Tracking uncertainty propagation from model to formalization: Illustration on trust assessment

VALENTINA DRAGOS
JEAN DEZERT
KELLYN REIN

This paper investigates the use of the URREF ontology to characterize and track uncertainties arising within the modeling and formalization phases. Estimation of trust in reported information, a real-world problem of interest to practitioners in the field of security, was adopted for illustration purposes. A functional model of trust was developed to describe the analysis of reported information, and it was implemented with belief functions. When assessing trust in reported information, the uncertainty arises not only from the quality of sources or information content, but also due to the inability of models to capture the complex chain of interactions leading to the final outcome and to constraints imposed by the representation formalism. A primary goal of this work is to separate known approximations, imperfections and inaccuracies from potential errors, while explicitly tracking the uncertainty from the modeling to the formalization phases. A secondary goal is to illustrate how criteria of the URREF ontology can offer a basis for analyzing performances of fusion systems at early stages, ahead of implementation. Ideally, since uncertainty analysis runs dynamically, it can use the existence or absence of observed states and processes inducing uncertainty to adjust the tradeoff between precision and performance of systems on-the-fly.

Manuscript received January 17, 2018; revised October 2, 2018; released for publication March 20, 2019.

Refereeing of this contribution was handled by Anne-Laure Jousselme.

Authors' addresses: V. Dragos, ONERA—The French Aerospace Lab, Chemin de la Hunière, Palaiseau, France (E-mail: valentina.dragos@onera.fr). J. Dezert, ONERA—The French Aerospace Lab, Chemin de la Hunière, Palaiseau, France (E-mail: jean-dezert@onera.fr). K. Rein, Fraunhofer FKIE, Wachtberg, Germany (E-mail: kellyn.rein@fraunhofer.fkie.de).

I. INTRODUCTION

A key element when designing information fusion systems is the way the system designer isolates and analyzes real world phenomena. A model is abstracted into a simpler representation, in which components, modules, interactions, relationships and data flows are easier to express. Uncertainty tracking highlights approximations induced by model construction and its formalization, as well as providing a checklist to ensure that all uncertainty factors have been identified and considered ahead of system implementation.

This paper illustrates the use of the uncertainty representation and reasoning framework (URREF) ontology [13] to identify and assess uncertainties arising during the modeling and formalization phases of an information fusion system intended to estimate trust in reported information.

Trust assessment is a real-world problem grounded in many applications relying on reported items, with different persons observing and then reporting on objects, individuals, actions or events. For such contexts, using inaccurate, incomplete or distorted items can result in unfortunate consequences and analysts need to ensure the consistency of reported information by collecting multiple items from several sources.

From the perspective of an information analyst, trust can be analyzed along two dimensions: the subjective evaluation of items reported by the source itself, called self-confidence, and the evaluation of source by the analyst, called reliability. While self-confidence encompasses features of subjectivity, the reliability of a source is related to the quality of previously reported items, the competence of the source for specific topics, and the source's capacity for misleading intentions. Trust estimation aims at capturing, in an aggregated value, the combined effects of self-confidence and reliability on the perceived quality of information. The model is represented with belief functions, a formalism which offers a sound mathematical basis to implement fusion operators which estimate trust by combining self-confidence and reliability.

The model developed for trust assessment focuses on the global characterization of information and provides a better understanding of how trust is to be estimated from various dimensions. The overall process has humans as a central element in both the production and the analysis of information.

Trust in reported information offers a good illustration for tracking uncertainty: the phenomenon is complex, so any model adopted is generally a simplification of the real world interactions. Uncertainties can be made explicit not only for static elements of the model, such as sources or items, but also for the dynamic processes of combining items with one another. Moreover, adopting belief functions as representation formalism will have an impact on the way an information system could be implemented and on the accuracy of its results.

The contribution of this paper is twofold: first, it presents a trust estimation model which combines the reliability of sources and self-confidence of reported items, and, second, the paper analyzes types of uncertainty occurring during modeling and formalization by relating elements of the model to uncertainty criteria defined by the URREF ontology.

The remainder of this paper is divided into 8 sections: section II discusses related approaches for trust modeling and uncertainty assessment. The problem tackled in this paper in presented in section III. Section IV describes the model developed for trust estimation, while its implementation with belief functions is presented in section V. The analysis of uncertainty is discussed in VI, while examples and scenarios for trust assessment are presented in section VII. Strengths and limitations of belief-based formalization are discussed in section VIII and section IX concludes this paper.

II. RELATED APPROACHES

The work presented in this paper is related to approaches for trust modeling and assessment as well as solutions for uncertainty analysis for information fusion systems. Trust modeling is not a new research topic; it spans diverse areas such as agent systems [30] and logical modeling and argumentation [50]. The Internet and social media offer new application contexts for trust assessment; this topic is addressed in relation to service provision on the Internet [36], social networks analysis [57], and crowdsourcing applications [64]. Trust analysis is also of interest in the military field where techniques have been developed in order to identify clues of veracity in interview statements [63].

The concept of trust in these communities varies in how it is represented, computed and used. Although having an obvious social dimension, trust is not only understood with regard to other humans, but also towards information pieces [64], information sources [44], Internet sites [21], algorithms for data and knowledge fusion [20], intelligent agents [30], and services for the Internet of things [31].

While definitions of trust vary from one domain to another, there are some common elements. The first commonality for all research areas cited above is to consider trust as a user-centric notion that needs to be addressed in integrated human-machine environments which rely heavily on information collected by humans, even if further processing can be executed automatically. Moreover, all definitions associate some degree of uncertainty with trust, which is then captured by concepts such as subjective certainty [27] and subjective probability [10].

Trust goes hand in hand with the concepts veracity [4] and deception. [45] addresses veracity along the dimensions of truthfulness/deception, objectivity/subjectivity and credibility/implausibility. The authors developed a veracity index ranging from true/objective/credible to untrustworthy/subjective/implausible to characterize texts in the context of big data analysis. Deception is defined as a message knowingly transmitted with the intent to foster false beliefs or conclusions. The topic is addressed in studies from areas such as interpersonal psychology and communication [9], [33] and it is also considered in the field of natural language processing, as part of a larger research direction tackling subjectivity analysis and the identification of private states (emotions, speculations, sentiments, beliefs). These solutions stem from the idea that humans express various degrees of subjectivity [55] that are marked linguistically and can be identified with automatic procedures [54].

Contributions on trust estimation keep the distinction between analyzing the source of information, the item reported and reasoning about trust. Approaches developed for trust in information sources consider that trust is not a general attribute of the source but rather related to certain properties: competence [29], sincerity and willingness to cooperate [50]. On this basis, it becomes possible to consider the competence of a source not in general but with respect to specific topics [28]. Trust can be also analyzed in relation to roles, categories or classes [34].

Research efforts on reasoning about trust analyze information sources from past behaviors rather than directly from their properties [46], or they infer trust from estimations already computed for a set of properties [1]. These approaches generally focus on building trust by using argumentation [62] or beliefs functions [26], or investigating the joint integration of those techniques [52]. Taking this work a step further, [51] identified several patterns for reasoning about trust and its provenance while the notion of conflict in handling trust is discussed in [65].

As shown by approaches above, trust is a multifaceted concept and, in practice, this complex notion can be decomposed into two components: communication or interaction trust, and data trust [48]. The model developed deals with data trust and keeps the distinction between sources and items provided by those sources, although several approaches consider these elements as a whole [26], estimating the trust of information sources [1], [65] rather than information items. The model does not require statistical data to infer the behavior of the source [46] and introduces reliability to characterize the source. More specifically, reliability encompasses not only competence [34], [29] and reputation [28]—two attributes already considered by previous approaches—but also intentions which constitute an original aspect of the model. Intention is of important significance in the context of human-centered systems, including open-sources, and supports the analysis of emerging phenomena such as on-line propaganda or disinformation. Another original aspect of the model is consideration of the characterization of items by the source itself, thus overcoming a main limitation of the solution presented in [12]. Our approach can be considered as partially overlapping solutions investigating trust propagation in direct and indirect reporting [51], [62], and the model enables a particular kind of trust estimation, based both on more or less complete characterizations of the source by the analyst, and more or less accurate characterizations of the items by the source. The model also addresses disagreement and the fusion of diverging opinions, not in a panel of experts as described in [52], but rather between items showing high levels of confidence according to the source and sources having low reliability according to the analyst. By ascribing characterizations to both information sources and reported items, the model allows analysts to make use of both prior experience and their own beliefs in order to assess various degrees of trust.

From a different perspective, the evaluation of uncertainty regarding the inputs, reasoning and outputs of the information fusion is the goal of Evaluation Techniques for Uncertainty Representation Working Group[1] (ETURWG). The group developed an ontology for this purpose [13]. The URREF ontology defines the main subjects under evaluation [18], such as uncertainty representation and reasoning components of fusion systems. Furthermore, the frame also introduces criteria for secondary evaluation subjects: sources and pieces of information, fusion methods and mathematical formalisms. URREF criteria have generic definitions and therefore can be instantiated for applications with coarser or finer granularity levels. This means evaluation metrics can be defined for data analysis [17], increased particularity for data specific types [22] or attributes, reliability and credibility [7], self-confidence [8] or veracity [5].

In addition to allowing a continuous analysis of uncertainty representation, quantification and evaluation, as described in [15], URREF criteria are detailed enough to capture model-embedded uncertainties [37], imperfection of knowledge representations [25], and their propagation in the context of the decision loop [16]. The frame also offers a basis to compare different fusion approaches [24]. URREF criteria were used for uncertainty tracking and investigation in several applications: vessel identification for maritime surveillance [38], activity detection for rhino poaching [43] and imagery analysis for large area protection [6].

Beyond developing a model for trust estimation, this paper also fills a gap within the ETURWG community by illustrating how uncertainty analysis tracks imperfections occurring from problem definition to model abstraction and formalization.

III. HUMAN SOURCES AND REPORTED INFORMATION

Many applications rely on human sources which are used to continuously supply observations, hypotheses, subjective beliefs and opinions about what they sense or learn. In such applications reports are often wrong, due to environment dynamics, simple error, malicious act or intentions, [58]. From the analyst standpoint, decisions have to be made based on indirect reporting and trust relies upon the in-depth investigation of items and sources, thus the analysis of reported items is a critical step. This analysis is a multilevel process, relying on the ability of analysts to understand the content of messages and assess their quality from additional clues. The use cases described below highlight levels of indirection occurring when collecting information and their with impact on trust estimation.

Fig. 1. Assertions and opinions in human messages.

A. Assertions, opinions and reported information

For illustration, let's consider X, the analyst receiving information provided by a human source Y.

