

Geovisualization applications to examine and explore high-density and hierarchical critical infrastructure data

Robert Edsall,^a Harvey Hembree^a

^a *National and Homeland Security Division, Idaho National Laboratory*

Abstract: The geospatial research and development team in the National and Homeland Security Division at Idaho National Laboratory was tasked with providing tools to derive insight from the substantial amount of data currently available – and continuously being produced – associated with the critical infrastructure of the US. This effort is in support of the Department of Homeland Security, whose mission includes the protection of this infrastructure and the enhancement of its resilience to hazards, both natural and human. We present geovisual-analytics-based approaches for analysis of vulnerabilities and resilience of critical infrastructure, designed so that decision makers, analysts, and infrastructure owners and managers can manage risk, prepare for hazards, and direct resources before and after an incident that might result in an interruption in service. Our designs are based on iterative discussions with DHS leadership and analysts, who in turn will use these tools to explore and communicate data in partnership with utility providers, law enforcement, and emergency response and recovery organizations, among others. In most cases these partners desire summaries of large amounts of data, but increasingly, our users seek the additional capability of focusing on, for example, a specific infrastructure sector, a particular geographic region, or time period, or of examining data in a variety of generalization or aggregation levels. These needs align well with tenets of information-visualization design; in this paper, selected applications among those that we have designed are described and positioned within geovisualization, geo-visual analytical, and information visualization frameworks.

Keywords: geovisual analytics, homeland security, resilience, decision support, design

1. Introduction: Critical infrastructure protection and visual analytics

The Presidential Commission Report on Critical Infrastructure Protection (1997) highlighted the complexity, vulnerability, and interconnectedness of our nation's critical infrastructures (CI). An "infrastructure," according to this report, is a network of independent and mostly privately-owned systems that function to provide essential goods and services, and a critical infrastructure is one whose incapacity or destruction would have a debilitating effect on our defense and national security (Lewis, 2006). Currently, there are sixteen sectors that are included in the definition of CI of the Department of Homeland Security (DHS), including energy, food and agriculture, information technology, and transportation systems (PPD 21, 2013); such systems, at the time of the 1997 report, were becoming increasingly vulnerable to physical and cyber-attacks that could result in detrimental service disruptions. Following this report, and particularly after the September 11 attacks, a series of directives, executive orders, and policies were implemented to strengthen the security and resilience posture of the United States with regard to its critical infrastructure. At the forefront of these policies and efforts is a need to improve ability to make decisions and analyze risk associated with protective and support measures associated with man-made and natural threats to critical infrastructure (National Research Council, 2007).

The DHS is the primary government organization tasked with the analysis and protection of critical infrastructure of the United States. DHS provides resources for its own analysts – as well as public and private sector entities, organizations, and companies that own or manage critical infrastructure – to manage risk, mitigate hazardous impacts, and strengthen resilience of the national CI (Department of Homeland Security, 2017b). Decisions about CI security often must be made under constrained time, personnel, and monetary resources (Department of Homeland Security, 2016). In the past, decision making was often hampered by a lack of data; now, of course, the opposite is the case: we lack the tools to efficiently aggregate, analyze, and visualize the tremendous amount of data needed for decision making regarding critical infrastructure. Analysts, responders and decision makers have many disparate data sources associated with critical infrastructures available to them that include web-based data feeds associated with threats such as weather, earthquake, etc., static geospatial data layers, reports generated from assessments, and vulnerability information. At subsequent stages of the decision-making process, the original data are interpreted and analyzed to produce information useful to decision makers.

Given this general scenario requiring the time-critical understanding of, and decision-making about large, multi-dimensional data sets, the analysis of critical infrastructure is apt for the application of the theories of visual analytics (Thomas & Cook, 2005). More

specifically, because the understanding and visualization of the spatial component of these data sets is vital, our innovations adopt, and adapt, research in geovisual analytics, which melds practices of geographic visualization – including interactive interfaces, geographic data modeling, and statistical and map-based graphics – with current understanding of human reasoning, cognition, knowledge construction, and decision making ((G. Andrienko et al., 2007; G. Andrienko, N. Andrienko, Keim, MacEachren, & Wrobel, 2011; Tomaszewski & MacEachren, 2012). The development of capabilities in response to both the strategic directions of the DHS (Department of Homeland Security, 2016) and the technologies vision report of the National Geospatial-Intelligence Agency (National Geospatial-Intelligence Agency, 2014) represents an opportunity to connect theories of human-reasoning-centered GVA research with real-world analysis problems safeguarding the social and economic well-being of the United States.

