

Navigating Various Shades of Grey

Uncertainty Representation within the Threat Intelligence Domain

Ritten Roothaert¹

¹Vrije Universiteit Amsterdam, De Boelelaan 1111, 1081 HV, Amsterdam, The Netherlands

Abstract

The Threat Intelligence Decision Ontology (TIDO) was developed as a common vocabulary for Threat Intelligence (TI) analysts to structure and support their decision making. However, a major limitation of TIDO is the absence of a formal representation of the inherent uncertainties associated with TI data, which originate from various sources such as imperfect sensors, biases, and incomplete information. This ongoing research aims to address this critical gap by characterising the uncertainties involved in a TI decision process and comparing these characteristics to the capabilities of three selected uncertainty representations: Bayesian Probability Theory, Probability Bound Analysis, and Belief Function Theory.

Keywords

uncertainty representation, threat intelligence, interval probabilities, belief function theory

1. Introduction

Intelligence institutions, some of the most well known being the Mossad, CIA, MI5 and MI6, seek to access the secrets of their adversaries while preventing their adversaries from accessing their own secrets. They do so, at least normatively speaking, in the service of moral rights [1]. However, obtaining access to the secrets of adversaries inevitably results in the privacy of this adversary being violated to some extent. To ensure that this power to violate the privacy of others is not misused, the permissible actions and the conditions under which they may be applied are subjected to legislation and compliance to this legislation is monitored by independent institutions. The Threat Intelligence Decision Ontology (TIDO)¹ was designed to support these monitoring institutions by providing a unified vocabulary that can describe the decision process of Threat Intelligence (TI) analysts.

Currently, the situational understanding of the TI analyst, an integral part of the justification for the actions taken by intelligence institutions, is expressed in TIDO using pieces of information, hypotheses, and supporting or disputing the relationships between them. This representation resembles Thagard's Theory of Explanatory Coherence (TEC) [2] and captures the core argumentative structure of the analyst. However, it lacks the expressiveness to address the uncertainties that are inherently associated with the information used within the TI domain, as AI solutions and human experts have biases, models are imperfect, sources might be poor or misleading, and data-sharing might be limited for operational, strategic, or legal reasons [3]. Consequently, at each step within a TI decision process, an TI analyst must review this limited and uncertain information [4], which means that in order to fully understand *why* a decision was made, the explanation of the decision should also address the uncertainties involved in the decision process. The work presented in the poster consists of an investigation on the suitability of three different uncertainty representations within the TI domain and how these representations could be integrated into the TIDO ontology.

Ethical considerations The author is aware of the sensitivity of the subject and the potential implications with respect to the privacy of citizens. The motivation, though, is not to collect new data but to provide a vocabulary to enrich and structure data already obtained by the intelligence services.

K-CAP 2025 Companion Volume

✉ h.m.roothaert@vu.nl (R. Roothaert)

ORCID  0009-0008-7843-6513 (R. Roothaert)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

¹Accessible at: w3id.org/tido

This enables the decision makers to make their assumptions and considerations explicit, increasing their accountability and improving the testability of their decisions with respect to compliance with legislation and internal policies. The proposed research does not provide intelligence agencies with additional capabilities, nor justifications, to collect more data, and therefore, does not negatively impact the privacy of citizens.

2. Research Goals

The objective of the proposed research is to explore the suitability of three distinct quantitative uncertainty representations with respect to integration into the TIDO ontology. For our analysis, we include Bayesian probability theory (BPT, [5]), probability bound analysis (PBA, [6, 7]) and belief function theory (BFT, [8, 9]). Although this is not an exhaustive list of all possible uncertainty representations, it provides a manageable number of widely used approaches with which to start our study. With these representations in mind, we aim to answer the following research questions:

RQ1: What are the characteristics of the uncertainties associated with the data used for decision making within the TI domain?

RQ2: What are the benefits and limitations of the selected uncertainty representations?

RQ3: How can the data characteristics identified in RQ1 be captured by the representations analysed in RQ2?

RQ4: What tooling is available for the analysed uncertainty representations?

