International Journal of Approximate Reasoning ••• (••••) •••–•••

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Counter-deception in information fusion ☆

Johan Schubert

Department of Decision Support Systems, Division of Defence and Security, Systems and Technology, Swedish Defence Research Agency, Stockholm, Sweden

ARTICLE INFO

Article history:
Received 7 April 2017
Received in revised form 28 August 2017
Accepted 8 September 2017
Available online xxxx

Keywords:
Deception
Counter-deception
Information fusion
Entropy
Conflict
Belief function
Dempster–Shafer theory

ABSTRACT

In this article, we develop an entropy-based degree of falsity and combine it with a previously developed conflict-based degree of falsity in order to grade all belief functions. The new entropy-based degree of falsity is based on observing changes in entropy that are not consistent with combining only truthful information. With this measure, we can identify deliberately deceptive information and exclude it from the information fusion process. An experiment is performed comparing conflict and entropy measures and their combination. The effectiveness of the combination of the two measures is suggested.

1. Introduction

Managing false and possibly deliberately deceptive information is, in general, an important issue within an information fusion process. If false and deceptive information is not actively managed, it becomes impossible to trust any conclusions that is based on combining information from several different sources without knowing if one is deceptive. Conclusions that are drawn based on a combination of information from all sources may become degraded or false when truthful information is combined with deceptive information that supports untrue possibilities.

We previously developed methods within the theory of belief functions [1–6] to cluster information about several unrelated problems that should be handled separately when the information about different problems can be mixed up [7–11]. When we know that all information concerns only one problem at hand, this method could be used to identify false pieces of information and allow us to calculate a conflict-based degree of falsity for each piece of evidence [12]. These approaches use a function of the conflict [13,14] in Dempster's rule [2] as criterion function.

Smets [15] developed a methodology for managing a special case of deception where a deceiver may observe a truthful report and send the complement of a truthful belief function as deception instead of the truthful report itself. Pichon et al. [16] later developed a correction scheme that generalizes Shafer's discounting rule [4] by taking into account uncertain meta-knowledge regarding the source relevance and truthfulness. This model now subsumes Smets' model. Furthermore, they recently introduced a contextual correction mechanism [17] for [16].

However, the approach taken by Smets is a special case where the deceiver always sends the complement of what is observed from a truthful source. We think that this is not a realistic strategy by the deceiver, as it is easily countered by

☆ This paper is part of the Virtual special issue on Belief Functions: Theory and Applications, Edited by Jirina Vejnarova and Vaclav Kratochvil.

E-mail address: johan.schubert@foi.se.
URL: http://www.foi.se/fusion.

http://dx.doi.org/10.1016/j.ijar.2017.09.003
0888-613X/

the counter-deception technique developed in Smets' approach [15]. Instead, we would allow the deceiver to act in any way it chooses and assume it might want to deceive us by supporting some focal elements of the frame of discernment that are wrong but we already somewhat believe. We think that this might be a more realistic approach.

In this article, we develop an entropy-based measure of degree of falsity [18] based on the change in entropy when truthful belief functions are combined with a deceptive belief function. The aim is that this new approach should be able to manage more generic types of deception than Smets' approach. As we have previously developed a conflict-based measure of degree of falsity [12] we will here combine these two approaches into one method for recognizing and managing deceptive information.

In Section 2, we discuss approaches to analyzing belief functions for their likelihood of being false due to deception. In Section 3, we review a previous approach to grading pieces of evidence for their degree of falsity based on their contribution to the conflict [13,14] received from Dempster's rule [2]. We then develop a new complimentary approach for grading pieces of evidence based on such changes in entropy that are not consistent with adding truthful evidence into the combination of all belief functions (Section 4). In Section 5, we combine the previously developed conflict-based degree of falsity with the new entropy-based degree of falsity into a combined degree of falsity. We use this approach to reason about which pieces of evidence might be false and should be either discounted or eliminated from the combination of information from all sources. In Section 6 we conduct an experiment with different numbers of deceptive belief functions and study the performance of the conflict and entropy approaches, and their combinations. Finally, in Section 7, we present the study's conclusions.

2. Analyzing belief functions

A belief function that is constructed to be deliberately false may be discovered in two different ways. Such a belief function is aimed to change the conclusion when analyzing the combination of all belief functions. Thus, it must be different from truthful belief functions.

One way to find this is by observing the conflict when combining a new belief function with all previous belief functions. For each belief function at hand, we may observe the change in conflict if we remove this particular belief function from the entire set of all available belief functions χ [7,19]. This will either lower the conflict or leave it unchanged. From a change in conflict, we can derive a degree of falsity for the belief function in question and, for example, use that to discount this particular belief function [12]. For an alternative approach using discounting rates, see [20].

A second approach is to observe a change in entropy when receiving a new belief function. If we receive a good belief function about the problem at hand we should assume that it will further reduce both the scattering and the nonspecificity of the basic belief by focusing it on small focal sets containing the ground truth. Thus, the belief of the ground truth will gradually become more believed and the entropy of the combined belief function will approach zero. On the other hand, if we receive a false belief function that incrementally changes the belief function a small step towards a uniform mass function, then the entropy of the combined belief function will increase. A very strong false belief function may swap the preferred order of the focal sets and leave the entropy unchanged or increased.

We will use both of these approaches to identify which belief functions may be deceptive in order to manage or eliminate them completely from the combination. It is important to note that combining truthful information with deceptive information leads to high conflict and entropy, while the reverse is not true. High conflict and entropy can arise through misrepresentation, mixed up belief function from failed clustering, or measurement errors, etc. A prerequisite for this approach is that the number of deceptive belief functions is less than the number of true belief function, otherwise we will eliminate the truth. If so, we can observe an initial upturn followed by a fall in the conflict as more deceptive information is included.

3. Conflict-based degree of falsity

We interpret the conflict received when combining a set of basic belief assignments (bbas) χ, as if there is at least one bba in χ that violates the representation of the frame of discernment Ω. Such a bba is interpreted as if it does *not* belong to the evidence that refer to the problem at hand [19] described by Ω. Instead, it should be removed from χ.

A conflict when combining all bbas in χ may thus be interpreted as a piece of evidence on a metalevel stating that at least one bba that is placed in χ does not actually belong to χ. On the metalevel, we reason only about the inclusion of bbas in χ, the frame of discernment is $\Theta = \{\text{Adp}, \neg\text{Adp}\}$ where AdP is short for χ being an *adequate partition* [19] of all bbas (i.e., with all bbas in χ), which means we can have metalevel evidence that the partition is either adequate or not. This can be reformulated to $\Theta = \left\{\forall j.e_j \in \chi,\ \exists j.e_j \notin \chi\right\}$, where AdP is refined to the first element of Θ, and $\neg$Adp is refined to the second element. This indicates that we can have evidence that all bbas belong to χ or at least one bba currently in χ does not. In addition, it is possible to refine the frame on the metalevel as $\Theta = \left\{\forall j.e_j \in \chi, \left\{e_q \notin \chi\right\}_q\right\}$, but we will use the first formulation.

We represent the conflict as,

$$\begin{aligned} m_\chi\left(\exists j.e_j \notin \chi\right) &= c_0, \\ m_\chi\left(\Theta\right) &= 1 - c_0, \end{aligned} \tag{1}$$

Please cite this article in press as: J. Schubert, Counter-deception in information fusion, Int. J. Approx. Reason. (2017), http://dx.doi.org/10.1016/j.ijar.2017.09.003