2011 International Conference on Computer and Automation Engineering (ICCAE 2011)
IPCSIT vol. 44 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V44.24

Data Fusion of Distributed D-S Evidence Theory Based on Predicted Reliability

Liu Ping, Zhang Lingxiao and Yang Xinfeng[+]

Computer Science and Technology Department, Nanyang Institute of Technology,Nanyang,473000, China

Abstract. The available evidence on the basis of source credibility factor algorithm to predict the reliability coefficient proposed the concept and give a specific formula; and introduced the training method of the credibility factor balancing factor. Second, the reliability coefficient of the forecast given in the formula with the original DS theory of evidence combination rule to combine, to ensure all the properties of the original synthesis of the rules by adding the prediction based on the reliability coefficient to solve the conflict of evidence. Finally, Simulation proves effectiveness of the algorithm.

Keywords*:* Incomplete information systems; null value estimates; feature reduction; DS evidence theory

1. Introduction

Into over the years, whether it is military, medical, meteorological or geographic areas and other areas are beginning to detect widespread concern data fusion technology. Sensor data fusion technology can be collected from incomplete, uncertain, vague or even conflicting data processing and integration, eliminating redundant data, according to the internal rules for target identification and contact of the formation of a relatively complete and consistent description and interpretation of the perception . Currently existing data fusion algorithm, DS evidence theory can be expressed with a "do not know, uncertain" information to satisfy weaker than the condition of the Bayesian theory of characteristics, has been widely used.

Evidence theory is proposed by Dempster in 1967, Shafer-depth research on this theory and analysis, presented in 1976 for the general law of universal synthesis. Combination rule of DS evidence theory there is evidence of conflict resolution, many scholars have made in this respect a certain contribution. For example: Yager earlier proposed a rule to cancel the regular thought, Smets produced evidence that the main cause of conflict is not perfect due recognition framework is proposed TBM model and the open world thought. In addition, Zhang Bingand other scholars of the evidence to solve the distribution of source credibility on the issue of conflict of evidence also research.

However, the evidence in resolving the conflict has also brought new problems, increases the computational data processing, data fusion seriously reduced the efficiency of the process. According to currently available methods to resolve conflicts of evidence, the majority of methods require multiple sources of information on the data from the pretreatment, treatment and re-allocation or conflicting evidence or evidence obtained by calculating the source's reliability coefficient of each further integration calculation. The pretreatment process must in all data processing center only after arrival and can only proceed after completion of the pre-data fusion. Communicate this to the fusion center and the fusion center to open sales of computation and storage capacity required will correspondingly increase, etc., and the processor utilization of the fusion center is also greatly reduced. There are many areas, especially in the military field of real-time data fusion, integration of both speed and accuracy of higher demand. Given that military intelligence information gathered from huge, broad geographic distribution of the source of the evidence, the evidence on the

[+] Corresponding author. Tel.: +86-377-62076332
E-mail address: ywind2005@163.com

relationship between the complex and timeliness of data fusion for better circumstances, to build a distributed data fusion model for specific The applications will play a very significant effect. Not only able to adapt to large amounts of data processing and ensure accuracy in the premise of integration greatly enhanced the speed, the utilization of the processing center has been the overall improvement. Distributed fusion system to the high reliability, suitability, ability to survive in the modern war has been widely used.

2. The Calculation of Reliability Coefficients

For example, weather changes, and changes in the state of the sensor dynamic environment coupled with the data fusion in real time requirement of very high field, able to ensure the accuracy of the data fusion result reached under the premise of timely data fusion the field of data fusion is currently being studied by many scholars question.

In [1] presents a method of calculating reliability coefficients (referred to as HU Chang method), the method is Philippe Smets in [2] proposed pignistic probability function and Chen Lei and other in [3] proposed The concept of entropy in information theory formed the basis of a reliability coefficient to determine the method. The credibility of its calculation method is as follows:

(1) First, the probability function to calculate the evidence by pignistic set E, any two sources of evidence conflict between Ei and Ej level. Formula is

$$difBet_i^j = \max_{A\subseteq\Theta}(|BetP_{m_i}(A) - BetP_{m_j}(A)|)$$

Where $difBet_i^j$ represents the extent of conflict between Ei and Ej that evidence of the source; $BetP_{m_i}(A)$ that probability assignment function m_i of the probability of identifying the full value of the object A.

