Identifying Security Checkpoint Locations
to Protect the Major US Urban Areas

Daniel M. Watkins, Leticia Cuéllar, Deborah A. Kubicek, Erick Rodriguez, Phillip D. Stroud

ABSTRACT

Transit networks are integral to the economy and to society, but at the same time they could allow terrorists to transport weapons of mass destruction into any city. Road networks are especially vulnerable, because they lack natural checkpoints unlike air networks that have security measures in place at all major airports. One approach to mitigate this risk is ensuring that every road route passes through at least one security checkpoint. Using the Ford-Fulkerson maximum-flow algorithm, we generate a minimum set of checkpoint locations within a ring-shaped buffer area surrounding the 50 largest US urban areas. We study how the number of checkpoints changes as we increase the buffer width to perform a cost-benefit analysis and to identify groups of cities that behave similarly. The set of required checkpoints is surprisingly small (10-124) despite the hundreds of thousands of road arcs in those areas, making it feasible to protect all major cities.

INTRODUCTION

Road networks are integral to the economy and to society, but are also a source of vulnerability. The same streets that allow first responders to arrive quickly at the scene of a disaster can also allow a terrorist to bring a heavy weapon into the heart of a densely populated region. One approach to mitigate the risk of such an event is to deploy a system of checkpoints around locations we wish to protect in such a way that every incoming road route requires passing through one of these checkpoints. Depending on factors such as expected risk of attack and cost of implementation, checkpoints could be large radiation detector installations

where every suspicious vehicle is stopped and searched, locations of automated sensor systems that remotely alert authorities to suspicious vehicles, or, if the appearance of the perpetrator is known, locations to station police to visually inspect incoming traffic. The checkpoints should be far enough from the city center so that authorities have time to interdict an attacker before they reach the densely populated and developed city center. Thus the challenge is to identify locations that allow inspecting all incoming traffic, and hopefully to find that in large urban areas that have hundreds of thousands of road segments, the number of required checkpoints is small enough to make it practical to deploy the necessary security measures. In this paper, we address the problem of identifying the number and location of the needed security checkpoints, but do not elaborate on the specific security measures required at each checkpoint.

One approach to finding optimal locations for checkpoints is to search first for a minimum set of checkpoints for a selected region. That is, we seek to identify the smallest number of road segments such that every possible route into the region to be protected passes through a checkpoint. This set is known as the minimum cut set (MCS) in graph theory as cutting these links would completely disconnect the two portions of the network. As demonstrated by Barnett et al.[1], this problem can be solved on large urban road networks using the Ford-Fulkerson maximum flow algorithm.[2] Whereas the cost of an individual checkpoint depends on various factors including the road type, the number of lanes, and the checkpoint type, the MCS is a feature of the network structure and as such can be found without consideration of these factors.

Barnett et al. found that for New York City, more than 700 road arcs cross into a circle with a radius of 15 miles, centered at Times Square.[3] It is reasonable to expect that by taking advantage of natural features such as parks, bridges, and open spaces, the MCS would be smaller. By allowing the cut set to be located anywhere within a 30 mile wide band outside the 15 mile radius circle, Barnett et al. generated a cut set containing only 89 road segments.[4] This is a surprising and promising result, as it indicates that the size of the MCS of an urban road network is small relative to the number of incoming arcs. Thus the idea of protecting assets in a large urban area becomes feasible by securing access to a relatively small number of security checkpoints.

In this paper, we demonstrate empirically that the results from New York City are typical. Specifically, we present results on the size of cut sets of the 50 largest urban areas in the United States and on how the size varies as a function of the protected area and the size of the buffer zone.

We demonstrate how to use the minimum cut set results for protecting high value targets. We simulate two scenarios. In the first, the decision maker fixes the region to protect, and is free to place checkpoints anywhere outside that region. Extending the outer boundary of the buffer area in 1-mile increments, we observe the decrease in the number of checkpoints and see whether there is an optimum width for the buffer area. In the second scenario, the decision maker has a fixed jurisdiction, and would like to protect as many of its assets as possible while still only placing checkpoints within his/her jurisdiction. In this case, we fix the outer boundary of the buffer area and observe the increases in cost as the inner boundary is expanded in 1-mile increments. Results from both scenarios demonstrate that while there is a great deal of variety in the topology of urban road networks, even in the worst case the number of checkpoints required to defend the city is practicable.

