A Reliability Discounting Strategy Based on Plausibility Function of Evidence

Cheng Zeng  
School of Information Engineering  
Hebei University of Technology  
Tianjin, China  
zch@hebut.edu.cn

Pinghui Wu  
School of Information Engineering  
Hebei University of Technology  
Tianjin, China  
wph@hebut.edu.cn

Abstract - Evidence gathered from different sources may have different reliabilities. Such reliability should be integrated into corresponding evidence model to make the evidence combination result rational. In this paper, a novel discounting strategy is developed for the integration of evidence’s model and reliability. Dissimilar to the current one based on BPA, this strategy discounts the evidence’s plausibility function with power rather than discounts BPA function with coefficient. It can alter the evidence’s belief mass as well as its core simultaneously and hereby is more suitable for the discounting of evidence with unreliable core than the one based on BPA function. Furthermore, this new strategy is in some sense a linear discounting strategy which reduces the maximum possible support degree for each atomic proposition with an amount proportional to discount percent. Finally, two examples are presented to demonstrate the characteristics of the proposed discounting strategy.

Keywords: Dempster-Shafer theory, discounting strategy, reliability, plausibility function, uncertainty measure, open frame of discernment.

1 Introduction

As a promising approach to reason with uncertain or imperfect information, the Dempster-Shafer (D-S) theory of evidence has received much attention in many application areas such as data fusion [1, 2], pattern recognition[3] and data mining[4-6] etc. In the D-S theory, the inference is performed by aggregating independent evidences with the Dempster’s rule of combination. These evidences are gathered from different sources or by different ways and modeled by the belief function or its corresponding basic probability assignment (BPA) function. When combined, all of them are treated equivalently by the Dempster’s rule. Such a treatment is theoretically reasonable and feasible for ordinary cases. However, in some applications, the evidences gathered by the aggregating center may be nonequivalent since their sources may not have the same reliability or importance. If these evidences are combined by Dempster’s rule directly, the combination result may deviate from the true state and subsequently lead to improper decisions.

One of the effective ways to combined evidences with different reliabilities is to discount each evidence at a percent according to its reliability before combining. Its essence is to integrate evidence’s reliability into evidence BPA or belief function model. A widely accepted discounting strategy (we call it BPA discounting strategy in this paper) is defined as follows. Let the \( m(\cdot) \), defined on a frame of discernment \( \Theta \), be the BPA function of an evidence originating from source \( S \), if we discount \( m(\cdot) \) at a percent \( 1 - \alpha \in [0,1] \), then the resulting BPA function \( m'(\cdot) \) will be [7]

\[
\begin{align*}
  m'(A) &= \alpha m(A), \quad \forall A \subset \Theta, \\
  m'(\Theta) &= \alpha m(\Theta) + 1 - \alpha.
\end{align*}
\]

(1)

where \( \alpha \in [0,1] \) represents the confidence degree one has about the source \( S \). In addition to a means to discount unreliable evidence, the Eq. (1), as the only known evidence discounting strategy so far, is also adopted as a method to manage the conflicts between combined evidences [8], or as a decorrelation way in combining dependent evidences [9].

Nevertheless, among the two aspects of an evidence model: focal elements and the BPAs assigned to them, such a BPA discounting strategy mainly focuses on the latter while cares little about the former. For an unreliable evidence, this strategy discounts its BPA values but usually keeps its core (the set of all focal elements) unchanged (except for the core of dogmatic evidence to which a new focal element \( \Theta \) will be added after discounting). In other words, the core of an unreliable evidence is treated as reliable by the BPA discounting strategy. Therefore, such a strategy may not be suitable for discounting evidences with inexact cores.

The main objective of this paper is to develop a new discounting strategy which can discount evidence which is unreliable both in its core and its BPA values.
discounting these two kinds of uncertainty simultaneously. Then in section 3, a new plausibility discounting strategy which meets such requirements is proposed and its characteristics are discussed. Section 4 contrasts the proposed plausibility discounting strategy with the BPA one through two examples. Finally, some conclusions are drawn in section 5.

