Air Defense Threat Evaluation using Fuzzy Bayesian Classifier

Wei Mei
Dept. of Electronic Engr., Shijiazhuang Mech. Engr. College, 97 Hepingxilu Rd., Shijiazhuang, P.R. China

Keywords: Fuzzy Systems, Bayesian, Conditional Probability, Likelihood, Threat Evaluation.

Abstract: The connection between probability and fuzzy sets has been investigated among the community of approximate reasoning for decades. A typical viewpoint is that the grade of membership could be interpreted as a conditional probability. This note extend this viewpoint a step further by introducing the concepts of conditional probability mass function (CPMF) and the likelihood mass function (LMF). We draw the conclusion that conditional probability can be used for describing either randomness or fuzziness depending on how it is interpreted. If expanded to CPMF, then it can be used for modelling randomness; if expanded to LMF, then it can be a useful expression for modelling fuzziness. A fuzzy Bayesian theorem is derived based on the fuzziness interpretation of conditional probability. Its successful application to an example of target recognition demonstrates that the proposed fuzzy Bayesian theorem provides alternative approach for handling uncertainty.

1 INTRODUCTION

The operation of air defense is a time critical process, which includes a series of automation/semi-automation steps of information fusion and the final step of engagement. The process of information fusion may range from target tracking, target recognition, through to situation awareness, threat evaluation (TE) and weapon assignment. According to Paradis (Paradis et al., 2005), TE refers to "the part of threat analysis concerned with the ongoing process of determining if an entity intends to inflict evil, injury, or damage to the defending forces and its interests, along with the ranking of such entities according to the level of threat they pose." The difficulties of developing a TE system are evident, largely due to three factors (Roux and Vuuren, 2007; Steinberg, 2005): 1) Weak spatio-temporal constraints on relevant evidence. Many TE problems may involve evidence that is wide-spread in space and time, with no easily defined constraints. 2) Weak ontological constraints on relevant evidence. Evidence relevant to TE may be very diverse and may contribute to inferences in unexpected ways. 3) Weakly-modeled causality. TE involves inference of human intent and behavior. Models are extremely difficult to formulate, since sub-domains (individual minds) are unique and attributes may be very difficult to measure or even define.

The fundamental problem involved in information fusion of air defense is the need to deal with uncertainty. In the words of Von Clausewitz (Clausewitz et al., 2004), "war is the realm of uncertainty; three quarters of the factors on which action in war is based are wrapped in a fog of greater or lesser uncertainty. A sensitive and discriminating judgment is called for a skilled intelligence to scent out the truth." Up to now, a huge number of methods, such as Bayesian inference (Chen and Ho, 2008; Lane et al., 2010), fuzzy sets (Bailadora and Triviño, 2010; Xu et al., 2012), neural networks (Jan, 2004; Young et al., 1997), and evidential reasoning (Delmotte and Smets, 2004; Leung and Wu, 2000), have been promoted for handling uncertainty arising from applications of information fusion including TE. Though there are a variety of approaches as listed above for uncertainty inference, in our opinion the uncertainty involved in and of itself can be broadly categorized into (or interpreted as) two types, randomness and fuzziness. Randomness is usually measured by probability whereas fuzziness is often gauged by membership or possibility. It is worth noting that there is ongoing endeavor of connecting probability and possibility. Some works intend to unify them or interpret one uncertainty by another one (Cheeseman, 1988; Coletti and Scozzafava, 2004; Dubois et al., 1997), some works try to find out the relationship for probability-possibility transformation (Oussalah, 2000; Dubois et al., 2004; Mouchaweh and Billaudel, 2006).

Mei W..
Air Defense Threat Evaluation using Fuzzy Bayesian Classifier.
DOI: 10.5220/0004512602270232
In *Proceedings of the 5th International Joint Conference on Computational Intelligence* (FCTA-2013), pages 227-232
ISBN: 978-989-8565-77-8

This paper discusses the Bayes based TE method. Bayesian inference is an approach to statistics in which all forms of uncertainty are expressed in terms of probability. It has a large body of applications and is believed to be the most classic, rigorous and popular method for modeling uncertainty (Jain et al., 2000). Nevertheless, Bayesian method has always been criticized for lack of prior probability and being difficult to define the conditional probability. From the viewpoint of application, e.g. target recognition and TE, it is usually very inconvenient to build and maintain the knowledge database of the inference rule in form of conditional probability. Practitioners complain that whenever new inference rule is to be added to the knowledge database, all former defined inference rules have to be redefined to ensure the sum of corresponding conditional probabilities maintains one. We in this paper try to eliminate this problem by reinterpretation of the Bayes theorem, which can handle randomness and fuzziness simultaneously, and leads to an open structure of knowledge database for uncertainty inference.