Case 1: direct reporting X is an analyst collecting evidence in order to decide whether or not an individual is involved in terrorist activities. In particular, he takes into account reports submitted by Y, a human source. Those reports usually consist on a mixed set of assertions (e.g., descriptions of events or states observed by Y) and opinions (i.e., judgments, assessments, or beliefs) expressed by Y about assertion which give the analyst an insight into how strongly the source commits to the assertion, see Fig. 1.

In the statement contained in Fig. 1, the source Y lets us know that she does not commit her full belief to the assertion that *John is a terrorist*, otherwise the reporter would have used phrasing such as *I am completely convinced* or *it is without doubt* or simply reported *John is a terrorist* as an unadorned statement.

The information item is the sentence, which contains the assertion *John is a terrorist* and the uncertainty degree to be assigned because the analyst knows that Y is not completely certain about her own statements. The analyst must make a judgment about the veracity of John being a terrorist based upon factors such as previous experience with Y's assessments in the past, or, perhaps, on the fact that other sources are relating the same information.

Case 2: indirect reporting Again, let X be an analyst collecting evidence in order to decide whether or not an individual is involved in terrorist activities. In this case, he takes into account reports submitted by Y, a human source who is herself relating information obtained from a secondary source named Mary, see Fig. 2.

The source Y does not report on her direct observations or her deductions or beliefs, but conveys information received from a second source, in this case Mary, in the statement in Fig. 2.

[1] http://eturwg.c4i.gmu.edu/

Fig. 2. Hearsay, assertions and opinions in human messages.

In this report the information item is again the sentence containing the assertive part *John is a terrorist* but this use case introduces more levels of complexity in uncertainty to deal with. The information that the assertion comes from Mary, who has added her own opinion, is a distancing mechanism on the part of the source *Y* as (unlike in Fig. 1), she is neither claiming the opinion nor the assertion.

This case introduces yet more layers of uncertainty. How sure can we be that the reporter *Y* has accurately repeated what Mary said? For example, did Mary really say *it is likely* or did the reporter insert this (intentionally or unintentionally) based upon the reporter's assessment of the reliability of Mary as a source of information? Or perhaps, subtly, *Y* is expressing her own uncertainty by putting words in Mary's mouth. Furthermore, it is possible Mary made this statement under circumstances which would strengthen or weaken this statement, but those conditions have not been passed on by the reporter.

The goal of the analyst is to take this assertion into account, but also to encode his own belief about the quality of the source further in the analysis. All these different attitudes have to be evaluated by the analyst, who may have additional background information or prior evaluation of the source that have to be considered.

In both cases discussed above, the outcome of the analyst is the assertive part of the information item, augmented with a coefficient that helps to measure and track the different levels of trust for their future exploitation. For the purpose of this work, this quality is called *trust in reported information*.

B. Concepts and notions for trust assessment

This section introduces several notions that are relevant for trust analysis.

Trustworthiness of information sources is considered, for the purpose of this work, as confidence in the ability and intention of an information source to deliver correct information, see [3]. Trustworthiness is an attribute of information sources who have the competences to report information, and who can be relied upon to share sincerely and clearly their beliefs on the uncertainty level of reported information. An item provided by such a source is then trusted by analysts.

Self-confidence [8] captures the explicit uncertainty assigned to reported assertions by the source. Statements may include the source's judgments when lacking complete certainty; these judgments are generally identified through the use of various lexical clues such as *possibly, probably, might be, it is unlikely, undoubtedly*, etc., all of which signal the source's confidence (or lack thereof) in the veracity of the information being conveyed. It should be noted that self-confidence, in our usage understood as the linguistic dimension of the certainty degree that the source assigns to reported items, is an aspect exhibited by the source, but it will be considered from the analyst's standpoint during trust analysis.

Reliability of sources indicates how strongly the analyst is willing to accept items from a given source at their face-value. As an overall characterization, reliability is used in this work to rate how much a source can be trusted with respect to their reputation, competence and supposed intentions.

Reputation of sources [11] captures a commonly accepted opinion about how the source performs when reporting information, and is generally understood as the degree to which prior historical reports have been consistent with fact. For human sources, reputation is considered by the analyst for each source based on previous interactions with the source and on the source's history of success and failure in delivering accurate information. Reputation relies, to a large extent, upon negative and positive experiences provided to the analyst by the source in the past.

Competence of sources [29] is related to a source's possession of the skills and knowledge in reporting on various topics: This aspect defines to what extent a human source can understand the events they report on, whether the source has the ability to accurately describe those events, and how capable the source is of following the logic of processes producing the information.

Intentions correspond to specific attitudes toward the effect of one's actions or conduct. Reporting information can become *more a means to manipulate others than a means to inform them* [14] and thus can be carried out with the express purpose of inducing changes in another person's beliefs and understanding. Intentions are specific to human sources as only humans have the capacity to deliberately provide false or misleading information. Sensors may provide erroneous data due to a number of factors such as device failure or environmental conditions, but never due to intention.

In addition to the above facets, *credibility of information* and *reliability of sources* are two notions introduced by the STANAG 2511 [49], which standardizes the terminology used in analysis of intelligence reports used by NATO Forces with distinct focus on sources and information provided. STANAG reliability is understood with respect to the quality of information that has been delivered by sources in the past. STANAG credibility relies on the intuition that a joint analysis of items in combination with each other will likely reveal inconsistencies, contradictions or redundancies. Reliability and credibility are independent criteria for evaluation. Definitions for both reliability and credibility are in natural language.

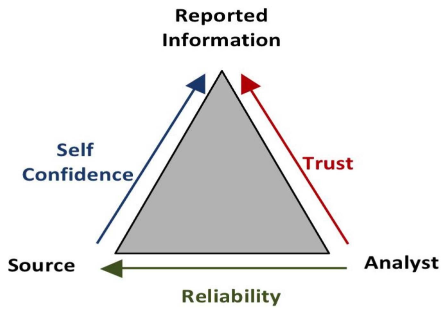


Fig. 3. Model for trust analysis.

Attributes of sources and information items adopted for the model of trust are related to the notions introduced by the STANAG 2511 but are addressed differently: reliability of sources is understood here in terms of source competence, reputation and intentions, while credibility is restricted to features of self-confidence as described above.

IV. A FUNCTIONAL MODEL OF TRUST

This section introduces the model developed to estimate trust in reported information by taking into account the reliability of the source and the source's own characterization of reported items. The advantage of this distinction is to better dissociate the impact of both beliefs of sources and opinions of analysts on the source on the information provided.

Even if the primary function of a source is to provide information, we keep the distinction between the source and the information by considering separate dimensions for each element. The rationale behind this is the observation that even reliable sources can sometimes provide inaccurate or imprecise information from one report to another, which is even more plausible in the case of human sources.

The model, illustrated in Fig. 3., is composed of a source which provides an information item augmented with a degree of uncertainty captured by self-confidence to an analyst. Based upon his direct assessment of the reliability of the source, the analyst constructs his own estimation of trust in the item reported.

In the following section, the model is discussed using a granularity that is detailed enough to describe its elements, but still rough enough to avoid the adoption of a representation formalism.

A. Elements of the trust model

The model is composed of two elements: an information source and reported items from that source. The analyst is considered to be outside the model, although she has multiple interactions with its elements.

Definition of information source: an information source is an agent who provides an information item along with a characterization of its level of uncertainty. "Source" is a relative notion, depending on the perspective of analysis. In general, information is propagated within a chain relating real world information to some decision maker, and agents along the path can be both trained observers, whose job is to provide such reports, as well as witnesses or lay observers who may add items, in spite of not being primarily considered as information sources, but rather as opportunistic ones.

The notion of source is central in many information fusion applications and numerous research efforts aimed at modeling the properties of those applications. A general analysis of sources is undertaken by [32], who identify three main classes: S-Space, composed of physical sensors, H-Space for human observers and I-Space for open and archived data on the Internet. In [39], a unified characterization of hard and soft sources is described, along with a detailed description of their qualities and processing capabilities.

Processing hard sensor information is widely covered [42] in the research community, and can be considered quite mature, while the integration of human sources brings many new challenges. Our model addresses human sources, and reported items can refer to actions, events, persons or locations of interest.

Information reported by humans is unstructured, vague, ambiguous and subjective, and thus is often contrasted with information coming from physical sensors, described as structured, quantitative and objective. While humans can deliberately change the information or even lie, sensors are also prone to errors and therefore hard information items are not always accurate.

For human agents, the source is part of the real world, (a community, a scene, an event) and can be either directly involved in the events reported, or just serving as a witness.

Definition of reported information: Reported information is a couple $(I, \chi(I))$, where I is an item of information and $\chi(I)$ the confidence level as assigned by the source. Items are information pieces that can be extracted from natural language sentences, although the extraction and separation from subjective content are out of the scope for the model developed. Each item I has assertive i_a and subjective i_s components conveying factual and subjective contents respectively.

The analysis of reported information continues to be an open topic as the fusion of information from soft sources receives increasing attention in recent years. Although some authors have developed logic-based approaches for modelling distortions of items exchanged between agents who have both the intention and the ability to deceive [12], there are still more challenges arising when the information is analyzed in its textual form.

Features of uncertainty, as expressed in natural language statements, are analyzed in [2] while [23] provides a broader discussion of pitfalls and challenges related to soft data integration for information fusion.

B. Functions of the trust model

The model introduces several functions estimating features of reliability, self-confidence and trust, as described hereafter.

Definition of a reliability function: a reliability function is a mapping which assigns a real value to an information source.

This real value is a quantitative characterization of the source, inferred with respect to the source's previous failures, its reputation and the relevance of its skills for specific domains. For this model, the reliability of human sources combines three features: competence, reputation and intention. Competence captures the intuition that the quality of information reported by a source depends on the level of training and expertise, which may be designated as satisfactory or not, depending upon the task. Reputation is the overall quality of a source, estimated by examination of the history of its previous failures. Intentions refer to attitudes or purposes, often defined with respect to a hidden purpose or plan to achieve.

Reliability is a complex concept and, from a practical standpoint, it is difficult to have complete information about the global reliability of a source. Thus, this model describes reliability along the three attributes (competence of a source, its reputation and its intentions) described above. In practical applications, this solution allows for compensation for insufficient information on one or several aspects of reliability and to conduct, if necessary, the analysis of reliability based on just one attribute.

Evaluation of reliability Assessing reliability is of real interest when opportunistic sources are considered because the analyst has neither an indication of how the source might behave nor the ability to monitor or control either the human providing the information or the environment in which the source operates. Various methods can be developed to estimate competence, reputation and intentions of the source. For example, competence is closely related to the level of training of an observer or can be defined by domain knowledge. Values can be expressed either in a linguistic form (*bad, good, fair, unknown*) or by a number. Reputation is an attribute which can be constructed not just by examining previous failures of the source but also by considering its level of conflict with other sources; this too can be expressed by numeric or symbolic values.