DHS officials and analysts recognize the limitations of the current tools for effectively utilizing the vast and disparate information stored in various digital repositories and in paper documents. Motivated by the possibilities of spatial, temporal, and statistical insight generation from the data, as well as the importance of facilitating rapid decision making, DHS supported the construction and customization of information and geographic visualization tools to explore, and “tell stories” with, the CI data under their management. These projects include the customization of a standard web GIS toolkit, the development of visual and computational links between spatial and graph-based databases to examine CI relationships and dependencies, the adaptation of open-source graphical and map-ping JavaScript libraries to perform basic data analytics of program- and mission-specific data, and the design of novel interactive mapping projects to view the results of risk and vulnerability assessments that until now have been printed in lengthy paper reports.

For the geospatial capabilities team at INL, the development of tools typically begins with a needs assessment among the target users of the products. This is frequently framed within a set of scenarios and use cases; subject-matter experts, analysts, DHS staff and leadership, and/or stakeholders associated with CI will come to our team with problems and situations that current tools fail to address effectively. The designs described in the following sections exemplify this approach; because our user base is specialized and very limited (special training and clearances are required to access much of the data), formal user testing is not practical, but iterative design based on frequent interactions with our customers leads us to successful end products that drive decisions about, and enable insight from, CI data.

The tools and techniques described in this paper represent a subset of the applications of (geo)visual analytics and dynamic geographic visualization developed by the geospatial research and development team in the National and Homeland Security Division at Idaho National

Laboratory to support the various CI protection efforts of the DHS. These applications span the spectrum from those designed for hypothesis generation and data exploration to those meant to communicate results of analyses and assessments in a concise but data-rich manner. We will introduce each specific application with a brief description of the framework or program the application is meant to support, an explanation of the problem or scenario the application is designed to address, and a demonstration of the typical design and development workflow we employ to deliver these tools to DHS for implementation in real-world scenarios.

2. Infrastructure visualization in the IP Gateway

The Infrastructure Protection (IP) Gateway is a web-based repository for CI information products in a wide variety of formats, secured and accessible only by designated users (in various roles) designated by DHS¹. The repository’s tools provide access to these products in several ways, including a searchable digital library of written digital documents, a dashboard for examining user log files, standard folder systems for file storage, and a customized ESRI-based mapping interface called Map View that allows customized access to spatially referenced information about CI facilities and associated map layers including some housed on the Gateway and others available to the general public. This disparate array of data formats provides challenges typical in the age of big data: rafts of data that are potentially vital to analysts but stored in relatively inaccessible forms within paper reports, gigantic data-bases, or cluttered diagrams.

The Map View of the IP Gateway is designed first and foremost as a viewer of the individual critical infrastructure assets across the nation. The locations and attributes for each asset – essentially the records in the databases visualized on the Map View – come from a variety of sources, but a primary and high-confidence source is the re-reports of DHS protective security advisors, local field agents who visit CI facilities to conduct interviews and assessments, ground-truth locations and attribute information, and subsequently contribute their findings to the database in the Map View. When the database was relatively small, the Map View was a Google Maps application with markers showing the location of the visited sites, and satisfied the initial requirements as a simple viewer of the facilities’ locations. However, as more facilities were added, as other databases were folded into the application, and as analyst, design, and leadership teams began to envision scenarios and uses of the Map View that exceeded the capabilities original design, our group adopted and customized an application based on ESRI technology that could handle these kinds of larger databases and more advanced analysis.

¹ For this reason, the examples and figures throughout this paper will be based on hypothetical situations, studies, and data sets only.

