, 2 hours, 1 day, 5 days, weeks, months. Most existing work involves activity durations of a few minutes. Our work addresses activity durations of about 1 day.

<u>The regularity of the activity:</u> The range of regularity includes: once, sporadically, regularly, continuously. Most existing work involves observing an activity once. Our work addresses continually repeated activity.

<u>The structural complexity of the activity:</u> There are many kinds of structure. One way to describe structure is via “threading”, which spans from no-thread (a point event), a single thread (a single sequence), multiple threads (parallel and branching), and more complex structure. Another way to characterize structure is the number of agent-object interactions, the number of agent-agent interactions, and the number of agent-area interactions. Our work addresses multiple parallel threads, large numbers of agent-object interactions, small numbers of agent-agent interactions, and large numbers of agent-area interactions.

<u>Whether the core signature of the activity occurs in the geospatial, temporal or interaction domain:</u> This does not apply to all activities, however some activities are strongly

defined or identified by a core geospatial, temporal or interaction pattern.

Whether the activity is embedded within clutter (meaning other activities occurring in the same space-time volume) and the amount of clutter: Most existing work (video) involves little or no clutter. Our work addresses activity embedded in significant clutter.

The degree to which details of the activity depend on the contextual environment.

Whether the activity is of interest to the military: Most existing work on recognizing activities uses video and focuses on activities of near-term interest for commercial applications. Military needs are sometimes distinct. Our work addresses two new activity categories of great relevance primarily to the military: wide-area activities and process-type activities.

Most existing work on recognizing activities has involved video and addresses the opposite from wide-area activities: small areas, few agents, short times, and lower-level activities. Following are some exceptions of note.

A large urban traffic network is characterized using several metrics calculated from a weighted transition matrix in [3]. Similar metrics were applied in [2] to the simulated movement of 1.6 million individuals in Portland, OR. The simulator, designed for urban traffic analysis professionals, uses the actual transportation network in Portland, individuals modeled using census and DMV records, and common daily (single person) activities. Using GPS tracks of individuals in another city, [9] learned a hierarchical Markov model for transportation routines and could detect deviations from the routines.

Anomalous ship behavior was detecting from fused tracks (AIS ship transponders, radar detections, video) using rule-based and simple pattern recognition methods in [12]. Military maneuvers were detected from GPS tracks of hundreds of vehicles engaged in large force-on-force military exercises using ad hoc pattern recognition methods in [4]. A visualization tool is presented in [1] to study patterns of movement in virtual environments and games, so designers can better understand user behaviors and improve the design of the game.

A large body of work addresses recognizing activities in video. Typically, this work uses relatively high-quality tracks and hidden Markov models (HMM) and dynamic Bayes nets (DBN). For example, multi-agent activity (football plays) was recognized from video tracks of people in [6], using belief networks to detect individual agent goals and the overall activities (plays). Activities involving video tracks of pedestrians were represented as a series of events modeled with HMMs in [10]. A large body of similar work applying HMMs and DBNs to video has occurred in the last ten years. A video event representation language (VERL) was presented in [5]. An example was presented where symbolic VERL description of events are extracted from video of people and then events are organized into time sequence threads and then into branching/parallel threads. Motion patterns are extracted from long-term (months) outdoor video of people in [14], using co-occurrence matrices calculated from codebook index sequences derived from tracks of people in the video. Methods to estimate sources, sinks and transits from short tracks in video are presented in [13].

Person-object interactions during routine household activities were sensed using RF id tags on the objects, and the activities were recognized using HMM and DBN models in [11]. RF tags can scale to building-sized areas. Binary detection of movement was sensed using infrared motion detectors in a busy office building, and the activities were recognized using HMMs in [15]. Noisy positional readings from id badge sensors were constrained to the links of the Vornoi graph for an office area in [8]. Their tracker algorithm also utilizes the graph.

3. No-Go Topology

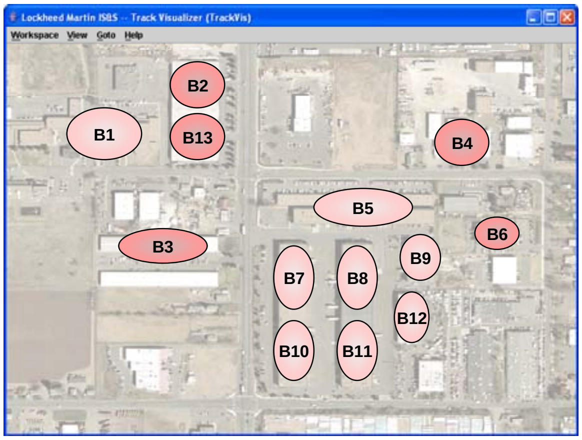
The idea behind *no-go topology* is that a built-up area is filled with “no can go” areas that a vehicle can not drive through, such as a fence, building, curb, ravine, stream. The no-go areas are typically obtained from maps, high-resolution downlooking imagery, elevation data, and possibly augmented by on-the-ground observations. The no-go areas define a topology and a graph structure, showing where vehicles *can* go, and this can be used to interpret movement evidence because the only possible movements are along the links in the graph. Figure 1 shows an example of the no-go areas and no-go graph for a facility.

