

Linear Referencing for Network Analysis of IED Events

K. M. Curtin, PhD

Department of Geography and GeoInformation Science
George Mason University
4400 University Drive (MS 6C3)
Fairfax, VA 22030
curtin@gmu.edu

Abstract—This paper outlines a motivation for associating IED events (and other significant physical and human geographic features) with the road network, describes the use of methods known as linear referencing in order to do so, and presents an example of how linear referencing of several types of events can occur. This is followed by a description of several measures of network density of events, and a demonstration of how linearly referenced events can be combined to analyze spatial coincidence of different event types. This is followed by suggestions for future research including the development of network based spatial statistics, optimization of network services based on the linearly referenced events, and geographic information system tool development to integrate these methods.

I. INTRODUCTION

Initial analysis of the placement of IEDs as a stochastic process has shown that over large areas of operation IED incidents have well-defined temporal patterns including distinct time-of-day and day-of-week patterns (Kolesar *et al.* 2008). It is well-accepted that IED countermeasures could be improved if both temporal and spatial patterns were better understood (Bartow 2008b).

For the purposes of this research it is asserted that since convoys and other Blue assets traveling along roads are one of the primary targets for IED attacks, the road network should be the spatial platform on which the description of these attacks is made, and with which the forecasting (or probability) of future attacks is associated. It has long been recognized that Geographic Information Systems (GIS) are a significant tool for transportation network modeling due to their ability to realistically model linear and network features (Miller *et al.* 2001, Curtin 2007, Curtin 2008b).

Moreover, there is a set of tools and methods in GIS designed to associate geographic locations with networks. These tools and methods, known collectively as linear referencing, comprise a process for locating a point or line along a network feature or features, rather than using classical geographic coordinate systems (Curtin *et al.* 2007, Curtin 2008a). Linear referencing refers to the point or line as an ‘event’ whether it represents an object or an occurrence. The process of linear referencing allows for more efficient data capture of events that occur along networks, for easier relocation of the event sites in subsequent time periods, and—

most importantly—for intuitive modeling of network based events.

The primary objective of this paper is to demonstrate how IED events – and the social and physical geographic components of the theater of operations – can be associated with network features in order to generate actionable descriptions of spatial IED density, probability, or risk. Moreover this paper will outline some of the ways in which those linearly referenced events can be analyzed, with particular attention to potential measures of network density. Finally, this paper will discuss possible future research into network based spatial statistics, optimization of IED countermeasures based on the linearly referenced events, and potential GIS tool development for in-theater use.

For the purpose of presenting our methodology in an unclassified forum, we constructed a dataset that includes a fictional Major Supply Route (MSR), termed MSR-Chicago (Figure 1), together with a set of randomly generated fictional IED events and associated features (e.g., tribal regions, intersections, etc.). MSR-Chicago is intended to be similar to what MSR-Tampa would look like. The entire MSR is evaluated in terms of the IED events along it, as are subsections of the entire route.

II. LINEAR REFERENCING BACKGROUND

The term “linear referencing” emerged from engineering applications where it was preferable to locate a point along a linear feature (such as a road) by referencing that location to some other well-defined location, rather than using classical geographic coordinate systems. The most familiar illustration of linear referencing is the mile markers along US highways (Federal Highway Administration 2001, Federal Transit Administration 2003).

For any network application, the use of linear referencing has several primary benefits. First, locations specified with linear referencing can be readily recovered in the field and are generally more intuitive than locations specified with traditional coordinates. Second, linear referencing removes the requirement of a highly segmented linear network based on differences in attribute values. The implementation of linear referencing allows an organization to maintain a network database with many different attribute events associated with a

