9-19-2011

Pre Incident Indicator Analysis (PIIA) System

Frank Gilfeather

Pre Incident Indicator Analysis (PIIA) System

Frank Gilfeather, Thomas P. Caudell, Mahmoud Reda Taha and Dave Weinberg[1]
University of New Mexico[2]

August 17, 2011

UNM Technical Report EECE-TR-1-0008

[1] Practical Risk, LLC
[2] The UNM team consists of four primary faculty, a number of students, postdocs and outside consultants.

Pre Incident Indicator Analysis (PIIA) System

Overview

Policy leaders have identified a need for a mathematical based computational system that would support requirements of situational awareness to analyze terrorist threats and risks given the vast amount of input data in the form of open source, intelligence as well as other indicators. In response to these expressed needs[3], the University of New Mexico (UNM) developed a proto-type system, the Pre Incident Indicator Analysis (PIIA) system, which is capable of combining open source and intelligence data with other indicators and information as inputs to provide a dynamic assessment of threats and risks along with a measure of the uncertainties inherent in that analysis. The object is to build a mathematical and then a prototype system to access large complex "faceted information ontologies" consisting of information and data available to analysts and in a fashion they can use. A framework has been established for such a system and demonstrated. In addition considerable progress has been made on the key challenge of creating faceted ontologies (combining specific views of data and information). This constitutes a data collection system that combines specific views or ontologies, allows for analyzing the integrated information and then ascertaining certain new views of the resultant information. This has the potential to evaluate specific threats for importance, provide temporal sensitivity and provide warnings for analysts to consider.

National Security Imperative

Based on recent events identifying terrorist risk through integrating and analyzing intelligence and open source data is seen as extremely important. The PIIA system started four years ago as a fresh look at this issue, using recent research advances and a multidisciplinary team. The system is a research project and is focusing its efforts in areas such as use of evidence and possibility theory, faceted ontologies, visualization techniques, IA and machine learning methods for classifying, understanding and representing exceptionally large data sets consisting of disparate indicators and information. Data points even if highly creditable by themselves are only the start of successful analysis when it takes many steps to constitute an event chain of a terrorist event. The intent of PIIA is to associate these data points with others and then integrate the evidence from all the linked data to rank order a set of high evidence scenarios for decision and warning analysts to consider. Insofar as chemical, biological, radiological, nuclear and explosive (CBRNE) terrorist attacks represent sufficiently-significant risks, a thorough analysis of pre-incident indicators of the potential risk for these events and a determination of the threat level is a necessary capability for national security.

PIIA Objectives and Structure

UNM has designed PIIA as an information-processing and analysis system that aggregates information or intelligence from several viewpoints or facets into larger faceted ontologies of information, then allows for projecting from such a combined faceted ontology potential events with specified

[3] These efforts are partially supported by several grants and contracts from DTRA, the DHS Office of Risk Management Analysis (RMA) through ANL and ORAU, and a CAE grant through ODNI.

characteristics such as the high plausibility of potential events. PIIA utilizes the mathematical model of faceted ontologies described in detail in [Caudell et al. 2011] and utilizes models within evidence theory to organize and aggregate evidence toward deriving the plausibility and uncertainty of specific events [Ross et al, 2011]. Fictionalized and open-source information has been created to provide test data for exercising the tool. Within the constraints of limited test data, the methodology and software developed in the PIIA project provides a defensible approach to connecting information in such a quantitative and reproducible way as to provide a mechanism to prioritize events such as attack scenarios by either their plausibility and possibility or their degree of uncertainty.

To be successful the system is dependent on gaining an understanding of how to:

- move information from an analyst view or other information source to a larger structured faceted information ontology incorporating multiple facets or views while preserving consistency,
- aggregate data associated with the information including measures of evidence and uncertainty,
- access the combined faceted information ontology to project onto a new view which is designed to expose certain information, i.e., likely events, and
- visualize in transparent ways the processes and results.

An approach via faceted ontologies to the data structures is under active research and development [Caudell et al, 2011]. This approach includes developing faceted ontology theory based on mathematical category theory to allow for creation from local ontologies of information a larger "faceted ontology" as well as projecting useful new faceted ontologies from the combined faceted ontology. Examples of views or facets into a combined faceted ontology might include information on potential target classes, specific weapon information, specific geographic areas and/or infrastructure elements along with information on adversarial groups or agent analysis. A projection facet or view from such a combined faceted ontology might be descriptions of highly likely group actions against certain types of targets. Such a facet might contain scenarios and measures of the accumulated evidence for them.

