Engineering Applications of Artificial Intelligence 43 (2015) 157–165

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

An interactive ontology-driven information system for simulating background radiation and generating scenarios for testing special nuclear materials detection algorithms

Alexandre Sorokine [a,*], Bob G. Schlicher [a], Richard C. Ward [a], Michael C. Wright [a], Kara L. Kruse [a], Budhendra Bhaduri [a], Alexander Slepoy [b,1]

[a] Oak Ridge National Laboratory, 1 Bethel Valley Rd., Oak Ridge, TN 37831-6085, United States
[b] Lawrence Berkeley National Laboratory, Berkeley, CA 94720-8099, United States

ARTICLE INFO

Article history:
Received 29 October 2014
Received in revised form
31 March 2015
Accepted 20 April 2015

Keywords:
Ontology
Ontology driven information system
Scenario
Geographical information system
Database

ABSTRACT

This paper describes an original approach to generate scenarios for the purpose of testing the algorithms used to detect special nuclear materials (SNM) that incorporates the use of ontologies. Separating the signal of SNM from the background requires sophisticated algorithms. To assist in developing such algorithms, there is a need for scenarios that capture a very wide range of variables affecting the detection process, depending on the type of detector being used. To provide such a capability, we developed an ontology-driven information system (ODIS) for generating scenarios that can be used for testing of algorithms for SNM detection. The Ontology-Driven Scenario Generator (ODSG) is an ODIS based on information supplied by subject matter experts and other documentation. The details of the creation of the ontology, the development of the ontology-driven information system, and the design of the web user interface (UI) are presented along with specific examples of scenarios generated using the ODSG. We demonstrate that the paradigm behind the ODSG is capable of addressing the problem of semantic complexity at both the user and developer levels. Compared to traditional approaches, an ODIS provides benefits such as faithful representation of the users' domain conceptualization, simplified management of very large and semantically diverse datasets, and the ability to handle frequent changes to the application and the UI. The approach makes possible the generation of a much larger number of specific scenarios based on limited user-supplied information.

1. Introduction

Detection of special nuclear material (SNM) represents one of the most important challenges of national security. Reliable SNM detection is critical for the enforcement of nuclear non-proliferation international agreements, preventing nuclear terrorism and illicit trafficking of SNM (ITDB, 2014; Runkle et al., 2009; Murauskaite, 2014). But as noted by Runkle et al. (2009) “it is a daunting technical challenge to identify special nuclear materials via gamma ray detection.” From the physics point of view detection of SNM poses a difficult task as its photon signature has to be extracted from the background radiation (Runkle et al., 2009). The background radiation field varies due to a very large number of factors such as geographic location, geology, human activity, weather, and others. The background radiation is generated by macroscopic objects, like the materials used in buildings or anthropomorphic sources such as individuals receiving cancer or diagnostic treatment, is highly situation specific, and cannot be well-represented by a generic instance. Algorithms used for extracting SNM signatures, such as GADRAS (Medalia, 2010), must be tested and validated to assure their quality, i.e., low rate of misdetections and false alarms under diverse background radiation conditions. Due to high variability of environments and high costs or regulatory impracticality of field radiation measurements with significant quantities of SNM, testing of the algorithms is most effectively achieved in simulated environments using synthetic data.

Separating the SNM signal from the background requires sophisticated algorithms because SNM signal may be masked by other sources of radiation such as medical isotopes (Wald, 2014). Search mode algorithms in an urban setting are hampered by a structured background radiation that resists simple noise models. To assist in developing such algorithms, there is a need for scenarios which capture a very wide range of variables affecting the detection process, depending on the type of detector being used. Each urban segment produces a significantly unique signature background, complicated by the dynamic entry and exit of cars, trucks, and pedestrians. A

* Correspondence to: One Bethel Valley Rd., P.O. Box 2008, MS-6017, Oak Ridge, TN 37831-6017, United States. Tel.: +865 574 4966.
E-mail addresses: sorokina@ornl.gov (A. Sorokine), aslepoy@lbl.gov (A. Slepoy).
[1] Deceased March 19, 2013

http://dx.doi.org/10.1016/j.engappai.2015.04.010
0952-1976/

particular instance of such a setting can only be described by a *scenario*, i.e., description of a sequence of events and a setting where these events occur. Our goal was to use ontologies to expand the number and variety of scenarios generated, and thus to provide algorithm developers and testers with an unprecedented level of completeness in covering the space of combinatoral variablity available.

As a result of our effort we have developed a system, the Ontology-Driven Scenario Generator (ODSG) that generates a very large number of scenarios from a small amount of user input, using parameters that describe a particular process of searching for SNM. To accomplish this task we based software development on the concept of an ontology-driven information system (ODIS) (Guarino, 1998; Fonseca and Egenhofer, 1999; Fonseca, 2007). The ODSG automatically converts a concise description by the user into a large number of scenarios consistent with the user-defined requirements. The system is based on a geographical information system (GIS, Wikipedia 2014) utilizing both top- and domain-level ontologies to facilitate the framework. The scenarios are based on both real and synthetic data, and represent a realistic sequence of events in a realistic urban setting. The collection of scenarios can be used to benchmark and train search algorithms, and evaluate their performance with respect to a particular scenario type.

