Operation Condition Monitoring using Temporal Weighted Dempster-Shafer Theory

Xiaoyun Wang[1], Tingdi Zhao[2]

[1,2] *School of System Engineering and Reliability, BeiHang University, Bejiing,100191, China*

wangxys@gmail.com
ztd@buaa.edu.cn

ABSTRACT

System operation is a real time, dynamic decision process, a continuous observation should be implemented to support timely decision. Real time condition monitoring and diagnosis is featured with ongoing event sequence. The more recent observation, the much detailed, accurate information, and the more obsolete observations with much weak correlation to current faults and errors vise versa.

Dempster-Shafer evidence theory is best suitable for the problem of redundant sensors, insufficient data reasoning. However, D-S base applications largely focused on causational relationship between symptoms and effects, and the fusion process of evidences was performed regardless whatever order observed. As an improvement to the frame of discernment of the D-S theory, we purposed a time weighted evidence combination method. Observed events were extracted from multiple time points to form a temporal evidence sequence. Basic probability assignment was altered by temporal weights in accordance with the time proximity between the observed events and current time. The temporal weights value set was in accordance with its occurring time point. Evidences with same timestamps should be allocated with the same temporal weights. An example was discussed to illustrate the temporal weight, D-S rule based assessment framework. In the framework, latest observed evidences stream were combined into the framework to improving fault recognition.

1. INTRODUCTION

Condition assessment for system operation is a real time, dynamic decision process, during that course, a continuous observation should be implemented to support timely condition assessment. Currently, as a method widely in the area of fault diagnosis applications, Dempster-Shafer evidence theory is best suitable for the problem of redundant sensors, reasoning of insufficient data which might be imprecise and incomplete(Yang, 2006)(Parikh,2001)

As an extension of traditional probabilistic theory, the Dempster-Shafer Theory (DST) of evidence provides beneficial approaches to uncertain reasoning. In the network security area, DST was used as a method for incursion detection (Lan, 2010), intrusion prioritizing (Zomlot, 2011). In ubiquitous network and pervasive computing, DST was applied to recognize situation and activities in smart environment (McKeever, 2009). It also play an important role in bank fraud detection applications (Beranek, 2013). Some of these researches concerned the temporal property of evidence to improve performance of detection, for examples, McKeever tried to use a duration measure to generate the belief of event and evidence.

Our research focused on the problem of temporal aspect of DST evidence. During the online condition monitoring, some observed information might not up to date sufficiently while others may appears better timeliness. Outdated information as one of three kind of major information problems (Garvin, 1988), is not sufficiently for the task of fault detection. The more recent observations could provide much detailed, accurate information about current condition.

This paper is organized as follows. In section 2, the classic Dempster-Shafer Theory of Evidence is introduced, and the problem of application DST to online diagnosis for operation condition monitoring and failure detection and recognition is analyzed. Here we purposed a temporal weighted evidence combination method together with the procedure of application. In section 3, an example is discussed to illustrate how the temporal weight D-S rule combination method can be applied to online failure identification. Also we compared the result of classic D-S rules of combination with our method.

2. METHODOLOGY

2.1. D-S evidence theory

The Dempster-Shafer Theory (DST) or D-S theory of evidence was first introduced at 1960s(Dempster, 1968)(Shafer, 1976). The DST is basically an extension of traditional probabilistic modeling of uncertainty. Currently, the D-S theory of evidence was applied widely in fault diagnosis and recognize for its effectiveness to incomplete, inaccurate or conflict data.

According to the classic D-S theory of evidence, the elements needed to model the problem could be summarized as following:

A frame of discernment Ω, which should be a finite set of all of the possible hypotheses that are mutually exclusive;

A mapping of $m: 2^{\Omega} \to [0,1]$, which defines the basic probability assignment(BPA) of each subset $A \subset \Omega$ of hypotheses and satisfying $m(\varnothing) = 0; \sum_{A \subseteq \Omega} m(A) = 1$. The BPA represents a certain piece of evidence.