The rest of this paper is organized as follows. Section 2 presents two interpretations of conditional probability, which are suitable for describing randomness and fuzziness, respectively. Section 3 revisits the well known Bayesian theorem by applying these two interpretations of conditional probability and derives two forms of Bayesian theorem, the usual one and the fuzzy Bayesian theorem. Section 4 proposes a probability-possibility conversion method through the bridge of Bayesian theorem but with specific interpretations of conditional probability. Section 5 introduce the application of the fuzzy Bayesian theorem to the problem of TE. Section 6 concludes the paper.

2 TWO INTERPRETATIONS OF CONDITIONAL PROBABILITY

The Bayesian theorem is a well-known mechanism for relating two conditional probabilities. This section gives two interpretations of conditional probability, based on which the Bayesian theorem can be reinterpreted as in the next section. Probability originally comes with randomness while possibility comes with fuzziness. Randomness is the uncertainty whether an event occurs, or the possible outcomes an event variable may take. Sometimes, the event itself is certain and you may be uncertain about it because of your lack of information of it. Fuzziness is the uncertainty whether a concept holds given its attribute values.

The chief similarity between probability and possibility is that both methods describe uncertainty with numbers in the unit interval [0, 1]. The key distinction concerns how they deal simultaneously with the outcome and its opposite of an event variable. Probability demands the sum of all possible outcomes of an event variable is one. Possibility has no additivity constraint as probability. Mathematically, a possibility on the finite set A is a mapping to [0, 1] such that

$$\pi(\phi) = 0 \tag{1}$$

$$\pi(A) = \mathrm{Max}(\pi(A = a_i)) = 1, \quad i = 1, 2, ..., n \tag{2}$$

where A is called event variable, and $A = a_i$ is one of n possible outcomes of event variable A (in short, event). Without lose of generality, this work only considers the case of discrete event to simplify the discussion. As we can see, possibility is similar to probability, but it relies on an axiom which only involves the operation “maximality” as shown in (2). In contrast, probability is additive which requires that probability sum of all possible outcomes of event variable is one. Though probability origins from randomness or frequency, it has been widely used in various applications for modeling different uncertainty that satisfies the additively constraint of probability. Likewise, possibility has been extensively used for formulating any uncertainty that satisfies (1, 2) besides fuzziness.

Conditional probability $p(A = a_i \mid B = b_j)$ is the occurrence probability of a conditional event $A = a_i \mid B = b_j$, which equals to the probability of $A = a_i$ given $B = b_j$. In order to completely formulate the randomness of the conditional event $A = a_i \mid B = b_j$, we need to use conditional probability mass function (CPMF), $\{p(A = a_i \mid B = b_j), \quad i = 1, 2, ..., m\}$ (in short, $p(A \mid B = b_j)$). Here event variable B is fixed at b_j and m is the number of possible outcomes a_is of event variable A. Now we see CPMF provides a complete description of the stochastics of the event variable A given conditioning event $B = b_j$. According to the property of probability, the sum of $p(A \mid B = b_j)$ across a_i is one. The randomness formulated by CPMF is here called probabilistic randomness.

If event variable A is fixed at a_i and let B take value from n possible outcomes b_j s, we then get the likelihood mass function (LMF), $\{p(A = a_i \mid B = b_j),\ \ j = 1, 2, ..., n\}$ (in short, $p(A = a_i \mid B)$), which is the likelihoods of the fixed a_i stemming from different b_j . Note that though $p(A = a_i \mid B = b_j)$ is a probability, there is no need that the sum of $p(A = a_i \mid B)$ across b_j should be one since it is actually not a probability mass function. Let $A = a_i$ be a fuzzy concept, then $p(A = a_i \mid B)$ naturally defines a membership function

$$\mu_{A=a_i}(B) = p(A = a_i \mid B) \tag{3}$$

As we can see from (3), LMF is a natural form of membership function for describing fuzziness. The only constraints on $p(A = a_i \mid B)$ is that its sum over a_i should be one for every $B = b_j$. The fuzziness formulated by (3) is here called probabilistic fuzziness since it is derived from a conditional probability, and $\mu_{A=a_i}(B)$ is called probabilistic membership function. If we let $\mu_{A=a_i}(B) = \alpha_0 p(A = a_i \mid B)$, where scale factor α_0 is applied so that the maximum value of $\mu_{A=a_i}(B)$ over B is one, then $\mu_{A=a_i}(B)$ is a standard membership function derived from conditional probability.