2.1 Problem: visualizing spatial variability in point density

An immediate concern as the number of records in the database grew in the Gateway was handling the large number of records in our databases. Not unlike a listing of many different kinds of phenomena at a national scale, a dot map of the facilities would amount to hundreds of thousands of point features with a great spatial variability of densities across the country (Figure 1, showing just one of the available databases of national CI). This is over-whelming for a user and allows for no other data layers to be visualized at the same time. Indeed, at the national scale, the variation in density of the points is both predictable and difficult to comprehend if each facility is represented by a single point marker (at that scale, the difference between 10 and 1,000 points in an urban are, for ex-ample, is visually indistinguishable). The spatial variation is dramatic: not surprisingly, facilities are clustered around population centers, with dozens of points in an urban center; in rural areas, only one or two facilities might be located across tens or hundreds of square kilometers. However, this variability results not only from population density changes: the relative completeness of the database also varies spatially, as some individual officials who contribute have chosen to enumerate facilities in different ways than other officials (or have simply been more diligent). This information may be more useful to an analyst in understanding a phenomenon or planning future assignments for further additions to the database.

Previous strategies for overview of the spatial distribution an aggregation and summary of the facilities through a so-called “cluster map,” which grouped facilities by enumeration units defined by political boundaries (multi-state regions, states, counties), displaying the aggregated data as circles in the approximate center of the region. The circles were simple and direct but were of course less than optimal to communicate detailed information about the underlying data. We also implemented a version of a “heat map” using one of the included “renderers” in the ArcGIS JavaScript API, which applies a combination of color and opacity of an interpolated surface to encode spatial intensity of points. The representation affords some (limited, somewhat ambiguous) querying, with a mouse-click displaying information about the facilities near the clicked pixel. Additionally, the image is redrawn with every zoom action, allowing multi-scale visualization of the density of the points. However, users found it difficult to understand the interpolation concepts behind the representation and the methods of pixel-coordinate (as opposed to projected-geographic coordinate) blurring that changes the appearance of the map with each zoom level: areas that appeared to be somewhat densely covered in one zoom level become absent of facilities on the next-closest zoom level. The spatial variability of the facilities, thus, remained an important characteristic of the data set that had eluded effective visualization.

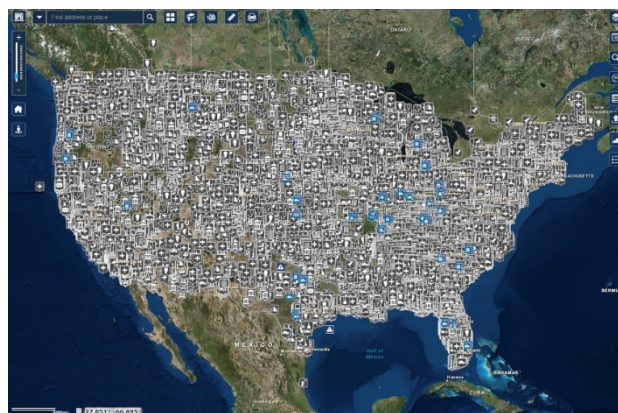


Fig. 1. Cluttered representation of dense and spatially variable national critical infrastructure.

2.2 Problem: visualizing attribute variability in sector dominance

There is also attribute variability of the points that is important to our end users – primarily a categorical attribute of infrastructure sector. Our database categorizes critical infrastructure into sixteen sectors, such as communications, emergency services, government, and municipal water facilities (the database also contains other categorical variables that might be of interest to an analyst; e.g. responsible agency or owner, presence or absence of vulnerability assessments for the facility, level of assessed resilience in various categories). There are regional variations in the dominance of a particular sector – for example, in the Gulf Coast, there are relatively more petroleum sector facilities, and in the Pacific Northwest, there is a relative overabundance of hydroelectric facilities. Once again, biases or lack of completeness in the database can also appear in this context, as some sectors are overrepresented because those facilities might be reported as “critical” by local officials in one area but not in another. In either case, important information is contained in a quantification of sector dominance – or lack thereof – in a given area.

Given this interest in sector information among the facilities in a region, and other attribute characteristics of the region, such as total number, average importance, dependency on other sectors or regions, or economic or societal value of the facilities, we needed to provide the capability to visualize summary statistics about regions. Once again, this is a capability that was lacking with the cluster maps, whose summary statistic was a simple count the total number of points in a region (regardless of the definition of the region). Categorical information such as a region’s dominant sector is appropriately encoded using hue differences. The heat map solution uses color (hue, value, or saturation, depending on the choice of the designer) to represent intensity but not a secondary variable, as we need. Again, both cluster maps and heat maps, while intuitive and familiar, fell short of the capabilities for visualization we were hoping to provide our end users.