METHODS

As in Barnett et al.[5], we only consider checkpoints on inbound road segments rather than checkpoints at intersections. Unlike checkpoints on intersections, checkpoints on road segments require stopping or otherwise investigating traffic in only one direction and therefore seem more practical.

Road network data was obtained from StreetMap North America, a dataset included in ESRI's ArcGIS 9.3 software. The data is based on 2003 TeleAtlas data. We followed the conventional method of encoding a road network as a graph by representing road segments as links and intersections as nodes. In the StreetMap data, some nodes exist only to record a change in road attributes, such as changes in speed limits or road types. While this affects some graph attributes (such as the degree distribution), it does not affect the size of the MCS.

We selected the 50 largest urban areas (UAs) in the US according to 2010 Census data.[6] The center for each UA was assigned to be the center of the most populous city included in the UA. For example, in the Dallas-Fort Worth-Arlington, Texas UA, Dallas has 1.24 million inhabitants, Fort Worth has 793 thousand, and Arlington has 376 thousand, hence the center of Dallas was chosen to be the center of the entire UA.

Deciding on where to place the inner and outer boundaries is a balancing act. In our study we use concentric circles to allow comparisons between the cities and with prior work.[7] As noted by Barnett et al., the methodology applies to more general buffer areas, such as county boundaries. A decision maker would be able to use the landscape to his advantage. Rivers, lakes, oceans, and mountains all provide natural barriers that limit the number of roads coming into a region. Boundaries of the buffer area can be placed intelligently to balance competing factors such as cost per checkpoint, travel times, political jurisdictions, economic importance of the region, locations of critical assets, and population density.

To observe the behavior of the size of MCSs with varying size of the buffer area, we

calculated the MCS on buffer zones with inner radius of 5, 10, and 15 miles, and increased the outer radius in 1-mile increments from the inner radius to a maximum of 45 miles.

Denote by $M(r_i, r_o)$ the size of the MCS in the buffer area with inner radius r_i and outer radius r_o. For fixed inner radius r_i, the size of the MCS decreases monotonically as the outer radius is expanded. To wit, if buffer area A is a subset of buffer area B, then any cut set for A is also a cut set for B. Hence the size of the cut set for A is an upper bound for the size of the cut set of B. Therefore, for our study, $M(r_i, 45)$ is the smallest MCS for buffer areas with fixed inner radius r_i. Furthermore, the MCS found when the inner and outer radii are equal is the number of inbound edges which gives an upper bound for the size of the MCS with fixed inner radius r_i.

To illustrate, Figure 1 displays for the city of Phoenix buffer areas bounded by circles with a 15 mile inner radius and 25 and 45 mile outer radii respectively. Inner and outer circles are colored in light blue and incoming road segments intersecting the circles are colored in bright blue. Road segments in the MCS are colored in red. The MCS for the buffer area with outer radius of 25 miles is 64 and it decreases to 36 by expanding the outer circle to a 45 mile radius.

Figure 1: MCSs for Phoenix for buffer zones with inner circle of 15 miles and outer circles of 25 (image 1) and 45 miles (image 2). Arcs crossing the inner and outer circles are colored in blue and arcs in the cut set are colored in red.

Image 1 - Phoenix M(15,25) = 64

Image 2 - Phoenix M(15, 45) = 36

We developed a tool using standard Java libraries to extract a graph representation of a city from a database and implemented the Ford-Fulkerson maximum flow algorithm to find the MCS. In our implementation of the Ford-Fulkerson algorithm, all nodes inside the inner radius are collapsed into a supersink, and all nodes outside the outer radius are collapsed into a supersource.

Even for the largest buffer area in our study (inner radius equal to 5 miles, outer radius equal to 45 miles) on the largest city (New York City), an area containing 560,051 nodes, the Java implementation took less than 10 minutes to find the MCS.

A cost-benefit analysis for checkpoint placement would weigh the benefits of expanding the size of the protected area against

the costs incurred by increasing the number of checkpoints.

To quantify the protected assets, we use the total street mileage. This measure implicitly weights more heavily those regions that are dense with homes and businesses. A simple metric for the benefit from choosing inner radius r_i is the proportion of roads within the disk with radius r_i, as measured by the total miles of roadway inside the disk divided by the total miles of roadway in the urban area.