3 REITERPRETAION OF BAYESIAN THEOREM

The well-known Bayesian theorem is as follow:

$$\begin{aligned} p(A = a_i \mid B = b_j) &= \frac{p(A = a_i)p(B = b_j \mid A = a_i)}{p(B = b_j)} \\ &= \alpha_1 p(A = a_i)p(B = b_j \mid A = a_i) \end{aligned} \tag{4}$$

where $p(A = a_i \mid B = b_j)$, the posterior, is the probability in $A = a_i$ after $B = b_j$ is observed; $p(A = a_i)$ is prior probability; conditional probability $p(B = b_j \mid A = a_i)$, is called the likelihood; and α_1 is a normalizing factor such that the sum of $p(A = a_i \mid B = b_j)$ over a_i is one.

In applications, the likelihood $p(B = b_j \mid A = a_i)$ is usually defined among the space of CPMF, $p(B \mid A = a_i)$, which means $p(B = b_j \mid A = a_i)$ represents randomness. Following the interpretation of conditional probability in Section 2, $p(B = b_j \mid A = a_i)$ can also be used to model fuzziness, only if it is defined among the space of LMF, $p(B = b_j \mid A)$. Let $\mu_{B=b_j}(A) = \alpha_0 p(B = b_j \mid A)$, we then have

$$p(A \mid B = b_j) = \alpha_2 p(A)\mu_{B=b_j}(A) \tag{5}$$

Note that (5) holds for any $\mu_{B=b_j}(A)$ proportional to $p(B = b_j \mid A)$ considering the effect of the normalization constant α_2 . Eq. (5) provides a mechanism to fusion randomness and fuzziness to arrive at a conclusion with uncertainty of randomness, and is called a fuzzy Bayesian theorem. Recall that probability and possibility can be used for modelling any uncertainty only if their specific constraints are satisfied. Mathematically, Eq. (5) could be used to fusion probability and possibility no matter fuzziness is involved or not, but the name of fuzzy Bayesian theorem always holds. The choosing of (4) and (5) for a certain application depends on our interpretation of $p(B = b_j \mid A = a_i)$.

4 PROBABILITY-POSSIBILITY TRANSFORMATIONS

Similarly, following different interpretations of conditional probability, we can derive transformations from possibility to probability and conversely. Let $p(B = b_j \mid A = a_i)$ in (4) be expanded to LMF and represent possibility, i.e., $\pi_{B=b_j}(A) = \alpha_0 p(B = b_j \mid A)$, we get

$$p(A \mid B = b_j) = \alpha_3 p(A)\pi_{B=b_j}(A) \tag{6}$$

where α_3 is a normalizing factor. Eq. (6) can be used for transformation from possibility to probability and is similar to Klir's normalized transformation from possibility to probability (Mouchaweh and Billaudel, 2006). The difference lies that in order to convert a possibility to a more specific probability, (6) suggests that the prior probability $p(A = a_i)$ should be used. Let $p(A = a_i)$ be a uniform distribution, then we have

$$p(A \mid B=b_j) = \alpha_4 \pi_{B=b_j}(A) \tag{7}$$

where α_4 is a normalizing factor. Eq. (7) is exactly the same as Klir's normalized transformation from possibility to probability.

Let $p(B=b_j \mid A=a_i)$ in (4) be expanded to CPMF $p(B \mid A=a_i)$ and represent probability, we get

$$\pi_{A=a_i}(B) = \frac{\alpha_5 p(B \mid A=a_i)}{p(B)} \tag{8}$$

where $\pi_{A=a_i}(B) = \alpha_0 p(A=a_i \mid B)$ is expanded to LMF and represents possibility; scale factor α_5 is such that the maximum value of $\pi_{A=a_i}(B=b_j)$ over b_j is one. Note that $p(A=a_i)$ is removed from (8), which is a constant as A is fixed at a_i, considering the effect of the factor α_5. Eq. (8) can be used for transformation from probability to possibility and is also similar to Klir's normalized transformation from probability to possibility (Mouchaweh and Billaudel, 2006).

5 APPLICATION TO THREAT EVALUATION

Factors considered in assessing target threat under the background of air defense may include target type, heading, velocity, altitude, distance with respect to the related high value defended assets, the detection of emissions from its fire control radar, and the estimation of its possible courses of attack action (Roux and Vuuren, 2007). In addition, peer supplied TE report may be used for own-ship TE update. The TE example introduced in this section considers two factors, i.e., target type and target distance.

