1-1-2014

Analyzing Current and Future Catastrophic Risks from Emerging-Threat Technologies

Anthony Michael Barrett
University of Southern California

Analyzing Current and Future Catastrophic Risks from Emerging-Threat Technologies

Project Technical Description

1. **Theme Areas:** Operations Research in Risk Management; Risk Analysis - Management of Risks from Intelligent, Adaptive Adversaries
2. **Principal Investigator:** Anthony M. Barrett
3. **Institution:** Global Catastrophic Risk Institute
4. **Co-Investigators:** Jun Zhuang (SUNY Buffalo), Seth Baum (Global Catastrophic Risk Institute)
5. **Keywords:** Emerging-threat assessment, intelligence analysis, risk and decision analysis

6. Background and Brief Description:

Events such as terrorist attacks and technological accidents can have a range of possible consequences. Potential events with extremely catastrophic consequences merit at least some risk analyses of greater scope, including longer time horizons, than for less-consequential events. Some rapidly developing technologies, such as in synthetic biology or critical infrastructure control systems, pose much greater risks of catastrophe in the future than they do today. These risks could be significantly reduced with enough foresight and advance warning. However, risk assessments that focus on what is possible in the relatively near term may overlook indicators of developments that would enable more-catastrophic events.

These risks are difficult to characterize, partly because of the general challenges of technological and long-term forecasting. Assessing the tradeoffs of risk-management options is also complicated by the difficulties and uncertainties of determining when some conditions for catastrophe are present.

In this project, we develop a methodology for analyzing risks and risk-management tradeoffs of potential emerging threats by systematically identifying potential catastrophe-enabling developments and indicators of precursor events; estimating probabilities of facing precursors; assessing tradeoffs of available options; and continually monitoring for potential indicators of catastrophe precursors and updating probability estimates with new information and new judgments. We apply this methodology to a potential emerging-threat technology area, such as the use of synthetic biology to produce bio-weapon agents [1]. While some methodology components have been applied in terrorism or technology-development assessments, to the best of our knowledge, they have not been developed into a single integrated methodology for use by risk practitioners, nor has such a methodology been applied to real-world emerging-threat problems.

7. Objectives:

The objectives of this research are to develop and apply an integrated methodology to use available information and expert judgment to: a) identify ways in which catastrophic terrorist attacks or accidents could occur with new technological developments, b) estimate risks, c) assess tradeoffs of options, and d) update assessments. The methodology will be designed for implementation by risk practitioners.

8. Interfaces to CREATE Projects:

This work will establish interfaces with CREATE's current projects on adaptive-adversary terrorism risk analysis, expert elicitation, and intelligence analysis, looking for opportunities for cross-fertilization. We will also seek to leverage elements of previous CREATE-funded work, such as on Markov decision process models for Bayesian updating and warning-time dynamics [2], and expert elicitation tools [3].

9. Major Deliverables, Research Transition Products and Customers:

Project deliverables will include (1) a report that presents the methodology developed in this project and the findings from the application of this methodology to the case study, and (2) software that analysts can use to apply the methodology to other cases. The report will include discussion of the methodologies' limitations, potential extensions, and steps for implementation by risk practitioners.

Products will include: A risk assessment methodology and software for assessing and managing emerging-threat catastrophe risks associated with developing technologies; research publications; and reports. Publications will appropriately protect sensitive information, while providing description of our methods sufficient to allow other researchers to examine and employ them.

The quantity of investigators' efforts will necessarily be limited in the proposed one-year project. We would propose more extensive related work in follow-on projects. The current year's methodology application case will be designed with the end customer. The scope of the case will be selected to maximize customer value within constraints, and to allow prototyping of all methods (e.g. for intelligence-monitoring or model updating) to some degree if possible.

We already have a DHS end customer for a methodology application case, have discussed the project with them, and have incorporated their feedback into this statement of work. The end customer is the National Biosurveillance Integration Center (NBIC) at the DHS Office of Health Affairs. Preliminary discussions with the DHS end customer, NBIC, have indicated that our project would enhance their ongoing efforts to create and implement a risk-based framework for biosurveillance by helping them to systematically focus on key risk indicators. This project's methodology components for analyzing tradeoffs of risk management options, such as intelligence resource allocations, is intended to support prioritization of NBIC's biosurveillance activities and associated DHS risk-reduction options. NBIC has also indicated that they are interested in emerging threats, which our project addresses. For methodology application case scenario sets, the project's current leading candidates involve use of synthetic biology technologies, which NBIC has indicated would be appropriate; the final choice of scenarios will be decided with NBIC after the period of performance begins. We also have discussed personnel needs with NBIC, and identified people that could serve as sources of experts to complement NBIC's own resources, e.g. on synthetic biology technologies.

10. Technical Approach:

We will begin with a standard event risk function in which the event's risk is the product of the probability and magnitude of that event. We will focus on scenarios with consequences above particular thresholds, as in definitions of global catastrophes [4]. (We will consider public policies, risk attitudes, and analytic constraints in setting thresholds.) We will identify and estimate the probabilities of combinations of events that would cause catastrophe scenarios, looking over specific time periods, e.g. the next 30 years.