A rule of D-S evidence combination, which could be used to yield a new BPA from two independent evidences and their BPAs. There are a number of possible combination rules in application (Sentz, 2002). One of them is the Dempster's Rule, that could be defined as follows

$$m(A) = m_1 \oplus m_2(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - K} \quad (1)$$

$$K = \sum_{B \cap C = \varnothing} m_1(B) m_2(C) \quad (2)$$

.

B and C are subset of hypothesis. K reflects the conflict between B and C, while the higher the K, the greater the conflict between the evidences. It was proven that the Dempster rule of combination meets the commutative and associative laws, which could be depicted as such:

$$(m_1 \oplus m_2) \oplus m_3 = m_1 \oplus (m_2 \oplus m_3)$$

and

$$m_1 \oplus m_2 = m_2 \oplus m_1 .$$

Therefore evidences are treated as equal, as well as the order of evidences dose not affect the result of evidence combination.

2.2. Temporal Weighted Evidence Combination

D-S rule of combination treat evidences equally from different sensor. However, that assumption generally cannot hold during an online condition monitoring. System data and evidence unveiled gradually, sequentially, as time lapsing. What we have identified is only a fraction of the facts. At the early stage of a fault or failure, the symptom could be dim and weak. As the system operation went on, the system performance appears variation, while some symptom may change as well, others could be expired or not valid any more. The creditability of past evidence is not static. Instead it should change in course of timeliness. Evidence that is up-to-date should be assessed as a strong sample. The more recent observations could provide much detailed, accurate information about current condition. At the same time, those past, obsolete evidences only have partial utility, appeared a weak correlation to current faults and errors(Garvin, 1988).

Based on the weighted view of evidence (Yu, 2005), we purposed a temporal weighted combination rules to solve the problems of timeliness of evidence. The weight of each evidences are based on their timestamp properties. The temporal weighted rule combination is:

$$m_{1,2}(A) = \frac{\sum_{B \cap C = A} m_{t1}(B)^{w1} \cdot m_{t2}(C)^{w2}}{\sum_{B \cap C \neq \varnothing} m_{t1}(B)^{w1} \cdot m_{t2}(C)^{w2}} \quad (3)$$

where w_1, w_2 is the temporal weights of time point t_1, t_2 for evidence B and C:

$$w_i = \exp(K(t_i - T)) \quad (4)$$

in which T is the current time (system time). K is a user predefined constant, $K \geq 0$.

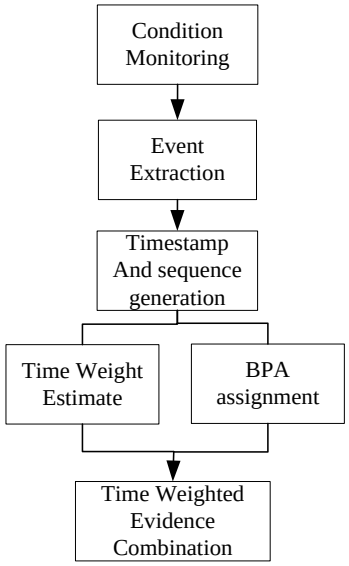

Figure 1. Schematic of the method for time weighted evidence combination.

From the equation (4), we could find some feature of w_i:

a) The temporal weight of the latest evidence is greater than that of those previously evidences.

b) The older of the evidence, the less timeliness and values of the belief, as well as its temporal weight.

c) Evidences with time proximity have similar temporal weights.

d)Temporal weight of on-going evidence has approximate value to 1, which represent it is the most up-to-date evidence.

The workflow of temporal weighted D-S evidence combination method is described as Figure 1. Observed events were extracted from multiple time point to form a temporal evidence sequence. Basic probability assignment is altered by temporal weights in accordance with the time proximity between the observed events and current time. The temporal weights setup is in accordance with its occurring time point. Lately observed evidence could have better influence and support to the hypothesis than those older evidences. Evidences with same timestamps should be allocated with the same temporal weights.

Considering the introduction of temporal weighted combination rules, the combination of multiple evidence is no longer commutative and equally treated, that means each time point we need to recalculate the set of temporal weight w_i, as shown in Figure 2.