Many activities involve specific sources and destinations for movements. Movements to/from a building are typically made to/from a portal, which is a doorway for a person or vehicle. If a vehicle is transporting materials to/from a building, then the vehicle is positioned right at a portal, whereas if the vehicle is simply transporting a person then the vehicle may be positioned in a parking area near the portal. New nodes corresponding to portals can be inserted into the graph. Vehicles moving between portals in the real world are equivalent to following a path between two portal nodes in the graph.

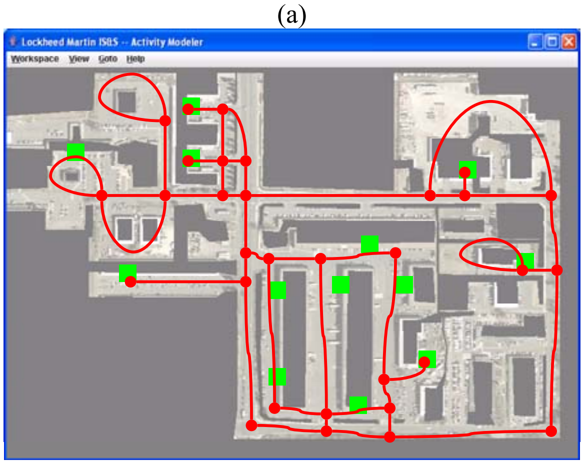
The no-go graph provides a constraint for the interpretation of movement evidence, for example to estimate frequencies of node-to-node transits from noisy movement evidence. No-go topology is an urban generalization of the classic technique of projecting GMTI dots onto road networks in the open countryside. Similar ideas were reported in [8] for noisy sensors in an office area.

4. Chokepoint-Observation Interaction

The width of paths in the no-go topology will vary. Some paths may go along a paved road with curbs, which is a relatively narrow path. Other paths may travel through large open areas, for example if several buildings are surrounded by paved areas, so in that case the width of a path may be the relatively large distance between the buildings. The width of a path often varies along the length of a path, for example a path can narrow when passing through a gate. Intervals where paths have narrow width are called chokepoints. Often these intervals are points or are very short. Some examples of chokepoints are: a gate, an entrance road, narrow corridors between buildings, designed intersections. Sources and destinations for movement are often chokepoints, for example a portal or a small parking area next to a portal.



(a)



(b)

Figure 1. (a) Buildings in the facility. A few buildings were divided into two virtual buildings. (b) No-go areas are shaded gray. Nodes in the no-go graph are shown as circles. Additional nodes associated with portals are shown as squares.

Observations of movement through chokepoints carry large amounts of information because the movement is constrained to pass through the small area of the chokepoint. We use the term *chokepoint-observation interaction* to refer to the class of algorithms that compute features from the interaction of observations and polygonal areas carefully placed around chokepoints. Multiple polygons can be placed at a single chokepoint for use in a feature calculation, for example concentric circles, adjacent rectangles, or multiple wedge shapes. Features can be computed from the interaction of the chokepoint polygon(s) with a single observation, or the interaction with a set of observations, for example all observations during a short time interval.

The interaction of track observations and polygonal areas is used in video surveillance systems as tripwires and for counting passages. Our approach extends that simple idea by (a) utilizing more complex features derived from the chokepoint-observation interaction, (b) using multiple chokepoints together to extract information about more complex patterns of movement, and (c) our formulation specifies the chokepoints as being located on links in the no-go graph. The next portion of this paper incrementally presents our solution – building on the ideas of the no-go graph and chokepoint-observation interaction – for recognizing activity-type processes using GMTI data. Some more general details are presented below, and the next section addresses activity-type processes.

Polygon-Tracklet Interaction. A GMTI sensor produces geo-registered dots over time that correspond with moving vehicles. A tracker is normally used to connect the dots into tracks. Long tracks are difficult to produce from sensor data of built-up environments. We begin with observation data that consists of more realistic short tracks, called tracklets. So the idea of chokepoint-observation interaction is instantiated as the interaction of a polygon (for the chokepoint) and a tracklet (for the observation), as illustrated in Figure 2. The ellipses overlayed on the tracklets represent the known positional error associated with those sensor measurements. The positional error ellipses shown in the figure are intentionally made small to make the illustration more clear. In practice the ellipses are much larger than typical chokepoint polygons.