While reputation and competence can be, at least in some cases, estimated from prior knowledge, characterizing the intentions of a source is subject to human perception and analysis. Judgment of human experts is needed not just because there usually is no *a priori* characterization of the source with respect to its intentions but also because it is important to assess those aspects from the subjective point of view of an expert in the form of binary values only.

From a practical standpoint, it is suitable to provide an expert with a description of source competence, reputation and intentions as assessed independently. This way, experts can have the opportunity to develop different strategies of using reliability: they can decide to assign different importance to those attributes under different contexts or can use their own hierarchy of attributes. For instance, an expert may consider as irrelevant the information provided by a source whose competences is lower than a specific threshold or if he suspects the source of having malicious intentions.

Definition of a self-confidence function: a self-confidence function is a mapping linking a real value and an information item. The real value is a measure of the information credibility as evaluated by the sensor itself and is of particular interest for human sources, as often such sources provide their own assessments of the information conveyed. Identifying features of self-confidence requires methods related to a research task of natural language processing: the identification of assertions and opinions in texts. In this field, the commonly adopted separation of those notions considers assertions as statements that can be proven true or false, while opinions are hypotheses, assumptions and theories based on someone's thoughts and feelings and cannot be proven.

Evaluation of self-confidence: Estimation of self-confidence aims at assigning a numerical value which captures how strongly the author stands behind assertions in the statement, on the basis of lexical clues he has included in the utterance. More generally, markers of an author's commitment are in the form of hedges, modal verbs and forms of passive/active language. A hedge is a mitigating word that modifies the commitment to the truth of propositions, i.e., certainly, possibly. Its impact can be magnified by a booster (highly likely) or weakened by a downtoner (rather certain).

Modal verbs indicate if something is plausible, possible, or certain (*John could be a terrorist, you might be wrong*). Moreover, in some domains sentences making use of the passive voice are considered as an indicator of uncertainty, in the sense that author seeks to distance himself from the assertions in the items reported through use of passive voice. Quantifying self-confidence is a topic of particular interest for intelligence analysis, and it was early addressed by Kent in 1962, [40] who created a standardized list of words of estimative probability which were widely used by intelligence analysts. This list has continued to be a common basis to be used by analysts to produce uncertainty assessments.

Kesselman describes in [41] a study conducted to analyze the way the list was used by analysts over the past, and identifies new trends to convey estimations and proposes a new list having the verb as a central element. Given the variety of linguistic markers for uncertainty, the estimation of a numerical value based on every possible combination seems unrealistic, as the same sentence oftencontains not just one but multiple expressions of uncertainty. Additionally, assigning numerical values to lexical expressions is not an intuitive task, and Rein shows that there are no universal values to be associated in a unique manner to hedges or other uncertainty markers, see [53]. As the author argues further, it is, however, possible to order those expressions and use this relative ordering as a more robust way to compare combinations of uncertainty expressions, and thus highlight different levels of uncertainty in natural language statements.

Using the model for trust analysis: The model proposed in this work proposed in this work combines various attributes of the source (discussed previously under "reliability") with "self-confidence" in order to capture trust of information as conveyed by the human. The model is source-centric predominantly focused on the source's ability to correct, alter or qualify the information report Although the rules for ranking, prioritizing and combining the attributes introduced by the model can be drafted empirically, the estimation of a trust value requires a formal representation of the model.

A possible solution for estimating a unified value for trust is to consider reliability and self-confidence within the framework of an uncertainty theory and to rely on the set of combination rules the theory defines—for example, those developed in probability theory, in possibility theory, or in belief functions theory. All these theories provide various operators to combine reliability and self-confidence in order to estimate trust.

In the following the model is represented by using belief functions and several scenarios are used to illustrate trust estimation.

V. TRUST FORMALIZATION WITH BELIEF FUNCTIONS

The aim of trust formalization is to provide a formal representation of the model, combining the capability to exploit the structure and relationship of elements of the model with the ability to express degrees of uncertainty about those elements. Of particular interest to this paper is the observation that the developed model introduces a cognitive view of trust as a complex structure of beliefs that are influenced by the individual's opinions about certain features and elements, including their own stances. Such a structure of beliefs determines various degrees of trust, which are based on personal choices made by analyst, on the one hand, and the source, on the other hand. Therefore, the formalization requires a formalism that is more general than probability measures or fuzzy category representation, which are more suitable for applications considering trust in the context of interactions between agents. Moreover, the limitations of using subjective probabilities to formalize trust from this cognitive standpoint are clearly stated in [10]. As a result, the model was represented with belief functions, a formalism that is consistent with the cognitive perspective of trust adopted by the model. This belief-based representation provides the most direct correspondence with elements of the model and their underlying uncertainty, while being able to quantify subjective judgments.

After introducing main concepts of belief functions, this section shows how the formalism is used to represent the trust model.

A. Basic Belief Assignment

Belief Functions (BF) have been introduced by Shafer in his his mathematical theory of evidence [56], also referred to Dempster-Shafer Theory (DST), to model epistemic uncertainty. The frame of discernment (FoD) of the decision problem under consideration, denoted Θ, is a finite set of exhaustive and mutually exclusive elements. The powerset of Θ denoted 2^Θ is the set of all subsets of Θ, empty set included. A body of evidence is a source of information characterized by a Basic Belief Assignment (BBA), or a mass function,which is the mapping $m(.): 2^\Theta \to [0,1]$ that satisfies $m(\emptyset) = 0$, and the normalization condition $\sum_{A\in 2^\Theta} m(A) = 1$. The belief (a.k.a credibility) Bel(.) and plausibility Pl(.) functions usually interpreted as lower and upper bounds of unknown (subjective) probability measure $P(.)$, are defined from $m(.)$ respectively by

$$\mathrm{Bel}(A) = \sum_{B\subseteq A|B\in 2^\Theta} m(B) \tag{1}$$

$$\mathrm{Pl}(A) = \sum_{B\cap A\neq\emptyset|B\in 2^\Theta} m(B) \tag{2}$$

An element $A \in 2^\Theta$ is called a focal element of the BBA $m(.)$, if and only if $m(A) > 0$. The set of all focal elements of $m(.)$ is called the core of $m(.)$ and is denoted $\mathcal{K}(m)$. This formalism allows for modeling a completely ignorant source by taking $m(\Theta) = 1$. The Belief Interval (BI) of any element A of 2^Θ is defined by

$$BI(A) \triangleq [Bel(A), Pl(A)] \tag{3}$$

The width of belief interval of A, denoted $U(A) = Pl(A) - Bel(A)$ characterizes the degree of imprecision of the unknown probability $P(A)$, often called the uncertainty of A. We define the uncertainty (or imprecision) index by

$$U(m) \triangleq \sum_{A\in\Theta} U(A) \tag{4}$$

to characterize the overall imprecision of the subjective (unknown) probabilities committed to elements of the FoD bounded by the belief intervals computed with the BBA $m(.)$.

Shafer proposed using Dempster's rule of combination for combining multiple independent sources of evidence [56] which is the normalized conjunctive fusion rule. This rule has been strongly disputed in the

BF community after Zadeh's first criticism in 1979, and since the 1990s many rules have been proposed to combine (more or less efficiently) BBAs; the reader is advised to see discussions in [59], in particular the proportional conflict redistribution rule number 6 (PCR6). To combine the BBAs we use the proportional conflict redistribution (PCR) rule number 6 (denoted PCR6) proposed by Martin and Osswald in [59] because it provides better fusion results than Dempster's rule in situations characterized by both high and low conflict as explained in detail in [19], [35].

The PCR6 rule is based on the PCR principle which transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved. The steps in applying the PCR6 rule are:

1) apply the conjunctive rule;
2) calculate the total or partial conflicting masses; and
3) redistribute the (total or partial) conflicting mass proportionally on non-empty sets.

The general PCR6 formula for the combination of $n > 2$ BBAS is very complicated (see [59] Vol. 2, Chap. 2). For convenience's sake, we give here just the PCR6 formula for the combination of only two BBAs. When we consider two BBAs $m_1(.)$ and $m_2(.)$ defined on the same FoD Θ, the PCR6 fusion of these two BBAs is expressed as $m_{PCR6}(\emptyset) = 0$ and for all $X \neq \emptyset$ in 2^Θ

$$m_{PCR6}(X) = \sum_{\substack{X_1,X_2\in 2^\Theta \\ X_1\cap X_2 = X}} m_1(X_1)m_2(X_2) + \sum_{\substack{Y\in 2^\Theta\setminus\{X\} \\ X\cap Y=\emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X)+m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X)+m_1(Y)}\right] \quad (5)$$

where all denominators in (5) are different from zero. If a denominator is zero, that fraction is discarded. A very basic (not optimized) Matlab code implementing the PCR6 rule can be found in [59] and [61], and also in the toolboxes repository on the web.[2]

Instead of working with quantitative (numerical) BBA, it is also possible to work with qualitative BBA expressed by labels using the linear algebra of refined labels proposed in Dezert-Smarandache Theory (DSmT), [59] (Vol. 2 & 3).

B. Trust formalization model

Because beliefs are well defined mathematical concepts in the theory of belief functions, we prefer to use self-confidence terminology to represent the confidence declared by a source Y on its own assertion A. Let's denote by A the assertion given by the source, for instance A =*John is a terrorist*. With respect to elements of the model, A (the assertion) corresponds to i_a, the assertive part of the item I and $v(A)$ is a numeric estimation of the subjective i_s component of I.

The valuation $v(A)$ made by the source Y about the assertion A can be done either quantitatively (by a probability or a BBA) or qualitatively (by a label associated to a linguistic form). This paper considers quantitative representation of $v(A)$ for simplicity.[3]

The basic information items provided by a source Y consists of A (the assertion), and $v(A)$ (its valuation). To be as general as possible, we suppose that $v(A)$ is a basic belief mass assignment defined with respect to the very basic frame of discernment $\Theta_A \triangleq \{A, \bar{A}\}$ where $\bar{A}$ denotes the complement of A in Θ_A, that is $v(A) = (m(A), m(\bar{A}), m(A\cup\bar{A}))$. Note that only two values of the triplet are really necessary to define $v(A)$ because the third one is automatically derived from the normalization condition $m(A) + m(\bar{A}) + m(A\cup\bar{A}) = 1$. So one could also have chosen equivalently $v(A) = [Bel(A), Pl(A)]$ instead of the BBA. In a probabilistic context, one will take $m(A\cup\bar{A}) = 0$ and so $v(A) = P(A)$ because $Bel(A) = Pl(A) = P(A)$ in such a case.