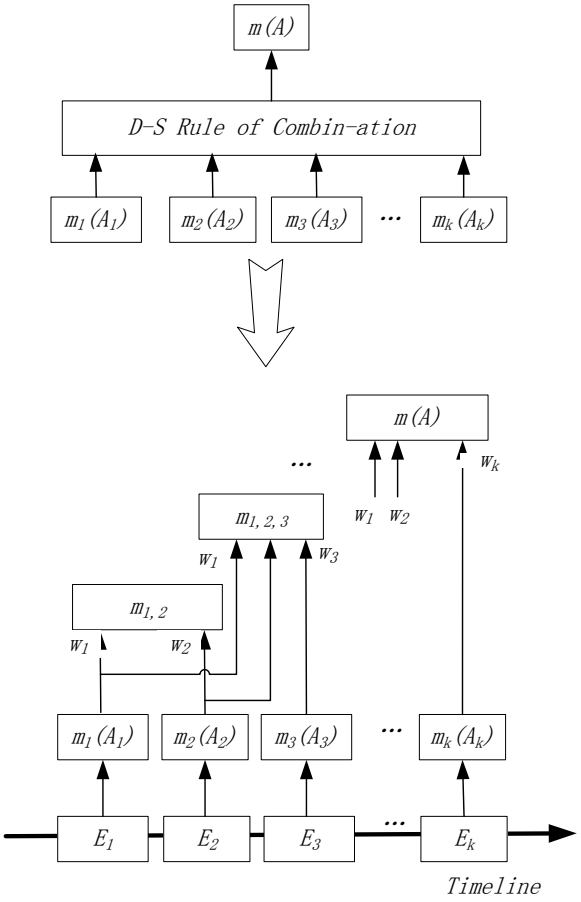

Figure 2. Time weighted D-S evidence combination.

However, this approach might be faced with time complexity for the calculation of w_i at each time point. To simplified the framework, we merged the past combination result into a new evidence at each time point, as shown in Figure 3. The improved framework has better time performance while yield approximately result as Figure 2.

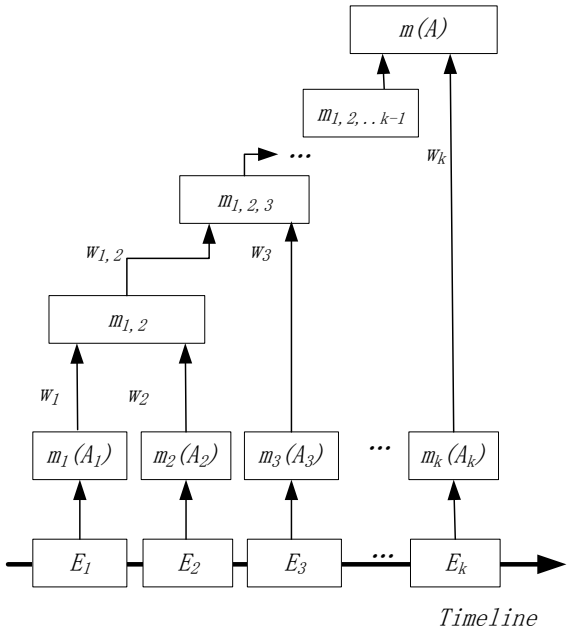

Figure 3. A improved framework of time weighted D-S evidence combination for multiple symptoms.

3. CASE STUDY

In this case, we adopted the dataset of a power generator (Ray, 2007) as an example to illustrate the temporal weight, D-S rule based assessment framework. During its operation, working condition and performance events was monitored periodically.

We need to assess the on-going events and symptoms to identify the type of possible failure(s) in a near real time manner. The challenge lies in that the observing events were ever changing and added in, while the early events and their information might expired, the latest evidence need be combined into the frame to improving the accuracy of results.

3.1. Dataset Preparations

There are three kinds of power generator failures given by domain specialist, namely, h_1, h_2, h_3. The corresponding frame of discernment could be given with

$$\Theta=\{h_1, h_2, h_3, \theta\}$$

where θ is the unknown type of failures.