Costs to deploy a checkpoint network depend on geography, property values, and the number of checkpoints required. All these attributes are city dependent. To enable comparisons across the 50 US urban areas, we instead focus on the number of checkpoints. The cost of protecting a larger region can be approximated by fixing the outer radius at 45 miles and by computing the number of checkpoints as the inner radius increases from 5 to 40 miles in increments of 1 mile.

To determine which cities behave similarly, we perform k-means clustering on three sets of data, corresponding to the three inner radius lengths of 5, 10, and 15 miles. For each of the 50 UAs, the variables are the sizes of the MCS on each of the possible buffer areas. Thus for r_i=5, there are 41 variables, for r_i=10 there are 36 variables, and for r_i=15 there are 31 variables. To prevent large numbers from skewing the results, we standardize each variable by subtracting the mean and dividing by the standard deviation.

K-means clustering is a simple and widely used method for identifying clusters in quantitative data. We refer the reader to references such as Hastie et al. for a detailed overview of the method.[8] There are a few drawbacks to using k-means.[9] Clusters that don't resemble an n-dimensional ball won't be recognized well. Centers for the k clusters are initialized with a random sample, so there is inherent variability in the produced clusters. To ameliorate the latter, it is common to generate many random initial cluster centers and choose the cluster that minimizes the overall within cluster sum of squares. Another issue is that the number of clusters, k, must be specified by the analyst. The algorithm will always sort the observations into k groups, regardless of whether the observations exhibit any real clustering. Many methods to estimate k have been proposed. We employ the Gap statistic, which has the advantage of being able to recognize when only one cluster is present in the data.[10] Discussion of the merits and weaknesses of a sampling of the other methods can be found in Walther and Tibshirani.[11]

Results

We first establish a context for the results on New York City found by Barnett et al.[12] Denote by $M(r_i, r_o)$ the size of the MCS in the buffer area with inner radius r_i and outer radius r_o. Barnett et al. used JServer data and found $M(15, 45)$ = 89. With the ESRI StreetMap data, we found $M(15, 45)$ = 87. This result shows remarkable agreement – the slight discrepancy is likely due to different choices for center locations (Times Square versus our choice of City Hall). The histogram in Figure 2 shows the distribution of $M(15, 45)$ over the 50 UAs. The median for $M(15, 45)$ is 55.5, and the quantity varied from 10 (Las Vegas) to 124 (Philadelphia). With $M(15, 45)$=87, New York requires more checkpoints than do most, but not all, other cities. This makes sense, for while New York City has more road segments than does any other city in our data, a significant portion of the boundary of the city is open water. Cities requiring a larger number of checkpoints than New York City include Atlanta (98), Pittsburgh (99), Indianapolis (107) and Philadelphia (124). None of these urban areas have significant natural barriers preventing expansion. The rivers cutting through Philadelphia are both within the 15-mile inner circle, and thus do not provide a natural barrier in this case. It is surprising and worrisome that for a few cities, $M(15, 45)$ is as small as 10. Although such a small cut set makes installation of a checkpoint network feasible, it also represents a major vulnerability. For such cities, damage to just a few roads would significantly interfere with regional traffic, and even isolate the city's road network from the rest of the country.

Figure 2: Distribution of the size of the MCS

Figure 3: Distribution of Decrease of the MCS by considering a 30 mile wide buffer zone

Distribution of M(15,45)

Distribution of M(15,45)/M(15,15)

The distribution of the ratio of the size of the MCS and the number of roads entering the protected area ($M(15, 45)/M(15, 15)$) is depicted by the histogram in Figure 3. Details for a selection of cities are shown in Table 1.

This ratio represents the relative decrease of the number of checkpoints needed compared to the baseline number of inbound roads into the inner circle of radius 15.

Table 1: MCS decrease by considering a 30 wide mile buffer area

City	M(15,15)	M(15,45)	M(15,15)/ M(15,45)	MCS Reduction
Miami	309	16	5%	95%
Virginia Beach	167	10	6%	94%
Los Angeles	650	61	9%	91%
New York City	759	87	11%	89%
Atlanta	478	98	21%	79%
Indianapolis	293	107	37%	63%
Oklahoma City	215	81	38%	62%
Columbus	188	91	48%	52%