We will use the following basic five-step approach:

1) Construct, populate and integrate the following types of models, using the literature, expert opinion about the systems involved, and other available information:
    a) Catastrophe fault-tree accident and attack path logic models, working backward from catastrophe scenarios [5, 6].
    b) Technology development event trees (or equivalent Bayesian networks / influence diagrams), working forwards from the current state of the world.
2) For catastrophe-enabling development events:
    a) Estimate the dates when those developments (i.e. public availability of a particular type of equipment) will occur, as in Baum et al. [7] and using standard methods for characterizing and propagating uncertainty [8, 9].

   b) Identify likely indicators of those developments, and estimate conditional probabilities of developments given presentation of those indicators.
3) Construct Bayesian models that incorporate the models and probability estimates in steps 1 and 2 to make projections of event probabilities and risk profiles extending some time into the future, which can be updated using new information about technological developments [10, 11].
4) Represent tradeoffs of risk management options, such as intelligence resource allocations or risk-reduction policies, in terms of option effects on risk model parameters (e.g. to reduce uncertainties or reduce catastrophe event probabilities), and in terms of option costs.
5) Integrate the basic risk models from steps 1-3 and the option-tradeoff models in step 4 to support decision analysis of risk management options [12, 13].

The methodology will aim to facilitate the following:

- First, to allow an intelligence or monitoring group to continually a) watching for signal events indicating that a catastrophe precursor had presented itself, and b) updating assessments of both risk profiles and tradeoffs of risk-reduction options, using Bayesian methods. [14]
- Second, to periodically (e.g. annually) update the model-building and analyses in steps 1 through 3, reflecting technological developments, as well as information gathered about base rates of indicator events, again using Bayesian methods.

As part of the project, we plan to develop and apply computational models using commercially available off-the-shelf (COTS) software. This software will be based on influence diagrams, i.e. Bayesian networks with decision and value nodes [10], and will provide stochastic simulation of uncertainty using Monte Carlo algorithms. Such software systems could include *Analytica* by Lumina Decision Systems or *Netica* by Norsys Software. Monte-Carlo based simulation within such models can allow use of fairly straightforward functions of risk of an event, i.e. Risk = Probability × Consequence, while accounting for correlation between risk function arguments, i.e. between event probability and event consequence [15]. The computational models will be provided to the application case research end customer to aid in prototyping of methods and implementation by risk practitioners.

The project management roles are as follows: The PI (Barrett) will have primary responsibility for the project's overall direction, application models, and interactions with the end customer. The Co-PI (Zhuang) and his graduate student will define and refine important elements of the mathematical methods, and contribute to other areas as appropriate. The third investigator (Baum) will primarily support expert elicitations.

References:

1. PCSBI, *New Directions: The Ethics of Synthetic Biology and Emerging Technologies*. 2010, Presidential Commission for the Study of Bioethical Issues: Washington DC.
2. Pate-Cornell, E. and D. Blum, *Warning for Generalized Dynamic Crisis Scenarios*, in *Current Research Project Synopses*. 2012, CREATE.
3. Hora, S.C., *Probability Modeling and Expert Judgment*, in *Research Project Summaries*. 2011, CREATE.
4. Bostrom, N. and M.M. Cirkovic, *Introduction*, in *Global Catastrophic Risks*, N. Bostrom and M.M. Cirkovic, Editors. 2008, Oxford University Press: New York. p. 1-29.
5. Rasmussen, N.C., *The Application of Probabilistic Risk Assessment Techniques to Energy Technologies.* Annual Review of Energy, 1981. **6**: p. 123-138.
6. Cheesebrough, A.J., *Risk Assessment Process for Informed Decision-making (RAPID): A Progress Report*, in *Fourth Annual Conference on Security Analysis and Risk Management*. 2010: Arlington, Virginia.

7. Baum, S.D., B. Goertzel, and T.G. Goertzel, *How long until human-level AI? Results from an expert assessment.* Technological Forecasting & Social Change, 2011. **78**(1): p. 185-195.
8. Hora, S.C., *Eliciting Probabilities from Experts*, in *Advances in Decision Analysis: From Foundations to Applications*, W. Edwards, R.F. Miles Jr., and D.v. Winterfeldt, Editors. 2007, Cambridge University Press: New York. p. 129-153.
9. Morgan, M.G. and M. Henrion, *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. 1990, Cambridge: Cambridge University Press.
10. Pearl, J., *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. 1988, San Mateo, California: Morgan Kaufmann.
11. Pate-Cornell, E., *Fusion of Intelligence Information: A Bayesian Approach.* Risk Analysis, 2002. **22**(3): p. 445-454.
12. Green, D.M. and J.A. Swets, *Signal Detection Theory and Psychophysics*. 1966, New York: John Wiley and Sons.
13. Paté-Cornell, M.E., *Warning Systems in Risk Management.* Risk Analysis, 1986. **6**(2): p. 223-234.
14. Ni, K.S., et al., *Exploitation of Ambiguous Cues to Infer Terrorist Activity.* Decision Analysis, 2013. **10**(1): p. 42-62.
15. Cox Jr., L.A., *Some Limitations of "Risk = Threat x Vulnerability x Consequence" for Risk Analysis of Terrorist Attacks.* Risk Analysis, 2008. **28**(6): p. 1749-1761.