2.3 Operational solution: Spyglass tool

In order to enhance the utility of the heat map representation, we chose to adapt the spyglass tool, which ESRI has implemented in their Story Maps applications (ESRI, 2016). In the traditional spyglass implementation, a second map, georeferenced to the default map, is revealed through a round tool that can be dragged across the map. Through this lens, a different theme, a different base map, a georeferenced image view, or perhaps a historic map can be seen, providing a second perspective on the region of interest.

Because our interest is a closer focus on the individual points (facilities) than is provided by the heat map, our spyglass is actually a magnifier, revealing a zoomed view of large enough scale to differentiate individual facilities (we also made it a rectangle to maintain a consistent look and feel to the rest of the application). In this sense, a more appropriate name for the tool might be a “loupe,” though we chose to retain the existing tool title for its familiarity (Figure 2).

Because the centers of the two views were identical, we added a small extent window on the tool as it is dragged around the map to help orient the user. We also added the capability to adjust the zoom level of the inset map, but we found the tool became difficult to use if the inset scale was more than 800% (three zoom levels) larger than the main map, lest a one-pixel move of the tool result in a disorienting large change on the inset. Thus, context of the smaller scale representation of spatial variability can be maintained while providing focus on individual facilities with the larger-scale representation.

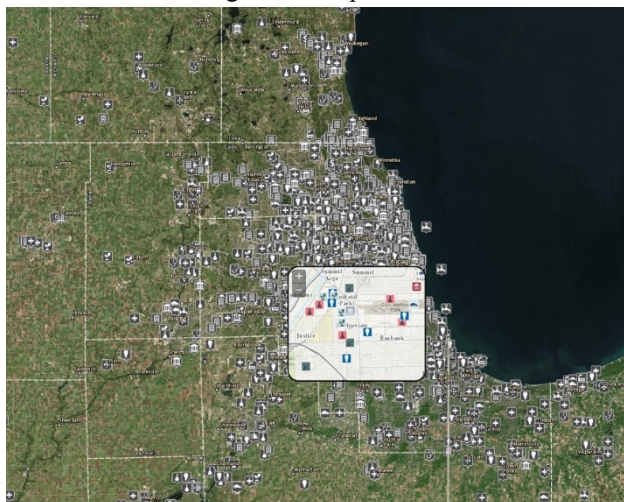


Fig. 2. Spyglass tool showing magnified area with detail about individual facilities while retaining regional context.

In order to examine sector dominance, the spyglass tool also includes a dynamic bar graph (histogram) window that displayed the frequency of the occurrences of facilities and assets belonging to one of several sectors. As a user moves the spyglass tool through space and scales, the bar graph updates with the frequency of facilities that are in the current view, allowing the visualization of the most common sector(s) of the facilities therein. We adopted this (rather basic) technique

rather than a more complex multivariate representation, such as a parallel coordinate plot, because the core users of the system are neither trained statisticians nor particularly sophisticated visualization users. In addition, the bar graph matches the style and complexity of the existing graphics and tools in the application, and we determined that its familiarity requires little training and enables rapid decision making.

Our use cases for this tool include the analysis of a multi-state event like a hurricane’s landfall or blizzard. In such a scenario, critical infrastructure at risk of impact from the event could be visualized in general using a spatial query using an existing polygon (e.g., an “uncertainty cone” of a hurricane track) or a user-generated polygon. This need to focus on individual facilities while maintaining context may also be important in a situational awareness scenario, when an analyst may wish to survey a nationwide state of affairs but quickly examine an area of concern (where, say, several events may be occurring or scheduled). In both cases, point data— relevant infrastructure, in our context — would appear aggregated and rendered on the small-scale map, and disaggregated — colored by sector, sized by importance, or otherwise varied visually to encode important attributes — through the use of the spy-glass tool in the inset window.

3. Pilot data analytics tools to explore infrastructure protection data

Given the success of the implementation of coordinated statistical graphics like the spyglass tool and histogram with the Map View, we were tasked with the development of a pilot set of demonstrations to more fully explore IP Gateway data — both “mission-specific,” influencing decisions about preparedness for and response to terror threats and natural hazards, and “programmatic” data, which concerns prioritization of resources and tracking of critical infrastructure protection programs in DHS.