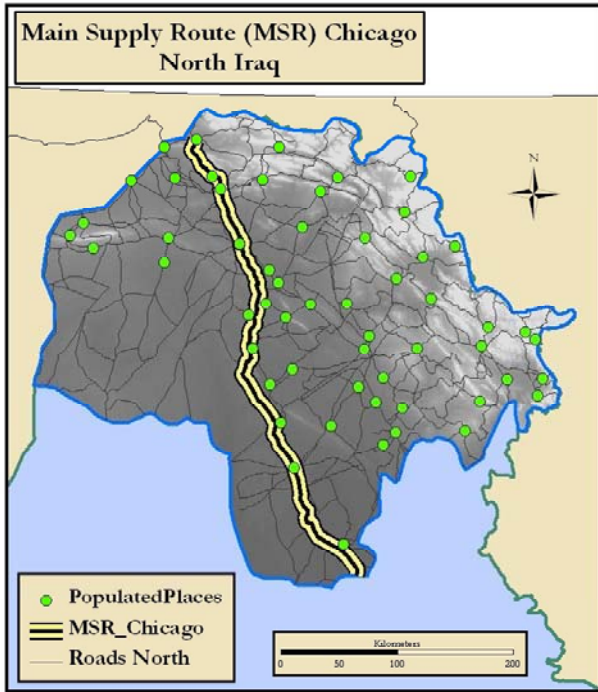


Figure 1: Major Supply Route (MSR) Chicago

single, reasonably small, set of network features. The implementation of linear referencing thus reduces the redundancy and potential error within the database, and it facilitates multiple cartographic representations of network attribute data (Curtin *et al.* 2007).

To begin the process of linear referencing of IED events several decisions need to be made regarding the nature of the network to which they are referenced and the way in which their linear measures are captured. The first issue is the determination of the underlying “routes” to which events are referenced. In the terminology of linear referencing a “route” is defined as the largest single network feature – or collection of network features – that is uniquely identified and along which events occur. In typical urban transportation contexts, routes may be defined as an entire named road – that is, all road features with the name “Main Street” can be a single route for linear referencing, while all road features with the name “Elm Street” would be another route. For our purposes, we may need to define several types of route. One option is to make every MSR or ASR (Alternate Supply Route) into its own route for purposes of linear referencing. In that case MSR-Chicago would be a single route that extends from the south end of MND-North to the north end of that region. Another option is to delimit routes based on the features that make up Targeted Areas of Interest (TAIs), or to delimit routes based on sections of road that correspond to patrol areas or command areas. This is an operational decision that can be changed as conditions or analytical needs change (Figure 2).

The second issue to address when initiating a process of linear referencing is to determine the characteristics of the measure values along the routes. There are three primary

considerations when setting measures along routes: 1) the unit of measure that is most appropriate, 2) the source for the measure values, and 3) the direction of increasing measure values. In the present case, the appropriate coordinate system for this area of operations is the Universal Transverse Mercator (UTM) coordinate system. The area is of the appropriate size for one UTM zone, and this coordinate system is also the basis for the Military Grid Reference System. By convention UTM locations and distances are given in meters, so the measures presented here will be in meters (sometimes converted to kilometers in the results). There is no known superior source for the measure values than the GIS data itself, so measure values will be calculated based on the feature length. If more detailed measure values from the field become available in the future these can be used to update measure values in the GIS. Lastly, there is no known basis to prefer one direction of measure over another (e.g. address ranges, historical convention) so measurements begin with 0 in the south, with measures increasing to the north. MSR-Chicago is 446,600 meters long.