(2) The conflict between the evidence source vector be normalized $\overline{difBet_i}$.

(3) Through the normalized entropy vectors after the value of conflict

(4) Entropy reliability coefficient $E_{n_i}^{-1}$ obtained after taking the countdown

The reliability coefficient method can effectively deal with the evidence conflicts, the integration process of convergence speed has been improved, but the algorithm is based on data from all data sources are to reach the case of data fusion center can be calculated This method of real-time data fusion require a higher speed requirements and the environment was not very fast and flexible and efficient. Therefore, this article on the basis of the method proposed a new method of calculating reliability coefficients, the method used by recursively all the evidence over the past data to predict the source of the credibility factor, calculated by the confidence method of distributed data fusion, we can improve the efficiency of data fusion and real-time also improved[4-7].

In order to achieve integration of distributed evidence needs to be calculated in advance based on historical data the reliability coefficient of each source of evidence, due to environmental or external factors such as the state of the sensor itself, evidence of the credibility of the source's credibility level with the first sub coefficient a great relationship, in addition to the previous situation of the actual reliability coefficient of the gradient coefficient of the credibility of the first meeting of a progressive, the past can be used to predict the expected reliability coefficient of reliability is an important factor basis. Therefore, the proposed method of forecasting reliability coefficient is evidence of the credibility of the source of the time the actual coefficient and reliability coefficient of 0 ... times the expected value of reliability through a balance between the factors to weigh and draw [8].

3. The Evidence of Distributed Fusion Algorithm DS Based on Forecast Credibility

The original theory of evidence combination formula of evidence in the case of conflict will become invalid, the synthesis of the results obtained is often inconsistent with the actual situation, and the combination rules are also robust enough. Has been studied by scholars of the methods can be summarized into two

categories: one is the combination of the original data before the change, the original data by pre-conflict evidence will be modified, and they think DS combination rule is no errors, such as ROCKETS methods. The second category is for the DS combination rule itself changes, mainly around the conflict to start the distribution and management improvements. In addition, how to effectively improve the efficiency and integration of evidence synthesis rate have also become issues of concern in many areas, and the current research on this area is still relatively small. Taking into account the inadequacies of the above two points, this chapter first reference ROCKETS solution to the conflict of evidence, put forward on this basis to predict the reliability coefficient of view of data fusion to improve efficiency. Thus, it realizes a distributed data fusion and the evidence of an effective solution to the conflict, but also maintained the original rules of DS combination of synthesis and exchange of good nature and so on. In a large number of training samples, based on the formula (1), Formulas (2) Trained out of a more stable value, and can be applied to predict the value of reliability coefficients in order to achieve the distributed data fusion . In practice, according to historical data from the previous formula (2) to calculate the evidence of the credibility of the source the first time the prediction coefficients. When all the evidence the source of the prediction reliability coefficient calculations are complete, according to the literature proposed the basic probability assignment function of the "discount rate" concept and the relative reliability coefficient vector is calculated:, the source of evidence can be adjusted reliability coefficient, which means all the source first time evidence of the maximum predicted reliability coefficient [9].

The improved basic probability assignment function into the original DS combination rules to get the following formula:

$$m(A)=\begin{cases}0 & A=\varnothing \\ k\sum\limits_{A_1\cap\ldots\cap A_n=A} m_{\varpi^n_{预i}}(A_1)\ldots m_{\varpi^n_{预n}}(A_n) & A\neq\varnothing\end{cases} \tag{1}$$

$$k=\left(\sum_{A_1\cap\ldots\cap A_n\neq\Phi} m_{\varpi^n_{预i}}(A_1)\ldots m_{\varpi^n_{预n}}(A_n)\right)^{-1}$$

4. Simulation

The main purpose of this chapter is to experiment with the literature [9] proposed weighted algorithm (ROCKETS method) were compared, the main focus of comparison and integration through the integration of accurate time and reflect both the performance. Experiments in Windows XP system, using stochastic simulation methods, and programming through Matlab simulation results validate the effectiveness of the proposed algorithm in this chapter[10]. Simulation experiment 3 were used as evidence of the source sensor for temperature data collection, identification framework U={High,Medium,Low}, which, Low=[0,15];Medium=[16,30];Hig=[31,45]; units: degrees Celsius.