2 Two kinds of uncertainty

In the D-S theory, the evidence model is defined on the power set \(2^\Theta\) of the frame of discernment \(\Theta\). In general, the information attached to an evidence has two aspects: the form of its core and the belief mass assigned to each focal element of the core. Accordingly, an evidence will include two kinds of uncertainty. The first one is the BPA distribution among evidence’s focal elements. Such an uncertainty is somewhat similar to the uncertainty characterized by the probability distribution, so it can be ranged to the category of Hartley uncertainty and measured by an uncertainty measure similar to the Shannon entropy

\[
S = - \sum_{A \subseteq \Theta} m(A) \log_2 m(A).
\]

The second one is the uncertainty about the number and structure of its focal elements. This kind of uncertainty belongs to the category of Hartley uncertainty and is usually measured by a so-called measure of nonspecificity (or nonspecificity for short) defined as [10]

\[
N = \sum_{A \subseteq \Theta, A \neq \emptyset} m(A) \log_2 |A|,
\]

where \(|A|\) is the cardinals of set \(A\). However, the nonspecificity is in fact a weighted average of Hartley measure of all focal elements weighted by their BPA values, so it also characterizes the evidence’s BPA distribution (i.e. Shannon uncertainty) to some extent. But it is invalid when the BPA distribution is a probability distribution since in such a case the nonspecificity is always zero.

Neither the Shannon entropy nor the nonspecificity can characterize the uncertainty of an evidence completely. Each of them emphasizes particularly on one aspect of the evidence’s total uncertainty. In order to measure the total uncertainty of an evidence, several other measures have been proposed and they are discussed in detail in [11]. Nevertheless, how to measure the evidence’s uncertainty properly is still an open question to be discussed. None of proposed measures can meet all the requirements that an ideal evidence uncertainty measure should satisfy. Thus it is difficult to select a proper measure as the guideline for the discussion of evidence discounting or for the developing of new discounting strategies. In this paper, we adopt the specificity as a referential measure to evaluate the variation of evidence’s uncertainty before and after discounting. It should be noticed that such an evaluation is only a rough description without rigorousness.

If an evidence is unreliable, it may have more uncertainty than what it exhibits currently. The more uncertain an evidence is, the larger its uncertainty measure should be. Therefore, when an evidence \(E\) is discounted, let \(N\) be its nonspecificity before discounting and \(N'\) the one after discounting respectively, a reasonable requirement is

\[
N \leq N'
\]

where the equation holds only when \(N = N_{\text{max}}\), i.e. \(E\) is a vacuous evidence (\(m(\Theta) = 1\)).

When we discount an unreliable evidence (it is assumed to be non-dogmatic firstly, i.e. the \(\Theta\) is the evidence’s focal element) with the BPA discounting strategy denoted by the Eq. (1), the result does satisfy the Ineq. (4). However, the increment of the nonspecificity results only from the variation of the BPA values \(m(\cdot)\), and the evidence’s core is not altered after all discounting, i.e. the BPAs are distributed among the same focal elements before and after discounting. In other words, with the BPA discounting strategy, the Shannon uncertainty of the unreliable evidence is discounted but its Hartley uncertainty remains unchanged. If the unreliable evidence is dogmatic, then after discounting, the only alteration of its core is the appending of a new focal element \(\Theta\).

Nevertheless, that an evidence is unreliable may imply that both its core and BPA values are probably inexact. Therefore, only discounting its BPA values is not enough to deliver the total unreliability. A strategy which can also discount its core (i.e. alter its focal element configuration) may be preferable, especially when the evidence’s focal elements are subjectively determined by someone. In such a case, we usually focus mainly on few noticeable propositions while some unnoticeable but probably important propositions may be neglected, which will lead to a deficient evidence model. Under such circumstances, if we discount this evidence with a strategy which is capable of altering evidence’s core, the deficiency of the evidence model can be complemented to a certain extent.

