

GIS-based Command and Control Infrastructure for Critical Infrastructure Protection

Stephen D. Wolthusen
Gjøvik University College, Gjøvik, Norway and
Fraunhofer-IGD, Darmstadt, Germany
swolthusen@ieee.org

Abstract

Critical infrastructure components are often dispersed over large areas; at the same time even an infrastructure individual component relies on a significant number of parameters that must be controlled and monitored in addition to interdependencies with other infrastructure components.

Modeling and simulation of infrastructure elements and particularly of interdependencies and risks to those elements can be performed on the basis of a geographical information system providing a common semantic basis for presentation and analysis as well as a mechanism for sharing only selected and where necessary downgraded information with other infrastructure operators.

1. Introduction

Command and control (C2) of critical infrastructure systems as well as their modeling and simulation faces challenges similar to that encountered in network-centric environments in the defense sector [15, 10]. The primary challenge lies in the requirement to integrate large and disparate volumes of data and present the resulting information in such a way to permit the identification of all relevant and critical items; while this problem is present in modeling and simulation environments, it becomes particularly important in case of emergency situations where the time available for reflection, analysis, and decisions is severely limited.

Even within a single organization (e.g. a telecommunication service provider), information that can contribute to an overall view of the infrastructure component's ability to provide services and for its proper internal operation need not be available in a usable and timely manner. Reasons for this include that some information is retained only locally in organizational subunits, is kept manually or semi-manually, or that the formats used for processing these data are not amenable for easy exchange with an organization's overall

C2 and emergency response systems. Moreover, the number of information sources to be integrated is a priori not bounded while the semantics of the individual data sources may also not be known in advance.

Given the interdependencies of critical infrastructure components, an isolated C2 facility is, however, insufficient to counter a number of threats and risks. Such threats may include external threats from natural disasters such as flooding or severe weather events but also from other infrastructure component failures. In the former case, a C2 system must be capable of incorporating external information sources in a one-way information flow whereas in the latter the bidirectional exchange of information will be necessary and must be subjected to strict security controls.

In addition to C2 environments, the ability to gather information on operational characteristics and threats into a single view also has significant applications for both modeling and planning as well as for simulation.

Depending on the application areas as well as individual requirements emerging during operation, user requirements may differ sharply for the type of data to be presented, its representation, and also the bandwidth and characteristics of the devices used for the human-computer interface.

In this paper, we present the design and analysis of a C2 environment for critical infrastructure modeling, simulation, analysis, and emergency management. The underlying security model and controls required for consolidation and exchange of data items among infrastructure operators as well as for the compartmentalization within an individual infrastructure operator organization has been described in detail earlier [21]; this paper is primarily concerned with the presentation and analytical capabilities of the C2 system. The remainder of the paper is organized as follows: Section 2 describes the mechanisms for the aggregation and normalization of data items within a unified and extensible framework while section 3 describes the graph-based model used to identify dependency structures within the model. Section 4 then discusses the mechanisms used

for collating and condensing information in the presentation layer with particular emphasis on the approaches for enforcing need to know and adaptations to display devices; section 5 then describes a selected number of application scenarios.

2. Aggregation Layer

Even within individual infrastructure components it is frequently the case that multiple mutually incompatible information systems are used to monitor and control aspects of planning, operation, and emergency handling of the infrastructure. Given the scope of data contained in such databases and information systems, a direct harmonization among such entities for information sharing and exchange is impractical even at the level of an individual infrastructure operator, and even more so when attempting to link multiple independent infrastructure operators with a joint information system.

As a result, a key requirement for information exchange is the use of an interoperable intermediate format of sufficient generality to contain not only the data elements but also the underlying ontological structures, which can then be used to aggregate data for a unified analysis and presentation. The latter requirement not only results from the need to translate data points and relational tuples, but also from the fact that infrastructure elements evolve over long time scales – during which the semantics of individual data points and metrics are likely to change. To this end, a common ontological model must provide a common abstraction layer for the plenitude of underlying data formats.