Similar to the development framework described above for the Map View application, we began our construction of the data analytic tools by formulating, through consultation with our target users, several hypothetical scenarios in which these sorts of tools would be useful, and built tools around those as examples of the potential of the tools for both the communication of stories within the data and the generation of hypotheses about the data. Our goal in this pilot project was to demonstrate the power of visual analytics rather than create a full-fledged exploration toolkit, which may become a priority in future engagements with DHS.

3.1 Filtering and details-on-demand based on spatial buffers around an event

We designed our first suite of tools for decision making based on a hypothetical hazard event that would disable nearby infrastructure. We envisioned a preparation training exercise for an explosion at a site of a major event; in such an exercise, analysts might be given a list of six likely targets — perhaps places with a sporting

event, celebration, or other large gathering. Particular interest would be in summaries of facilities (counts, sector dominance, economic impact) within a given buffer around each of the possible targets in order to ascertain which sector or which location would be most vulnerable to different levels of incidents.

Through our query tools on the Map View, we gathered data about CI locations within 10 miles of each site and built tools to examine the quantitative and qualitative aspects of the infrastructure. We implemented the D3 JavaScript library (Bostock, 2015) and specifically a sub-library of open-source data analytics tools called dc.js for this new implementation. These flexible libraries allow the creation of coordinated SVG objects for a web-browser display which listen for events (hover, click, etc.) and can be programmed to react in built-in or custom ways to responsively transform, aggregate, filter, or re-symbolize representations. In our implementation, charts and maps that illustrate the six cities' infrastructure, and that each of these charts is interactive and responsive, such that we can isolate or filter the data to just highlight dimensions we are interested in.

We built the tools in accordance with information visualization and visual analytic theories and frameworks, allowing for multi-scale exploration of data through a variety of dimensions and transformations. Our pilot application involved relatively familiar statistical and geographic representations, coordinated through interaction. Clicking any piece of any of the representations (a pie wedge, a map glyph, or a bar in the bar chart) filters the data so that only corresponding data is shown; for example, we can isolate a sector of interest such as "Government Facilities" across our six study regions by clicking on a pie wedge, then we can isolate only those within two miles by setting the buffer accordingly, and then we can further isolate by selecting one of the six cities on the map.

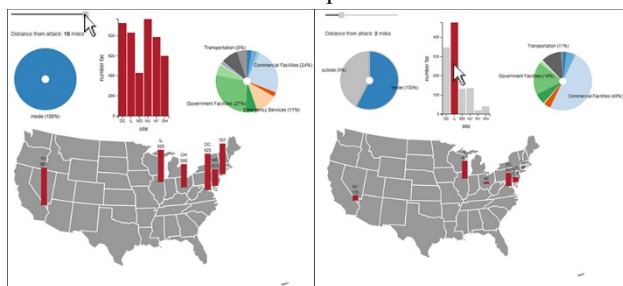


Fig. 3. Overall and filtered view of CI data around within a 10-mile and 2-mile buffer from an incident; the filtered view also displays attribute data (sector, in this case, in the pie graph) updated to the selected site (Illinois)

In the demonstration, of the six sites, it first appears (on the bar graph and the proportional-bar map) that the Las Vegas site has the most CI facilities within a 10 mile radius, and that simple fact alone may help drive decision and limited resources there, as opposed to, say, Baltimore, which has only half as many. We can get details about the sectors represented in each of the buffer zones via the pie graphs, which may also influence allocation of re-sources or assessments of vulnerabilities.

We can filter these facilities in space as well by isolating those facilities within two, rather than ten, miles (perhaps based on information about the damage expected with the attack). The story changes significantly in this case; overall 25% of the facilities in our data set are within 2 miles of the site, but this varies, as you would expect, from city to city. If we click to show only those that are inside the new radius, we see interesting new patterns – Chicago's CI is densely packed around the location we chose, whereas Las Vegas and New York's CI is relatively distant from the predicted attack location in their cities.

Naturally, the simple number of CI assets might not be the most actionable information for a decision maker in this scenario – likely he or she would like information about financial consequence of a disablement, about resilience, about impacted population in the area, or about loss of life. Additionally we might be interested in response resources already on the ground, perhaps just outside of the incident zone, that might be available to be deployed quickly. We might select a different variable, say, from a pull-down list, to rescale the symbols according to a different variable than simple counts; that has not been made functional here pending more conversations with potential users of the system to determine the most effective statistics for such decisions.

