Assessment of Trust in Opportunistic Reporting using Belief Functions

Valentina Dragos
ONERA - The French Aerospace Lab
Palaiseau, France
valentina.dragos@onera.fr

Jean Dezert
ONERA - The French Aerospace Lab
Palaiseau, France
jean.dezert@onera.fr

Kellyn Rein
ITF, Fraunhofer, FKIE
Wachtberg, Germany
kellyn.rein@fraunhofer.fkie.de

***Abstract*—This paper addresses the assessment of trust in items reported by opportunistic sources and develops a model taking into account the reliability of sources and the credibility of information. Opportunistic sources are witnesses describing events, scenes or actions of interest and often such sources may also indicate confidence, doubt, skepticism or conviction about the items they report. Reliability captures how trustworthy the analyst believes someone is, by considering that opportunistic sources are not always competent and honest but also erroneous or even malicious. Credibility relies upon self-confidence, a measure of how strongly the source believes what he is uttering. Reliability is furthered decomposed according to competences and skills of the source, intentions and reputation while self-confidence varies in a specified interval. Belief functions are used to formalize the model and trust assessment is implemented thanks to discounting operators. The paper also discusses several experiments designed to investigate the effects of reliability and self-confidence on the quality of reported items, as perceived by the analyst.**

***Index Terms*—trust, reliability, belief functions, information fusion, DSmT**

I. Introduction

In practical applications, integration of human reporting has much to offer for tasks where subtle interactions and associations are difficult to detect with physical sensors, such as intelligence analysis. Those applications combine semantic information with numerical data and often use opportunistic sources, as different persons observe and then report on actions or events of interest and the information is instantly available.

Systems able to properly utilize information reported by opportunistic sources face two main challenges: understanding the meaning of items [1] and assessing the trust or confidence degree with respect to previous failures, known skills and supposed intentions of the source.

This paper tackles the problem of trust assessment for opportunistic sources and adopts a model considering not the information itself, but the quality of sources and their perspective towards reported statements. Trust is then discounted along two dimensions, capturing the reliability of sources, as stated by the analyst, and the credibility of assertions, as stated by the source.

When using human sources, reports are of variable quality, as individuals have their own skills, intentions, objectives and subjective standpoints. Processing inaccurate or distorted items can result in unfortunate consequences and analysts have to estimate the quality of reported information before further aggregation. Another reason for trust assessment is related to source's ability to distort the facts, as humans can deliberately misreport observations, with malicious intentions to commit a wrongful act. How strong the analyst trusts the items is then assessed on the basis of several criteria, including whether the source is considered as being competent or honest.

Against this background, this paper presents a model to assess trust by taking into account the subjective evaluation of information by the source, called self-confidence, and its evaluation by the analyst, called reliability. While self-confidence covers intention aspects, reliability is assigned to sources according to the quality of their previously reported items, their competence and supposed intentions.

Previous work has focused on analyzing the elements of the model [2]. This paper completes the previous contribution and describes the estimation of trust with belief functions, a formalism which offers sound mathematical basis to implement specific fusion operators. Several scenarios illustrate how the model performs to estimate trust in the case of incomplete, misleading or ambiguous reporting situations.

The remainder of the paper is divided into six sections : the next section briefly describes the model for trust estimation; model formalization with belief functions is presented in section III. A running example and scenarios for trust assessment are illustrated in section IV. Section V discuses related approaches for trust assessment and section VI presents concluding remarks.

II. A functional model of trust

A. Trust modeling for opportunistic reporting

The model for trust assessment is presented in details in [2]. This paper builds upon this initial contribution and briefly discusses the model for trust assessment before addressing its formalization with belief functions.

The model keeps the distinction between source and information by considering separate dimensions for each element. The rationale behind this approach is the observation that even reliable sources can provide inaccurate or imprecise information from one report to another, which is even more plausible in the case of opportunistic sources.

By focusing on the global characterization of reported items, the model aims at providing a better understanding of how

trust is to be constructed from various dimensions. The model consists of several elements related by functions, see fig. 1.

Fig. 1. Model for trust analysis.

B. Elements of the trust model

The model is composed of two elements: an information source and reported items.

Definition of information source: information source is a system producing or containing information intended for transmission. Several research efforts aimed at modeling properties of information sources. A general analysis of sources is undertaken by Hall and Jordan [3], 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. For this work, we consider opportunistic sources, which is to say agents who can be either directly involved in the events reported, or just uttering as witnesses. Information reported by such sources is unstructured, vague, ambiguous and subjective and is often contrasted with information coming from physical sensors, described as structured, quantitative and objective. Opportunistic sources can deliberately distort the information or even lie.