Data in this format must have well-defined semantics that can be retained over changes in underlying representations and storage and be archivable. This represents a particular challenge since the lifetimes of many infrastructure components encompass a large number of information system generations. Moreover, in many cases the full semantics is not fully contained in data repositories but only accessible through interpretative logic layers. Given the cost associated with re-acquiring all data (in addition to direct sensor measurements) as well as the danger of inconsistencies among such parallel repositories, it becomes necessary to interpose an interpretative layer that can isolate the underlying data sources from its interpretation and provide adaptation where necessary.

Such an intermediate layer can be provided based on open, interoperable standards defined by the World Wide Web Consortium (W3C) in the form of the Resource Description Framework (RDF) as described in [21]. The RDF represents predicates (e.g. data points) over entities as a directed graph with vertices representing entities and edges annotated with properties and property values [12, 9, 3] and is encoded syntactically in extensible markup language

(XML). A basic RDF graph can therefore be considered a superset of the dependency graphs discussed in section 3; several syntactical features such as RDF containers (bags, sequences) can be normalized and decomposed into regular directed graphs for this purpose. Within RDF, both entities (vertices) and properties (edges) are represented by Uniform Resource Identifiers (URI); while a basic descriptive format exists, this can be extended arbitrarily using the RDF schema mechanism including RDF reification [4]; this definition includes a semi-rigorous model-theoretic definition of the formal semantics of RDF [8].

It should be noted that the use of URIs for representing underlying representations provides a natural solution for satisfying the requirement for real-time data access and mediation to existing data repositories; this mechanism also permits natural interaction e.g. with web service-based architectures such as those found in geographical information systems [11, 5, 7, 20]. The actual ontological representation [18] can also be accomplished using open standards, in this case using the W3C Web Ontology Language (OWL) [13, 17, 2, 14], which can be considered a syntactical and semantic extension of RDF. These standards define descriptions of classes, properties and their instances and, more importantly, semantic entailments which can be used for reasoning within the ontological model. In the critical infrastructure model described here, the restricted OWL descriptive logic subset was chosen; this permits the nevertheless permits the efficient computation of entailments that are important to the automated analysis of infrastructure dependencies as described in [21].

3. Dependency Layer

As described in [21], the dependencies of infrastructure elements can be described using a multigraph model as the mathematical framework. The basic elements of the model are reviewed in the following section.

Definition 1 *Infrastructure components are separated into entities \mathcal{E} ($\mathcal{E} = \{e_1, \dots, e_k\}$) represented as vertices and dependencies \mathcal{D} ($\mathcal{D} = \{d_1^1, \dots, d_n^m\}$) among entities represented as directed edges where the set of edges is partitioned into m dependency types, resulting in a graph $\mathcal{G} = (\mathcal{E}, \mathcal{D})$. \mathcal{G} may contain parallel edges, but may not contain self-loops.*

Edges between two given vertices e_a, e_b are not uniquely identified by the 2-tuple (e_a, e_b) as is the case in simple graphs since they may differ in their dependency type:

Definition 2 *For two given vertices e_a, e_b within \mathcal{G} , the set of edges must not contain two edges of the same dependency type.*

The set of all dependencies between given vertices e_a, e_b is denoted as (e_a, e_b) and abbreviated (a, b) . By collecting

edges of different dependency type, a directed simple graph \mathcal{G}^s is produced and referred to as the aggregate dependency graph.

For a given dependency type t , a t -dependency path is a sequence $P = \{e_1, d_1^t, e_2, d_2^t, \dots, d_{i-1}^t, e_i\}$ of alternating vertices and edges such that for $1 \leq j \leq i$, d_j^t is incident with e_j and e_{j+1} .

Two paths are t -edge disjoint if they do not have an edge of type t in common.

Dependency paths and connectivity properties are preserved by the aggregate dependency graph, edge disjointness is defined analogously to t -edge disjointness. For a given graph with edges e_k and dependency types t_j , a relation $(e_k \times t_j) \mapsto \mathbb{N}$ is defined. The range of this relation is referred to as the *dependency strength* and denoted s_{e_k} for a given edge e_k .