The setting for this analysis of IED events is a Major Supply Route (MSR) traversing a theater of operations. Many types of Blue activity (e.g. supply convoys, troop transports, vehicles engaged in active operations) occur along the MSR. Different types of Blue activities may have different operational value, and different characteristics (e.g. number of vehicles, typical travel speed, typical time-of-day or day-of-week patterns). There may be areas of the MSR that have been identified in the field as Targeted Areas of Interest (TAI) due to either their importance for Blue operations along the MSR,

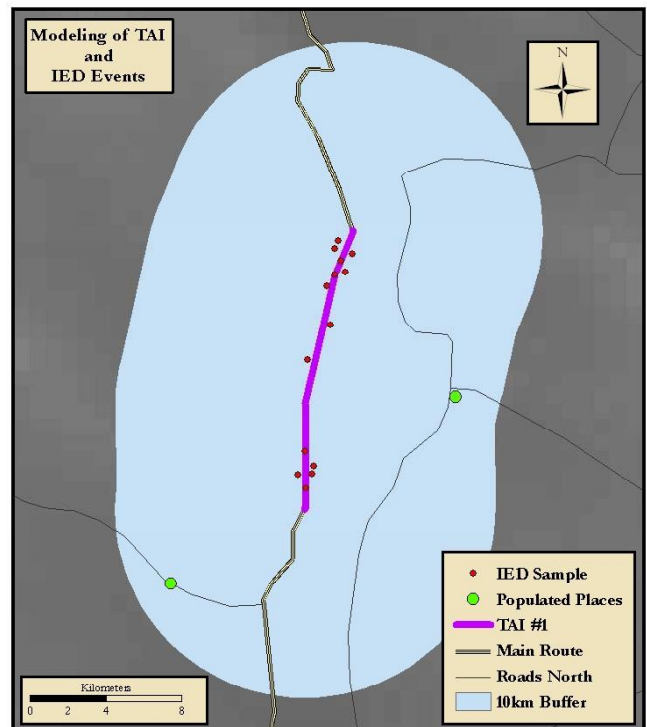


Figure 2: TAI as a Linear Referencing Route

or due to the prevalence of Red operations (including IED attacks) in that area.

Although not all IED attacks are focused on roads, both an inspection of the spatial distribution of past events, and descriptions of IED attacks from the theater demonstrate that a large majority of attacks are directed at Blue (or perhaps some Green) activity that occurs on or very near roads.

IED events are assigned geographic coordinates in the field. There are several potential sources of error in these coordinates, not the least of which is the likelihood that precise locations are difficult to capture at a scene of an explosion where there is active Blue-Red interaction. Moreover, the spatial representation of the road itself may contain errors, and is certainly subject to the limits of the precision with which it was captured. Regardless of its positional accuracy, linear features such as roads are generally represented in the GIS as single centerlines with no explicit width (although the feature can be drawn with a line of any user specified width). This single line may be used to represent a real-world highway feature that may have a large and variable width. Typical minimum lane widths on U.S. highways are 12 ft., so a typical highway with 2 lanes in each direction will have a minimum width of 48 ft. With the addition of shoulders and a median between the directions of travel the width of a road can exceed 150 ft.

Therefore, although IED event locations captured in the field will not precisely coincide with the spatial representation of the road (Figure 3), it is safe to assume that any IED event that is located within a 100 ft. buffer of the road is very likely to have been directed at Blue activity occurring on the road. This assumption may fail if there are areas very near to the road which contain buildings that are also the sites of IED activity, and in this case more careful examination of the attributes of the IED in the SIGACTS database may be necessary to determine whether or not it should be included in the network analysis of road events. Those events falling within the buffer are forced onto the road centerline with linear referencing (although their offset distance from the line can be retained). The events are assigned a measure value along the network feature to which they are referenced.

The following five sets of events are referenced to the network (Examples of the IED and Tribal Affiliation events are shown in Figure 4). These were chosen based on their relationship to terrain and demographic features that may contribute to IED activity:

- IED events – set of 1000 point events randomly distributed over the measure extent of MSR-Chicago (0-446,600 meters);
- Road Intersection events – point events at any intersection of MSR-Chicago with any other road;
- Culvert events – a point event at any intersection of MSR-Chicago with a river feature;
- Urban/Rural events – Linear events with an attribute that either defines each event as “Rural”, or specifies the urban area to which that section of MSR-Chicago is adjacent; and