This ability to see overall patterns then drill down to records of particular interest is, naturally, a key component of analytics of large multivariate databases.

3.2 Temporal exploration of programmatic data

Where the prior tool showed mostly attribute information about CI data as it stands currently, we developed tools and demonstrations based on scenarios where temporal analysis (associated with space), examining historical patterns, would be of primary concern. For example, we developed a demonstration that examines a related data set – a list of the number and types of interactions, visits, and assessments for CI assets within a 10-mile radius of one of the sites queried in the scenario above. We identified an interest in understanding the frequency, recentness, and type(s) of interactions with these assets, perhaps to prioritize resources or focus on neglected sectors (Figure 4).

For this visualization case, we provide a row chart indicating categories types of assessments or interactions for the overall data set (which starts in 2003) and the breakdown of CI types that have been the focus of the visits. This data can be visualized temporally using a frequency-based temporal bar chart that shows the number of facilities visited by month. Brushing of the timeline is supported, allowing visualization of a time window and filtering on data within it.

A second application that employed this brushable time line allows visual analysis of the logs of the use of the Gateway. In one hypothetical scenario, we imagined interest in the changes in the number and type of users of the Gateway before, during, and after a hurricane (Matthew, from October 2016). We associate the timeline with both a state-level choropleth map (showing the

locations of the users that accessed the Gateway during this time) and allow a data-rich stacked bar graph as the brushable timeline. The categories within each of the stacked bars can be, of course, any categorical variable in the user logs; we chose to toggle between the type of the users (Federal vs. State employees, e.g.) and the type of tools on the Gateway (including the digital library, Map View, among many others on the website) that were heavily used in the logged interactions for each day.

Importantly, when subject-matter experts (in critical infrastructure or in the programs and branches within DHS that the IP Gateway supports) examined these demonstrations, there were many more questions and hypotheses than answers and conclusions. Like many visual analytic tools, these demonstrations were proven successful by their ability to make sense of raw data, observing patterns, trends, and outliers in which that might otherwise have gone undetected. These pilot demonstrations represent only starting points; with each use new questions and hypotheses about the data arise; further development of rich and effective data analytic tools is best accomplished with regular communication with subject-matter experts who can tell us the stories that they hope to tell using available data.

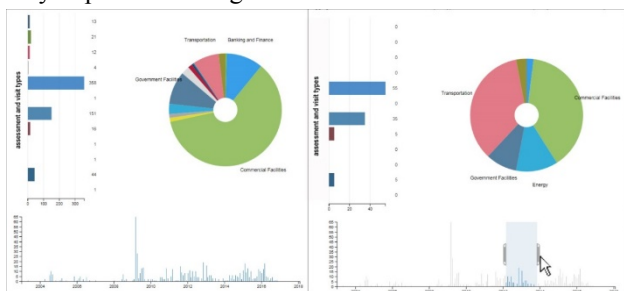


Fig. 4. Programmatic data visualized with a brushable timeline. Dragging the handles on the moving window changes its duration; moving the window changes the time span visualized in associated charts.

4. Hierarchical data visualization of resilience assessments

Interactive data visualization is also employed in a more traditional communication sense in presentation materials and reports for the DHS's Regional Resiliency Assessment Program (RRAP), in which the cyber and/or physical infrastructure of a US region (typically of the scale of a city or part of a state) is analyzed, mapped, and assessed for potential weaknesses or vulnerabilities. These presentations are designed as replacements or supplements to lengthy written final reports, which have the reputation of sitting on a shelf unused, or only available to a limited number of readers. The "big data" in this case are tables, appendixes, maps and graphics, and text content that we wish to present to users – who might still be interested in deeper exploration of the information than is presented in the written reports – in dynamic web-based products that enable a concise but complete summary of the information within.

Each year the RRAP focuses on a selected set of geographic areas, and within each, focuses on a set of

infra-structure systems and addresses a range of hazards that could have regionally and nationally significant consequences (Department of Homeland Security, 2013). These often require assessments of the current or planned security policies and procedures of cooperating private industry partners or public municipal authorities who own and manage the infrastructure systems. Because these partners are the eventual audience for the assessments and their related reports, the products we build for them must be relatively intuitive, but data-rich, such that the partners can plan for hazards in the future and take appropriate actions as recommended in the reports to enhance CI resilience to such hazards.