For the initial implementation[4] of PIIA, we focused on identifying highly important attack scenarios from identifying the adversary's intent including methods and targets as well as data on targets and attack event. Thus we aggregated data and evidence to form event chains and then to select important scenarios.

Figure 1: A basic top level event chain representation of a scenario

A scenario is an ordered set of events, actions or states that leads to an attack on a target. In PIIA, a scenario is a description of the process an adversary uses to carry out an attack, including type, agents and geospatial and temporal information. PIIA architecture will use the large faceted ontology

[4] Use of multiple facets is a feature to be inserted into the PIIA implementation. The initial test implementation considered just an attack type facet.

including information and evidence on all the needed steps to project out a view or facet consisting of such scenarios. The various ontologies can be local or scale to a large perspective so that the scenarios are determined by the ontologies that go into making the combined faceted ontology used to project from. This also allows for scaling of the problem, e.g., a facet may represent a kind of target or a geographical area, an intent group or a class of groups. Scenarios will be generated based on the facets or ontologies contributing to the combined faceted ontology being built and analyzed.

The key advantage of using such ontologies is that various parameters (e.g., source credibility, evidence credibility, evidence applicability, ambiguity, attractiveness and meta data) can be assigned to information within the data structure for later computation. Using Fuzzy Set Theory, class membership values are assigned to these parameters using linguistic values associated to the evidence indicators. Using evidence theory, possibility theory and other related aggregation methods we assign levels of belief and plausibility to information in the combined faceted ontology that includes all the relevant indicator parameters such as various creditability factors, as described below. Subsequently the projected ontology can be "solved" in various ways depending on the analysis being queried and such a solution produces scenarios and provides values for the "belief" and "plausibility" for attacks as derived from the indicators (open or intelligence evidence) provided to the model.

Faceted Ontologies

Faceted ontology theory we feel offers a new way to view large data sets such as the massive amounts of open source and IC data and information. The theory arose in library information science and has become an important feature of web search and data management such as in the medical and health science fields. Basically it refers to using multi approaches to large data sets and organizing collections (e.g., books, web catalogues, patient data, etc.) through a common set of characteristics.

Within PIIA we seek to link evidence in each facet to that in the larger combined faceted ontology and also link evidence across facets to be identified and evaluated [Caudell et al, 2011]. Then using forms of cross analysis, links from some evidence will trigger further links, reinforce evidence within facets, and lead to identifying a set of high evidence scenarios containing information from all facets. Ideally one should be able to query any of the terms through any of these facets with a system

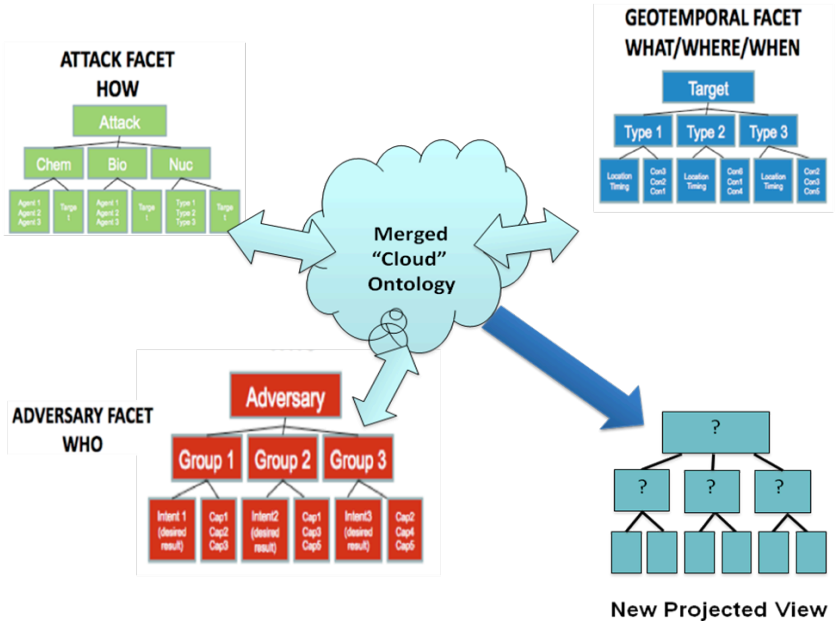

Figure 2: Ontologies are combined and then projected to a new view.

such as PIIA. Then using the theory of faceted ontologies and including artificial intelligence (AI), machine learning and other internal analysis, PIIA will locate all the links and arrive at a set of high evidence scenarios. The structure is illustrated in Figure 2.