For the 50 cities studied, the gain in a reduced number of checkpoints ($1-M(15, 45)/M(15, 15)$) ranged from 52% to 95%, with median 81%. For New York City, the reduction gain is 89%, placing it in the upper quartile. At the high end, we find that Miami (95%) and Virginia Beach (94%) benefit the most by considering a 30 mile wide buffer area. At the low end, the cities that benefit the least are Oklahoma City (62%),

and Columbus (52%). The range of the quotient $M(15, 45)/M(15, 15)$ values is somewhat surprising, but an inspection of each of the cities helps explain it. For example, for both Miami and Virginia Beach it makes sense that the ratios $M(15, 45)/M(15, 15)$ are so much smaller than those for Oklahoma City and Columbus. Both Columbus and Oklahoma City are small enough and lie almost completely within the

15-mile-radius disk, leading to a small value of $M(15, 15)$. Since neither Oklahoma City nor Columbus has any substantial natural barriers nearby, there are roads entering the cities from every direction, increasing $M(15, 45)$. Thus the ratio $M(15, 45)/M(15, 15)$ is especially large.

For Miami and Virginia Beach, the opposite is true. The buffer zone of Miami, being squeezed between the Everglades and the Atlantic, is long and narrow. Hence the 15-mile circle intersects dense urban areas in the north and the south of the city, increasing $M(15, 15)$. Since so few roads enter Miami, $M(15, 45)$ is especially low. Similarly, although 75% of the 15-mile circle centered at Virginia Beach is over water, the remaining portion intersects the densely populated area near Norfolk. The presence of large natural barriers including the Atlantic, two wildlife refuges, and Hampton Roads allows for $M(15, 45)$ to be extremely low.

Next, we investigate the overall behavior of the size of the MCS as we increase the buffer zone. We fix the inner circle and increase the buffer zone by adding one mile at the time to the outer circle radius. Figure 4 shows the decrease in the size of the MCS as a function of the outer radius for four representative cities and for each of the three choices for inner radius (5, 10, and 15 miles). Plots for the full set of 50 cities are included in the Appendix. As expected, the size of the MCS decreases as the size of the buffer area increases, but in most cases, it becomes constant rapidly. The largest decrease happens in the initial one-mile wide buffer area that is not shown in the plots since the size of the MCS $M(15, 15)$ is too large to be displayed, but Table 2 shows the initial mean reduction in the size of the MCS for the one mile wide buffer area (1- $M(15, 16)/M(15, 15)$) for the three inner radius values.

Figure 4: Decrease in MCS as outer radius increases

Table 2: Percent change in the size of the *MCS* from baseline achieved via 1-mile buffer area

Inner radius ri (mi)	Mean Percent decrease	Std. dev.
5	62%	8.5%
10	66%	5.4%
15	63%	7.0%

As we increase the buffer zone (by choosing a larger outer radius), the size of the minimum cut set decreases at a lower rate and in most cases, it rapidly levels off (see Figure 4, Figure 9, and Figure 10). Summary statistics for that distance are shown in Table 3.

Table 3: Mean distance at which the size of the *MCS* levels off

Inner Radius ri (mi)	Distance (miles)	Std. dev. (miles)
5	17.0	10.1
10	16.3	9.4
15	16.2	7.8

We find that for each choice of r_i, by the time the buffer area is about 16 miles wide, most of the decline in the size of the MCS has occurred. There are, of course, a few cities where this distance is much larger. For r_i=5 miles, these cities are New York City (r_o=45), Houston (r_o=42), Phoenix (r_o=42), and Seattle (r_o=42). In each of these cases, there are geographic features such as parks, mountains, and large bodies of water that allow further drops in the size of the MCS past the initial leveling out. For example, in Houston, the size of the MCS first flattens out at r_o≈30 miles, but beginning at r_o=40 miles the outer radius intersects the ocean, producing a sudden small drop.

Next we investigated the benefit of increasing the outer radius. We fixed the inner radius r_i and found the size of the MCS $M(r_i, r_i)$, namely the baseline number of inbound roads into the inner circle of radius r_i. Then we followed how the size of the MCS decreases (compared to $M(r_i, r_i)$) as we increased the outer radius and calculated the percent decrease in the size of the MCS. We performed this calculation for all the cities and kept track of how many cities had an MCS size decrease of at least *p* percent. To summarize, for each choice of outer radius and a fixed *p*, we calculated the number of cities for which the percent decrease in the size of the MCS was at least *p* (compared to $M(r_i, r_i)$). We made this calculation for all p values between 50% and 99% using 1% increments. Figure 5 depicts this situation: in the figure, the y-axis provides the outer radius r_o, the x-axis provides the percent reduction p, and for each combination of *(x, y)* values, the shade of blue indicates the number of cities for which $M(15, r_o)$ is p percent smaller than $M(15,15)$. For example, by extending the outer radius to 30 miles, we find that for more than 45 cities, $M(15, 30)$ is 64% smaller than $M(15, 15)$, for between 30 and 35 cities, $M(15, 35)$ is 78% smaller than $M(15, 15)$, and for between 5 and 10 cities, $M(15, 45)$ is 88% smaller than $M(15, 15)$. The plot demonstrates that for most of the 50 cities, reductions in the size of the MCS from increasing r_o occur early on – only for a few cities does an expansion of the outer radius past 30 miles provide any further reduction in the number of cuts.