By using the ODIS development paradigm we addressed the problem of semantic complexity both at the developer and user level. A description of the problem in terms of numeric parameters of the setting alone limits the description to a particular narrow semantic frame. For example, to set the positions of trees, or the make and model of vehicles, such entities must be postulated as present or absent in the scene. The existence of these in the scenario must be provided by a semantic description. It is conceivable that one could include all possible objects and all possible relationships between objects in an over-arching model, and then use this model to create a strictly parametric description, but such a practice requires large numbers of constraints to be useful and such constraints would constitute a de facto ontology.

ontology-driven information systems (ODIS) and Ontology-Driven Architecture for Software Engineering (ODASE) have been discussed in the literature for at least a decade (Guarino, 1998; Fonseca and Egenhofer, 1999; Fonseca, 2007; Bossche et al., 2007; Akerman and Tyree, 2006; others). The distinctive property of these paradigms is the reliance on ontologies to accomplish multiple goals during system development and operation. Recent advances in ontologies and other kinds of semantics technology are turning ODIS and ODASE into practical tools capable of addressing a number of challenges of GIS development. Placing ontologies at the core of GIS and GIS development offers a promise of more effective geographic information integration, reduced costs of development, streamlined and coherent user interfaces, and, most important, adequacy of the information system to the users' conceptualization of the domain.

Ontologies played a pivotal role at all stages of the development of ODSG and are indispensable for its operation. During system development we relied on ontologies and ontology-based techniques to formalize application domain and expert knowledge. At the system configuration stage, ontologies are used to link disparate datasets and simulation models and to generate the web user interface (UI). At run time ontologies are used for database schemata translation, workflow planning, reasoning about missing data, and presentation of the results. The overall architectural approach can be described as mapping between ontologies and database schemata and traversing ontology categories to pull the most appropriate data from known datasources.

An important part of the process of developing software to generate scenarios is gathering expert opinion on the variables to be included in the simulated environment. The variables should be extensive enough to properly define the complexity of radiation detection in urban environments accounting for the presence of legitimate sources such as construction materials, patients who have received medical isotopes for diagnosis or treatment, shielding, natural sources, and enhanced NORM in all likely geometric configurations.

We discuss existing solutions to automated generation of scenarios in Section 2. Our proposed solution, increasing the number of scenarios by using an ODIS framework, is presented in Section 3. Section 4 discusses results of example scenarios generated. Finally in Section 5, we conclude with the lessons learned and potential for application of ODGS in other subject areas.

2. Existing solutions for automated scenario generation

For purposes of examining the existing solutions for automated scenario generation, we define scenarios as *a time sequence of events describing an activity or process*. In this section we will review common approaches to automated generation of scenarios. Scenarios are used in many areas to generate synthetic sequences of events including training/planning, resource allocation, estimating risks, evaluating physical models, etc. Martin and Hughes (2010) describe various approaches to implementing automated scenario generation including: 1) seeded approaches that blend human-created elements and system-automated processes; 2) heuristic approaches that create scenarios and their subcomponents by randomly selecting components and then comparing the scenario against the heuristics (Grois et al., 1998); 3) enumeration approaches which are a variant of the heuristic approaches in which all possible scenarios are enumerated; and 4) procedural modeling. The last approach was first popularized in computer graphics to create visual models, textures, and/or animations where rules and symbols were used to represent the elements being modeled. Examples of the procedural rules encoding are “shape grammars”, L-systems (Lindenmayer systems) or Functional L-systems (Martin and Hughes, 2010).

An example of the use of procedural rules is the generation of scenarios for training emergency rescue crews (Hullett and Mateas, 2009). The specific goal was to generate scenarios for collapsed structure rescue. To develop the domain model, the authors used a hierarchical task network (HTN), which was derived from procedural methods such as L-systems to display geometric models of plants and buildings. As in the earlier reference to work by Martin and Hughes, the approach of Hullet and Matea builds concepts into rules that can quickly implement the conceptual objects as needed. While the goal was to generate a large number of “levels by randomly combining elements,” the authors found that “some plans result in the same scenario by performing the same steps in different order.” Furthermore, as the size of the scenario grows, “the combinatorial number of possible variations grows…”. The researchers found that an approach of creating variation by “changing the specified goals” created a wider variety of scenarios than using random generation techniques. The authors also note that a “human author is responsible for the knowledge engineering work required to specify both methods (tasks) and operators” (Hullett and Mateas, 2009).

Grois et al. (1998) used a Bayesian-network based model for generating scenarios for crisis management training. The software developed, called *ScenGen*, was based on Noisy-OR Bayesian Networks (NOBNs). The authors then compared their software to three conventional alternative approaches: 1) manual generation of scenarios by a human expert; 2) randomly-instantiated scenarios created by selecting a random initial event (corresponding to a root node of the NOBN net and propagating the network forward); and 3) case-based stochastic perturbation, which is a hybrid of the first two approaches. In this third traditional approach, human subject matter experts create seeds for the network, perturbing them “so as to obtain variability.” The authors comment that in using their new approach “the very selection bias that gives *ScenGen* its power also significantly limits its