Given a dependency graph, the graph can be partitioned into vertex subsets $\mathcal{E} = \mathcal{E}_1 \cup \dots \cup \mathcal{E}_k$ (where $k \leq |\mathcal{E}|$) called *partitions* (\mathcal{P}_i). For observing dependencies at higher levels of abstraction, theorem 1 provides a justification for coalescing graphs (see [21] for a proof of the following theorem).

Theorem 1 For a given dependency graph $\mathcal{G} = (\mathcal{E}, \mathcal{D})$ and a partitioning over the vertices $\mathcal{E} = \mathcal{E}_1 \cup \dots \cup \mathcal{E}_k$, each partition (\mathcal{P}_i) can be substituted by a single coalesced vertex.

For *edge coalescion* two or more edges with different types t_i and t_j incident with vertices e_k, e_l are coalesced by forming the set union over the types with the derived type $t_{i,j}$. The coalesced edges are then removed from the graph; no self-loops can occur in this step. If all edges are coalesced, the result is a *typeless dependency graph*. The dependency strength of a coalesced edge created from k individual edges is trivially defined as

$$\frac{1}{k} \sum_{1 \leq i \leq k} s_{e_i}$$

if a normalized dependency strength is used.

4. Presentation Layer

A multimodal visual presentation mechanism using both topological and topographical information can provide situational awareness by presenting objects, dependencies, and interrelations within a common situational view that also incorporates background information from various sources.

The presentation of the collected data on infrastructure elements and environmental conditions as well as the integration into relevant information that is immediately required by decision makers can occur in a number of different views depending on the task at hand.

Since a significant number of infrastructure elements have network characteristics (i.e. depend on edges connecting individual vertices, as in the case of telecommunication

or power transmission lines), a topological view provides key insights in an abstract format which permits dependencies to be identified and analyzed.

However, a purely topological view may omit highly significant data points that can be critically important to know for decision makers. Particularly when multiple infrastructure elements or external threats such as weather events are to be considered, a topological view does not provide the appropriate context to judge such influences. Moreover, while some mutual and transitive dependencies can be identified automatically [21], these may dependencies and interactions may not all be known in advance and can only be derived intuitively given an appropriate presentation of the data.

Geographical information systems [11] can provide this contextualization as well as a foundation for integrating the varied types of information that must be aggregated and selectively displayed for decision makers. A particular challenge for the presentation mechanism is tightly coupled with the usage patterns likely to be found in all application areas from planning to emergency management, namely that the information presented is likely to be shared visually (e.g. in the same situation room or in the field) with individuals for which the security controls have no information.

Since the security model and technical controls can obviously only control information flow and display for entities and individuals known (e.g. logged into the C2 system), the visual information sharing problem can only be addressed through procedural controls. However, the user interface can support these procedural controls by allowing the presenter in a visual sharing environment to selectively downgrade or de-select certain aspects of the display prior to sharing. The general logical architecture distinguishes four layers for the presentation, namely the physical and logical models, stored in distributed databases providing the basis for upper layers on which the conceptual model is based, which also correlates the topological and topographical data. The topmost layer incorporates the semantics of the application layer. Each of the layer components is associated with security properties as described in [21]. The actual software representation of these layers is provided by a typical three-tier application architecture with multiple (potentially distributed) database backends, separated into geospatial data, application data, and metadata that are accessed by a web-server based application layer which provides the application semantics through an web-service based GIS application server in combination with servlet containers that can also perform direct rendering (see section 4.2) depending on services requested by the client side, which is represented by a web-based OpenGIS consortium (OGC) client with minimum client-side requirements.

4.1. Local Presentation

Local presentation of the C2 system data is characterized by high bandwidth, computational and display capabilities. While most of the information sources in a purely local display will not carry security constraints, there may still be restrictions on the use of certain data sets that must be controlled for need to know. For data sets obtained externally from other sources, these controls must also be applied prior to aggregation and rendering

For a topological view of a given configuration or dependency, the display mechanism can omit elements during the aggregation process. However, in actual operation and particularly during emergency management, it is necessary to notify an operator that the C2 system has withheld certain data items from the view. This way, out-of-band communication between infrastructure operators and organizational units of an individual operator or, in particularly severe cases where national security may be at stake, government intervention can lift the need-to-know restrictions on the data sets that have been withheld.