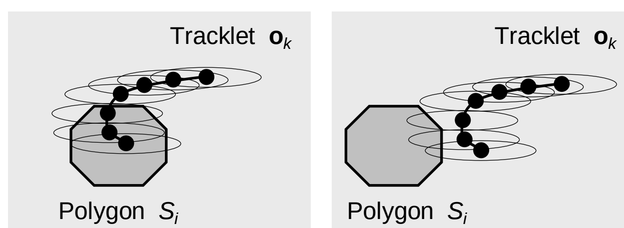


Figure 2. Example interactions between a polygon and a tracklet.

Estimating Enter and Exit Events. Many different features could be calculated from polygon-tracklet interactions, but for our example we need to detect when a vehicle enters or exits a chokepoint. An enter event occurs iff a tracklet is outside the polygon and then at a subsequent time is inside the polygon. The probability that a tracklet entered a polygon can be estimated as the product of two probability masses, corresponding to the trailing section of the tracklet being outside the polygon (while accounting for the positional error of the tracklet points) and then the leading section of the tracklet being inside the polygon. The time that an enter event occurs is more difficult to estimate because of the complex interaction of the tracklet error ellipses with the polygon. We calculate it from the point where a linear path derived from the tracklet crosses the boundary of the polygon.

Many tracklets $O=\{o_k\}$ can be near a polygon S_i at any point in time t_k, so we really want the values of $P(enter,S_i,t_k|O)$, the probability of an enter event occurring at each time instant while considering all nearby tracklets, which can be calculated (we use an approximation) from the tracklet-specific quantities described above. The equations for an exit event are similar to those described for an enter event.

Estimating Transfer Events. A transfer event is the combination of an exit event from one polygon followed by an entry event at another polygon. In our example of manufacturing processes, this corresponds with a vehicle making a trip to transfer material between two buildings. We want to estimate $P(transfer,S_i,S_j,t_k|O)$, the probability that a transfer occurs from polygon S_i to polygon S_j at time step t_k, which we calculate from three core elements: the probability that an exit event occurred at polygon S_i, the probability that an enter event occurred at polygon S_j, and the probability distribution for the transfer time between the two polygons. The basic idea, depicted in Figure 3, is that a transfer occurs if a tracklet departs polygon S_i at some time, and then a consistent time later (determined by the length of the paths between the two polygons in the no-go topology) another tracklet enters polygon S_j. If a third chokepoint existed on the transfer path, tracklets interacting with that chokepoint could be incorporated to improve the quality of the estimates. A method similar to ours for estimating exit, entry and transfer events, but for less noisy video tracklets, was reported in [13].

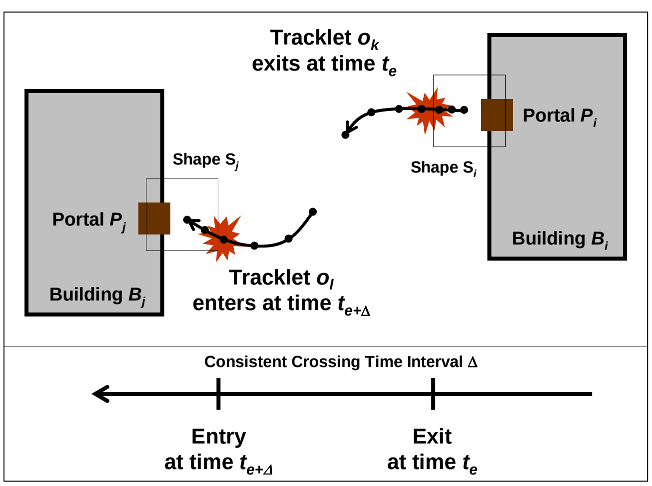


Figure 3. The basic idea for detecting a transfer event.

5. Direct Modeling and Recognition of Process-Type Activities

A *process-type activity* is an activity that involves regularly repeated actions. Those actions cause materials (raw materials or intermediate components) to move through a fixed, structured process. Obvious examples of process-type activities are various manufacturing processes. Less obvious examples are the operation of a retail store, a pizza delivery store, and security operations around an office building. Many activities that are not processes include components that are processes, and sometimes the larger activity can be detected or classified from that component alone.

The following subsections present our mathematical model for a process-type activity, our methods for estimating model parameters and for classifying observations, and lastly our experimental results. The example problem used in this paper involves a manufacturing facility that usually produces "benign" products but there is suspicion that is sometimes produces "bad" products. Process models for the "benign" and "bad" products are known. Persistent GMTI observations are made of the facility. The problem is to estimate the parameters of the process model from those observations and to decide whether the facility is manufacturing the benign or good product.