Evidence Theory

For PIIA we consider several methods for evidence aggregation to use in order to determine say the level of risk (or threat) for scenarios. All these methods allow the use of crisp (certain) or fuzzy (uncertain) open source and intelligence inputs. The methods used in PIIA provide the platform for aggregating all types of evidence and indicators of an event. In developing PIIA we made several assumptions all of which are subject to modification and revision once specific agency or user requirements are determined[5].

In the initial PIIA implementation, it is assumed that analysts can assign to any piece of information or evidence two metrics: the source credibility (a parameter that captures analysts' judgment of reliability and accuracy of source's information), and the evidence credibility (a parameter that captures analysts' judgment of credibility of specific information included in evidence). Once the aggregated faceted ontology is built from the constituent views then we propagate all pieces of the evidence and their uncertainty. Once a scenario faceted onthology is determined then we can evaluate the evidence, including the uncertainty of any possible attack scenario that can be generated within that ontology based on the event chain for such scenarios. This allows various attack scenarios to be rank-ordered based on their level of evidence with the risk and/or the uncertainty associated with that scenario. This is the essence of PIIA – determine scenarios of special interest to the analyst along with giving overall threat levels for terrorist attack types based on available indicators and intelligence.

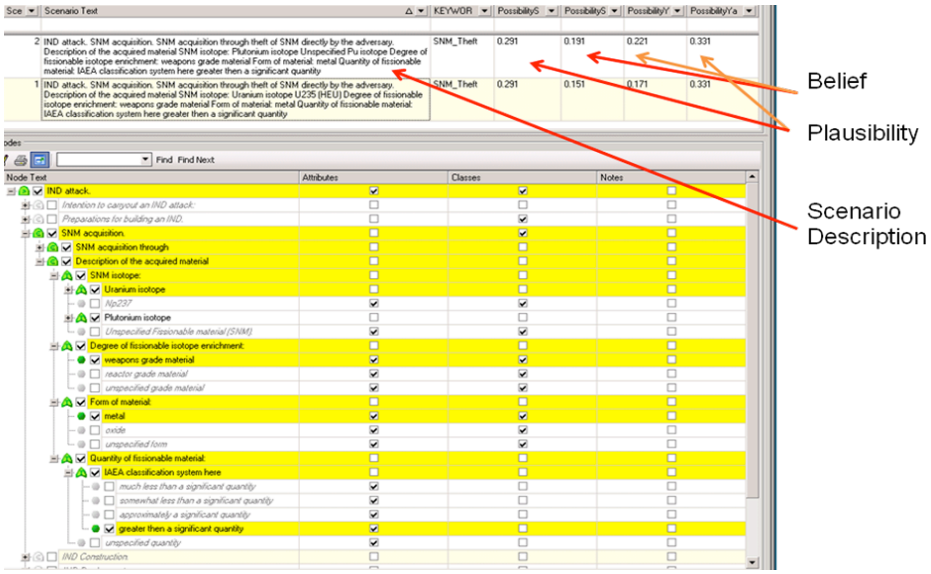

Figure 3: The scenario ontology displays a scenario and the accumulated evidence for it- the first is using DS and then the Yeager algorithm

The evidence analysis methods are based on fuzzy set theory [Ross 2004, Grabisch et al. 1994] which allows computing evidence mass and an overall source credibility for each evidence element of the

[5] Our assumptions were considered reasonable by consultants who have or are security analysts and are placeholders for other types of analysis needs.

attack scenario trees. The use of Dempster-Shafer (DS) evidence [Sentz et al. 2002] to aggregate all evidence at each element of the scenario facet is used. DS evidence theory establishes the aggregated evidence interval between 0 to 1 corresponding to belief and to plausibility. Six methods for evidence aggregation within the DS framework of evidence theory have been examined and several suitable methods have been selected as shown in Figure 3. A fused interval of belief and plausibility is established for each branch of the scenario tree. The use of possibility theory is then used to aggregate the evidence intervals at the different branches. The proposed methods have been examined with mock-up data and have proved to work effectively.