The self-confidence of the source Y is an extra factor $\alpha_Y \in [0,1]$ which characterizes the self-estimation of the quality of the piece of information $(A, v(A))$ provided by the source itself. $\alpha_Y = 1$ means that the source Y is 100% confident in his valuation $v(A)$ about assertion A, and $\alpha_Y = 0$ means that the source Y is not at all confident in his valuation $v(A)$. In the theory of belief functions, this factor is often referred as the discounting factor of the source because this factor is usually used to discount the original piece of information $(A, v(A))$ into a discounted one $(A, v'(A))$ as follows [56]:

$$m'(A) = \alpha_Y \cdot m(A) \quad (6)$$

$$m'(\bar{A}) = \alpha_Y \cdot m(\bar{A}) \quad (7)$$

$$m'(A\cup\bar{A}) = \alpha_Y \cdot m(A\cup\bar{A}) + (1-\alpha_Y) \quad (8)$$

The idea of Shafer's discounting technique is to diminish the belief mass of all focal elements with the factor α_Y and redistribute the missing discounted mass $(1-\alpha_Y)$ to the whole ignorance $A\cup\bar{A}$. Note that the valuation of the discounted piece of information is always degraded because its uncertainty index is always greater than the original one, that is, $U(m') > U(m)$, which is normal.

The reliability factor r estimated by the analyst X on the piece of information $(A, v(A))$ provided by the source Y must take into account both the competence C_Y, the reputation R_Y and the intention I_Y of the source Y. A simple model to establish the reliability factor

[2]http://bfaswiki.iut-lannion.fr/wiki/index.php/Main_Page

[3]Without loss of generality one can always map a qualitative representation to a quantitative one by a proper choice of scaling and normalization (if necessary).

r is to consider that C_Y, R_Y and I_Y factors are represented by numbers $[0,1]$ associated to select subjective probabilities, that is $C_Y = P(Y \text{ is competent})$, $R_Y = P(Y \text{ has a good reputation})$ and $R_Y = P(Y \text{ has a good intention (i.e. is fair)})$. If each of these factors has equal weight, then one could use $r = C_Y \times R_Y \times I_Y$ as a simple product of probabilities. However, in practice, such simple modeling does not fit well with what the analyst really needs to take into account epistemic uncertainties in Competence, Reputation and Intention. In fact, each of these factors can be viewed as a specific criterion influencing the level of the global reliability factor r. This is a multi-criteria valuation problem. Here we propose a method to solve the problem.

We consider the three criteria C_Y, R_Y and I_Y with their associated importance weights w_C, w_R, w_I in $[0,1]$ with $w_C + w_R + w_I = 1$. We consider the frame of discernment $\Theta_r = \{r, \bar{r}\}$ about the reliability of the source Y, where r means that the source Y is reliable, and $\bar{r}$ means that the source Y is definitely not reliable. Each criteria provides a valuation on r expressed by a corresponding BBA. Hence, for the competence criteria C_Y, one has $(m_C(r), m_C(\bar{r}), m_C(r \cup \bar{r}))$, while for the reputation criteria R_Y, one has $(m_R(r), m_R(\bar{r}), m_R(r \cup \bar{r}))$ and for the intention criteria I_Y, one has $(m_I(r), m_I(\bar{r}), m_I(r \cup \bar{r}))$.

To get the final valuation of the reliability r of the source Y, one needs to efficiently fuse the three BBAs $m_C(.)$, $m_R(.)$ and $m_I(.)$, taking into account their importance weights w_C, w_R, and w_I. This fusion problem can be solved by applying the importance discounting approach combined with PCR6 fusion rule of DSmT [60] to get the resultant valuation $v(r) = (m_{PCR6}(r), m_{PCR6}(\bar{r}), m_{PCR6}(r \cup \bar{r}))$ from which the decision (r, or $\bar{r}$) can be drawn (using BI distance, for instance). If a firm decision is not required, an approximate probability $P(r)$ can also be inferred with some lossy transformations of BBA to probability measure [59]. Note that Dempster's rule of combination cannot be used here because it does not respond to the importance discounting, as explained in [60].

The trust model consists of the piece of information $(A, v(A))$ and the self-confidence factor α_Y provided by the source Y, as well as the reliability valuation $v(r)$ expressed by the BBA $(m(r), m(\bar{r}), m(r \cup \bar{r}))$ to infer the trust valuation about the assertion A. For this, we propose using the mass $m(r)$ of reliability hypothesis r of the source Y as a new discounting factor for the BBA $m'(.)$ reported by the source Y, taking into account its self-confidence α_Y. Hence, the trust valuation $v_t(A) = (m_t(A), m_t(\bar{A}), m_t(A \cup \bar{A}))$ of assertion A for the analyst X is defined by

$$m_t(A) = m(r) \cdot m'(A) \tag{9}$$

$$m_t(\bar{A}) = m(r) \cdot m'(\bar{A}) \tag{10}$$

$$m_t(A \cup \bar{A}) = m(r) \cdot m'(A \cup \bar{A}) + (1 - m(r)) \tag{11}$$

or equivalently by

$$m_t(A) = m(r)\alpha_Y \cdot m(A) \tag{12}$$

$$m_t(\bar{A}) = m(r)\alpha_Y \cdot m(\bar{A}) \tag{13}$$

$$m_t(A \cup \bar{A}) = m(r)\alpha_Y \cdot m(A \cup \bar{A}) + (1 - m(r)\alpha_Y) \tag{14}$$

The DSmT framework using the PCR6 fusion rule and the importance discounting technique provides an interesting solution for the fusion of attributes having different degrees of importance while making a clear distinction between those attributes.

The discounting method proposed in this work is directly inspired by Shafer's classical discounting approach [56]. In our application, the classical discounting factor that we propose integrates both the mass of reliability hypothesis $m(r)$ and the self-confidence factor α_Y. It is worth noting that more sophisticated (contextual) belief discounting techniques [47] exist and they could also have been used, in theory, to refine the discounting but these techniques are much more complicated and they require additional computations. The evaluation of contextual belief discounting techniques for such types of application is left for further investigations and research works.

VI. UNCERTAINTY ANALYSIS UNDER URREF CRITERIA

Tracking uncertainties from problem description to model construction and formalization is done under criteria of the uncertainty representation and reasoning evaluation framework.

The goal of URREF is to place the focus on the evaluation of uncertainty representation and reasoning procedures. The URREF ontology defines four main classes of evaluation criteria: Data Handling, Representation, Reasoning and Data Quality. These criteria make distinctions between the evaluation of the fusion system, the evaluation of its inputs and outputs, and the evaluation of the uncertainty representation and reasoning aspects.

Listing all criteria is an extensive task and in this paper the authors will provide one piece of the puzzle by considering criteria that relate to the evaluation of uncertainty induced by the proposed model. In the model developed in this paper, uncertainty is due to imperfections of information gathering and reporting as well as constraints of the representation formalism.

Uncertainty analysis is carried out by assigning uncertainty criteria to elements and functions of the trust model in order to make explicit the uncertainty arising when the problem is abstracted by the model and the model is then simplified in order to fulfill constraints of specific formalism, Fig. 6.

The URREF criteria selected are subclasses of two main concepts: *Credibility*, a subconcept under *DataCriteria*, and *EvidenceHandling*, a subconcept of *RepresentationCriteria*.

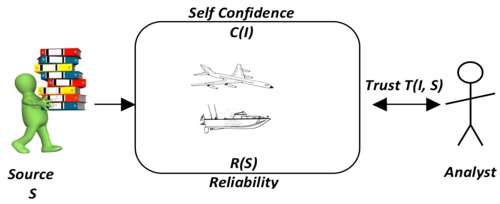


Fig. 4. Trust estimation from source to analyst

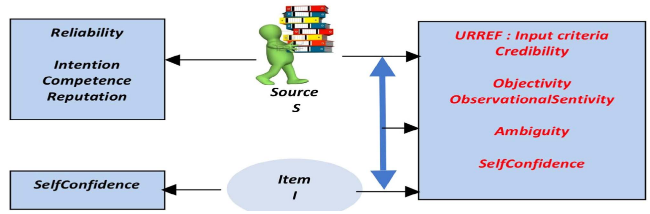


Fig. 5. Mapping of model attributes to URREF criteria

To summarize, uncertainties of the model will be captured by the following URREF criteria:

- **Objectivity**, subconcept of **Credibility**: indicates a source providing unbiased information;
- **ObservationalSensitivity**, subconcept of **Credibility**: characterizes the skills and competences of sources;
- **SelfConfidence**, subconcept of **Credibility**: measures the certainty degree about the piece of information reported, according to the source;
- **Ambiguity**, subconcept of **EvidenceHandling**: captures if the sources provide data supporting different conclusions;
- **Dissonance**, subconcept of **EvidenceHandling**: captures the ability of formalism to represent inconsistent evidence;
- **Completeness**, subconcept of **EvidenceHandling**: is a measure of how much is known given the amount of evidence; and
- **Conclusiveness**, subconcept of **EvidenceHandling**: indicates how strong the evidence supports a conclusion;

Besides selecting uncertainty criteria relevant for trust estimation, the analysis also discusses the mapping of URREF criteria to attributes of the model and sheds a light on imperfect matchings. This mapping offers a basis for identifying the limitations of the URREF ontology, by emphasizing those elements whose characterizations in terms of uncertainty are out of the ontology's reach or beyond the ontology's intended scope.

A. Uncertainties from problem definition to model abstraction

Let M be the model for trust estimation, with elements introduced in paragraph IV: the source Y, the reported item I with its assertive i_a and subjective i_s parts, and $\chi(I)$ the confidence level assigned by the source Y to I.

From an information fusion standpoint, inputs of the model are the source and the information items, along with their uncertainty, captured with the following URREF criteria: *Objectivity, ObservationalSensitivity and SelfConfidence*. These criteria are subclasses of the concept *InputCriteria*.

Objectivity is an attribute of the source, related to its ability to provide factual, unbiased items, without adding their own points of view or opinions. For a source Y providing information item i, having i_s and i_a as the subjective and factual parts respectively, objectivity can be expressed as:

$$Objectivity(Y,I) = \psi_o(i_s,i_a) \tag{15}$$

where $\psi_o(i_s,i_a)$ represents the mathematically quantified expression of the subjective over the factual content of i.

ObservationalSensitivity is an attribute of the source which represents the source's ability to provide accurate reports. In the proposed model, this criterion is an aggregation of competence C and reputation R, two attributes of the model.

$$ObservationalSensitivity(Y,i) = \psi_{os}(C,R) \tag{16}$$

where $\psi_{os}(C,R)$ is a function aggregating values of competence and reputation.

Information items entering the system are described by *SelfConfidence*. Again, considering i_s and i_a as the subjective and factual items conveyed by I, *SelfConfidence* can be expressed as:

$$SelfConfidence(I) = \psi_{sc}(i_s) \tag{17}$$

with $\psi_{sc}(i_s)$ a function quantifying the subjective content of item I.