Assume a missile approaching the defended assets belongs to two possible types of target, combat aircraft and missile denoted by $C=\{c_1, c_2\}$. Target distance is supposed to be classified as three levels, close (<20km), medium (<100km & >20km), far (e.g., >100km), denoted by $D=\{d_1, d_2, d_3\}$. Let target threat be three levels, low, medium and high, denoted by $T=\{t_1, t_2, t_3\}$. At consequent times k_1, k_2, k_3, the TE system receives target type probability $p(c_i \mid e)$ (*e* is the raw observation) given in Table 1 from a classifier, and target distance data given in Table 2 from a tracker. Note that in Table 2, e.g., at time k_1, $p(d_2|e) = 1$ while $p(d_1|e) = p(d_3|e) = 0$, which is due to the fact that current target distance is medium (d_2).

Table 1: Target type probability.

	k_1	k_2	k_3
$p(c_1\|e)$	0.5	0.2	0.2
$p(c_2\|e)$	0.5	0.8	0.8

Table 2: Target distance.

	k_1	k_2	k_3
Distance	90km (medium)	50km (medium)	18km (close)
$p(d_1\|e)$	0	0	1
$p(d_2\|e)$	1	1	0
$p(d_3\|e)$	0	0	0

The threat level of the approaching missile is evaluated by using a classifier based on the Bayesian theorem or the fuzzy Bayesian theorem. The posterior probability of target threat could be calculated as follows:

$$p(t_i \mid e) = \sum_{c_j}\sum_{d_s} p(t_i \mid c_j, d_s) p(c_j \mid e) p(d_s \mid e) \tag{9}$$

$$p(t_i \mid c_j, d_s) = \alpha p(t_i) p(c_j \mid t_i) p(d_s \mid t_i) \tag{10}$$

where $p(t_i)$ is the prior probability of target threat with assumed uniform distribution; conditional probabilities $p(c_j \mid t_i)$, $p(d_s \mid t_i)$ define the uncertain mapping between the threat category space and the threat factor space; and α is a normalization constant such that values of $p(t_i \mid c_j, d_s)$ over t_i sum to one. Traditionally, $p(c_j \mid t_i)$, $p(d_s \mid t_i)$ are usually defined from the threat category space to the threat factor space as in Table 3, meantime Eqs. (9, 10) is called the Bayesian classifier. Always, practitioners are hesitated in assigning an appropriate value for $p(c_j \mid t_i)$ or $p(d_s \mid t_i)$. It looks somewhat strange, e.g., that a certain level of threat will produce a certain type of target with a certain probability. In contrast, it is more reasonable to say that a certain type of target will exhibit a certain level of threat with a certain possibility. Therefore $p(c_j \mid t_i)$, $p(d_s \mid t_i)$ need to be defined from the threat factor space to the threat category space as in Table 4, meantime a fuzzy Bayesian classifier (9, 11) can be applied with (11) given bellow, where $p(c_j \mid t_i)$, $p(d_s \mid t_i)$ are rewritten as $\mu_{c_j}(t_i)$, $\mu_{d_s}(t_i)$.

$$p(t_i \mid c_j, d_s) = \alpha p(t_i)\mu_{c_j}(t_i)\mu_{d_s}(t_i) \tag{11}$$

Table 3: $p(c_j \mid t_i)$, $p(d_s \mid t_i)$ For Bayesian method.

		t_1	t_2	t_3
$p(c_j \mid t_i)$	c_1	0.70	0.50	0.10
	c_2	0.30	0.50	0.90
$p(d_s \mid t_i)$	d_1	0.10	0.20	0.80
	d_2	0.10	0.50	0.10
	d_3	0.80	0.30	0.10

Table 4: $\mu_{c_j}(t_i)$, $\mu_{d_s}(t_i)$ For fuzzy Bayesian method.

		t_1	t_2	t_3
$\mu_{c_j}(t_i)$	c_1	0.10	1.00	0.50
	c_2	0.10	0.50	1.00
$\mu_{d_s}(t_i)$	d_1	0.00	0.50	1.00
	d_2	0.10	1.00	0.50
	d_3	1.00	0.50	0.00

The results of TE are given in Table 5. As we can see, e.g., at k_1, the TE results of the Bayesian classifier is (0.15, 0.72, 0.13), which means $p(t_1 \mid e)$ =0.15, $p(t_2 \mid e)$ =0.72, $p(t_3 \mid e)$ = 0.13. It is shown to the user in a simpler and more intuitive form as medium (0.7), which means the current threat level is medium with a confidence of 0.7. The overall performances of the two methods are competing, though the fuzzy Bayesian classifier is easier to implement due to the easiness of defining $\mu_{c_j}(t_i)$, $\mu_{d_s}(t_i)$. For example, we need not to make sure the sum of $\mu_{c_j}(t_i)$ over t_i is one when using fuzzy Bayesian classifier, but we need to make sure values of $p(c_j \mid t_i)$ over c_j sum to one when using the conventional Bayesian classifier.