5.1. Model for Process-Type Activities

The model for a process-type activity consists of a graph $G=(\{B_i\},\{L_{ij}\})$ and the associated variables $\{w_{ij}\}$ and $\{t_{ij}\}$. Figure 4 shows an example. Each node B_i represents location i, a source and/or destination location for material transfers. Each link L_{ij} represents a transfer process of materials between two locations. The variable w_{ij} denotes the time interval between material transfers, and the variable t_{ij} denotes the time needed to perform one material transfer. Both these variables have probability distributions. If a vehicle performs a material transfer and then does something else not involving transfer of materials, that something else is not described by our process model, and that subsequent movement is treated as yet another of the many vehicles driving around that are not involved in the modeled activity.

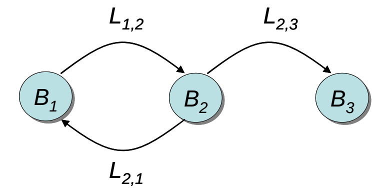


Figure 4. Graph for the model of a process-type activity.

The graph for a process model is related to the no-go graph for the facility where the process is performed. In our example, each node B_i is a portal to a building, and the transfer process represented by L_{ij} involves vehicles driving along any of the paths in the no-go graph that link locations *i* and *j* (for simplicity we assume only one such path exists). The probability distributions for $\{t_{ij}\}$ can be estimated from the path lengths in the no-go graph and the types of vehicles and materials involved in the transfers.

The model parameters that are specific to an activity are the $\{w_{ij}\}$ variables, which can be arranged into a matrix $W=[w_{ij}]$. In general, all the other quantities above are common to all process models within the area of operations. The structure of the process graph can be characterized using various quantities [2][3] calculated from the weighted connection matrix *W*.

5.2. Model Estimation and Recognition of Process-Type Activities

Given two models, W^A and W^B, for the process-type activities named A and B, and given a set of observations $O=\{o_k\}$, we want to decide whether the observation data is most consistent with process A or process B.

First we need to estimate the process model parameters W^* given the observations $O=\{o_k\}$. Recall that w_{ij} denotes the time interval between material transfers. Each value w^*_{ij} in the matrix W^* is estimated independently. Section 4 described how we estimate $P(transfer,S_i,S_j,t_k|O)$, the probability that a transfer occurred along link L_{ij} at time t_k. Our current method for estimating w w^*_{ij} thresholds the $P(transfer,S_i,S_j,t_k|O)$ values and calculates the frequency of transfers over time in the resulting binary data. That method has been sufficient for our experiments, but a better estimate would utilize Fourier transforms. Our experiments utilized only the tracklet type of observation, but our approach can easily be extended to utilize additional observation types, such as vehicle identity evidence, GMTI dots, imagery-derived change events, and evidence about the types of materials being transferred.

Given the process models, W^A and W^B, and the process model W^* estimated from the set of observations, the decision about which process model α is best supported by the data is made using the following rule.

$$\alpha = \arg\min_{\alpha=A,B} \sum_{i,j} \left\| w^*_{ij} - w^\alpha_{ij} \right\|$$

5.3. Experiments

The real-world industrial area depicted in Figure 1 was used for our experiments to ensure a realistic facility layout, including building portals, fences, gates, parking areas, movement routes, etc. The operational area spans 700 by 600 meters. An understanding of typical geospatial routes through the facility was obtained by recording once-per-second GPS positions of five cars emulating a manufacturing process. These field experiments were invaluable and enabled us to simulate realistic data for the facility. The area is real, but the facility is hypothetical.

This facility can manufacture product A or B. The models for manufacturing process A and B, which we also call *foreground* activities, are shown in Figures 5 and 6. Markov models were also defined for three *clutter* (2 security and 1 maintenance) activities. Thus, the foreground activities are embedded in the same space-time volume as the clutter activities. Our simulator utilizes the geospatial paths associated with the no-go graph links, models for the foreground and clutter activities, and a model for the GMTI sensors. Our simulator generates GMTI dots and assembles sets of 5-15 dots into tracklets. One typical dataset is illustrated in Figure 7. Manufacturing process A was performed during Day 1 and manufacturing process B was performed on Day 2. GMTI observations of the facility were made continuously. The 2σ positional error ellipses for the GMTI dots here are 150 by 15 meters.

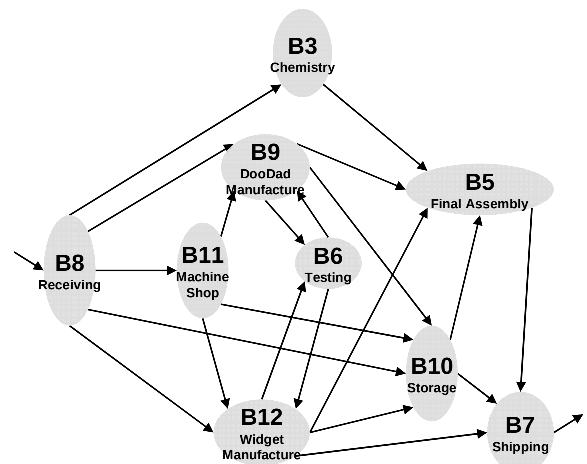


Figure 5. Graph that combines nodes and links from all process models at the facility.