Definition of reported information: Reported information is a couple (I, C), where I is an item of information and C the certainty level as assigned by the source. For this work reported information is understood as natural language assertions, in their textual form.

C. Functions of the trust model

The model introduces several functions to estimate reliability, self-confidence and trust described hereafter.

Definition of a reliability function: a reliability function is a mapping associating a numerical value to information sources. This can be a description can real value, e.g. probability of providing accurate information or failure rate, or as a mass distribution expressing to what extent the source is considered as reliable, unreliable or having unknown status.

Reliability is a complex concept and, from a practical standpoint, it is difficult to have information about the global reliability of a source. Thus, the model describes reliability along three attributes: the competence of a source, its reputation and intentions, and seeks to estimate each attribute. This solution allows us to compensate for insufficient information on one or several aspects of reliability and to conduct, in some cases, the analysis of source's reliability based on a single attribute.

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 appreciations of the information conveyed. For this work, we adopt the notion of self-confidence as introduced by the homonym concept in the URREF (Uncertainty Representation and Reasoning Evaluation Framework) ontology developed by the Evaluation of Technologies for Uncertainty Representation Working Group (ETURWG) [1]. Thus *self-confidence is intended to shed light on the subtle transposition of source's qualities into attributes of provided items*, see [4].

Using the model for trust analysis: The model proposed in this work allows assessing levels of trust in reported information by combining various attributes of the source, introduced under the hat of reliability, and self-confidence, capturing the credibility of information as stated by the human.

The model is source-centric and considers aspects of reliability as main factors having the ability to correct, alter, or qualify the information reported by the source. If several rules to rank, prioritize, or combine 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 to estimate 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 we formalize the model using belief functions and illustrate how trust is estimated using operators for three practical scenarios. We adopted belief functions because they offer a general theoretical framework to deal with various aspects of uncertainty, while having a sound mathematical ground, flexible representation capabilities and a variety of operators to combine uncertainty representations.

III. Trust formalization with belief functions

A. Basic Belief Assignment

Belief Functions (BF) have been introduced by Shafer in his mathematical theory of evidence [5], also called 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

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

evidence is a source of information characterized by Basic Belief Assignment (BBA), or a mass function, which is a 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

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

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

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 to model a full 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 imprecision (or uncertainty) of the probability of any subset A of the frame of discernment Θ is characterized by the width $U(A) = Pl(A) - Bel(A)$ of the belief interval of A.

Shafer did propose Dempster's rule of combination to combine multiple independent sources of evidence [5], which is the normalized conjunctive fusion rule. This rule has been strongly disputed in BF community since Zadeh's first criticism in 1979, and since the 1990's many rules have been proposed to combine (more or less efficiently) BBAs, see discussions in [6], in particular the proportional conflict redistribution rule no 6 (PCR6).

Instead of working with quantitative (numerical) BBA, it is also possible to work with qualitative BBA expressed by labels with the linear algebra of refined labels proposed in Dezert-Smarandache Theory (DSmT), [6] (Vol. 2 & 3).

B. Trust formalization model

In order to avoid confusion with *belief* which are well defined mathematical concepts in the theory of belief functions, in the following we will use the term *self-confidence* to name the confidence declared by a source Y on its own assertion A. Moreover, the self-confidence is considered as a piece of information for the analyst X.

Let's denote by A the assertion given by the source, for instance $A =$ *John is a terrorist*. 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).

Let's consider the quantitative representation of $v(A)$ for simplicity[2].

[2]In practical applications and without loss of generality one can always map a qualitative representation to a quantitative one by a proper choice of scaling and normalization.

The basic piece of information 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 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 case.

The self-confidence of the source Y is a 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 confident at all in his valuation $v(A)$. In the theory of belief functions, this factor is referred as the reliability discounting rate factor of the source.