Selection of Dempster-Shafer (DS) and possibility theories and avoiding classical probability theory for evidence propagation in terrorist attack scenarios is attributed to two reasons: First, all evidence on terrorist events is typically uncertain and such uncertainty is related to epistemic and aletoric (random) uncertainty. Additional knowledge can reduce uncertainty that is calculated naturally in PIIA but more difficult to handle using probability theory. Second, information collected as evidence cannot be connected using the principle of insufficient reasoning, which requires all possible outcomes to be equally likely. Therefore, we suggest that evidence of terrorism activities be handled best within the framework of epistemic uncertainty using DS evidence theory. As opposed to probability theory, Dempster-Shafer theory admits the set of evidence is incomplete (allows ignorance). DS theory also allows the allocation of the evidence mass, *m*, to sets rather than the allocation of probabilities to singletons. However, DS theory assumes independence of evidence sources and dependencies need special treatment [Ross et al, 2011]. As applied, DS evidence theory allows the calculation of an evidence interval that ranges between belief and plausibility. The difference between these bounds of the evidence interval is equal to the level of ignorance (uncertainty) we have about the attack scenario. A case study on information about the computation of the possibility of a nuclear material attack was demonstrated in an early PIIA exercise as shown in Figure 3. This approach can be used to rank order possible terrorist attack scenarios while allowing the propagation of information uncertainty.

Visualization in PIIA

The faceted ontologies used and generated by PIIA will grow very complex and large. The projected facet ontology mentioned above consisting of scenarios are themselves a case in point and visualizing them or any faceted ontology becomes a challenge. Since the scenario faceted ontology with its evidence accumulated is of special interest, we developed a visualization tool for it. The Figure 4 below illustrated part of this visualization for an early version of PIIA. Here scenarios are shown in a graph structure with resulting possibility and plausibility range shown for each of them. The visualization allows one to magnify a portion of the graph and scroll through all scenarios of high interest. Details of each would be reveled with drill down to the level of the evidence used to specify it as a scenario of interest. The visualization of the faceted ontologies is an active area of research with lots of potential for systems providing transparency and drill down capabilities.

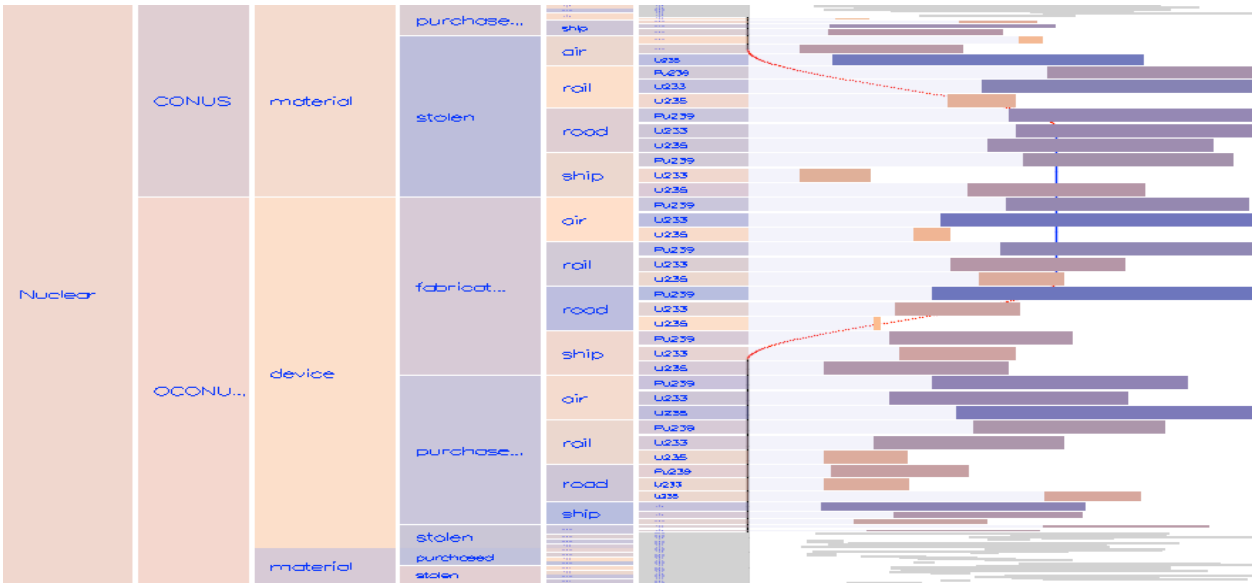

Figure 4: Scenario facets are generated showing aggregation of evidence - uncertainty is represented by the length of the bar.