	B6	B7	B8	B9	B10	B11
B6				0.28		
B7						
B8				0.14		0.14
B9	0.26				0.14	
B10		0.14				
B11				0.42		

(a)

	B6	B7	B8	B9	B10	B11
B6						
B7						
B8						1.39
B9						
B10		0.14				
B11					0.28	

(b)

	B6	B7	B8	B9	B10	B11
B6				36.1		
B7						
B8				9.4		16.9
B9	36.7				18.9	
B10		8.6				
B11				9.7		

(c)

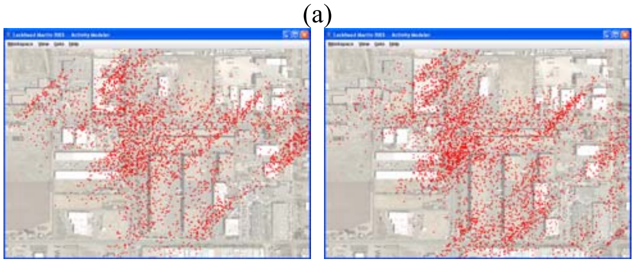
	B6	B7	B8	B9	B10	B11
B6						
B7						
B8						13.3
B9						
B10		7.5				
B11					10.6	

(d)

Figure 6. Process model parameters. Values for w_{ij}, interval in hours between transfers between location B_i and B_j, are show for activity (a) A and (b) B. Values for t_{ij}, duration of transfer in 10^{-3} hours, are shown for activity (c) A and (d) B.

	Day 1	Day 2
Foreground		
# Transfers	55	17
# Tracklets	184	31
# Dots	3115	502
Background		
# Transfers	50	50
# Tracklets	122	127
# Dots	2256	2272

(a)



(b) (c)

Figure 7. (a) Size of one typical dataset. GMTI dots for a small subset of the (b) Day 1 and (c) Day 2 datasets.

The polygons shown in Figure 8 were manually created around portals. One portal was identified for each building. Figure 9 depicts the material transfers detected between all portal pairs. This figure depicts the thresholded value of $P(transfer,S_i,S_j,t_k|O)$. All the (S_i,S_j) pairings are listed along the y-axis. Following any single horizontal line across the chart shows a series of line segments, and each line segment denotes the duration of one material transfer. Next, the process model parameters were estimated from the observation datasets. Our decision rule correctly classifies the Day 1 dataset as being most consistent with manufacturing process A, and the Day 2 dataset as being most consistent with process B.

Figure 8. Polygons created around portals and chokepoints in the facility.

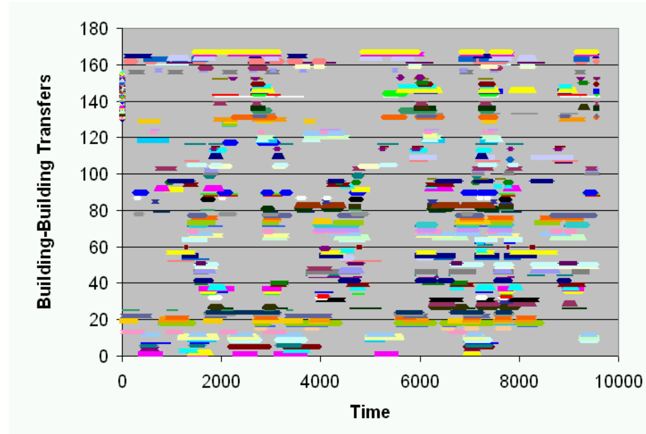


Figure 9. Material transfers detected.

We generated several versions of the Day 1 dataset, each with a different size for the positional error ellipse associated with the GMTI sensor observations. The major axis of the error ellipse was varied from 10 meters to 150 meters. The minor axis was one tenth of the major axis. Figure 10 shows how the correct classification performance for those Day 1 datasets varies as the positional error ellipse becomes larger. The y-axis is a matching metric reflecting how well the model for manufacturing process A is supported by a dataset. Each point along the y-axis denotes a different Day 1 dataset generated using the specified positional error ellipse size. These initial experimental results use tracks rather than tracklets and no clutter activities.

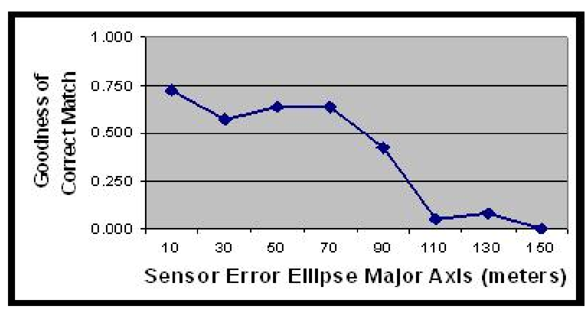


Figure 10. Relationship of correct classification for Day 1 datasets with different positional error ellipse sizes.