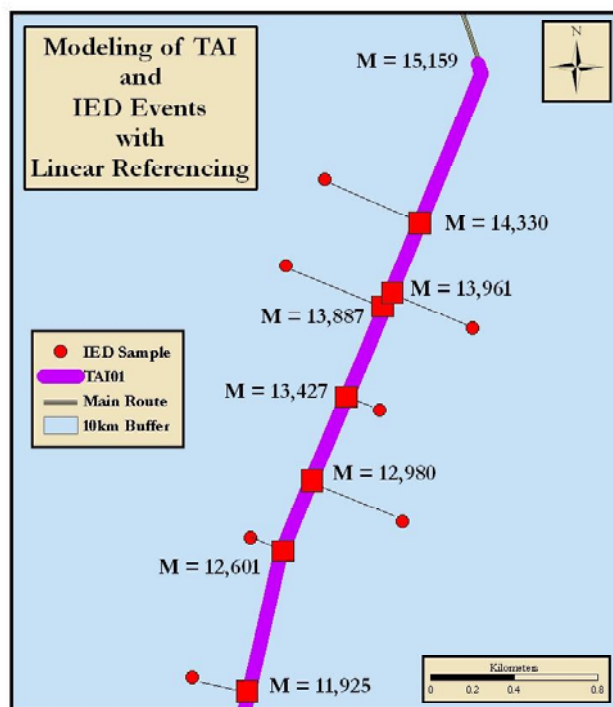


Figure 3: Moving events to the Network

- Tribal Affiliation events – Linear events based on polygon maps of tribal affiliation, where the event corresponds to the section of MSR-Chicago that intersects with the tribal area. MSR-Chicago appears to frequently be the border between tribal areas and therefore right and left side events were both created. The understanding of Tribal influence including the spatial component has been identified as an important issue in U.S. policy decision making and military operations (Al-Shahery *et al.* 2008).

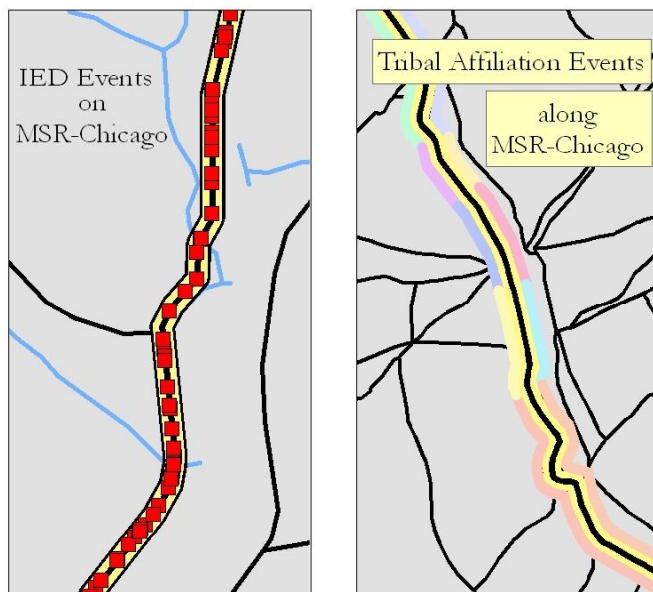


Figure 4: Events (IED and Tribal) along MSR

III. ANALYSIS OF LINEARLY REFERENCED EVENTS

The linear referencing of events to the road network is not an end in and of itself. Rather it is a means to the end of performing analysis on the events in order to better plan and execute operations on the road, such as convoy movements, RCT allocation, or perhaps even Red-Blue interaction operations. However, since there has been very little use of linear referencing techniques by the GIS community in general, and almost no use of these methods outside of transportation planning and management, there are limited analytical tools available. More generally, while the field of network analysis does have a solid theoretical background in the mathematical sub-discipline of graph theory, only a few of the more basic network algorithms have found their way into common GIS practice. A greater interest in networks of all types, particularly analysis of the characteristics of social networks as a means of identifying command structures may lead to increased research and application in this area in the near term (Curtin 2007). Below is a demonstration of how to compute network densities via linear referencing, and a discussion of overlay operations on events.