While we have developed a range of interactive diagrams, maps, and applications that tell various stories associated with the RRAPs, we wish to highlight one particularly successful application that illuminated a (previously) particularly opaque data set: the results of the Cyber Resiliency Review (CRR), an assessment of an organization's operational resilience and cyber security practices (Department of Homeland Security, 2017a). For the purposes of this description, the domain of the assessment (cyber security) is relevant only in that the assessment itself is organized hierarchically, with groupings and domains that can be separately weighted for their importance. The assessment is a set of over 100 questions (with three-value ordinal answers such as "not started," "incomplete," and "complete"), grouped into sets of two to seven questions, and the resulting groups are further grouped into ten domains of several groups each.

By assigning the three levels of responses above to quantitative scores of 0, 0.5, and 1, respectively, we can derive mean scores for the various groupings and domains, and onward through the assessment hierarchy to overall mean scores for each assessed organization, and then even more generally, a mean score for an entire region of organizations. These mean scores can be weighted by various statistics about each organization (e.g. population served, economic consequence of a cyber disruption to that facility) or importance of the domain or group in the assessment.

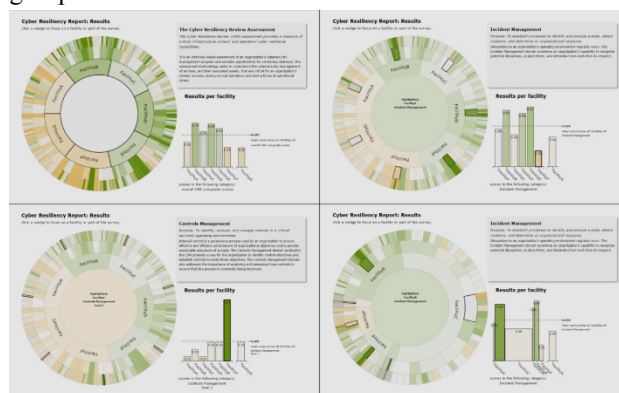


Fig. 5. upper-left: initial sunburst/bar representation of hierarchical data; upper-right: selecting a wedge alters the bar chart and highlights corresponding wedges for each of the seven facilities; lower-left: note selection of question wedge with low scores except for in one case, the center circle is

colored to the mean score for that question; lower-right: wedges and bars are resized according to demographic variable, with Facility C's relevant statistic the highest value.

To visualize the responses, scores, and weightings, we adapted a sunburst diagram (Andrews & Heidegger, 1998; Chuah, 1998; Stasko & Zhang, 2000), a radial space-filling diagram for hierarchical visualization, available with customization options in D3, and a standard bar graph (Figure 5). The sunburst diagram's center circle is a root node of a hierarchy, with wedges arranged in concentric rings whose distance from the center of the display is proportional to the depth of the node represented. The "leaf" nodes in our case are the individual questions in the assessment, and are represented as tiny wedges around the outer ring. In our case, we chose the root node to represent the regional mean: that is, the overall mean score of all questions of each of the seven organizations assessed (while the seven first-ring wedges represent the overall assessment score for each of the organizations).

The wedges of the sunburst diagram are filled with a color along a diverging scheme from tan to green (through gray, representing the overall mean of the region) representing the score in each wedge from 0 to 1. We added labels and various interactions allowing filtering and detail; hovering over a wedge, for example, will highlight the wedges of the corresponding question (or group or domain) for each of the other six organizations represented, and clicking on the wedge will alter bar graph to display the scores for that particular datum for comparison across the organizations. Finally, we allow the weighting of the diagrams and the statistics therein by scaling the size of the wedges and the widths of the bars in the bar chart according to the relative proportion of the weighting variable chosen. By doing so, we allow visualization of each organization or grouping as equally – or, alternatively, proportionally – important, which allows the diagram to be versatile according to the user's needs (for example, the presenter may wish to highlight organizations that serve the greatest number of people or highlight questions that speak for a specific kind of cyber vulnerability).