A downward trending plot is observed, which makes sense. A manufacturing process should be easy to identify given simulated GMTI sensor data that has very small positional error ellipses, and the manufacturing process should be harder to identify as the positional error ellipse gets larger. Classification performance is of course also dependent on the distance separating buildings in a facility. The building separation values in our facility are: minimum 63, ad hoc typical 189, average 309, maximum 701 meters.

This type of plot shows that *it is possible* to correctly classify a complex activity in GMTI data, although the size of the required positional error ellipse is smaller than most

GMTI sensors today. Simulation experiments like this are helpful to designers of future GMTI sensor systems, because they indicate how much positional error can be tolerated while still allowing complex activities to be recognized. Distance between buildings is also an important parameter, and so it will be easier to recognize activities in a facility with widely separated buildings relative to the GMTI sensor system's positional error ellipse size.

6. Indirect Detection and Analysis of Unknown Activities Using Clustering and Markov Models

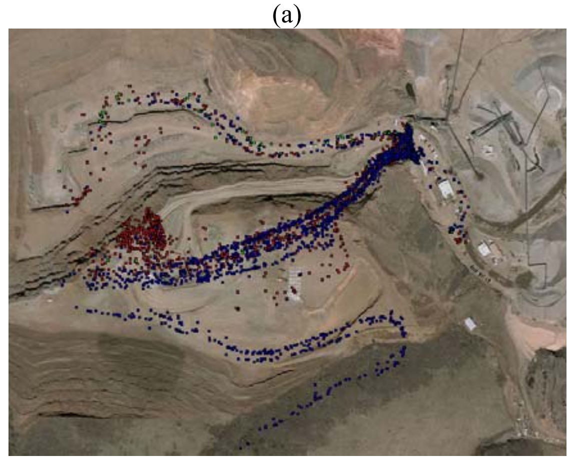
Often a high-level model for a wide-area activity is not available (such as the process-type activity model in the previous section) because not enough is known beforehand about the activities occurring in an observation area. In these situations a low-level approach is required. Our approach provides a generic low-level model for the overall activity in an area, which we refer to as a *normalcy model*. Unknown but interesting activities can be discovered as deviations from a learned normalcy model. Our approach can be used to discover activity patterns, characterize an unknown activity, and generally to help a human develop a good understanding of the activity, which can eventually lead to the creation of an explicit high-level model for the activity.

Following is a quick summary of our approach [7]. This approach was originally developed for tracks of people playing sports in video. Our purpose here is to show the benefits of applying those methods to analysis of wide-area activity patterns. The approach uses track or tracklet observations that have small positional errors. If tracks are used, each track is broken into short segments (tracklets) that have a fixed time duration, typically a few seconds. Several features are calculated for each tracklet (e.g., world position, speed, direction, acceleration, curvature, shape primitives) and packed into a feature vector. A clustering algorithm is applied to the set of feature vectors in an observation dataset, resulting in a set of *prototype* vectors, which are the centers of the resulting clusters. Since the feature vector includes the world position of a tracklet, the clustering can be biased towards geo-spatial clusters that also share some other common features. A movement through space can then be represented as a sequence of cluster indices. A Markov model, with weighted transition matrix, is learned from those sequences. This assumes enough information is available to reliably connect either a few or several tracklets into a longer track.

The rock quarry depicted in Figure 11 was used for the experiments here. The operational area spans 3000 by 2000 meters. Many activities in the quarry have a repetitive or process-like nature, but we do not explicitly use that knowledge. The observation data consists of GPS tracks for three mining vehicles, sampled once per second, all day, for 25 days. These tracks are broken into fixed length tracklets that are a few seconds long. We believe the large area and relative separation of vehicles in this facility would enable a GMTI tracker to produce reasonably good tracklets.

Figure 12 shows the clusters derived from the entire observation dataset. The ellipses represent the covariance of the points within each cluster, and the ellipses have been projected onto the same latitude-longitude coordinates as the overhead view in Figure 11(b).

(a)



(b)

Figure 11. (a) Side and (b) overhead view of quarry.