3 Plausibility discounting strategy for unreliable evidence

Wherever it comes from, almost none information can be considered as absolutely reliable. Same information with different reliabilities ought to act differently in approximate reasoning. Thus, the information reliabilities of different evidences should be integrated into their uncertainty description models. However, how to measure the reliability of information is still an open question to be discussed. Sometimes maybe each piece of information of one evidence has its own reliability which will further complicate the measuring of the reliability of the whole evidence. A comparatively thorough discussion about evidence’s reliability measure will be presented in another paper. Here we roughly treat it as a whole and measure it by a parameter \(\alpha \in [0,1]\). The more unreliable the evidence is, the smaller the \(\alpha\) will be. \(\alpha=1\) means that the evidence is absolutely reliable while \(\alpha=0\) completely unreliable.
3.1 Reliability discounting strategy

In order to develop a new discounting strategy that is capable of discounting Hartley uncertainty and Shannon uncertainty simultaneously, instead of discounting the BPAs with a percent $1 - \alpha$, we consider its plausibility functions with a power $\alpha$. Let $P(\cdot)$ be the plausibility function of the original evidence, then the plausibility function of the evidence discounted at $\alpha$ is defined as

$$P'(A) = [P(A)]^\alpha, \quad \forall A \subseteq \Theta \quad \text{and} \quad A \neq \emptyset. \quad (5)$$

It is clear that $\forall A \in \Theta$, $P'(A) \geq P(A)$ since $1 \geq P(A) \geq 0$ and $1 \geq \alpha \geq 0$. It also can be justified that the $P'(\cdot)$ is still a plausibility function measure which satisfies [12]

1. $P'(\emptyset) = 0$,
2. $P'(\Theta) = 1$,
3. $P'\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \sum_{j|i \subseteq \Theta} (-1)^{j-i} P'(\bigcup_{j=i}^{n} A_j)$, \quad $\forall A_i(i = 1, 2, \ldots, n) \subseteq \Theta$.

Hereby, the BPAs after discounting can be obtained by performing an inverse Möbius transformation from $P'(\cdot)$:

$$m'(A) = \sum_{B \subseteq A} (-1)^{|B|} [1 - P'(\overline{B})], \quad \forall A \subseteq \Theta. \quad (6)$$

3.2 Properties

It is not difficult to see that after discounting both the BPAs and the core of the evidence are usually different from the ones before discounting, i.e. such a strategy can discount the evidence’s Hartley uncertainty and Shannon uncertainty simultaneously. But this does not necessarily mean that the resulting evidence will have more focal elements than the original one. As a counterexample, if we discount any evidence with $\alpha = 0$, then the result is a vacuous evidence which represents a completely unknown state. The only focal element of its BPA function is $\Theta$, i.e., $m(\Theta) = 1$. However, just like the BPA discounting strategy, the evidence’s nonspecificity measure after plausibility discounting does not get less than the one before discounting (i.e. satisfying the Ineq. (4)).

Another advantage of the plausibility function discounting strategy is that it discounts evidence in a linear manner. Such linearity can be interpreted by examining the evidence’s potential function defined on an open frame of discernment [13].

Suppose we can merely know part of the atomic propositions in a discourse universe. In order to construct a complete frame of discernment $\Theta$, a $\overline{\Theta}$ which denotes the set of all unknown atomic propositions can be introduced as an atomic set of $\Theta$. Such a $\Theta$ in which $\overline{\Theta}$ is included is called an open frame of discernment. Based on the open frame, a generalized D-S theory named as open frame D-S theory (OFDST) was proposed in [13,14]. In the OFDST, the potential function is defined as

$$W(A) = \lim_{a \to P_1(A)} \log \frac{a}{b}, \quad \forall A \in \Theta. \quad (7)$$

Here right limits are adopted to make sense the logarithmic operation when $P(A) = 0$ and to make the potential function applicable to closed-frame evidence (i.e. $P(\overline{\Theta}) = 0$). Similar to plausibility function, $W(A)$ measures the maximum possible support degree an evidence may have for the atomic proposition $A$, while $W(A) < 0$ means that the evidence rejects $A$ in fact.

One good property of the potential function is that when we combine evidences with Dempster’s rule, their potential functions add up.

**Lemma 1**: Let $W_i(i = 1, 2, \ldots, N)$ be the potential functions of $N$ independent evidences, these evidences be combined with Dempster’s rule of combination and $W_z(\cdot)$ be the potential function of combination result, then

$$W_z(A) = \sum_{i=1}^{N} W_i(A), \quad \forall A \in \Theta. \quad (8)$$

**Proof**: Since Dempster’s rule of combination is associative and commutative, $N=2$ can be assumed here for simplicity.