A. Measurements of Network Density

There is no well-defined notion within the GIS community of network density in the sense that we intend here, as a correlation to two-dimensional spatial density. In a spatial context, density is the number of observations per unit area. The network domain does not have an area, thus there is no perfectly equivalent network density. The term “network density” has been used in several different contexts, however. It has been defined as the distance between nodes in a network (Villafuerte 2008), the level of connectivity between nodes in a network (Levis *et al.* 2008, Mizruchi *et al.* 2008), and even as a measure of the number of units of network links, per square units of surface area (Rodrique *et al.* 2006). None of these definitions captures the phenomenon we are trying to develop here – essentially the distribution of points over a network space.

Some work on measurements of network density in GIS has attempted to use traditional 2-dimensional kernel density estimators on the locations of nodes in networks (Borruso 2003), or to replace Euclidean distances with network distances in the computation of density surfaces (Borruso 2005). Both these methods, however, are accepting inputs from the network structure but not constraining the outputs to the network as well. That is, when two-dimensional density surfaces are generated based on the locations of network based phenomenon, the results state that there is a density of that phenomenon (perhaps small based on the kernel bandwidth) occurring across the entire space. If the phenomenon is restricted to the network (such as IED attacks) these results are specious at best.

This is demonstrated by the example shown in Figure 5, where IEDs have historically occurred in spatial concentrations along several roads in the network, in particular along two segments of the road network identified as TAI #1

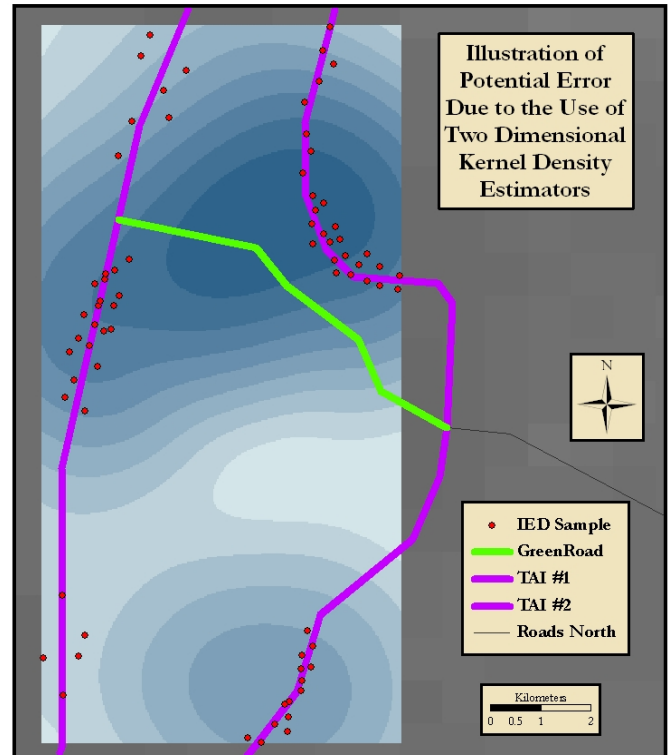


Figure 5: Traditional Spatial Density Problem on Network Events

and TAI #2. There are also segments of the road network that have had no IED activity, in particular note the segment of the network termed the GreenRoad. This segment of the road has (for some reason) been completely immune to IED attacks and therefore the probability of future attacks is likely to be low. It would be operationally fruitless to concentrate RCT resources on the GreenRoad where there is little to no expectation of finding and disarming an IED. However, when a two-dimensional kernel density estimation is made, the GreenRoad is identified as a part of the network most in need of RCT coverage. Any automated procedure that sought to cover portions of the network most at risk for IED attack would cause of waste of resources by covering the GreenRoad. An estimation of density across two-dimensional space should not be used for events that occur along a network.