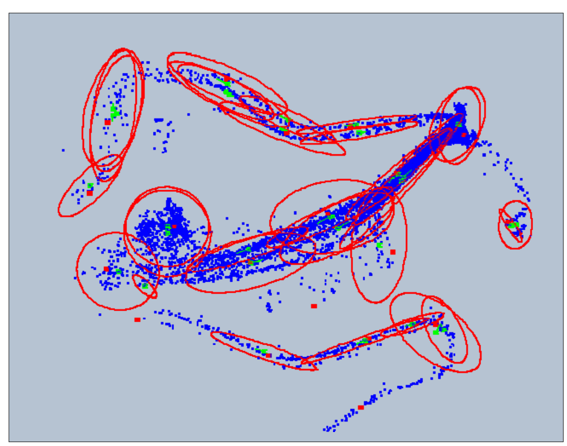


Figure 12. Feature vector clusters.

Careful study of the clusters can provide useful information about the activities being performed in the facility. Figure 13 shows two clusters, which cover approximately the same area, and a histogram of the data in those clusters over the hours of the day. The histograms show large peaks at 5am and 10pm, which suggests that this is a possible location of a garage. Figure 14 shows another pair of clusters (it is just a coincidence that there are again two clusters at a location), and a histogram of the data in these clusters over the Julian day of the year. The histogram shows large peaks at days 53-55, which suggests that this is the opening of a new dig face.

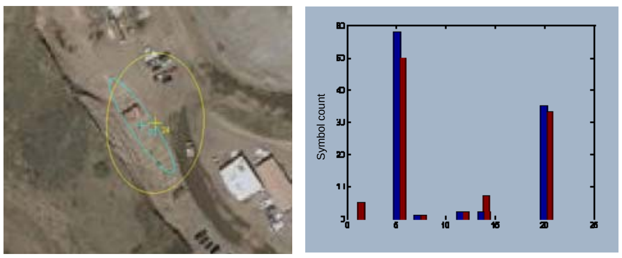


Figure 13. These two clusters are correlated with the time of day.

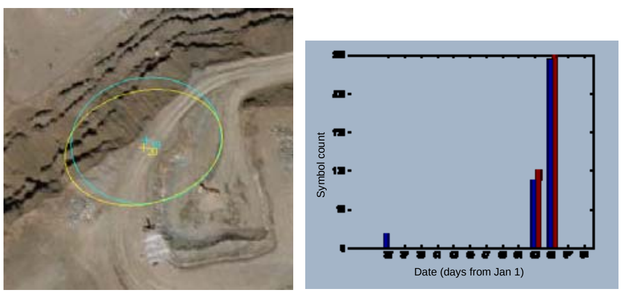
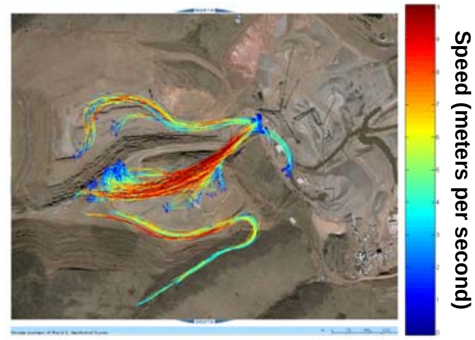


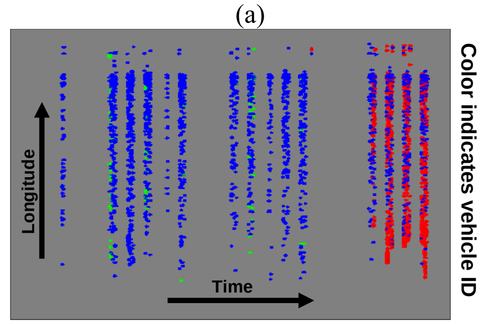
Figure 14. These two clusters are correlated with the Julian day of year.

Relatively simple visualizations of the feature vectors can also provide useful information about the activities being performed. For example, Figure 15(a) shows that high speeds occur along specific paths. Figure 15(b) shows periodic activity corresponding to days of the week, no activity at night, and that certain vehicles are only active during portions of the three-week period.

The probability of an observation sequence can be calculated from the learned Markov normalcy model. Figure 16 shows two such sequences (meaning the tracklet points associated with the cluster index sequences). Track 248 has a relatively high probability of 0.3, which means this track is highly consistent with the Markov normalcy model. Track 247 has probability 0.01 and is an example of a deviation from the normalcy model. While this is simply a deviant track, more complex types of deviations could be calculated from normalcy models. Track 247 is this example drives on the left side of the path, has unusual curved motion, and the vehicle stops and waits for a short time during the transit. Note that the terrain of the quarry is continually changing so the orthoimage shown in the figure is old relative to the track data.



(a)



(b)

Figure 15. (a) Dataset colored-coded to indicate speed. (b) Dataset plotted on longitude-time axes.