Fig. 5 shows the mapping between the elements of the model and the set of relevant URREF uncertainty criteria. The mapping shows a perfect match between *SelfConfidence* as introduced by the model and the eponymous URREF criterion as well as several imperfect matches described later in this paper.

At source level, URREF criteria are not able to capture in a distinct manner the features of competence, reputation and intentions, the main attributes of the sources added by the model under Reliability. To some extent, competence and reputation can be related to *ObservationalSensitivity*, but intentions clearly remains out of reach for URREF criteria.

B. Uncertainties from model to formal representation

Let F be the DST formalization of the trust estimation model, with parameters introduced in paragraph V. The formalism induces two types of uncertainty related to its capacity to handle incomplete, ambiguous or contradictory evidence. The uncertainty of evidence handling is captured by *Ambiguity*, *Dissonance*, *Conclusiveness* and *Completeness*. Those criteria are subclasses of the concept *EvidenceHandling*.

Ambiguity measures the extent to which the formalism can handle data sets which support different conclusions.

$$Ambiguity(F) = \phi_a(\alpha_Y, R_Y) \tag{18}$$

where the function $\phi_a(\alpha_Y, R_Y)$ considers the self-confidence factor α_Y provided by the source Y and the reliability of Y provided by the analyst R_Y to estimate the degree of ambiguity. The measure is of particular interest in the case where items having high values of self-confidence are provided by unreliable sources.

Dissonance captures the ability of the formalism to represent inconsistent evidence. For BBA representations, dissonance can be related to the capacity of the formalism to assign belief mass to an element and its negation, and can therefore be assessed for every BBA representation build for the model. For example, the dissonance for a source's competence can be in the form:

$$Dissonance(F) = \phi_d(m_C(r), m_C(\bar{r})) \tag{19}$$

where $\phi_d(m_C(r), m_C(\bar{r}))$ is a function combining the belief mass assigned to whether the source is considered to be competent or incompetent, respectively.

Dissonance is useful for highlighting situations in which there are significant differences in belief masses assigned at the attribute level, such as when a source is considered to be incompetent (low $m_C(r)$, high $m_C(\bar{r})$) but has a good reputation (high $m_R(r)$, low $m_R(\bar{r})$).

Conclusiveness is a measure expressing how strongly the evidence supports a specific conclusion or unique hypothesis:

$$Conc.(F) = \phi_{cc}(m_t(A), m_t(\bar{A}), m_t(A \cup \bar{A})) \tag{20}$$

where $\phi_{cc}(m_t(A), m_t(\bar{A}), m_t(A \cup \bar{A}))$ is a function combining the belief masses estimated for truthful, untruthful and unknown qualifications of assertion A respectively. This measure indicates to which extent the result of inferences can support a conclusion, in this case whether the hypothesis that the assertion under analysis is trustworthy or not. It can be used during the inference process to show how taking into account additional elements such as the competence of the source, its reputation or intentions impact the partial estimations of trust.

Completeness is a measures of the range of the available evidence, and captures the ability of formalism to take into account how much is unknown. The measures is somewhat similar to *Dissonance*, as is can be assessed for every BBA representation build for the model. Thus, completeness of source's reliability is described as:

$$Completeness(F) = \phi_{cp}(m_{(}r \cup \bar{r})) \tag{21}$$

where $\phi_{cp}(m_{(}r \cup \bar{r}))$ is a function depending on the belief mass assigned to unknown.

The measure is used for estimation and analysis before entering the fusion process, in order to have a picture of how complete the evidence describing the various elements of the model is, and to avoid performing

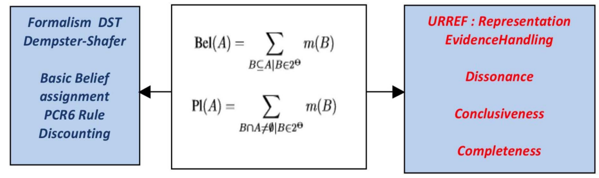


Fig. 6. Mapping of formalism uncertainties to URREF criteria

fusion on highly incomplete data sets. Both *EvidenceHandling* and *KnowledgeHandling* are subclasses of *RepresentationCriteria*.

This section has analyzed the nature of uncertainties arising when going from problem to model definition and then on to formalization with belief functions. The next section shows how uncertainties can be highlighted for particular scenarios of trust estimation.

VII. UNCERTAINTY ANALYSIS FOR TRUST ESTIMATION

A. Running example and method for uncertainty tracking

As a running example, let's consider an assertion A and its valuation $v(A)$ provided by the source Y as follows: $m(A) = 0.7$, $m(\bar{A}) = 0.1$ and $m(A \cup \bar{A}) = 0.2$. Its self-confidence factor is $\alpha_Y = 0.75$. Hence, the discounted BBA $m'(.)$ is given by

$$m'(A) = 0.75 \cdot 0.7 = 0.525$$

$$m'(\bar{A}) = 0.75 \cdot 0.1 = 0.075$$

$$m'(A \cup \bar{A}) = 1 - m'(A) - m'(\bar{A}) = 0.4$$

Let's assume that the BBAs about the reliability of the source based on Competence, Reputation and Intention criteria are given as follows:

$$m_C(r) = 0.8, m_C(\bar{r}) = 0.1, m_C(r \cup \bar{r}) = 0.1$$

$$m_R(r) = 0.7, m_R(\bar{r}) = 0.1, m_R(r \cup \bar{r}) = 0.2$$

$$m_I(r) = 0.6, m_I(\bar{r}) = 0.3, m_I(r \cup \bar{r}) = 0.1$$

with importance weights $w_I = 0.6$, $w_R = 0.2$ and $w_C = 0.2$.

After applying the importance discounting technique presented in [60] which consists of discounting the BBAs with the importance factor and redistributing the missing mass onto the empty set, then combining the discounted BBAs with PCR6 fusion rule, we finally get, after normalization, the following BBA

$$m(r) = 0.9335$$

$$m(\bar{r}) = 0.0415$$

$$m(r \cup \bar{r}) = 1 - m(r) - m(\bar{r}) = 0.025$$

The final trust valuation of assertion A reported by the source Y taking into account its self-confidence

$\alpha_Y = 0.75$ and the reliability factor $m(r) =$ is therefore given by Eqs. (12)–(14) and obtaining

$$m_t(A) = 0.4901$$

$$m_t(\bar{A}) = 0.0700$$

$$m_t(A \cup \bar{A}) = 1 - m_t(A) - m_t(\bar{A}) = 0.4399$$

Note that if $m_C(r) = m_R(r) = m_I(r) = 1$, then we will always get $m(r) = 1$ regardless of the choice of weightings factors, which is normal. If there is a total conflict between valuations of reliability based on Competence, Reputation and Intention criteria, then Dempster's rule cannot be applied to get the global reliability factor $m(r)$ because of 0/0 indeterminacy in the formula of Dempster's rule. For instance, if one has $m_C(r) = m_R(r) = 1$ and $m_I(\bar{r}) = 1$, then $m(r)$ is indeterminate with Dempster's rule of combination, whereas it corresponds to the average value $m(r) = 2/3$ using PCR6 fusion rule (assuming equal importance weights $w_C = w_R = w_I = 1/3$), which makes more sense.

The following subsections explore several scenarios for trust assessment, corresponding to different situations of BBAs distributions, and track the uncertainty according to URREF criteria. Each scenario illustrates specific instances of the model developed for trust estimation.

The method adopted to track uncertainty defines the following measures to estimate URREF criteria:

$$SelfConfidence = \alpha_Y$$

$$Ambiguity = |\alpha_Y - m(r)|$$

$$Objectivity = m_I(r)$$

$$ObservationalSensitivity = min(m_C(r), m_R(r))$$

As shown in previous formulas, URREF criteria are estimated based on features of the BBA formalization and are assigned to the static elements of the model, i.e., the source and the information item. While *Objectivity* and *ObservationalSensitivity* captures imperfections of observations, *SelfConfidence* and *Ambiguity* reflect inaccuracies in reporting information to analysts. These criteria are assessed before entering the fusion phase, and describe the initial uncertainty present in the system before inferences.

In addition, *Dissonance*, *Conclusiveness* and *Completeness* will be estimated at the scenario level by adopting the following formulas:

$$Dissonance = 1 - |m_t(A) - m_t(\bar{A})|$$

$$Conclusiveness = |m_t(A) - m_t(\bar{A})|$$

$$Completeness = 1 - m(A \cup \bar{A})$$

Criteria above will be assessed for elements impacted by the fusion process: the reliability of the source, the updated BBAs of the initial assertion and estimated trust. In the following subsection we illustrate several scenarios for trust estimation and the uncertainty analysis underlying each scenario.

TABLE I.
Consensus: input uncertainty

Uncertainity of inputs		
Observation	Objectivity ObservationalSensitivity	1 1
Reporting	SelfConfidence Ambiguity	1 0

TABLE II.
Consensus: fusion uncertainty

Fusion uncertainty	Dissonance	Conclusiv.	Complet.
Updated BBAs	0	1	1
Reliability	0	1	1
Trust	0	1	1

B. Scenarios for trust assessment and uncertainty analysis

Scenarios introduced below provide examples of trust construction using various operators and highlight the uncertainty assigned to elements of the model and its propagation during the fusion process.

Scenario 1—Consensus: Suppose that Y provides the assertion A, while stating that A certainly holds and that X considers Y to be a reliable source.

In this case, the trust will be constructed on the basis of two consensual opinions: the analyst X that considers Y as a reliable source, and the source's conviction that the information provided is certain. In this case, $m(A) = 1$, $\alpha_Y = 1$ and $m(r) = 1$, so that $m'(A) = 1$ and $m_t(A) = m(r) \cdot m'(A) = 1$. The result will be in the form $(A, v(A))$ initially provided by the source.

This scenario illustrates an ideal situation for trust assessment, where the source is trustworthy and well known to the analyst, and observations are reported in perfect conditions. As shown in table I, there is no uncertainty induced by the source, and once fusion is performed the items impacted show high values for conclusiveness and completeness, while dissonance is 0 for the updates BBAs for values, source's reliability and estimated trust, as shown in table II.