Figure 5: Distribution of the number of cities that have an x% size reduction for the minimum cut as the outer radius increases (with fixed 15 miles inner radius).

Reductions in cut set size, inner radius 15 mi

Percent reduction in cut set size

CLUSTERING

As mentioned in the methodology section, for each of the three different values for the inner radius (5, 10, and 15) we have a sequence of MCSs that were obtained by increasing the outer radius one mile at a time until it reached 45 miles. We used each of these sequences of MCS sizes to cluster the 50 cities. From the Gap statistic we find that the only dataset that appears to exhibit a clear degree of clustering is when the inner radius is equal to 5. We ran the *k*-means algorithm with 25 starts and chose the clustering that minimized the within cluster sum of squares.

Figure 6: Outer radius vs. within-cluster mean cut set size

Figure 6 plots on the left hand side the individual MCSs colored by cluster, and on the left hand side the resulting within-cluster means, plus or minus one standard deviation. Most of the variation is in the first few miles. This is apparent from the overlap in error bands for the within cluster mean in the first ten miles, and from how much larger those error bands are in that range. Note that after about 16 miles, the expected distance at which the size of the cut set levels off, the different cluster bands for the within cluster means do not intersect anymore. For the most part, the clustering algorithm appears to have sorted cities by the size of the MCS. The exact composition of the clusters is contained in Table 4.

Table 4: Results of clustering algorithm.

Cluster 1	Cluster 2	Cluster 3	Cluster 4
San Francisco	Miami	New York City	Los Angeles
San Diego	Tampa	Washington, DC	Chicago
Las Vegas	Baltimore	Boston	Philadelphia
Virginia Beach	Riverside	Detroit	Dallas
Jacksonville	Portland	Phoenix	Houston
Salt Lake City	Sacramento	Seattle	Atlanta
New Orleans	San Jose	Denver	Minneapolis
	Orlando	St. Louis	Pittsburgh
	Austin	Cleveland	Kansas City
	Charlotte	San Antonio	Indianapolis
	Memphis	Cincinnati	Oklahoma City
	Louisville	Milwaukee	
	Nashville	Columbus	
	Richmond	Providence	
	Buffalo	Hartford	
	Bridgeport		
	Raleigh		

A few comments on the clusters:

- The first cluster contains cities where the MCSs are unusually small. These cities are characterized by having major natural barriers nearby, such as mountain ranges near San Diego and Salt Lake City, and large bodies of water near San Francisco, New Orleans, and Virginia Beach.

- The cities in clusters two and three behave similarly. For most of these cities, the minimum size for the cut set is reached quickly. For the second cluster, the size of the MCS is between 25 and 45. For the third cluster, the minimum tends to be between 45 and 65.

- The fourth cluster contains cities whose MCSs are unusually large. For the most part, these are cities that have very few natural barriers inhibiting their growth. The exceptions to this are Los Angeles and Chicago, both of which border a large body of water. These cities have much larger road networks than do the other cities in the cluster, which may be the reason that their MCSs are so large.

We expect that not all cities fit perfectly within their assigned cluster. The k-means algorithm requires that every observation be assigned a cluster, even if the connection is tenuous. However, as long as our interpretation of the resultant clusters allows for porous boundaries, the algorithm helps us make rough groupings of cities and discern similarities.

A few cities in particular don't seem to fit well in their assigned (or any) cluster. For example, whereas the decline in MCS size flattens out eventually for most cities, the curve for Houston has two distinct drops after the initial decline. These correspond to the distances where the inner radius first intersects Trinity Bay and the Gulf of Mexico, respectively. Miami has a similar drop corresponding to the southern extent of the city.