Similarly, an aggregation and selection mechanism can, particularly in case of information sharing, abstract from the internal, detailed information source of an information provider and yield only the relevant abstracted data for the infrastructure operator depending on this information. Spatial selection can be applied in addition to this to constrain viewers to the minimum area of interest required.

For topographical information, additional options for security-controlled aggregation and presentation of information are available. In addition to selecting only certain layers consistent with a viewer's need to know and restricting the viewing area to a given area of interest where a need to know is encoded in the applicable security policy as in the case of topological information, it is also possible to use a deliberately degraded rendering of certain layers and data in their presentation. This generally permits viewers to obtain sufficient situational awareness without necessarily disclosing e.g. the precise location of sensitive equipment or transmission lines.

While the modeling and simulation engine must have full access to all data points to be integrated (cf. [21]) and hence can apply security controls only at the boundaries between data sources and the modeling and simulation engine, it is possible to limit the risk of inadvertent exposure of sensitive information to operators by performing the actual rendering in a trusted environment where it is less likely that the computer performing the rendering based on sensitive data is compromised by the operator or through malware such as Trojan horses.

Given the high bandwidth available, high-resolution rendered visualizations can be displayed remotely at operator consoles with sufficient speed. This, however, is not neces-

sarily the case for mobile devices and implies different presentation mechanisms for remote data sharing and visualization as discussed in the following section 4.2.

4.2. Remote and Mobile Presentation

Information sharing with remote entities such as other infrastructure operators or providing situational data to engineers, security forces and first responders in emergency management situations cannot occur using direct access to all relevant databases and sources. While the primary reason for this, particularly for mobile devices, is the limited bandwidth available, security considerations preclude this arrangement for most remote information sharing with other organizations and infrastructure operators as well. While, as noted in section 4.1 the computation of certain aspects of dependency models requires selective information sharing at the individual data item level, the sharing of pre-rendered and suitably abstracted or sanitized presentations and visualizations provides for both the reduction of the need to share such primary data and also alleviates the end device from the burden of rendering.

Given the small screen sizes for mobile devices (a typical display device may have a resolution of 320 by 240 pixels with a maximum color depth of 16 bit, resulting in a worst-case uncompressed memory and bandwidth requirement of 150 kBytes) and the good compressibility of most presentation data as well as the ability to dynamically merge presentation layers at the device end, even full redispays can be accomplished in less than 10 seconds based on cellular radio or telephony connections.

The limited screen size imposes a further limitation on mobile devices; this problem can only be addressed by reducing the number of layers presented at any given time. However, since one can assume that use of mobile presentation mechanisms will generally be goal-oriented (i.e. to address a specific situation or for local exploration of infrastructure elements), this is not a critical limitation by itself. It is, however, generally advisable to limit the data provided to such devices since they may be compromised far more easily than

5. Application Scenarios

The following section provides selected sample scenarios for an interactive, GIS-based command and control system for critical infrastructure protection¹.

¹ The scenarios as well as data points and screenshots in the following section have been sanitized or created specifically for this purpose

5.1. Modeling and Planning

One of the primary uses of GIS-based systems is the ability to visually correlate information in such a way that decision makers can quickly judge a situation or potential risk without having to codify a precise analysis before analyzing data sets.

An example of such a risk analysis, infrastructure operators must be capable of checking and verifying that wiring and transmission lines that are nominally provided redundantly and are categorized as such by the service provider are in fact separated by sufficient space that an accident or sabotage event cannot disrupt these redundant circuits simultaneously.