Table 5: Threat Evaluation Results.

$p(t_i \mid e)$	k_1	k_2	k_3
Classifier 1[a]	(0.15, 0.72, 0.13) medium (0.7)	(0.11, 0.69, 0.20) medium (0.7)	(0.09, 0.17, 0.74) high (0.7)
Classifier 2[a]	(0.01, 0.64, 0.35) medium (0.6)	(0.01, 0.56, 0.43) medium (0.6)	(0.00, 0.26, 0.74) high (0.7)

[a] Classifier 1: Bayesian classifier, Classifier 2: fuzzy Bayesian classifier

6 CONCLUSIONS

It is more natural and convenient to model the uncertainty involved in threat evaluation using the so called fuzzy Bayes' Theorem, which has competitive performance with the conventional Bayesian method and the merit of an open structure of rule database.

ACKNOWLEDGEMENTS

This work was supported in part by the Chinese NSFC under Grant 61141009.

REFERENCES

Bailadora, G., Triviño, G., 2010. Pattern recognition using temporal fuzzy automata. *Fuzzy Sets and Systems*, 161, 37–55.

Clausewitz, C. V., Graham, J. J., Honig, J. W., 2004. *On War*, Barnes & Noble Publishing.

Cheeseman, P., 1988. Probabilistic versus fuzzy reasoning. *Uncertainty in Artificial Intelligence*, North-Holland, Amsterdam, Vol. 1, 85-102.

Chen, C. H., Ho, P. P., 2008. Statistical pattern recognition in remote sensing. *Pattern Recognition*, 41(9), 2731-2741.

Coletti, G., Scozzafava, R., 2004. Conditional probability, fuzzy sets, and possibility: a unifying view. *Fuzzy Sets and Systems*, 144, 227–249.

Delmotte, F., Smets, P., 2004. Target identification based on the transferable belief model interpretation of Dempster-Shafer model. *IEEE Trans. On Systems, Man, and Cybernetics - Part A: Systems and Humans*, 34, 457–471.

Dubois, D., Moral, S., Prade, H., 1997. A semantics for possibility theory based on likelihoods. *Journal of Mathematic Analysis and Applications*, 205, 359–380.

Dubois, D., Foulloy, L., etc, 2004. Probability-possibility transformations, Triangular fuzzy sets, and probabilistic inequalities, *Reliable Computing*, 10, 273-297

Jain, A. K., Duin, R. P. W., Mao, J., 2000. Statistical pattern recognition: a review. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, 22, 4–37.

Jan, T., 2004. Neural network based threat assessment for automated visual surveillance. *Proc. of 2004 IEEE International Joint Conference on Neural Networks*, Vol.2, 1309 – 1312.

Lane, R.O., Nevell, D.A., Hayward, S.D., Beaney, T.W., 2010. Maritime anomaly detection and threat assessmen. *Proc. of the 13th Conference on Information Fusion*, QinetiQ, UK, pp1–8.

Leung, H., Wu, J., 2000. Bayesian and Dempster-Shafer target identification for radar surveillance. *IEEE Trans. on Aerospace and Electronic Systems*, 36, 432–447.

Mouchaweh, M. S., Billaudel, P., 2006. Variable Probability-Possibility Transformation for the Diagnosis by Pattern Recognition. *International Journal of Computational Intelligence: Theory and Practice*, 1(1).

Oussalah, M., 2000. On the probability/possibility transformations: a comparative analysis. *Journal of General Systems*, 29(5), 671-718.

Paradis, S., Benaskeur, A., Oxenham, M.G., Cutler, P., 2005. Threat evaluation and weapon allocation in network-centric warfare. *Proc. of the 7th International Conference on Information Fusion*, Stockholm, 1078–1085

Roux, J. N., Vuuren, J. H., 2007. Threat evaluation and weapon assignment decision support: a review of the state of the art. *ORiON*, 23(2), 151–187.

Steinberg, A. N., 2005. An approach to threat assessment. *Proc. of the 8th International Conference on Information Fusion*, Vol. 2, Philadelphia, USA, 1256–1263.

Xu, Y., Wang, Y., Miu, X., 2012. Multi-attribute decision making method for air target threat evaluation based on intuitionistic fuzzy sets. *Journal of Systems Engineering and Electronics*, 23(6), 891–897.

Young, S.S., Scott, P.D., Nasrabadi, N.M., 1997. Object recognition using multilayer Hopfield neural network. *IEEE Trans. On Image processing*, 357–372.