Suppose that the BPA functions of these two combined evidences are $m_1(\cdot)$ and $m_2(\cdot)$ respectively. Then for $\forall A \subseteq \Theta$,

$$P_{\ell_2}(A) = \sum_{B \subseteq A, B \subseteq A} m_2(B) = \sum_{B \subseteq A, B \subseteq A} m_2(B)$$

$$= \frac{1}{K} \sum_{X \subseteq \Theta, X \subseteq A} \sum_{Y \subseteq \Theta, Y \subseteq A} m_1(X) m_2(Y)$$

$$= \frac{1}{K} \sum_{X \subseteq \Theta, X \subseteq A} m_1(X) \sum_{Y \subseteq \Theta, Y \subseteq A} m_2(Y) = \frac{1}{K} P_{\ell_1}(A) P_{\ell_2}(A),$$

where

$$K = 1 - \sum_{X \subseteq \Theta, X \subseteq A \cap Y \subseteq \Theta} m_1(X) m_2(Y)$$

$$= \sum_{X \subseteq \Theta, X \subseteq A \cap Y \subseteq \Theta} m_1(X) m_2(Y)$$

is the normalization coefficient of Dempster’s rule of combination.

Hence,

$$W_z(A) = \lim_{a \to P_{\ell_2}(A)} \log \frac{a}{b} = \lim_{a \to P_{\ell_1}(A)P_{\ell_2}(A)} \log \frac{a}{b}$$

$$= \lim_{a \to P_{\ell_1}(A)P_{\ell_2}(A)P_{\ell_3}(A)} \log \frac{a}{b}$$

$$= \lim_{a \to P_{\ell_1}(A)P_{\ell_2}(A)P_{\ell_3}(A)} \log \frac{a}{b}$$

$$= W_1(A) + W_2(A),$$

$$W_3(A) = W_1(A) + W_2(A).$$
i.e. for any finite $N$,
$$W_A = \sum_{i=1}^{N} W_i(A), \quad \forall A \in \Theta.$$ 

On the other hand, if we discount any evidence at a percent $\alpha$, the potential function of each atomic set will decrease in proportion to $\alpha$. 

**Lemma 2:** Let $W(A)$ be the evidence’s potential function before and $W'(A)$ the one after plausibility discounting, respectively, then
$$W'(A) = \lim_{a \to (A/A)^*} \log \frac{\alpha}{b} = \alpha W(A), \quad \forall A \in \Theta. \tag{10}$$

**Proof:** According to the definitions of plausibility discounting and potential function (i.e. equation 5 and equation 7), for any $A \subseteq \Theta$,
$$W'(A) = \lim_{a \to (A/A)^*} \log \frac{\alpha}{b} = \lim_{a \to (A/A)^*} \log \frac{\alpha}{b}$$
$$= \lim_{a \to (A/A)^*} \log \frac{\alpha}{b} = \alpha \cdot \lim_{a \to (A/A)^*} \log \frac{\alpha}{b}$$
$$= \alpha W(A).$$

Based on the lemma 1 and lemma 2, following theorem can be directly derived.

**Theorem 1:** Let $W_i(A)$ and $W_i'(A)$, $i = 1, 2, \ldots, N$ be $N$ independent evidences’ potential functions before and after plausibility discounting respectively, and $W'_N(A)$ be the potential function of combination result of discounted evidences, then
$$W'_N(A) = \sum_{i=1}^{N} W'_i(A) = \sum_{i=1}^{N} \alpha W_i(A), \quad \forall A \in \Theta. \tag{11}$$

where $\alpha_i$ is the reliability of $i$th evidence.

This theorem implies that the potential function of combination result of discounted evidences is a linear sum of the potential functions of undiscounted evidences. The coefficients of this linear combination are the reliabilities of combined evidences. It is obvious that if all of the $N$ combined evidences have same reliability $\alpha$, then after combination, the total potential function will also decrease in proportion to $\alpha$, i.e.
$$W'_N(A) = \sum_{i=1}^{N} W'_i(A) = \alpha \sum_{i=1}^{N} W_i(A), \quad \forall A \in \Theta. \tag{12}$$

Therefore, from the viewpoint of potential function, the plausibility function discounting strategy proposed here can discount evidence linearly and thus affect the combination result in a linear manner. Such a linearity makes the discounting process more intuitive and is helpful in applications to determine the discount percent $\alpha$. However, the BPA discounting strategy does not possess such a property, for the same $\alpha$, its discounting percent of potential function will vary with atomic set and differ from one evidence to another.