Cost-Benefit Analysis

We selected one city from each cluster to illustrate the cost-benefit analysis. Figure 7 depicts the increase in number of checkpoints for each mile added to the radius of the protected area and the corresponding increase in the percentage of road-miles within the inner circle. Analogous to what we observed with expanding the outer radius, there is an uptick in cost when the inner radius approaches the outer radius. The figure shows the costs up to r_i=44, so as to allow some reduction in cost from $M(45,45)$. The colors in the plots represent the percentage of road-miles in the 45-mile-radius disk that are within the inner radius.

Figure 7: Number of checkpoints versus inner radius for four cities. Percentage of road miles protected is indicated by color.

Table 5 reports the number of checkpoints for a selection of choices for r_i.

Table 5: Total checkpoints for representative cities

City	M(5,45)	M(15,45)	M(25,45)	M(35,45)	M(40,45)
San Francisco	17	39	46	84	95
Austin	33	46	65	88	109
New York City	46	87	109	128	168
Philadelphia	88	124	143	177	204

From these plots it is clear that the cost per mile increase of the inner radius is highly variable. Interestingly, we see a few places where a free increase of coverage is possible; that is, where the inner radius can be increased without increasing the number of checkpoints. For example, in New York City, it takes the same number of checkpoints (83) to protect an area with radius 8 miles as it does to protect an area with radius 12 miles – more than doubling the protected surface area.

By increasing the number of checkpoints to 89, an area with radius 19 miles can be protected, representing more than fivefold increase of protected area for a relatively small additional cost. Looking at the percentage of road-miles protected, the change from r_i=8 to r_i=19 provides an increase from 6% to 32% – a substantial gain.

Whereas both San Francisco and New York City can benefit from these free coverage increases, for Austin and Philadelphia, the cost increases with almost every increase in the inner radius.

Figure 8: Total number of checkpoints for the 25 most populous cities. Changes in color represent increases in the percentage of protected road-miles within the inner radius.

Finally, we look at the problem of protecting a large set of cities. We focused only on the 25 most populous cities: for each choice of inner radius we calculated the total number of checkpoints required to protect them and the percentage of road miles within each inner circle. The results are displayed in Figure 8. We find that 50% of the road miles in these cities can be protected by setting $r_i = 26$ on each city, corresponding to a total of 1,841 checkpoints. With $r_i = 5$, only 5% of road miles are protected, at a cost of 1,155 checkpoints. Expanding the inner radius to 44 miles, while covering 97% of road miles, has a cost of 4,602 checkpoints.

Conclusions

By implementing the procedure outlined in Barnett et al. for the 50 most populous urban areas in the United States we confirmed that the problem of finding a minimum cut on urban road networks can be solved using publicly available Java libraries on standard desktop computers. To our knowledge, the behavior of MCS within buffer areas on urban road networks has not previously been investigated. We discovered that there are substantial similarities among the 50 urban road networks in our study, particularly in that the number of checkpoints required to protect urban areas drops precipitously as the size of the buffer area increases until a minimum value is reached. This value is usually reached within a 16-mile-wide buffer area, although geographical features such as mountains and lakes can allow for further decreases. We found that for the 50 most populous regions of the US, the size of the MCS ranged from 10 to 124, with median value 55.5. This indicates that for even very large urban areas, it is feasible to use a checkpoint network to protect against terrorism. We also see that a few of the largest cities in the United States are particularly vulnerable to disruptions in the road network, in that damages to a small number of roads could severely limit connections to the regional road network, or even completely prevent all road-based transportation to and from the city.

Through our cost-benefit analysis we found that although in general the number of checkpoints rises with each increase in the size of the protected area, for many cities, it is possible to greatly increase the length of the inner radius without incurring additional cost. This demonstrates that there is no one-size-fits-all solution for an optimum buffer area configuration. Rather, unique features of road network topology and of the landscape surrounding cities allow for some cities to protect a larger area at little or no increase of cost, whereas for other cities, every increase in the inner radius requires a larger number of checkpoints.

Future Work

Road segments do not have equal importance. By design, urban road networks have a hierarchy of road types, from low-speed low-throughput residential streets to freeways and major arterials. Thus some road segments will make installation of checkpoint systems more expensive. The MCS is based on the sum of the weights assigned to the edges. In our study, each road segment was given weight 1, so that the sum of the edge weights gave the number of segments contained in the cut set. By encoding each street type with the approximate dollar cost to deploy a checkpoint and by assigning that cost as edge weights for the road network, a minimum cut set can be calculated that minimizes the estimated dollar cost to install a checkpoint network.