Knowing of no previous research into the specific problem being addressed, this paper starts from first principles and describes each way that the point pattern on a network can be described. First a summary of events across an entire route is made, then those results are compared to the values for specified segments of a route.

i. Events along the Entire Length of the Route

The most intuitive network correlate to standard density measures is simply to take the number of events on a route and divide by the length of the route. In the current case (randomly generated data) the network density of the entire route is 1000/446 km or 2.24 IED events per km over a three year period. Temporally this distribution can be sliced in any way to determine annual, monthly, weekly, daily or even hourly

network densities. Below are yearly and Day-of-Week distributions for MSR-Chicago as examples:

Year	# of IED Events
2006	328
2007	355
2008	317

Day of Week	# of IED Events
Sunday	156
Monday	142
Tuesday	141
Wednesday	135
Thursday	143
Friday	147
Saturday	136

But these measures are only slightly more specific than simple totals for an entire region or country. They are totals for a route in the network. A benefit of linear referencing is that it is possible to summarize IED activity along any subset of a route that is of interest.

ii. *Events Only Along a Targeted Area of Interest (TAI)*

In the current case a single TAI extends from measure 273,602 to 288,558 (nearly 15 km) along MSR-Chicago (Figure 6). Below are the yearly and Day-of-Week breakdowns for this TAI:

Year	# of IED Events
2006	18
2007	19
2008	6

Day of Week	# of IED Events
Sunday	7
Monday	6
Tuesday	3
Wednesday	5
Thursday	6
Friday	7
Saturday	9

Even this simple example demonstrates how the distribution of points along a segment of the network can differ from the network as a whole. The relative drop in IED activity in 2008 on this TAI is much greater than for the entire MSR. More work must be done to determine more rigorously if this drop is statistically significant. However, this dramatic year over year change leads to obvious questions about what might have caused the drop (RCT activity, political reconciliation, tribal cooperation, etc.). Moreover, this drop in activity may lead the commanders in the field to alter their route clearance activities to cover other areas of the MSR that are in more immediate danger of IED attack.

Perhaps most importantly, this description of the network point patterns can be made for any specific segment of the MSR, or for the entire MSR broken into regular or irregular segments. That is, the “network density” of points for every 1 km, 5 km, 10km, or other regular division of the network can be determined, to differentiate between areas of high and low IED activity. The immediate task is to determine the best way to segment the network for operational goals, and to generate data for further analysis from the linearly referenced events along those segments.

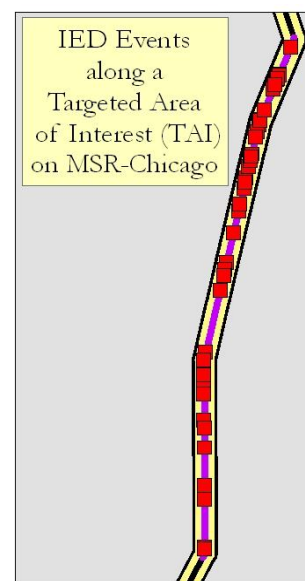


Figure 6: IEDs on a TAI

B. *Overlay Operations on Linearly Referenced Events*

Although a measure of “network density” or a related description of the network point pattern of IED is of primary concern, the spatial expression of the IEDs through linear referencing allows comparisons of those IEDs against other characteristics of the road that may have a bearing on the prevalence or absence of IEDs in certain areas. The IED events can be overlaid with other events to see their coincidence or lack thereof.

For example the IEDs can be intersected with tribal boundaries to determine which tribal areas contain a greater incidence of IEDs. The IED locations can be compared to the location of culverts or road intersections, giving a quantitative measure of how much more frequently these areas are the targets of IED activity. IEDs can be associated with demographic zones such as Urban, Suburban, or Rural areas, to name only a few. A secondary goal is to develop an overall measure of network risk for IEDs, or conversely, demand for RCT services, by spatially combining many factors that may contribute to the presence of an IED. Although this may not meet one’s specific definition of “prediction” such information could at least contribute to thoughtful operations planning.