4.1. Training balanced factorial reliability prediction

According to the actual characteristics of the sensor in accordance with the analog sensor output in the form of Gaussian distribution of random numbers as the simulation temperature, the Gaussian distribution in the form such as:

$$f(x)=\exp(-(x-\mu)^2/2\sigma^2)$$

Simulated in the experiment of the three sensors A,B,C corresponding Gaussian distribution function as follows:

$$\begin{aligned} f_A(x)&=\exp(-(x-22)^2/18)\\ f_B(x)&=\exp(-(x-24)^2/18)\\ f_C(x)&=\exp(-(x-20)^2/18)\end{aligned} \tag{2}$$

Simulation using Matlab, according to (2) in the three expressions, respectively, the output of the sensor data values generated for each sensor generates 100 to obey the Gaussian distribution of its own random

number, random number from this group of temperature and recognition framework matching interval to generate a set of basic probability assignment values as a set of training samples. In order to improve the accuracy of the experimental group that the Chinese Communists have 10,000 training data, and finally take expectations as to the actual application of distributed data fusion. Here only some examples of experimental data and the obtained values.

The first set of experiments generated random numbers are as follows:

$$A = \{26,22,16,20...\}$$
$$B = \{23,26,19,25...\}$$
$$C = \{19,20,20,29...\}$$

The random numbers obtained by the basic probability assignment values are as follows:

$$m_{A1}(\text{Low}) = 0.14,\ m_{A1}(\text{Medium}) = 0.73,\ m_{A1}(\text{High}) = 0.13$$
$$m_{B1}(\text{Low}) = 0.10,\ m_{B1}(\text{Medium}) = 0.82,\ m_{B1}(\text{High}) = 0.08$$
$$m_{C1}(\text{Low}) = 0.09,\ m_{C1}(\text{Medium}) = 0.77,\ m_{C1}(\text{High}) = 0.14$$

The credibility of this set of data coefficients from Table 1 .

Table 1. Results of the First Group of Training Samples

Sensors	Conflict Vector	Normalized vector of conflict	Predicted reliability coefficient	α
A	（0.14，0.06）	（0.7000，0.3000）	0.3538	0.7273
B	（0.14，0.08）	（0.6364，0.3636）	0.3297	0.3404
C	（0.06，0.08）	（0.4286，0.5714）	0.3165	1.3600

Table 1 is given in the first set of training sample data, according to the values described in section 2, the training process we can see, the first two groups can not calculate the value of training samples, the second set of training samples is calculated from the first group the same reliability coefficients were obtained: 0.3576,0.3268,0.3156. Starting from the third set of data values can be calculated, a random number from the third group received the basic probability assignment is:

$$m_{A1}(\text{Low}) = 0.07,\ m_{A1}(\text{Medium}) = 0.85,\ m_{A1}(\text{High}) = 0.08$$
$$m_{B1}(\text{Low}) = 0.21,\ m_{B1}(\text{Medium}) = 0.74,\ m_{B1}(\text{High}) = 0.05$$
$$m_{C1}(\text{Low}) = 0.13,\ m_{C1}(\text{Medium}) = 0.79,\ m_{C1}(\text{High}) = 0.08$$

The following results are given training in the third set of data, as shown in Table 2.