Summary

PIIA is thus a framework and a set of algorithms for building and analyzing information and generating views such as a scenario ontology for analysis. Its goal is to extend that capability to allow methods of accessing and providing portals or facets into a combined large faceted ontology of data and information in a consistent manner. PIIA implementation is thus dependent on the ongoing research into the structure of faceted ontologies that is aimed specifically at ways to organize these viewpoint-specific semantic data structures and combine them into a larger consistent data structure useful for analysis. The result will be structures that are more readily interpretable, robust, and provide a perspective neutral representation. These simpler semantic structures or facets are generated from various sources focusing on, for example, socio-cultural networks, geo-spatial distributions, or threat event scenario trees. When synthesized into a logical whole, the resulting structure will produce a common situational picture that, for example, will give decision makers insight into the roles, goals, relationships, and rules of behavior of relevant groups or individuals. Starting from the neutral representation, and based on rational assumptions, a semantic projection into possibly a specific or new facet is then possible, aiding in the discovery of missing informational clues and possibly obscure clandestine activities of outside groups or terrorists. Our PIIA research efforts are aimed at creating and using new knowledge in the area of ontology merging, faceted ontologies, inference across these knowledge structure and the visualization of their complex inter-relationships to generate programs useful in situational awareness and analysis of threats.

References

Ayyub, Bila (2001). Elicitation of Expert Opinions for Uncertainty and Risks, CRC Press, Washington DC.

Caudell, T. P., Healy, M. J., Sanchez-Silva, R. C., ”A Categorical Model for Faceted Ontologies with Data Repositories”, UNM Technical Report: EECE-TR-11-0002, March 21, 2011.

Dubois, D. and Prade, H. (1986), "A Set-Theoretic View on Belief Functions: Logical Operations and Approximations by Fuzzy sets." International Journal of General Systems 12: 193-226

Dubois, D. and Prade, H. (1992), "On the Combination of Evidence in Various Mathematical Frameworks." Reliability Data Collection and Analysis. J. Flamm and T. Luisi. Brussels, EEC, EAFC: 213-241.

Grabisch, M., Nguyen, H. T. and Walker, E. A. (1984), *Fundamentals of Uncertainty Calculi with Applications to Fuzzy Inference,* Kluwer Academic Press, Boston, USA.

Grabo, Cynthia M., (2004), *Anticipating Surprise – Analysis for Strategic Warning*, University Press, Lanham, Maryland, USA.

Hoffman, B., (2006), *Inside Terrorism* (Revised and Expanded Edition), Columbia University Press, New York.

Inagaki, T. (1991), "Interdependence between Safety-Control Policy and Multiple-Sensor Schemes Via Dempster-Shafer Theory." IEEE Transactions on Reliability 40(2) 182-188

Joslyn, C. (1997), "Measurement of Possibilistic Histograms from Interval Data," International Journal of General Systems, Vol. 26, Issue (1-2), 9-33.

Klir, G. J. (2006), *Uncertainty and Information*, John Wiley and Sons, Hoboken, NJ.

Klir, G. J. and Yuan, B. (1995), Fuzzy Sets and Fuzzy Logic, Prentice Hall, Upper Saddle River, NJ, USA.

Oberkampf, W.L., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F. and Ferson, S. (2004), "Challenge Problems: Uncertainty in System Response Given Uncertain Parameters". Reliability Engineering and System Safety, 85, 11-19.

Ross, T. J., (2004), *Fuzzy Logic with Engineering Applications*, John Wiley & Sons. UK.

Ross, T., Taha, M. R., Kim, J. J. and Gilfeather, F., "Logical models for the propagation of disparate information and uncertainty across effectivity trees", Integrated Computer-Aided Engineering 18 (2011) 251–264

Sentz, K., Ferson,S. (2002), "Combination of Evidence in Dempster-Shafer Theory", SAND 2002-0835. April.

Shafer, G. (1976), *A Mathematical Theory of Evidence*. Princeton University press. Princeton, NJ.

Shafer, G. (1986), *Probability Judgment in Artificial Intelligence*. Uncertainty in Artificial Intelligence. N. Kanal and J. F. Lemmer. New York, Elsevier. 4.

Soundappan, P., Nikolaidis, E., Haftka, R.T., Grandhi, R.V. and Canfield, R.A. (2004), "Comparison of Evidence Theory and Bayesian Theory for Uncertainty Modeling," Reliability Engineering and System Safety, 85, 295-311.

Yager, R. (1986), Arithmetic and Other Operations on Dempster-Shafer Structures. International Journal of Man-Machine Studies 25: 357-366.

Yager, R.(1987a), "On the Dempster-Shafer Framework and New Combination Rules." Information Sciences 41: 93-137.

Yager, R. (1987b), "Quasi-Associative Operations in the Combination of Evidence." Kybernetes 16: 37-41.

Zadeh, L. A. (1984), "Review of Books: A Mathematical Theory of Evidence." The AI Magazine 5(3): 81-83.