IV. CONCLUSIONS AND FUTURE RESEARCH

We envision three primary areas of future research based on our current activities and perception of overall JIEDDO needs: 1) the advancement of network based spatial statistics to evaluate IED incidence and to make descriptive and perhaps predictive statements regarding IED occurrences, 2) the optimization of route clearance activities on the network based on the linearly referenced events and the need to cover Blue or Red activities most efficiently, and 3) the development of GIS based tools to allow users in the theater of operations to perform spatial analyses and make operational decisions.

C. Network Based Spatial Statistics

Several documents have discussed the use of spatial statistical analysis in the determination of trends of IED emplacement and effectiveness, and as potential tools for the prediction or forecasting of future IED locations. The use of these statistics has been described as “rudimentary” and includes “hot-spot” analysis (Bartow 2008a) that typically involves visual analysis of groupings of IED events on a map. Even if formal spatial statistics are being used to determine hot spots, depending on the particular statistic used, this may or may not be the most appropriate statistic for description or prediction of IED locations. Moreover, the results of many of these statistics are very sensitive to the initial parameters chosen prior to their execution, and therefore their implementation should be rigorous and justifiable.

Moreover, there is very little extant research into the field of network-based spatial statistics. Since IEDs are most frequently associated with road networks and the potential targets that traverse those roads operate on networks, new statistics may need to be developed to address our analytic needs. We intend to pursue these statistics in our efforts to describe and analyze network based IED attacks.

D. Optimization of Route Clearance Activities Based on Linearly Referenced IED Events

The ability to locate events (IEDs, Red activities, Green locations, terrain conditions) on a network through linear referencing as shown above, combined with a stated need to optimize Route Clearance activities suggests that an integration of these methods could provide more robust answers to operational questions regarding RCT deployment on MSRs. During discussions with TRAC-WSMR personnel, the subject of “covering” TAIs with RCTs was a recurring theme. This suggests that a covering formulation of a linear program may be appropriate in this instance. Consider the following formulation of a variant of the Maximal Covering Location Problem cast in the context network patrol. We begin by defining some notation:

- i, j = the indices of nodes in the network
- E = the set of edges (i, j) in the network
- k, K = index and set of base locations for Network Patrol Teams (NPTs)
- P = the number of NPTs available to patrol the network
- S = the acceptable service distance (covering distance) between a NPT base location and section of the network to be covered
- N_{ij} = $[k \text{ in } K \mid \min(d_{ki}, d_{kj}) \leq S]$
- y_{ij} = the number of NPTs covering the arc from i to j
- x_k = the number of NPT assigned to base location k
- a_{ij} = the measure of priority for covering the arc from i to j

Given these notational conventions, we formulate the optimization problem as:

$$\begin{aligned} \text{Max } & \sum_{(i,j) \in E} a_{ij} y_{ij} \\ & \sum_{k \in K} x_k = P \\ & \sum_{k \in N_{ij}} x_k \geq y_{ij} \quad \forall (i, j) \in E \end{aligned}$$

The measure of priority for covering any given section of the network would be a function of the risk of IED, or demand for RCT services generated by the analysis of the linearly referenced events discussed above. It may be possible to add a temporal component to both meet any time patterns observed in the data, or to meet scheduling needs of the RCTs. We intend to pursue solution of instances of this problem in order to allow more specific operational plans to be generated using the linearly referenced IED events combined with Blue, Red, and Green activities.