This factor allows to adjust the value of the piece of information $(A, v(A))$ into a discounted one $(A, v'(A))$. The basic idea is to keep the piece of information unchanged if the source considers herself as totally reliable, and, in the extreme case, to not take into account the information if the source considers herself as totally unreliable. In this extreme case, the BBA $m(\cdot)$ provided by the source must naturally be transformed into the vacuous BBA defined by $m'(\Theta) = m'(A\cup\bar{A}) = 1$. This very simple and natural discounting technique has been presented in [5]. We recall its mathematical formulation for convenience

$$m'(A) = \alpha_Y \cdot m(A) \tag{4}$$

$$m'(\bar{A}) = \alpha_Y \cdot m(\bar{A}) \tag{5}$$

$$m'(A\cup\bar{A}) = \alpha_Y \cdot m(A\cup\bar{A}) + (1-\alpha_Y) \tag{6}$$

One can easily verify the belief mass of all focal elements are reduced with the factor $\alpha_Y \in [0,1]$ and all the missing discounted mass $1-\alpha_Y\cdot m(A)-\alpha_Y\cdot m(\bar{A})-\alpha_Y\cdot m(A\cup\bar{A}) = 1-\alpha_Y$ is transferred to the whole ignorance $A\cup\bar{A}$. In the extreme case when $\alpha_Y = 0$, one gets $m'(A\cup\bar{A}) = 1$ which corresponds to the vacuous BBA (i.e. the uninformative piece of information). Note that the valuation of this discounted piece of information is always degraded if $\alpha_Y < 1$ and $m(A\cup\bar{A}) < 1$ because[3] $m'(A\cup\bar{A}) > m(A\cup\bar{A})$, 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 r is to consider that

[3]Indeed, to prove $m'(A\cup\bar{A}) > m(A\cup\bar{A})$, it suffices to prove $\alpha_Y\cdot m(A\cup\bar{A})+(1-\alpha_Y) > m(A\cup\bar{A})$, or equivalently to prove $(1-\alpha_Y)\cdot m(A\cup\bar{A}) < (1-\alpha_Y)$, but the latter inequality is obviously true because $m(A\cup\bar{A}) \in [0,1[$ and $(1-\alpha_Y) > 0$ Q.E.D.

C_Y, R_Y and I_Y factors are represented by numbers $[0,1]$ associated to chosen 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 this factor has equal weight, then one could use $r = C_Y \times R_Y \times I_Y$ as simple product of probabilities. However in practice, such simple modeling does not fit well with what the analyst really needs for taking 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 it.

C. Trust estimation by discounting

We consider the three criteria C_Y, R_Y and I_Y with some associated importance weights w_C, w_R, w_I in $[0,1]$ with $w_C+w_R+w_I = 1$. Using those criteria allows us to emphasize different application scenarios, for which various aspects of reliability are important. For instance, an analyst considering the source as competent will increase the corresponding coefficient, although the same coefficients will be deceased for new, unknown sources.

Moreover, 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 $(m_C(r), m_C(\bar{r}), m_C(r \cup \bar{r}))$, 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 reliability r of the source Y one needs to fuse efficiently 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 in applying the importance discounting approach combined with Proportional Conflict Redistribution Rule # 6 (i.e. PCR6) fusion rule [6]. The importance discounting is the dual form of Shafer's (i.e. reliability) discounting technique. It has been presented in details in [7]. Its basic principle is to discount a BBA $m(.)$ by multiplying the mass of all its focal elements by the chosen importance discounting factor $w \in [0,1]$. A source of evidence is considered as not important when $w = 0$, fully important when $w = 1$, and any intermediate value of w in $[0,1]$ can be used to represent any (possibly subjective) other degree of importance of a source in the fusion process. Unlike Shafer's discounting, all the missing mass $1 - \sum_{A \in 2^\Theta} w \cdot m(A) = 1 - w$ is not redistributed back to the whole uncertainty as done in (6), but to the empty set. This importance discounting technique allows to make a clear distinction between reliability discounting and importance discounting of a source of evidence in the processing of information in the framework of belief functions. It is worth to note that Shafer's rule of combination is not responding to importance discounting (see example in [7]), and that is why we propose to make the combination of importance-discounted BBA $m_C(.)$, $m_R(.)$ and $m_I(.)$ with PCR6 fusion rule. As we see, the importance discounting technique is as simple as Shafer's reliability discounting and allows to model the discounting of the sources with their importance factors, which is particularly useful and appealing for multi-criteria decision-making problems. In summary, in our particular context we proceed as follows:

- Step 1: Importance discounting of $m_C(\cdot)$ by $w_C \in [0,1]$ to get

$$m'_C(r) = w_C \cdot m_C(r) \tag{7}$$
$$m'_C(\bar{r}) = w_C \cdot m_C(\bar{r}) \tag{8}$$
$$m'_C(r \cup \bar{r}) = w_C \cdot m_C(r \cup \bar{r}) \tag{9}$$
$$m'_C(\emptyset) = 1 - w_C \tag{10}$$