One potential disadvantage of the plausibility function discounting strategy is that there are usually much more focal elements in the evidence’s core after discounting than the one before discounting. This will increase the computational complexity of evidence discounting or evidence combination notably. In general, the computational loads can be reduced somewhat with the fast Möbius transformation (FMT) [15]. However, such a method will become impractical when $\Theta$ has more than 15 to 20 atomic sets [16].

### 4 Examples

In this section, two evidences, one is non-dogmatic evidence and the other is dogmatic evidence, are adapted to illustrate the characteristics of the proposed plausibility discounting strategy and the differences between it and the BPA one. These two evidences are separately discussed firstly and then they are combined to demonstrate the linearity of proposed discounting strategy. Additionally, open frame evidences are adopted here to facilitate the comparison of evidences’ potential functions before and after discounting. It should be pointed out that the plausibility discounting strategy is also suitable for closed-frame evidences.

#### 4.1 Non-dogmatic evidence

Suppose an unreliable evidence defined on an open frame $\Theta = \{a, b, c, \bar{a}\}$ has following BPA distributions:
$$m_1(ab) = 0.5, \quad m_1(c) = 0.25, \quad m_1(\Theta) = 0.25.$$ 

Such an evidence is non-dogmatic since $m_1(\Theta) \neq 0$. Its nonspecificity $N_1 = 1$. This evidence is discounted with $\alpha = 0.8$ by different strategies. The respective resulting BPAs and nonspecificity measures are as follows.

**BPA discounting strategy:**
$$m'_{1}(ab) \approx 0.4257, \quad m'_{1}(c) \approx 0.2056, \quad m'_{1}(\Theta) \approx 0.3299; \quad N'_{1} \approx 1.147.$$ 

From these results we can see that both discounting strategies alter the evidence’s BPA values and nonspecificity measures in similar manners and $N'_{1} > N_1$ holds for each. However, a new focal element $abc$ is added to the resulting core of plausibility discounting (i.e. the evidence’s Hartley uncertainty is altered), while in the result of BPA discounting, the evidence’s core remains unchanged.

The potential functions before and after discounting are shown in table 1. The plausibility function of $\bar{a}$ is always zero and it is omitted here. In the plausibility discounting, an identical potential function discount percent $1 - \alpha_i = 0.2$ holds for any atomic set $A$ (i.e.}
\[ W'_i(A) = \alpha W_i(A) = 0.8 W_i(A), \ \forall A \in \Theta \]; while in the BPA discounting, the potential function discount percent usually varies with the atomic set (e.g. \( W'_i(a)/W_i(a) \approx 0.63 \) but \( W'_i(c)/W_i(c) \approx 0.59 \)), and obviously it will also differ from one evidence to another. Therefore, from the viewpoint of potential function, the plausibility discounting is a linear discounting strategy whereas the BPA one is not.

### 4.2 Dogmatic evidence

Suppose an unreliable dogmatic evidence defined on an open frame \( \Theta = \{a, b, c, \overline{\Delta}\} \) has following BPA distributions:

\[
\begin{align*}
    m_2(a) &= 0.4, \quad m_2(b) = 0.25, \\
    m_2(c) &= 0.25, \quad m_2(\overline{\Delta}) = 0.1.
\end{align*}
\]

Its nonspecificity \( N_2 = 0 \). When it is discounted by \( \alpha_2 = 0.6 \) with different strategies, the resulting evidences’ BPAs and nonspecificity measures will vary in a manner similar to the non-dogmatic evidence. BPA discounting strategy:

\[
\begin{align*}
    m_2'(a) &= 0.24, \quad m_2'(b) = 0.15, \quad m_2'(c) = 0.15, \\
    m_2'(\overline{\Delta}) &= 0.06, \quad m_2'(\Theta) = 0.4; \quad N'_2 = 0.8.
\end{align*}
\]