5. Conclusions

Our designs and applications described in this paper represent a subset of tools and techniques developed by the team at INL for infrastructure visualization; we have selected a sample to demonstrate the application of geovisualization and visual analytics techniques for a limited audience of domain (but not visualization) experts and practitioners. The tools described here range from those that communicate large amounts of data, previously stored in paper documents, reports, and appendixes, in interactive dynamic visual representations, to those that allow the visualization of spatial and attribute variability in large nation-scale databases, to those that enable data exploration to generate hypotheses and aid decision making. We measure success in the construction of these tools in their ability to start conversations and promote communication among stakeholders and government

agencies in the mission of enhancing the resilience and protection of the critical infrastructure of the United States.

6. References

- Andrews, K., & Heidegger, H. (1998). Information slices: Cascading, Visualising and exploring large hierarchies using semi-circular discs. In IEEE Information Visualization Symposium (InfoVis '98), Late Breaking Hot Topics (pp. 9–12).
- Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M. J., MacEachren, a., & Wrobel, S. (2007). Geovisual analytics for spatial decision support: Setting the research agenda. *International Journal of Geographical Information Science*, 21(8), 839–857. <http://doi.org/10.1080/13658810701349011>
- Andrienko, G., Andrienko, N., Keim, D., MacEachren, A. M., & Wrobel, S. (2011). Challenging problems of geospatial visual analytics. *Journal of Visual Languages and Computing*, 22(4), 251–256. <http://doi.org/10.1016/j.jvlc.2011.04.001>
- Bostock, M. (2015). D3: Data-Driven Documents. Retrieved from <https://d3js.org/>
- Chuah, M. (1998). Dynamic aggregation with circular visual designs. In IEEE Symposium on Information Visualization (InfoVis '98) (pp. 35–43, 151). <http://doi.org/10.1109/INFVIS.1998.729557>
- Department of Homeland Security. (2013). Regional Resiliency Assessment Program. Arlington VA. Retrieved from https://www.dhs.gov/sites/default/files/publications/RR-AP-Fact-Sheet-508_0.pdf
- Department of Homeland Security. (2016). Science and Technology Visionary Goals. Retrieved February 15, 2016, from <http://www.dhs.gov/science-and-technology/visionary-goals>
- Department of Homeland Security. (2017a). Stakeholder Engagement and Critical Infrastructure Resilience. Retrieved from <https://www.us-cert.gov/sites/default/files/c3vp/crr-fact-sheet.pdf>
- Department of Homeland Security. (2017b). Topics: Critical Infrastructure Security. Retrieved January 25, 2017, from <https://www.dhs.gov/topic/critical-infrastructure-security>
- ESRI. (2016). Story Map Swipe and Spyglass. Retrieved October 20, 2016, from <https://storymaps.arcgis.com/en/app-list/swipe-spyglass/>
- Lewis, T. (2006). *Critical Infrastructure Protection in Homeland Security: Defending an Networked Nation*. Hoboken, NJ: John Wiley & Sons.
- National Geospatial-Intelligence Agency. (2014). 2020 Analysis Technology Plan. Springfield, VA. Retrieved from https://www.nga.mil/MediaRoom/PressReleases/Documents/NGA_Analysis_Tech_Plan.pdf

- National Research Council. (2007). *Successful Response Starts with a Map: Improving Geospatial Support for Disaster Management*. Washington, DC: The National Academies Press. <http://doi.org/10.17226/11793>
- PPD 21. (2013). *Presidential Policy Directive: Critical Infrastructure Security and Resilience*. Washington, D.C. Retrieved from <https://www.whitehouse.gov/the-press-office/2013/02/12/presidential-policy-directive-critical-infrastructure-security-and-resil>
- President's Commission on Critical Infrastructure Protection. (1997). *Critical Foundations: Protecting America's Infrastructures*. Washington, D.C. Retrieved from <http://www.fas.org/sgp/library/pccip.pdf>
- Stasko, J. T., & Zhang, E. (2000). Focus+context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In *IEEE Symposium Information Visualization (InfoVis '00)* (pp. 57–66).
- Thomas, J. J., & Cook, K. a. (2005). Illuminating the path: The research and development agenda for visual analytics. *IEEE Computer Society*, 54(2), 184. <http://doi.org/10.3389/fmicb.2011.00006>
- Tomaszewski, B. M., & MacEachren, A. M. (2012). Geovisual analytics to support crisis management: Information foraging for geo-historical context. *Information Visualization*, 11(4), 339–359. <http://doi.org/10.1177/1473871612456122>