- Step 2: Importance discounting of $m_R(\cdot)$ by $w_R \in [0,1]$ is done similarly as for $m'_C(\cdot)$ above to get $m'_R(r)$, $m'_R(\bar{r})$, $m'_R(r \cup \bar{r})$ and $m'_R(\emptyset) = 1 - w_R$.
- Step 3: Importance discounting of $m_I(\cdot)$ by $w_I \in [0,1]$ is done similarly as for $m'_C(\cdot)$ above to get $m'_I(r)$, $m'_I(\bar{r})$, $m'_I(r \cup \bar{r})$ and $m'_I(\emptyset) = 1 - w_I$.
- Step 4: Fusion of the importance-discounted BBAs $m'_C(\cdot)$, $m'_R(\cdot)$ and $m'_I(\cdot)$ with PCR6 rule of combination to get unnormalized BBA $m'_{PCR6}(r)$, $m'_{PCR6}(\bar{r})$, $m'_{PCR6}(r \cup \bar{r})$ and $m'_{PCR6}(\emptyset)$.
- Step 5: We normalize $m'_{PCR6}(\cdot)$ to get the final normalized BBA. This is done by dividing $m'_{PCR6}(r)$, $m'_{PCR6}(\bar{r})$, and $m'_{PCR6}(r \cup \bar{r})$ by $1 - m'_{PCR6}(\emptyset)$. That is by setting $m_{PCR6}(\emptyset) = 0$ and taking

$$m_{PCR6}(r) = m'_{PCR6}(r)/[1 - m'_{PCR6}(\emptyset)] \tag{11}$$
$$m_{PCR6}(\bar{r}) = m'_{PCR6}(\bar{r})/[1 - m'_{PCR6}(\emptyset)] \tag{12}$$
$$m_{PCR6}(r \cup \bar{r}) = m'_{PCR6}(r \cup \bar{r})/[1 - m'_{PCR6}(\emptyset)] \tag{13}$$

This importance discounting process followed by PCR6 fusion rule leads to the valuation $v(r) = (m_{PCR6}(r), m_{PCR6}(\bar{r}), m_{PCR6}(r \cup \bar{r}))$ from which either 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 as discussed in [6].

The trust model consists in using both the piece of information $(A, v(A))$ and self-confidence factor α_Y provided by the source Y, and 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, one proposes to use the mass $m(r)$ of reliability hypothesis r of the source Y as a new discounting factor of 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

$$m_t(A) = m(r) \cdot m'(A) \tag{14}$$
$$m_t(\bar{A}) = m(r) \cdot m'(\bar{A}) \tag{15}$$
$$m_t(A \cup \bar{A}) = m(r) \cdot m'(A \cup \bar{A}) + (1 - m(r)) \tag{16}$$

or equivalently by

$$m_t(A) = m(r)\alpha_Y \cdot m(A) \quad (17)$$
$$m_t(\bar{A}) = m(r)\alpha_Y \cdot m(\bar{A}) \quad (18)$$
$$m_t(A \cup \bar{A}) = m(r)\alpha_Y \cdot m(A \cup \bar{A}) + (1 - m(r)\alpha_Y) \quad (19)$$

Some strategies based on the level of $m(r)$ can be developed in order to avoid the automatic application of this discounting technique, depending on the problem under analysis.

IV. Scenarios for trust assessment

Let's adopt intelligence analysis as application task, see fig. 2, and consider an assertion A and its valuation $v(A)$ provided by the source Y as follows:

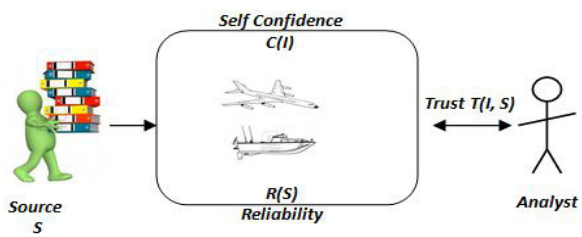

Fig. 2. Reported Information

$m(A) = 0.7$, $m(\bar{A}) = 0.1$ 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 [7] which consists to discount the BBAs with the importance factor and redistribute the missing mass onto the empty set, and combining the discounted BBAs with PCR6 fusion rule, we get after normalization the following BBA:

$$m(r) = 0.9335$$
$$m(\bar{r}) = 0.0415$$
$$m(r \cup \bar{r}) = 1 - m(r) - m(\bar{r}) = 0.025$$

Note that if we $m_C(r) = m_R(r) = m_I(r) = 1$, then we will always get $m(r) = 1$ whatever is the choice of weightings factors, which is normal. If one has a total conflict between valuations of reliability based on Competence, Reputation and Intention criteria then Dempster's rule cannot be applied to get global reliability factor $m(r)$ because of 0/0 indeterminacy in 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.