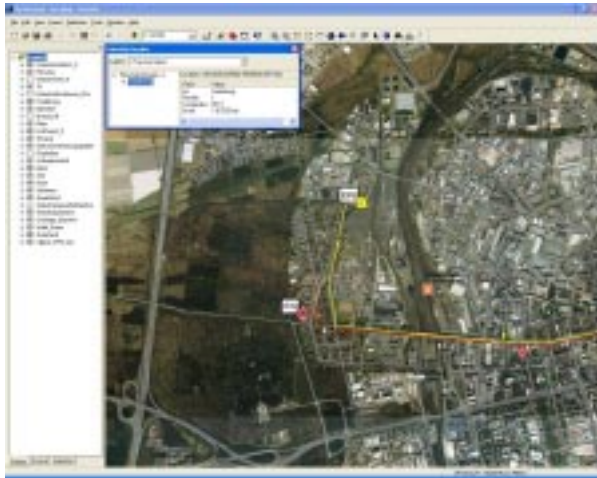


Figure 1. Identification of risks to power transmission line by parallel gas pipeline during planning

Figure 1 provides another example of exploratory use during planning and analysis stages of critical infrastructures; by combining planned electrical power transmission lines with layer information for other infrastructure components, it is possible to visually identify a gas pipeline that is already running in the same location as the planned transmission line. This information can now be taken into consideration e.g. in cooperating with the pipeline operator to identify potential service disruptions or other risks to the transmission line (e.g. during construction activities) or may lead to relocation of the transmission line if the risk is deemed to be unacceptable.

While the information for deriving such decisions is typically available, in many cases even for the general public, the effort to identify and assess such potential hazards based on correlating charts can be assumed to be sufficiently

large that not all possible considerations are covered because of fiscal and time constraints. Given the C2 system described here, however, correlation and aggregation of all relevant layers and visual identification of hazards (possibly followed by further investigation) is a relatively quick and simple operation.

In case of cooperation between multiple infrastructure operators, it is then also possible to avoid the inadvertent construction of critical locations (e.g. by the proximity of several infrastructure components that could be damaged or destroyed by a single event such as a natural disaster or terrorist attack); where such locations already have been created previously it is nevertheless possible to subject these locations, as shown in figure 2, to particular attention.

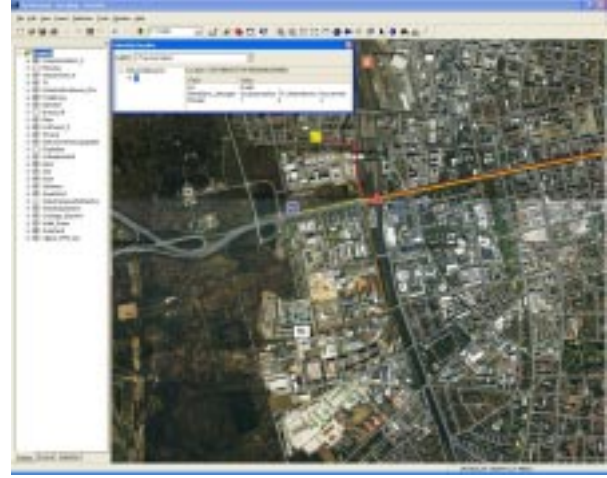


Figure 2. Identification of critical junction point

5.2. Operational Use

Based on the information sketched in the preceding section but also in combination with sensor data both from an infrastructure operator and also from external sources for monitoring as well as operational prediction.

Figure 3 provides a typical example of such operational use; here, topographical information is coupled with information on dependencies of several transportation and energy infrastructure elements. As can be seen in figure 3, the terrain contours in combination with water levels and predicted precipitation will result in severe flooding, which in turn will affect both a local railway line and, more importantly, will also result in a air traffic radio beacon having to be deactivated since the power transmission line leading to the radio beacon will be in the flooded area.

Given such predictive information it is relatively straightforward to mitigate the threat (e.g. by shoring up the beacon with sandbags and supplying a generator) or take other corrective measures.



Figure 3. Flooding threat analysis for air traffic radio beacon

The terrain information contained in the underlying GIS can also provide valuable predictive information that is generally not available to infrastructure operators, e.g. flooding prediction for cable ducts and tunnels, power substations or local telecommunication exchanges and, based on this, a prioritization for securing facilities or establishing alternative service facilities.