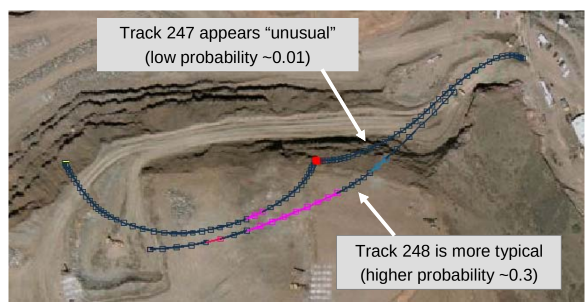


Figure 16. Identifying interesting activities as a deviation from the Markov normalcy model.

7. Summary

This paper introduced new methods to model, visualize and automatically recognize wide-area activities, which commonly occur in urban and built-up areas (e.g., a facility). A new category of wide-area activity is identified, process type activities, which involve structured repetitive material transfer movements.

We introduced the *no-go topology* method and the *chokepoint-observation interaction* method, and then showed how new algorithms can be built on them to recognize process models in observation data. Experimental results were presented for a manufacturing facility observed using persistent GMTI data.

This paper demonstrates that wide-area activities can theoretically be identified in built-up environments using GMTI data. This is a new concept and still a very hard problem. We essentially begin with the easiest possible theoretical situation (continuous GMTI observations, small positional error ellipses, repetitive activities) and want to move toward realistic parameters in operational situations (fewer observations, larger error ellipses, figuring out how to exploit additional kinds of activities). Simulation results were presented to illustrate how much positional error can be tolerated in the GMTI observations while still obtaining good discrimination of activities. This kind of analysis result is helpful for selecting design parameters for future GMTI systems that can address future mission needs.

In order to illustrate all major approaches for addressing wide-area activities, the paper quickly presents some results where explicit high-level activity models are not used. A combination of clustering methods and Markov models are used to represent low-level patterns of activity in a wide area. Then we presented experimental results illustrating how an interesting activity can be detected as a deviation from the normalcy model, and how new activity patterns can be discovered using simple visualizations of the results.

Acknowledgements

This work was supported by an Independent Research and Development (IR&D) project in CY2005 within Lockheed Martin IS&S (Integrated Systems & Solutions) and an associated IR&D subcontract to the Colorado School of Mines.

References

[1] L. Chittaro, L. Ieronutti, “A Visual Tool for Tracing User's Behavior in Virtual Environments,” *International Working Conference on Advanced Visual Interfaces (AVI),* 2004.

[2] G. Chowell, J.M. Hyman, S. Eubank, C. Castillo-Chavez, “Scaling Laws for the Movement of People Between Locations in a Large City,” *Physical Review E,* 68, 2003.

[3] A. De Montis, M. Barthelemy, A. Chessa, A. Vespignani, “The Structure of Inter-Urban Traffic: A Weighted Network Analysis,” 2005.

[4] M. Devaney, A. Ram, “Needles in a Haystack: Plan Recognition in Large Spatial Domains Involving Multiple Agents,” *National Conference on Artificial Intelligence (AAAI),* 1998.

[5] S. Hongeng, R. Nevatia, F. Bremond, “Video-Based Event Recognition: Activity Representation and Probabilistic Recognition Methods, *Computer Vision and Image Understanding,* 96(2):129-162, 2004.

[6] S.S. Intille, A.F. Bobick, “Recognizing Planned, Multiperson Action,” *Computer Vision and Image Understanding,* 81(3):414-445, 2001.

[7] J.Y. Lee, W. Hoff, “Activity Identification Utilizing Data Mining Techniques,” *IEEE Workshop on Motion and Video Computing (WMVC),* 2007.

[8] L. Liao, et al., “Vornoi Tracking: Location Estimation Using Sparse and Noisy Sensor Data,” *International Conference on Intelligent Robots and Systems (IROS),* 2003.

[9] L. Liao, D. Fox, H. Kautz, “Learning and Inferring transportation Routines,” *National Conference on Artificial Intelligence (AAAI),* 2004.

[10] N.M. Oliver, B. Rosario, A.P. Pentland, “A Bayesian Computer Vision System for Modeling Human Interactions,” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 22(8):831-843, 2000.

[11] D. Patterson, D. Fox, H. Kautz, M. Philipose, “Fine-Grained Activity Recognition by Aggregating Abstract Object Usage,” *International Symposium on Wearable Computers (ISWC),* 2005.

[12] M. Seibert, et al., SeeCoast Port Surveillance,” *Proceedings of SPIE Vol. 6204: Photonics for Port and Harbor Security II,* 2006.

[13] C. Stauffer, “Estimating Tracking Sources and Sinks,” *Workshop on Event Mining,* 2003.

[14] C. Stauffer, W.E.L. Grimson, “Learning Patterns of Activity Using Real-Time Tracking,” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 22(8):747-757, 2000.

[15] C.R. Wren, E.M. Tapia, “Toward Scalable Activity Recognition for Sensor Networks,” *International Workshop on Location and Context Awareness (LoCA),* 2006.