Plausibility discounting strategy:

\[
\begin{align*}
    m_2'(a) \approx 0.2640, \quad m_2'(b) \approx 0.1585, \quad m_2'(c) \approx 0.1585 \\
    m_2'(\overline{\Delta}) \approx 0.0613, \quad m_2'(ab) \approx 0.0448, \quad m_2'(ac) \approx 0.0448 \\
    m_2'(a\overline{\Delta}) \approx 0.0150, \quad m_2'(b\overline{\Delta}) \approx 0.0232, \\
    m_2'(b\Delta) \approx 0.0080, \quad m_2'(c\overline{\Delta}) \approx 0.0080, \\
    m_2'(abc) \approx 0.0549, \quad m_2'(ab\overline{\Delta}) \approx 0.0131, \\
    m_2'(ac\overline{\Delta}) \approx 0.0131, \quad m_2'(bc\overline{\Delta}) \approx 0.0055, \\
    m_2'(\Theta) \approx 0.1272; \quad N'_2 \approx 0.5355.
\end{align*}
\]

Since the evidence before discounting is dogmatic here, both discounting strategies can alter the evidence’s core and BPA values simultaneously. However, for the BPA discounting, only one new focal element \( \Theta \) is appended to the resulting core. While in the core after plausibility discounting, besides the \( \Theta \), many other new focal elements are also included. Such a notable increment in the size of the core may characterize the evidence’s unreliability about its Hartley uncertainty more appropriately. But it will also bring much more difficulty on the implementation of evidence discounting and evidence combination.

<table>
<thead>
<tr>
<th>Atomic set</th>
<th>Before discounting</th>
<th>BPA discounting</th>
<th>Plausibility discounting</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0.4771</td>
<td>0.3010</td>
<td>0.3817</td>
</tr>
<tr>
<td>b</td>
<td>0.4771</td>
<td>0.3010</td>
<td>0.3817</td>
</tr>
<tr>
<td>c</td>
<td>0.3010</td>
<td>0.1761</td>
<td>0.2408</td>
</tr>
</tbody>
</table>

### 4.3 Combination result of two evidences

Two evidences employed before are combined here to demonstrate the linearity of plausibility discounting strategy further. Table 3 lists the potential functions of combination results before and after discounting.

<table>
<thead>
<tr>
<th>Atomic set</th>
<th>Before discounting</th>
<th>BPA discounting</th>
<th>Plausibility discounting</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>1.0792</td>
<td>0.4444</td>
<td>0.7429</td>
</tr>
<tr>
<td>b</td>
<td>0.8750</td>
<td>0.3786</td>
<td>0.6205</td>
</tr>
<tr>
<td>c</td>
<td>0.6989</td>
<td>0.2537</td>
<td>0.4796</td>
</tr>
</tbody>
</table>

Based on the potential function listed in table 1, 2 and 3, we can see that all of following three equations hold:

\[
\begin{align*}
    W_{12}(A) &= W_1(A) + W_2(A), \quad \forall A \in \Theta, \\
    W'_{12}(A) &= W'_{1}(A) + W'_{2}(A), \quad \forall A \in \Theta, \\
    W''_{12}(A) &= W''_{1}(A) + W''_{2}(A), \quad \forall A \in \Theta.
\end{align*}
\]

Moreover, for the plausibility discounting,

\[
W''_{12}(A) = \alpha W'_1(A) + \alpha_2 W'_2(A), \quad \forall A \in \Theta.
\]

But such a linear relation between potential functions before and after discounting does not hold for BPA discounting strategy, i.e. \( W''_{12}(A) = \alpha W_1(A) + \alpha_2 W_2(A) \) is usually not true.

### 5 Conclusion

The currently widely adopted BPA discounting strategy for unreliable evidence focuses mainly on the discounting of evidence’s Shannon uncertainty but treats its core (i.e. the Hartley uncertainty) as reliable in general. In this paper, through discounting evidence’s plausibility function instead of BPA function, a new discounting strategy which discounts Shannon uncertainty and Hartley uncertainty simultaneously is proposed. Such a
plausibility discounting strategy is more suitable for evidence with unreliable core. It also discounts evidence's potential function linearly and consequently affects evidence combination results in a linear manner. However, such a discounting strategy may become computationally impractical when the cardinal number of the frame of discernment is too large.

References