The latter is particularly important in case other critical infrastructure elements are affected by a potential disruption of service; given a common operational picture and situational awareness of all parties involved, it is then possible to ensure continued contingency services to critical infrastructures such as hospitals or emergency services.

5.3. Transitive Dependency Analysis

While transitive dependencies can be analyzed from topological data and determined reliably, such information may not be available under all circumstances, either because the requisite data is not available a priori or because need to know and security concerns prevent the requisite data sets from being exchanged.

Such information may, however, still be derived visually from the C2 tool by visual inspection or by a combination of automated tools and visual inspection depending on the quality and extent of the data available. Figure 4 shows an example of a defective power substation, over which a (simplified) power outage radius has been projected in which

service availability can no longer be guaranteed in case of failure. In the given application scenario, there exists a redundant telecommunication and data link between a financial institution and a computing center; while each of the redundant transmission lines is spatially separated, the large extent of the power outage nevertheless would lead to service unavailability. Given this information, the financial institution and the telecommunication provider may rapidly establish alternate service schemes and thereby mitigate the failure caused by a transitive dependency that would have been too costly to prevent at the planning stage by employing a more circuitous transmission path between the facilities involved.

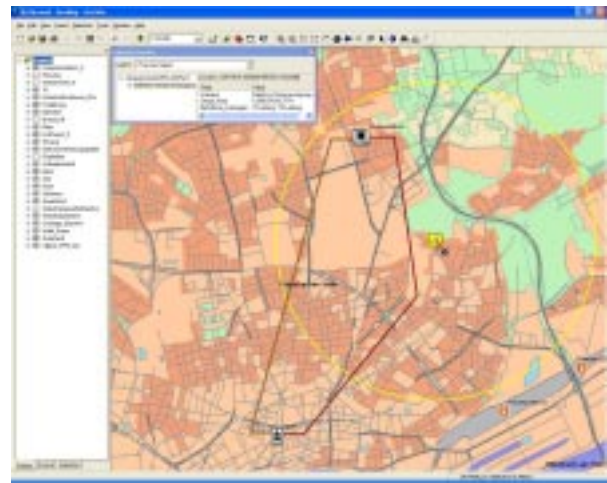


Figure 4. Transitive vulnerability of dispersed redundant transmission lines

Similarly, figure 5 demonstrates the potential for transitive vulnerabilities resulting from flooding damage; in this scenario flood damage will result in the forced shutdown of a power substation that is serving an airport facility. While this facility is both served by another redundant power substation and transmission line as well as by internal generating capacity, the vulnerability analysis and operational picture also identifies a potential vulnerability at a chemical factory caused by the impending power failure that may indirectly cause disruptions at the airport because of safety precautions or actual damage.

6. Related Work

A number of approaches have been proposed for modeling and simulation of critical infrastructures [1, 16] and vary considerably in the level of detail considered, ranging from simple dependency analyses to elaborate models containing



Figure 5. Transitive indirect vulnerability

continuous physical submodels (e.g. for pipelines and electrical grid systems) as well as behavioral models.

A number of vendors in the GIS community are providing GIS-based information systems for mapping and visualization; such general systems have also been used in assessment, preparedness, and emergency planning environments (e.g. for hazardous material plume analysis or evacuation routing) [6, 19]. The Open GIS Consortium has formed a pilot initiative for exploring the use of GIS for critical infrastructure protection in a limited regional area (northeastern US and southeastern Canada) in 2003 while the Geospatial Information and Technology Association has held workshops on requirements for GIS in the CIP application area in 2004 and 2005.

7. Conclusions

This paper has presented a command and control architecture for planning, modeling, and operation of critical infrastructure elements with particular emphasis on information sharing among infrastructure operators and the ability to deploy a broad range of presentation mechanisms ranging from command posts to emergency management in the field on handheld devices.

The representation of complex interrelations among infrastructure elements and the possibility of correlating and visualizing information from various sources both in topological and in topographical representations can provide important benefits for improving the quality of both planning and operation of infrastructure elements, particularly in situations where joint situational awareness and rapid analysis are essential, e.g. in emergency situations.