Table 2. Results of the Third Group of Training Samples

Sensors	Conflict Vector	Normalized vector of conflict	Predicted reliability coefficient	α
A	（0.12，0.05）	（0.7059，0.2941）	0.3532	—
B	（0.12，0.06）	（0.6667，0.3333）	0.3362	—
C	（0.05，0.06）	（0.4545，0.5455）	0.3106	—

Value method in accordance with such training, group training samples from the 10,000 three sensor values are obtained as follows:

$$\alpha_A = \frac{0.7273 + 0.8524 + ...}{10000} = 0.8051$$
$$\alpha_B = \frac{0.3404 + 0.3547 + ...}{10000} = 0.6437$$
$$\alpha_C = \frac{1.3600 + 1.2541 + ...}{10000} = 0.9384$$

4.2. Distributed Fusion Algorithm Based on Predictive Reliability of performance comparison

The value of the training into practice from the data fusion to obtain a set of experimental results shown in Table 3.

Table 3.Comparison of the Reliability Coefficient

Sensors	Predicted reliability coefficient	HuChanghua method	Deviation
A	0.3566	0.3588	0.0022
B	0.3279	0.3247	0.0032
C	0.3155	0.3165	0.0010

ROCKETS method is characterized by the conflict can be resolved on the basis of evidence converge faster, but the results can be seen in the above table, by value and credibility of historical data, the prediction coefficients obtained with the ROCKETS value calculated between the error is very small, so the prediction reliability using the proposed data fusion coefficients also can solve the problem of conflict of evidence also has good convergence. In addition, based on the predicted reliability of the data fusion is the fusion of data before the arrival of the next can be the source of the data to calculate the reliability coefficient, calculated in the integration of the credibility factor when saving time efficiency significantly in the improvement, especially in a large number of data fusion effect more visible.

5. Summary

Based on DS evidence theory proposes a distributed data fusion theory to the premise of ensuring the accuracy and the ability to improve the efficiency and effective integration of DS evidence theory there is evidence to solve the problem of conflict. The distributed idea is to predict the reliability coefficient proposed the concept to achieve, then the credibility of the prediction coefficient synthesis of DS evidence theory to combine the rules change in the nature of assurance on the basis of the original creates a new synthetic formula, and then gives concrete evidence fusion algorithm. Predict in advance the source of the evidence by the credibility of coefficients, which can effectively solve the conflict of evidence and can achieve the idea of distributed data fusion to improve the fusion efficiency. Finally, the simulation shows that the algorithm to achieve effectiveness.

Reliability coefficient based on forecasts made on the basis of distributed data fusion theory and the reliability coefficient of the other algorithms, not only can effectively solve the conflict of evidence, and can effectively handle large amounts of data, evidence of significantly increased integration efficiency and save time.

6. References

[1] Lefevre E, Colot O. Belief function combination and conflict management[J]. Information Fusion, 2002, 3(2): 149-162P

[2] Smets P. Decision making in the TBM: the necessity of the pignistic transformation[J]. International Journal of Approximate Reasoning，2004，38：133-147P

[3] Ziarko W. Rough Set Approaches for Discovering Rules and Attribute Dependencies. USA, Oxford: Oxford University Press, 2002: 328-339P.

[4] R.Mahler. Random sets: unification and computation for information fusion– a retrospective assessment," in the 7th International Conference on Information Fusion, 2004, 1-20P.

[5] Mourad Ossalah. Some notes on fusion of uncertainty information. International Journal of Intelligent Systems. 2004, 19(6): 457-471P

[6] Fu Yao-wen, Li Xiang, Zhuang Shao-wen. A decisionlevel fusion algorithm for target recognition with fuzzy densities determined adatively. Acta Electronica Sinica. 2004, 32(9):1433-1435P

[7] Delmotte Francois, Smets Philippe. Target inedtification based on the transferable belief model interpretation of Dempster-Shafer model. IEEE Trans. On SMC-Part A: Systems and Humans. 2004, 34(4): 418-434P

[8] Richard T. Principles of effective multisensor data fusion. Military Technology. 2003,27(5): 29-37 P

[9] HF Durrant-Whyte. Sensor models and multi-sensor integration. The International Journal of Robotics Research. 2001,7(6): 87-92P

[10]Megalooikonomou V, Yesha Y. Space efficient quantization for distributed estimation by a multi-sentor fusion system. Information Fusion. 2004, 5(5): 299-308P