Geography plays a major role in the defensibility of urban areas. Future work could include optimizing the buffer area based on landscape features such as coastlines, mountain ranges, rivers, and protected areas such as parks and wildlife refuges. Although some (particularly Midwestern) cities are not surrounded by natural barriers, large cities on the coasts are well-suited to such an approach. Many developed areas have oblong shapes, or even multiple dense centers, such as the Los Angeles region with centers in Los Angeles, Riverside, and Santa Ana, or the Washington, DC – Baltimore, MD region. Rather than arbitrarily choosing one size of circle to compare cities, it would be valuable to develop criteria for selecting protected areas based on population density, economic value, and landscape features. This would better match the approach that would be used in practice.

Further research should also incorporate water, rail, and air transportation, including existing border and port defenses.

Appendix

Figure 9 and Figure 10 show for fixed inner radius (5, 10,15) plots for the size of the minimum cut set (MCS) as a function of the outer radius for the 50 most populous urban areas (UAs). UAs are ranked by population size in decreasing order.

Figure 9: MCS size as the outer radius increases for the 25 most populous UAs.

Figure 10: MCS size as the outer radius increases for UA ranked 26th to 50th.

About the Authors

Daniel Watkins received his M.S. in mathematics from Drexel University. He has worked in homeland security topics at Los Alamos National Laboratory (LANL) and Pacific Northwest National Laboratory, and is currently pursuing a Ph.D. in Atmospheric Science at Oregon State University

Leticia Cuéllar received her Ph.D. in Statistics from the University of California at Berkeley and is a scientist at LANL. She has worked for various programs for the Department of Homeland Security (DHS) including the National Infrastructure Simulation and Analysis Center (NISAC) and risk analysis of nuclear smuggling. She may be contacted at leticia@lanl.gov

Deborah Kubicek received her Masters of Computer Engineering from the National Technological University and her Bachelors of Computer Science from the University of New Mexico. She is software engineer at LANL where she has worked on various projects for the DHS including the protection of national infrastructure and risk analysis of nuclear smuggling.

Erick Rodriguez is studying Mechanical Engineering at the New Mexico Institute of Mining and Technology and has worked at LANL.

Phillip Stroud is a scientist at LANL. He received his Ph.D. in Nuclear Engineering from the University of Illinois. Research areas include missile defense, radiation detection systems, and reducing risk from weapons of mass destruction.

Acknowledgement

Disclaimer

The views expressed here are solely the authors' and do not reflect the view of the Domestic Nuclear Detection Office.

Notes

1.   Robert L. Barnett, D. Sean Bovery, Robert J. Atwell and Lowell Bruce Anderson, "Application of the Maximum Flow Problem to Sensor Placement on Urban Road Networks for Homeland Security," *Homeland Security Affairs* 3, no. 3 (2007).

2.   L.R Ford and D. R. Fulkerson. "Maximal flow through a network," *Canadian Journal of Mathematics* 8 (1956): 399-404.

3.   Barnett et al., "Application of the Maximum Flow Problem," (2007)

4.   Barnett et al., "Application of the Maximum Flow Problem," (2007)

5.   Barnett et al., "Application of the Maximum Flow Problem," (2007)

6.   U.S. Census Bureau, "2010 Census Urban Area List," 2010, http://www2.census.gov/geo/ua/ua_list_all.xls (accessed October 22, 2014).

7.   Barnett et al., "Application of the Maximum Flow Problem," (2007)

8.   Trevor Hastie, Robert Tibshirani, and Jerome Friedman, *The Elements of Statistical Learning*, (Springer, 2009).

9.   Anil K. Jain, "Data clustering: 50 years beyond K-means." *Pattern Recognition Letters* 31, no. 8 (June 2010), 651-666.

10.   Robert Tibshirani, Guenther Walther, and Trevor Hastie, "Estimating the Number of Clusters in a Data Set Using the Gap Statistic," *Journal of the Royal Statistical Society* B 63 (2001), 441-423.

11.   Guenther Walther and Robsert Tibshirani, "Cluster Validation by Prediction Strength," *Journal of Computational and Graphical Statistics* 14, no. 3 (2005), 511-528.

12.   Barnett et al., "Application of the Maximum Flow Problem," (2007)