E. Based Tool Development for Improved Operations

We believe the work described above, when complemented with the suggested future research will result in a methodology for associating IED incidents to a road network and determining intensity along those networks using existing and new models in a GIS environment. When combined with optimization techniques, we believe that efficient solutions for operational questions can be generated. Since GIS is well-accepted in military contexts, and since a call has been made to provide access to mapping software (specifically ArcGIS) to the ORSAs in the theaters of operation who are collecting and maintaining the databases of IED locations (Bartow 2008a), the next logical step is to develop a set of GIS Based tools to be used to implement the research advances made here in an operational environment. Our research group has extensive experience in developing such tools, and we look forward to pursuing this avenue of future research.

ACKNOWLEDGMENTS

Thanks to the entire Mason JIEDDO team for allowing a geographer in their midst, and their support of these ideas. Special thanks to Fred Woodaman with whom many of the ideas in this paper were initially generated. Many thanks to Nancy Perry for her tireless work.

REFERENCES

- Al-Shahery, O., Karasik, T., Schbley, G. and Sollinger, J., 2008, Tribes in Iraq: Implications for U.S. Military Operations. p. 141 (Santa Monica: RAND Institute).
- Bartow, T., 2008a, After Action Report: Deployment to Afghanistan 28 January 2008 - 8 June 2008. p. 12 (Santa Monica: RAND Corporation).
- Bartow, T., 2008b, Counter-IED Predictive Analysis in Afghanistan. (Santa Monica: Rand Corporation).

- Borruso, G., 2003, Network Density and the Delimitation of Urban Areas. *Transactions in GIS*, **7**, pp. 177-191.
- Borruso, G., 2005, Network Density Estimation: Analysis of Point Patterns over a Network. In *Computational Science and Its Applications - ICCSA 2005*, O. Gervasi (Ed.), pp. 126-132 (Berlin: Springer-Verlag).
- Curtin, K.M., 2007, Network analysis in geographic information science: Review, assessment, and projections. *Cartography and Geographic Information Systems*, **34**, pp. 103-111.
- Curtin, K.M., 2008a, Linear Referencing. In *The Encyclopedia of Geographic Information Science*, K.K. Kemp (Ed.), pp. 261-264 (Los Angeles: Sage Publications).
- Curtin, K.M., 2008b, Network Data Structures. In *The Encyclopedia of Geographic Information Science*, K.K. Kemp (Ed.), pp. 314-317 (Los Angeles: Sage Publications).
- Curtin, K.M., Nicoara, G. and Arifin, R.R., 2007, A Comprehensive Process for Linear Referencing. *URISA Journal*, **19**, pp. 41-50.
- Federal Highway Administration, 2001, Implementation of GIS Based Highway Safety Analysis: Bridging the Gap. p. 104 (McLean, VA: U.S. Department of Transportation).
- Federal Transit Administration, 2003, Best Practices for Using Geographic Data in Transit: A Location Referencing Guidebook. p. 211 (Washington, DC: U.S. Department of Transportation).
- Kolesar, P.J., Woodaman, R.F. and Kleister, K.G., 2008, A Time Series Analysis of Improvised Explosive Device Incidence. In *C4I Center Technical Report*, p. 36 (Joint Improvised Explosives Device Defeat Organization).
- Levis, P., Brewer, E., Culler, D., Gay, D., Madden, S., Patel, N., Polastre, J., Shenker, S., Szewczyk, R. and Woo, A., 2008, The emergence of a networking primitive in wireless sensor networks. *Communications of the Acm*, **51**, pp. 99-106.
- Miller, H.J. and Shaw, S., 2001, *Geographic Information Systems for Transportation* (New York: Oxford University Press).
- Mizuchi, M.S. and Neuman, E.J., 2008, The effect of density on the level of bias in the network autocorrelation model. *Social Networks*, **30**, pp. 190-200.
- Rodrique, J., Comtois, C. and Slack, B., 2006, *The Geography of Transport Systems* (New York: Routledge).
- Villafuerte, F.L., 2008, Using Network Density as a New Parameter to Estimate Distance. In *Seventh International Conference on Networking*, pp. 30-35.