In the next subsections, we explore several typical scenarios for trust assessment corresponding to very different situations of distributions of BBAs.

A. Scenario 1 - Reputation

Suppose that Y provides A and $v(A)$ and X has no global description of Y in terms of reliability. As reliability of Y is not available, the 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

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

Therefore, one finally obtains as trust valuation

$$m_t(A) = 0.7559, m_t(\bar{A}) = 0.0945, m_t(A \cup \bar{A}) = 0.1496$$

For this scenario the source is confident about its 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. Reliability of the source is build 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 shows the impact of using incompletely 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

$$m_t(A) = 0.7559, m_t(\bar{A}) = 0.0945, m_t(A \cup \bar{A}) = 0.1496$$

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

B. Scenario 2 - 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

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

and

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

Therefore, one finally obtains as trust valuation

$$m_t(A) = 0.0040, m_t(\bar{A}) = 0.0013, m_t(A \cup \bar{A}) = 0.9946$$

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 skills or competence to provide accurate descriptions of events. In this case, the reports is incomplete, ambiguous, or even irrelevant. In addition to low competence and reliability, the source 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 an highly uncertain.

C. Scenario 3 - Misleading report

In this case, Y provides the assertion A, while stating that it certainly holds and X considers Y as a completely unreliable source. For this case, the analyst knows that the report is inaccurate, for example, it cannot be corroborated or it contradicts fully or partially information from other (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 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

$$m(r) = 0.0237, m(\bar{r}) = 0.9109, m(r \cup \bar{r}) = 0.0654$$

Therefore, one finally obtains as trust valuation

$$m_t(A) = 0.0237, m_t(\bar{A}) = 0, m_t(A \cup \bar{A}) = 0.9763$$

Values for this scenario show 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 wile be ignored, as the reliability of the sources is drastically decreased by a high mass assignment on misleading intentions.

As highlighted by the examples above, there is a need for a more detailed investigation of trust in reported information able to make the distinction between items considered as untrustworthy because the source reports on topics out of its depth, makes observations under difficult conditions or has the intention to deceive.

V. Related approaches

The concept of trust is directly related to actions of individuals [8]. Although having an obvious social dimension, trust is not only understood towards other humans, but also towards information pieces [9], information sources [10], Internet sites [11], data and knowledge fusion algorithms [12], intelligent agents [13] or services in the Internet of Things [14].

Trust assessment is not a new research topic, spanning areas as diverse as agent systems [13], logical modeling and argumentation [15], service provision on the Internet [16], decision making under uncertainty or social networks analysis [17]. Applications of trust analysis are also of interest in the military field, where techniques were developed in order to identify clues of veracity in interview statements [18].

While definitions of trust vary from one domain to another, they all highlight 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. Second, all definition associate some degree of uncertainty with trust, which is then captured by concepts such as subjective certainty or agent's beliefs [19], subjective probability [20] or the feeling of security [21].

Contributions on trust estimation keep the distinction between analyzing the source of information, the item reported or reasoning about trust. Approaches developed for trust in information sources considers that trust is not a general attribute of the source but rather related to some properties like competence in [22], sincerity, willingness to cooperate [15], validity or vigilance. On this basis, it becomes possible to consider the competence of a source not in general but with respect to specific topics [23], or trust is further analyzed in relation to roles, categories or classes [24].

Works on reasoning about trust focus in modeling trust on information sources not directly from its properties, but from analyzing past behaviors of sources [25] or inferring trust in some properties from already estimated trust in other properties [26]. Among them, some focus on building trust by using argumentation [27], beliefs fusion or judgment aggregation [28] or investigate their joint integration [29]. [30] developed various patterns to reason about trust ans its provenance while the notion of conflict in handling trust is discussed in [31].

Solutions addressing trust in reported information investigates the way items are distorted while passing from one source to another before being considered for decision [32], but generally remains a topic that has received few attention. Most approaches developed logic-based solutions to model distortions of items represented as logical structures that are exchanged between agents having intentions and the ability to deceive. However, there are more challenges arising when the information is analyzed in its textual form: uncertainty as expressed in natural language statements is investigated in [33] while [34] provides a broader discussion of pitfalls and challenges related to soft data integration.