Scenario 2—Uncertain utterances: Y is considered by X to be a reliable source and reports the assertion A, while showing a low level of certainty $v(A)$ about the veracity of A. This example is relevant for situations where a reliable source provides (possibly) inaccurate descriptions of events due to, say, bad conditions for observation. This scenario corresponds by example to

TABLE III.
Uncertain uttering: Input uncertainty

Uncertainty of inputs		
Observation	Objectivity ObservationalSensitivity	0.3 0.9
Reporting	SelfConfidence Ambiguity	0.6 0.38

TABLE IV.
Uncertain utterance: fusion uncertainty

Fusion uncertainty	Dissonance	Conclusiv.	Complet.
Updates BBAs	0.3	0.7	0.9
Reliability	0.02	0.98	0.98
Trust	0.59	0.41	0.54

the following case for inputs: $\alpha_Y = 0.6$

$$m(A) = 0.8, m(\bar{A}) = 0.1, m(A \cup \bar{A}) = 0.1$$

$$m_C(r) = 0.9, m_C(\bar{r}) = 0, m_C(r \cup \bar{r}) = 0.1$$

$$m_R(r) = 0.9, m_R(\bar{r}) = 0, m_R(r \cup \bar{r}) = 0.1$$

$$m_I(r) = 0.3, m_I(\bar{r}) = 0.3, m_I(r \cup \bar{r}) = 0.6$$

and $w_C = 0.5$, $w_R = 0.5$ and $w_I = 0$. This results in

$$m'(A) = 0.48, m'(\bar{A}) = 0.06, m'(A \cup \bar{A}) = 0.46$$

and

$$m(r) = 0.9846, m(\bar{r}) = 0, m(r \cup \bar{r}) = 0.0154$$

Therefore, one finally obtains the trust valuation

$$m_t(A) = 0.47, m_t(\bar{A}) = 0.05, m_t(A \cup \bar{A}) = 0.46$$

This case shows that self-confidence has an important impact on the values of discounted BBA, as $m'(A)$ is decreased from 0.8 to 0.48, and thus the remaining mass is redistributed on $m'(A \cup \bar{A})$.

The combination of competence, reliability and intention are in line with the assumption of the scenario, which states that Y is a reliable source. After normalization, values for trust assessment clearly highlight the impact of uncertain utterances, as the BBA shows a mass transfer from $m_t(A)$ to $m_t(A \cup \bar{A})$. Still, values of trust are close to BBA integrating the self-confidence, which confirms the intuition that when the analyst X considers Y to be a reliable source, the assertion A is accepted with an overall trust level almost equal to the certainty level stated by the source.

This scenario illustrates uncertainty induced by observations failures, as *Objectivity*, and *SelfConfidence* are low, see table III.

While the quality of the source is highlighted by high values of *Conclusiveness* and *Completeness*, showing the analyst's confidence in the reports analyzed, the impact of imperfect observation is shown in the overall estimation of trust, through a combination of *Dissonance*, *Conclusiveness* and *Completeness* which have values close to 0.5, see table IV.

Scenario 3—Reputation: Suppose that Y provides A and $v(A)$ and X has no global description of Y in terms of reliability. As the reliability of Y is not available, Y's reputation will be used instead, as derived from historical data and previous failures. This scenario corresponds by example to the following case for inputs: $\alpha_Y = 1$

$$m(A) = 0.8, m(\bar{A}) = 0.1, m(A \cup \bar{A}) = 0.1$$

$$m_C(r) = 0.1, m_C(\bar{r}) = 0.1, m_C(r \cup \bar{r}) = 0.8$$

$$m_R(r) = 0.9, m_R(\bar{r}) = 0.1, m_R(r \cup \bar{r}) = 0$$

$$m_I(r) = 0.1, m_I(\bar{r}) = 0.1, m_I(r \cup \bar{r}) = 0.8$$

and $w_C = 0.1$, $w_R = 0.8$ and $w_I = 0.1$. Hence, one gets

$$m'(A) = 0.8, m'(\bar{A}) = 0.1, m'(A \cup \bar{A}) = 0.1$$

and

$$m(r) = 0.94, m(\bar{r}) = 0.01, m(r \cup \bar{r}) = 0.03$$

Therefore, one finally obtains the trust valuation

$$m_t(A) = 0.75, m_t(\bar{A}) = 0.09, m_t(A \cup \bar{A}) = 0.14$$

For this scenario, the source is confident about their own assertions, and therefore

$$m(A) = 0.8, m(\bar{A}) = 0.1, m(A \cup \bar{A}) = 0.1$$

and

$$m'(A) = 0.8, m'(\bar{A}) = 0.1, m'(A \cup \bar{A}) = 0.1$$

have identical BBA distributions. The reliability of the source is built namely on its reputation, as there are clues about the competence and intentions of the source. Hence, the overall BBA

$$m(r) = 0.9449, m(\bar{r}) = 0.0196, m(r \cup \bar{r}) = 0.0355$$

is close to the initial reputation distribution

$$m_R(r) = 0.9, m_R(\bar{r}) = 0.1, m_R(r \cup \bar{r}) = 0$$

Values of trust show the impact of using not completely reliable sources, which decreased the certainty level of the initial BBA

$$m'(A) = 0.8, m'(\bar{A}) = 0.1, m'(A \cup \bar{A}) = 0.1$$

to

$$m_t(A) = 0.75, m_t(\bar{A}) = 0.09, m_t(A \cup \bar{A}) = 0.14$$

TABLE V.
Reputation: input uncertainty

Uncertainty of inputs		
Observation	Objectivity ObservationalSensitivity	0.10 0.10
Reporting	SelfConfidence Ambiguity	1 0.60

TABLE VI.
Reputation: fusion uncertainty

Fusion uncertainty	Dissonance	Conclusiv.	Complet.
Updated BBAs	0.30	0.70	0.90
Reliability	0.07	0.93	0.95
Trust	0.34	0.66	0.84

They also support the intution that the trust assigned by the analyst to A will have an upper limit equal to the reputation of the source.

This scenario is similar the previous one as, in both cases, there are incomplete descriptions of the source. For this particular case, a historical recording of source's failures offers a basis to overcome the missing pieces and, in spite of low values for *Objectivity* and *ObservationalSensitivity* (see table V), the final trust evaluation is improved with respect to the previous scenario and shows a better combination of *Dissonance*, *Conclusiveness* and *Completeness*, as shown in table VI.

Scenario 4—Misleading report: In this case, Y provides the assertion A, while stating that it certainly holds and X considers Y to be a completely unreliable source. For this case, the analyst knows that the report is somehow inaccurate, for example, it cannot be corroborated or it contradicts, at least in part. information from other (more reliable) sources. The analyst suspects the source of having misleading intentions, and can therefore assign a maximal uncertainty level to the information reported. This scenario corresponds by example to the following case for inputs: $\alpha_Y = 1$

$$m(A) = 1, m(\bar{A}) = 0, m(A \cup \bar{A}) = 0$$

$$m_C(r) = 0.1, m_C(\bar{r}) = 0.1, m_C(r \cup \bar{r}) = 0.8$$

$$m_R(r) = 0.1, m_R(\bar{r}) = 0.1, m_R(r \cup \bar{r}) = 0.8$$

$$m_I(r) = 0.1, m_I(\bar{r}) = 0.8, m_I(r \cup \bar{r}) = 0.1$$

and $w_C = 0.1$, $w_R = 0.1$ and $w_I = 0.8$. Hence, one gets

$$m'(A) = 1, m'(\bar{A}) = 0, m'(A \cup \bar{A}) = 0$$

and

$$m(r) = 0.02, m(\bar{r}) = 0.91, m(r \cup \bar{r}) = 0.06$$

Therefore, one finally obtains as trust valuation

$$m_t(A) = 0.023, m_t(\bar{A}) = 0, m_t(A \cup \bar{A}) = 0.976$$

TABLE VII.
Misleading report: input uncertainty

Uncertainty of inputs		
Observation	Objectivity ObservationalSensitivity	0.10 0.10
Reporting	SelfConfidence Ambiguity	1.00 0.97

TABLE VIII.
Misleading: fusion uncertainty

Fusion uncertainty	Dissonance	Conclusiv.	Complet.
Assertion	0	1	1
Source	0.11	0.89	0.93
Trust	0.76	0.23	0.03

The values for this scenario reflect the high self-confidence of the source and high accuracy of the assertion provided; therefore, the initial BBA is unchanged after fusion with self-confidence. Nevertheless, the impact of having misleading intention is visible first on the mass distribution assigned to reliability and then on the overall values of trust. With respect to the initial values

$$m(A) = 1, m(\bar{A}) = 0, m(A \cup \bar{A}) = 0$$

and the partially fused ones

$$m'(A) = 1, m'(\bar{A}) = 0, m'(A \cup \bar{A}) = 0$$

the integration of a misleading source transfers the mass assignation almost exclusively to $m_t(A \cup \bar{A})$. Intuitively, the assertion A will be ignored, as the reliability of the source is dramatically decreased by a high mass assignment on misleading intentions.

This scenario illustrates the impact of misleading sources on trust estimation. Hence, the use case has very good values for reporting induced uncertainty, with high *SelfConfidence* and low *Ambiguity* (see table VII)), but the overall trust characterization shows strong *Dissonance*, corroborated with low *Conclusiveness* and near zero *Completeness*, as shown in table VIII.

Scenario 5—Ambiguous report: The source Y provides A and $v(A)$, the uncertainty level. Suppose that $v(A)$ has a low value, as the source is not very sure about the events reported, and that X considers Y to be unreliable. This scenario corresponds by example to the following case for inputs: $\alpha_Y = 0.3$

$$m(A) = 0.6, m(\bar{A}) = 0.2, m(A \cup \bar{A}) = 0.2$$

$$m_C(r) = 0.1, m_C(\bar{r}) = 0.8, m_C(r \cup \bar{r}) = 0.1$$

$$m_R(r) = 0.1, m_R(\bar{r}) = 0.8, m_R(r \cup \bar{r}) = 0.1$$

$$m_I(r) = 0.1, m_I(\bar{r}) = 0.1, m_I(r \cup \bar{r}) = 0.8$$

and $w_C = 0.2$, $w_R = 0.4$ and $w_I = 0.4$. Hence, one gets

$$m'(A) = 0.18, m'(\bar{A}) = 0.06, m'(A \cup \bar{A}) = 0.76$$

TABLE IX.
Ambiguous report: input uncertainty

Uncertainty of inputs		
Observation	Objectivity ObservationalSensitivity	0.10 0.10
Reporting	SelfConfidence Ambiguity	0.30 0.27

TABLE X.
Ambiguous report: fusion uncertainty

Fusion uncertainty	Dissonance	Conclusiv.	Complet.
Assertion	0.6	0.4	0.8
Source	0.583	0.417	0.47
Trust	0.973	0.027	0.006

and

$$m(r) = 0.02, m(\bar{r}) = 0.43, m(r \cup \bar{r}) = 0.53$$

Therefore, one finally obtains the trust valuation

$$m_t(A) = 0.0040, m_t(\bar{A}) = 0.0013$$

and

$$m_t(A \cup \bar{A}) = 0.9946$$

This scenario is an illustration for the worst practical case and is relevant when the analyst receives a report provided by a source that lacks the skills or competence to provide accurate descriptions of events. In this case, the reports are incomplete, ambiguous, or even irrelevant. In addition to low competence and reliability, the source himself is also unsure about the statement.