Acknowledgments The author would like to thank S. Ritter, W. Stein, and J. Weber of the BSI (German Informa-

tion Security Agency), Bonn, Germany for discussions and comments. Visual display prototypes were constructed by A. Mrotzeck and B. Koch of GISec, Darmstadt, Germany. Portions of the research reported in this paper were funded by the BSI.

References

- [1] M. Amin. Toward Self-Healing Infrastructure Systems. *IEEE Computer*, 33(8):44–53, Aug. 2000.
- [2] S. Bechhofer, F. van Harmelen, J. Hendler, I. Horrocks, D. L. McGuinness, P. F. Patel-Schneider, and L. A. Stein. OWL Web Ontology Language Reference. W3C Recommendation, Feb. 2004.
- [3] D. Beckett. RDF/XML Syntax Specification (Revised). W3C Recommendation, Feb. 2004.
- [4] D. Brickley and R. V. Guha. RDF Vocabulary Description Language 1.0: RDF Schema. W3C Recommendation, Feb. 2004.
- [5] J. de la Beaujardière. Web Map Service Implementation Specification. Technical Report OGC 01-068r2, Open GIS Consortium, Wayland, MA, USA, Nov. 2001.
- [6] M. de Zuviria and S. McClure. Comparative Study of GIS Data Products Used in Various-Sized Municipalities for Emergency Management and Critical Infrastructure Protection across Canada. Technical Report PS48-6/2004E, Public Safety and Emergency Preparedness Canada, Ontario, Canada, Dec. 2003.
- [7] J. D. Evans. Web Coverage Service. Technical Report OGC 03-065r6, Open GIS Consortium, Wayland, MA, USA, Aug. 2003.
- [8] P. Hayes. RDF Semantics. W3C Recommendation, Feb. 2004.
- [9] G. Klyne and J. J. Carroll. Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C Recommendation, Feb. 2004.
- [10] A. Kott, editor. *Advanced Technology Concepts for Command and Control*. Xlibris, Philadelphia, PA, USA, 2004.
- [11] P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, editors. *Geographical Information Systems*. John Wiley & Sons, New York, NY, USA, 2nd edition, 1999. Two volumes.
- [12] F. Manola and E. Miller. RDF Primer. W3C Recommendation, Feb. 2004.
- [13] D. L. McGuinness and F. van Harmelen. OWL Web Ontology Language Overview. W3C Recommendation, Feb. 2004.
- [14] P. F. Patel-Schneider, P. Hayes, and I. Horrocks. OWL Web Ontology Language Semantics and Abstract Syntax. W3C Recommendation, Feb. 2004.
- [15] D. E. Pearson, editor. *The World Wide Military Command and Control System: Evolution and Effectiveness*. Air University Press, Maxwell Air Force Base, AL, USA, 2000.
- [16] S. M. Rinaldi. Modeling and Simulating Critical Infrastructures and Their Interdependencies. In *Proceedings of the*

37th Annual Hawaii International Conference on System Sciences (HICSS'04), Big Island, HI, USA, Jan. 2004. IEEE Computer Society Press.

- [17] M. K. Smith, C. Welty, and D. L. McGuinness. OWL Web Ontology Language Guide. W3C Recommendation, Feb. 2004.
- [18] J. F. Sowa. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing, Pacific Grove, CA, USA, 2000.
- [19] M. Terner, R. Sutton, B. Hebert, J. Bailey, H. Gilbert, and C. Jacqz. Protecting Critical Infrastructure. *GeoIntelligence*, 1(1):8–12, Mar. 2004.
- [20] P. A. Vretanos. Web Feature Service Implementation Specification. Technical Report OGC 02-058, Open GIS Consortium, Wayland, MA, USA, Sept. 2003.
- [21] S. Wolthusen. Modeling Critical Infrastructure Requirements. In *Proceedings from the Fifth Annual IEEE SMC Information Assurance Workshop, United States Military Academy*, pages 258–265, West Point, NY, USA, June 2004. IEEE Press.