This paper fills the gap by addressing trust estimation for opportunistic reporting. The paper is similar to the approach developed in [32] and it takes a step further by taking into account both self-confidence and intentions of opportunistic sources.

VI. Conclusions and Future Work

Trust assessment plays a crucial role for systems and processes relying on information reported by human sources.Trust estimation is also of particular interest for applications integrating assertions that are constantly transferred from humans to automatic processing and back.

This paper tackles the problem of trust in information reported by opportunistic sources and presents a model that explores different attributes of human sources and their ability to assess the items provided. Belief functions offers a useful frame to implement the model and allows the definition of specific operators to combine elements of the model.

The paper also discuss several scenarios in an attempt to illustrate the estimation of trust within a system composed of sources and analysts under different circumstances. Experimental results shows that variations of trust when several features are altered are consistent with commonsense assumptions, and, intuitively, the model behaves correctly.

References

[1] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *Journal of machine learning research*, vol. 12, no. Aug, pp. 2493–2537, 2011.

[2] V. Dragos and K. Rein, "What's in a message? exploring dimensions of trust in reported information," in *2016 19th International Conference on Information Fusion (FUSION)*. IEEE, 2016, pp. 2125–2132.

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

[4] 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*. IEEE, 2012, pp. 2301–2308.

[5] G. Shafer *et al.*, *A mathematical theory of evidence*. Princeton university press Princeton, 1976, vol. 1.

[6] F. Smarandache and J. Dezert, *Advances and applications of DSmT for information fusion*. American Research Press (ARP), 2015, vol. 1-4.

[7] 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*. IEEE, 2010, pp. 1–8.

[8] T. Grandison and M. Sloman, "A survey of trust in internet applications," *IEEE Communications Surveys & Tutorials*, vol. 3, no. 4, pp. 2–16, 2000.

[9] 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*. International Foundation for Autonomous Agents and Multiagent Systems, 2013, pp. 829–836.

[10] 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*, vol. 48, no. 2, pp. 219–235, 2017.

[11] 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*, vol. 8, no. 9, pp. 938–949, 2015.

[12] 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*, vol. 7, no. 10, pp. 881–892, 2014.

[13] 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)*, vol. 48, no. 2, p. 27, 2015.

[14] 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*. IEEE, 2015, pp. 324–331.

[15] 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*, vol. 20, no. 2, pp. 176–194, 2014.

[16] A. Josang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," *Decision Support Systems*, vol. 43, no. 2, pp. 618–644, 2007.

[17] W. Sherchan, S. Nepal, and C. Paris, "A survey of trust in social networks," *ACM Computing Surveys (CSUR)*, vol. 45, no. 4, p. 47, 2013.

[18] 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)*, vol. 1. IEEE, 2006, pp. 17a–17a.

[19] R. Falcone and C. Castelfranchi, "Social trust: A cognitive approach," in *Trust and deception in virtual societies*. Springer, 2001, pp. 55–90.

[20] 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*. IEEE, 2000, pp. 10–pp.

[21] D. H. McKnight and N. L. Chervany, "The meanings of trust," 1996.

[22] R. Falcone, A. Sapienza, and C. Castelfranchi, "The relevance of categories for trusting information sources," *ACM Transactions on Internet Technology (TOIT)*, vol. 15, no. 4, p. 13, 2015.

[23] 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)*, vol. 4, no. 2, p. 27, 2013.

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

[25] 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*. International Foundation for Autonomous Agents and Multiagent Systems, 2010, pp. 209–216.

[26] L. Amgoud and R. Demolombe, "An argumentation-based approach for reasoning about trust in information sources," *Argument & Computation*, vol. 5, no. 2-3, pp. 191–215, 2014.

[27] Y. Tang, K. Cai, P. McBurney, E. Sklar, and S. Parsons, "Using argumentation to reason about trust and belief," *Journal of Logic and Computation*, vol. 22, no. 5, pp. 979–1018, 2011.

[28] 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*. International Foundation for Autonomous Agents and Multiagent Systems, 2015, pp. 999–1007.

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

[30] 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*, vol. 5, no. 2-3, pp. 160–190, 2014.

[31] 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*. Springer, 2011, pp. 74–85.

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

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

[34] 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*. IEEE, 2014, pp. 1–8.