The first modification of BBA shows the strong impact of self-confidence, which changes drastically the BBA of the initial assertions, from

$$m(A) = 0.6, m(\bar{A}) = 0.2, m(A \cup \bar{A}) = 0.2$$

to

$$m'(A) = 0.18, m'(\bar{A}) = 0.06, m'(A \cup \bar{A}) = 0.76$$

Unsurprisingly, the overall reliability is low:

$$m(r) = 0.0223, m(\bar{r}) = 0.4398, m(r \cup \bar{r}) = 0.5379$$

and the results of the final combination show an important mass assigned to $m_t(A \cup \bar{A}) = 0.9946$. Intuitively, the information provided is useless, and considered as highly uncertain.

This scenario shows the combined effects of uncertain reporting and incomplete source description for trust estimation. First, the outcome is affected by high values of uncertainty induced during observation and reporting passes, table IX. Then, fusion leads to a trust estimation having high values of *Dissonance*, and very low values of *Conclusiveness* and *Completeness*.

The same criteria estimated for reliability show the main difference with respect to the previous case, which was also based on unreliable sources. While in scenario 4 the source still has important *Completeness*, this measure is drastically decreased for this scenario, as shown in table X.

VIII. STRENGTHS AND LIMITATIONS OF BELIEF-BASED FORMALIZATION FOR TRUST ASSESSMENT

This section discusses the strengths and limitations of the belief-based perspective in trust modeling in the light of results shown by previous scenarios. The main advantage of using belief functions is that the formalism is consistent with the cognitive perspective of trust adopted by the model, thanks to the notion of belief. It also captures uncertainties both of the analyst with respect to the source and of the source with respect to their own statements with different mechanisms. First, self-confidence is implemented thanks to a discounting coefficient, as, in practice, the values of self-confidence may rely upon linguistic clues of certainty/uncertainty that can be translated into numerical values. Second, the formalization introduces weighting factors in order to offer a flexible solution, which allow for situations in which the analyst has more or less complete knowledge about distinct attributes of the source, or wishes to emphasize one particular attribute. Moreover, the formalization is able to handle ignorance on various aspects, including missing data. The overall fusion mechanism performs trust estimation in several steps, which allows for a better traceability of the outcome and the mapping at different processing stages using URREF criteria. The results of these scenarios are in line with their specific hypotheses, reflecting the intuition that the fusion technique is appropriate for estimating trust.

As with any user-centric approach, the main limitation of the solution discussed in this paper is the lack of guidance for choosing the set of numerical values with which to instantiate the model. For example, two different analysts may choose differing mass distribution and weight coefficients with respect to the same source, and they may also use slightly different approaches to infer a numerical value from linguistic clues when handling self-confidence. Thus, the outcome depends crucially on the interventions of users and their ability to build a model able to capture the situation under analysis. Also, the solution requires preexisting knowledge about the source's reputation, competence, and intention, indeed, in practice, it is difficult to have access to information on those aspects. Provided that there is no other metadata or domain knowledge available for use, the model is likely to fail to produce an accurate trust evaluation in some contexts due to the shortage of knowledge on critical aspects.

As such, the belief-based formalization has limited capabilities to explain the outcome. To overcome this limitation, a mapping to URREF uncertainty criteria is used. The mapping highlights when uncertainties are added into the system and which partial results and affected. It facilitates the interpretation of results by adding additional information as to why the item is to be trusted or no; for example, whereas the fusion process outputs low values of trust for a given item, the mapping to URREF criteria allows to underline problems related to evidence collection or reporting, dissonance or incompleteness during the fusion stages.

As shown in previous scenarios, using a belief-oriented formalism and URREF criteria mapping offers a pragmatic approach to develop a more comprehensive and easy to interpret solution for trust estimation.

IX. CONCLUSION

This paper presents a computational model by which an analyst is able to assess trust in reported information based on several possible unknown attributes of the source as well as additional characterization of the informational content by the source itself. The paper also illustrates the use of URREF criteria to track uncertainty affecting the results, from model construction to its formalization with belief functions. First, a model for trust estimation has been developed that combines several attributes of sources and their own assessment of the items reported. The model is implemented using belief functions, and takes advantage of its mathematical background to define fusion operators for trust assessment. Several scenarios are presented to illustrate uncertainty analysis, illustrating when uncertainty occurs and how it affects partial results for different applications.

Tracking uncertainty is suitable for fusion systems in which various human sources send observations of questionable quality and there is a need to continuously update the trust associated with reports to be analyzed. The set of URREF criteria offers a unified basis to analyze inaccuracies affecting trust estimation during different phases: observation, reporting, and fusion. Select use cases clearly illustrated the benefits of managing uncertainties arising during the modeling and formalization phases, with the twofold analysis offering additional details on results and improving their interpretation.

The general approach taken in this paper could be adapted to investigate the general mechanisms by which fusion processes integrate information from multiple sources. The solution is especially useful for comparing different fusion approaches with respect to their implications for uncertainty management.

REFERENCES

[1] L. Amgoud and R. Demolombe
An argumentation-based approach for reasoning about trust in information sources.
Argument & Computation, 5(2–3):191–215, 2014.

[2] A. Auger and J. Roy
Expression of uncertainty in linguistic data.
In *Information Fusion, 2008 11th International Conference on*, pages 1–8. IEEE, 2008.

[3] K. S. Barber and J. Kim
Belief revision process based on trust: Simulation experiments.
In *Proceedings of Autonomous Agents '01 Workshop on Deception, Fraud, and Trust in Agent Societies*. Citeseer, 2001.

[4] L. Berti-Équille and J. Borge-Holthoefer
Veracity of data: From truth discovery computation algorithms to models of misinformation dynamics.
Synthesis Lectures on Data Management, 7(3):1–155, 2015.

[5] E. Blasch and A. Aved
Urref for veracity assessment in query-based information fusion systems.
In *Information Fusion (Fusion), 2015 18th International Conference on*, pages 58–65. IEEE, 2015.

[6] E. Blasch, P. C. Costa, K. B. Laskey, H. Ling, and G. Chen
The urref ontology for semantic wide area motion imagery exploitation.
In *Aerospace and Electronics Conference (NAECON), 2012 IEEE National*, pages 228–235. IEEE, 2012.

[7] E. Blasch, K. B. Laskey, A.-L. Jousselme, V. Dragos, P. C. Costa, and J. Dezert
Urref reliability versus credibility in information fusion (stanag 2511).
In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 1600–1607. IEEE, 2013.

[8] E. Blasch, A. Jøsang, J. Dezert, P. C. Costa, and A.-L. Jousselme
Urref self-confidence in information fusion trust.
In *Information Fusion (FUSION), 2014 17th International Conference on*, pages 1–8. IEEE, 2014.

[9] D. B. Buller and J. K. Burgoon
Interpersonal deception theory.
Communication theory, 6(3):203–242, 1996.

[10] C. Castelfranchi and R. Falcone
Trust is much more than subjective probability: Mental components and sources of trust.
In *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on*, pages 10–pp. IEEE, 2000.

[11] C. Castelfranchi, R. Falcone, and G. Pezzulo
Trust in information sources as a source for trust: a fuzzy approach.
In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, pages 89–96. ACM, 2003.

[12] L. Cholvy
How strong can an agent believe reported information?
In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 386–397. Springer, 2011.

[13] P. C. Costa, K. B. Laskey, E. Blasch, and A.-L. Jousselme
Towards unbiased evaluation of uncertainty reasoning: The urref ontology.
In *Information Fusion (FUSION), 2012 15th International Conference on*, pages 2301–2308. IEEE, 2012.

[14] L. Cronk
Communication as manipulation: Implications for biosociological research.
In *Proceedings of the American Sociological Association Annual Meetings*, 1991.

[15] J. P. de Villiers, K. Laskey, A.-L. Jousselme, E. Blasch, A. de Waal, G. Pavlin, and P. Costa
Uncertainty representation, quantification and evaluation for data and information fusion.
In *Information Fusion (Fusion), 2015 18th International Conference on*, pages 50–57. IEEE, 2015.

[16] J. P. de Villiers, A.-L. Jousselme, A. de Waal, G. Pavlin, K. Laskey, E. Blasch, and P. Costa
Uncertainty evaluation of data and information fusion within the context of the decision loop.
In *Information Fusion (FUSION), 2016 19th International Conference on*, pages 766–773. IEEE, 2016.

[17] J. P. de Villiers, R. Focke, G. Pavlin, A.-L. Jousselme, V. Dragos, K. Laskey, P. Costa, and E. Blasch
Evaluation metrics for the practical application of urref ontology: An illustration on data criteria.
In *Information Fusion (Fusion), 2017 20th International Conference on*, pages 1–8. IEEE, 2017.

[18] J. P. de Villiers, G. Pavlin, P. Costa, A.-L. Jousselme, K. Laskey, V. Dragos, and E. Blasch
Subjects under evaluation with the urref ontology.
In *Information Fusion (Fusion), 2017 20th International Conference on*, pages 1–8. IEEE, 2017.

[19] J. Dezert, W. Pei, and A. Tchamova
On the validity of dempster-shafer theory.
In *Information Fusion (Fusion), 2012 15th International Conference on*. IEEE, 2012.

[20] X. L. Dong, E. Gabrilovich, G. Heitz, W. Horn, K. Murphy, S. Sun, and W. Zhang
From data fusion to knowledge fusion.
Proceedings of the VLDB Endowment, 7(10):881–892, 2014.

[21] X. L. Dong, E. Gabrilovich, K. Murphy, V. Dang, W. Horn, C. Lugaresi, S. Sun, and W. Zhang
Knowledge-based trust: Estimating the trustworthiness of web sources.
Proceedings of the VLDB Endowment, 8(9):938–949, 2015.

[22] V. Dragos
An ontological analysis of uncertainty in soft data.
In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 1566–1573. IEEE, 2013.

[23] V. Dragos and K. Rein
Integration of soft data for information fusion: Pitfalls, challenges and trends.
In *Information Fusion (FUSION), 2014 17th International Conference on*, pages 1–8. IEEE, 2014.

[24] V. Dragos, X. Lerouvreur, and S. Gatepaille
A critical assessment of two methods for heterogeneous information fusion.
In *Information Fusion (Fusion), 2015 18th International Conference on*, pages 42–49. IEEE, 2015.

[25] V. Dragos, Z. Juergen, and J. P. de Villiers
Applicaton of urref criteria to assess knowledge representation in cyber threats models.
In *Information Fusion (FUSION), 2018 21 International Conference on*. IEEE, 2018.