3. Current progress and upcoming steps

The first preliminary results were presented at the RuleML+RR 2025 Doctoral Consortium track [10]. These results focused on **RQ1**, where an analysis of a Dutch military doctrine [11] and the recent literature characterising the TI domain [1, 3] are combined with an uncertainty characterisation originating from the information fusion domain [12]. A start is made with answering **RQ2**, but at the moment the analysis is based heavily on a single review by Zio and Pedroni [13]. To answer **RQ3**, the plan is to create a collection containing practical examples that highlight the uncertainty characteristics identified in **RQ1** and showcase how the benefits and limitations identified in **RQ2** would express themselves in these examples. Finally, **RQ4** will be addressed by creating an inventory of available software and reasoners that can process data represented according to the uncertainty representations analysed.

Acknowledgments

This PhD project is supervised by Professor Fabio Massacci, and co-supervised by Associate Professor Stefan Schlobach and Assistant Professor Lise Stork. We gratefully acknowledge the funding support by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) under the KIC HEWSTI Project with grant no. KIC1.VE01.20.004.

Declaration on Generative AI

During the preparation of this work, the author used the Writefull extension of overleaf for grammar and spelling checking. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

- [1] S. Miller, *On National Security Intelligence*, 1 ed., Routledge, London, 2024, pp. 33–50. doi:10.4324/9781003106449-3.
- [2] P. Thagard, Explanatory coherence, *Behavioral and Brain Sciences* 12 (1989) 435–467. doi:10.1017/S0140525X00057046.
- [3] S. Van Gerwen, J. Constantino, R. Roothaert, B. Weerheijm, B. Wagner, G. Pavlin, B. Klievink, S. Schlobach, K. Tuma, F. Massacci, To Know What You Do Not Know: Challenges for Explainable AI for Security and Threat Intelligence, in: T. Sipola, J. Alatalo, M. Wolfmayr, T. Kokkonen (Eds.), *Artificial Intelligence for Security*, Springer Nature Switzerland, Cham, 2024, pp. 55–83. doi:10.1007/978-3-031-57452-8_4.
- [4] F. Massacci, A. Papotti, B. Weerheijm, S. A. M. van Gerwen, E. Mezzi, J. Constantino Torres, R. Roothaert, Hybrid Explainable Workflows for Security and Threat Intelligence, 2023. URL: <https://www.nwo.nl/en/projects/kich1ve0120004>.
- [5] B. De Finetti, *Theory of Probability: A Critical Introductory Treatment*, John Wiley & Sons, Chichester, UK Hoboken, NJ, 2017.
- [6] J. A. Cooper, S. Ferson, L. Ginzburg, Hybrid Processing of Stochastic and Subjective Uncertainty Data, *Risk Analysis* 16 (1996) 785–791. doi:10.1111/j.1539-6924.1996.tb00829.x.
- [7] S. Ferson, V. Kreinovich, L. Ginzburg, F. Sentz, Constructing Probability Boxes and Dempster-Shafer Structures, Technical Report SAND2002-4015, 809606, 2003. doi:10.2172/809606.
- [8] A. P. Dempster, Upper and Lower Probabilities Induced by a Multivalued Mapping, *The Annals of Mathematical Statistics* 38 (1967) 325–339. doi:10.1214/aoms/1177698950.
- [9] G. Shafer, *A Mathematical theory of Evidence*, Princeton University Press, 1976.
- [10] R. Roothaert, Uncertainty Representation and Reasoning Within the Threat Intelligence Domain, in: *Companion Proceedings of the 9th International Joint Conference on Rules and Reasoning*, Istanbul, Türkiye, in press.
- [11] Dhr. P.A. Brouwer, Maj M. Scholten, *Joint Doctrine Publicate 2: Inlichtingen*, 2012.
- [12] E. Blasch, A.-L. Joussetme, K. B. Laskey, P. C. Costa, J. P. De Villiers, G. Pavlin, C. Laudy, URREF Risk analysis towards Data Fusion Certification, in: *2023 26th International Conference on Information Fusion (FUSION)*, IEEE, Charleston, SC, USA, 2023, pp. 1–8. doi:10.23919/FUSION52260.2023.10224171.
- [13] E. Zio, N. Pedroni, Literature Review of Methods for Representing Uncertainty, Technical Report, Fondation pour une culture de sécurité industrielle, 2013.