[26] P. Everaere, S. Konieczny, and P. Marquis
Belief merging versus judgment aggregation.
In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 999–1007. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[27] R. Falcone and C. Castelfranchi
Social trust: A cognitive approach.
In *Trust and deception in virtual societies*. Springer, 2001.

[28] R. Falcone, M. Piunti, M. Venanzi, and C. Castelfranchi
From manifesta to krypta: The relevance of categories for trusting others.
ACM Transactions on Intelligent Systems and Technology (TIST), 4(2):27, 2013.

[29] R. Falcone, A. Sapienza, and C. Castelfranchi
The relevance of categories for trusting information sources.
ACM Transactions on Internet Technology (TOIT), 15(4):13, 2015.

[30] J. Granatyr, V. Botelho, O. R. Lessing, E. E. Scalabrin, J.-P. Barthès, and F. Enembreck
Trust and reputation models for multiagent systems.
ACM Computing Surveys (CSUR), 48(2):27, 2015.

[31] J. Guo and R. Chen
A classification of trust computation models for service-oriented internet of things systems.
In *Services Computing (SCC), 2015 IEEE International Conference on*, pages 324–331. IEEE, 2015.

[32] D. L. Hall and J. M. Jordan
Human-centered information fusion.
Artech House, 2010.

[33] J. T. Hancock, L. E. Curry, S. Goorha, and M. Woodworth
On lying and being lied to: A linguistic analysis of deception in computer-mediated communication.
Discourse Processes, 45(1):1–23, 2007.

[34] R. Hermoso, H. Billhardt, and S. Ossowski
Trust-based role coordination in task-oriented multiagent systems.
Knowledge-Based Systems, 52:78–90, 2013.

[35] D. Jean, W. Pei, and A. Tchamova
On the validity of dempster's fusion rule and its interpretation as a generalization of bayesian fusion rule.
International Journal of Intelligent Systems, Special Issue: Advances in Intelligent Systems, 29(3):223–252, 2014.

[36] A. Josang, R. Ismail, and C. Boyd
A survey of trust and reputation systems for online service provision.
Decision Support Systems, 43(2):618–644, 2007.

[37] A.-L. Jousselme
Semantic criteria for the assessment of uncertainty handling fusion models.
In *Information Fusion (FUSION), 2016 19th International Conference on*, pages 488–495. IEEE, 2016.

[38] A.-L. Jousselme and G. Pallotta
Dissecting uncertainty-based fusion techniques for maritime anomaly detection.
In *Information Fusion (Fusion), 2015 18th International Conference on*, pages 34–41. IEEE, 2015.

[39] A.-L. Jousselme, A.-C. Boury-Brisset, B. Debaque, and D. Prevost
Characterization of hard and soft sources of information: A practical illustration.
In *Information Fusion (FUSION), 2014 17th International Conference on*, pages 1–8. IEEE, 2014.

[40] S. Kent
Words of estimative probability.
Studies in Intelligence, 8(4):49–65, 1964.

[41] R. F. Kesselman
Verbal probability expressions in national intelligence estimates: a comprehensive analysis of trends from the fifties through post 9/11.
PhD thesis, MERCYHURST COLLEGE, 2008.

[42] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi
Multisensor data fusion: A review of the state-of-the-art.
Information Fusion, 14(1):28–44, 2013.

[43] H. Koen, J. P. de Villiers, G. Pavlin, A. de Waal, P. de Oude, and F. Mignet
A framework for inferring predictive distributions of rhino poaching events through causal modelling.
In *Information Fusion (FUSION), 2014 17th International Conference on*, pages 1–7. IEEE, 2014.

[44] A. Koster, A. L. Bazzan, and M. de Souza
Liar liar, pants on fire; or how to use subjective logic and argumentation to evaluate information from untrustworthy sources.
Artificial Intelligence Review, 48(2):219–235, 2017.

[45] T. Lukoianova and V. L. Rubin
Veracity roadmap: Is big data objective, truthful and credible?
Advances in Classification Research Online, 24(1):4–15, 2014.

[46] P.-A. Matt, M. Morge, and F. Toni
Combining statistics and arguments to compute trust.
In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1–Volume 1*, pages 209–216. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[47] D. Mercier, B. Quost, and T. DenÅ\uxa
Refined modeling of sensor reliability in the belief function framework using contextual discounting.
Information Fusion, 9(2):246–258, 2008.

[48] M. Momani, S. Challa, and R. Alhmouz
Bnwsn: Bayesian network trust model for wireless sensor networks.
In *Communications, Computers and Applications, 2008. MIC-CCA 2008. Mosharaka International Conference on*, pages 110–115. IEEE, 2008.

[49] NATO
STANAG 2011.
Technical report, Intelligence reports, 2003.

[50] F. Paglieri, C. Castelfranchi, C. da Costa Pereira, R. Falcone, A. Tettamanzi, and S. Villata
Trusting the messenger because of the message: feedback dynamics from information quality to source evaluation.
Computational and Mathematical Organization Theory, 20(2):176–194, 2014.

[51] S. Parsons, K. Atkinson, Z. Li, P. McBurney, E. Sklar, M. Singh, K. Haigh, K. Levitt, and J. Rowe
Argument schemes for reasoning about trust.
Argument & Computation, 5(2–3):160–190, 2014.

[52] G. Pigozzi
Belief merging and judgment aggregation.
Stanford Encyclopedia of Philosophy, 2015.

[53] K. Rein
I think it is likely that it might be so—Exploiting Lexical Clues for the Automatic Generation of evidentiality Weights for Information Extracted from English Text.
Dissertation, University of Bonn, 2016.

[54] V. L. Rubin
Stating with certainty or stating with doubt: Intercoder reliability results for manual annotation of epistemically modalized statements.
In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers*, pages 141–144. Association for Computational Linguistics, 2007.

[55] R. Saurí and J. Pustejovsky
Are you sure that this happened? assessing the factuality degree of events in text.
Computational Linguistics, 38(2):261–299, 2012.

[56] G. Shafer
A mathematical theory of evidence,
volume 1. Princeton university press Princeton, 1976.

[57] W. Sherchan, S. Nepal, and C. Paris
A survey of trust in social networks.
ACM Computing Surveys (CSUR), 45(4):47, 2013.

[58] M. B. Sinai, N. Partush, S. Yadid, and E. Yahav
Exploiting social navigation.
arXiv preprint arXiv:1410.0151, 2014.

[59] F. Smarandache and J. Dezert
Advances and applications of DSmT for information fusion, volume 1–4. American Research Press (ARP), 2015.

[60] F. Smarandache, J. Dezert, and J.-M. Tacnet
Fusion of sources of evidence with different importances and reliabilities.
In *Information Fusion (FUSION), 2010 13th Conference on*, pages 1–8. IEEE, 2010.

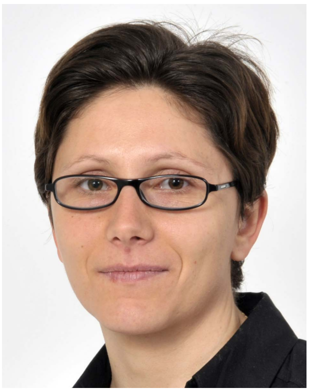
[61] F. Smarandache, J. Dezert, and J.-M. Tacnet
Fusion of sources of evidence with different importances and reliabilities.
In *Information Fusion (Fusion), 2012 15th International Conference on*. IEEE, 2010.

[62] Y. Tang, K. Cai, P. McBurney, E. Sklar, and S. Parsons
Using argumentation to reason about trust and belief.
Journal of Logic and Computation, 22(5):979–1018, 2011.

[63] D. P. Twitchell, D. P. Biros, M. Adkins, N. Forsgren, J. K. Burgoon, and J. F. Nunamaker
Automated determination of the veracity of interview statements from people of interest to an operational security force.
In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, volume 1, pages 17a–17a. IEEE, 2006.

[64] M. Venanzi, A. Rogers, and N. R. Jennings
Trust-based fusion of untrustworthy information in crowdsourcing applications.
In *Proceedings of the 2013 international conference on autonomous agents and multi-agent systems*, pages 829–836. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[65] S. Villata, G. Boella, D. M. Gabbay, and L. Van Der Torre
Arguing about the trustworthiness of the information sources.
In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 74–85. Springer, 2011.



Dr. Valentina Dragos is a research scientist, member of the Department of Information Modeling and Systems at ONERA, The French Aerospace Lab in Palaiseau, France. Valentina received Master and PhD degrees in Computer Science from Paris V University and her research interests include artificial intelligence, with emphasis on natural language processing, semantics technologies and automated reasoning. Since joining ONERA in 2010, Valentina contributed to several academic and industrial security-oriented projects, addressing topics such as: semantic interoperability for command and control systems, heterogeneous information fusion, and exploitation of semantic data (HUMINT, OSINT) for situation assessment.

Jean Dezert was born in l'Hay les Roses, France in 1962. He received D.E.A. degree from University Paris VII (Jussieu), Paris, in 1986, and the Ph.D.degree from University Paris XI, Orsay, France, in 1990, all in automatic control and signal processing. He is a Senior Research Scientist with the Information and Fusion Systems Research Team, in Information Processing and Systems (DTIS), ONERA, Palaiseau, France. He gave several invited plenary talks on information fusion in Europe, America, Australia, and China, during latest years. His current research interests include autonomous navigation, estimation, stochastic systems and their applications to multisensormultitarget tracking, information fusion, and plausible reasoning. He is cofounder with Prof. Smarandache of Dezert-Smarandache Theory (DSmT) for information fusion and has published around 180 papers in his research fields in scientific journals and in proceedings of international conferences, and five books. Dr. Dezert has served as a Local Arrangements Organizer for the 2000 3rd International Conference on Information Fusion (Fusion) in Paris, and a Secretary, an Executive Vice-President, and the President for the International Society of Information Fusion (ISIF) in 2001, 2004, and 2016, respectively. He has been involved in the Technical Program Committee of Fusion International Conferences.

Kellyn Rein is a Research Associate at the Research Institute for Communication, Information Processing Ergonomics (FKIE) at Fraunhofer Gesellschaft, Europe's largest applied research organization. She received her PhD from the University of Bonn in Linguistics, after a BA at Michigan State University, and two masters, one in Management and one in Computer Information Systems, both with joint conferral from Boston University (1984) and Vrije Universiteit Brussel in Belgium. Her area of research interest is in the analysis of uncertainty in natural language information, information fusion and information quality. She has been active in EU projects focused variously on crisis management, cross-border organized crime, smart cities, and CBRN terrorism. She has also been involved in numerous NATO Research Task Groups, including one on analysis and communication of uncertainty in intelligence, mission assurance and cyber risk assessment for unmanned autonomous system, and currently chairing a task group on multi-level, multi-source information fusion.