Event E_1, E_2 and E_3 were reported by system monitoring function, with the corresponding timestamps as $t_1=1, t_2=2, t_3=3$. So we have a sequence of symptom as $\{<E_1,1>,<E_2,2>,<E_3,3>\}$.

To simplify the calculation, we choose the time-weighted constant $K=\ln 2$. As a result, the time weight turns into

$$w_i = \exp(K(t_i - T)) = 2^{(t_i - T)} \quad (5)$$

The BPAs for the hypothesizes of evidence E_1, E_2 and E_3 were given.

Event E_1 was the symptom to failure types h_1, h_2, h_3 with a BPA of 0.7, so:

$$m_{t_1}\{h_1, h_2, h_3\} = 0.7 \text{ And } m_{t_1}\{\theta\} = 0.3$$

E_2 was the symptom of failure h_1 with the belief of 0.9:

$$m_{t_2}\{h_1\} = 0.9 \text{ And } m_{t_2}\{\theta\} = 0.1$$

For event E_3 which was a evidence for failures of h_2, h_3, the BPA is 0.8, so that:

$$m_{t_3}\{h_2, h_3\} = 0.8 \text{ and } m_{t_3}\{\theta\} = 0.2$$

3.2. Temporal Combination of Evidence Sequence

According to equation (5), the time weight could be given for sequence $\{<E_1,1>,<E_2,2>,<E_3,3>\}$.

Event E_1 was detected at $t_1 = 1$. With the new event E_2 was detected at $t_2 = 2$, evidence E_1 and evidence E_2 need to be fused. Table 1 shows the combination rules for m_{t_1,t_2}:

Table 1. Combination of E_1 and E_2

$m_{t_1}(B)^{w_1} \cdot m_{t_2}(C)^{w_2}$	$\{h_1\}$	$\{\theta\}$
$w_1 = 0.5$ $w_2 = 1$	$m_{t_2}\{h_1\} = 0.9$	$m_{t_2}\{\theta\} = 0.1$
$\{h_1, h_2, h_3\}$	$\{h_1\}$	$\{h_1, h_2, h_3\}$
$m_{t_1}\{h_1, h_2, h_3\} = 0.7$	0.753	0.084
$\{\theta\}$	$\{h_1\}$	$\{\theta\}$
$m_{t_1}\{\theta\} = 0.3$	0.493	0.055

According to equation (3),

$$m_{t_1,t_2}\{h_1\} = \frac{\sum_{B\cap C=h_1} m_{t1}(B)^{\frac{1}{2}} \cdot m_{t2}(C)^1}{\sum_{B\cap C\neq\varnothing} m_{t1}(B)^{\frac{1}{2}} \cdot m_{t2}(C)^1} = 0.9$$

$$m_{t_1,t_2}\{h_1, h_2, h_3\} = 0.06$$

$$m_{t_1,t_2}\{\theta\} = 0.04$$

With the event E_3 was detected at time $t_3 = 3$, new evidence added in and the result reflect the influence of up-to-date information. Table 2 shows the combination rules for m_{t_1,t_2,t_3}:

Table 2. Combination of E_1, E_2 and E_3

$m_{t_1,t_2}(B)^{w_{1,2}} \cdot m_{t_3}(C)^{w_3}$	$\{h_2, h_3\}$	$\{\theta\}$
$w_{1,2} = 0.5$ $w_3 = 1$	$m_{t_3}\{h_2, h_3\} = 0.8$	$m_{t_3}\{\theta\} = 0.2$
$\{h_1\}$	$\varnothing$	$\{h_1\}$
$m_{t_1,t_2}\{h_1\} = 0.9$	0	0.19
$\{h_1, h_2, h_3\}$	$\{h_2, h_3\}$	$\{h_1, h_2, h_3\}$
$m_{t_1,t_2}\{h_1, h_2, h_3\} = 0.06$	0.196	0.049
$\{\theta\}$	$\{h_2, h_3\}$	$\{\theta\}$
$m_{t_1,t_2}\{\theta\} = 0.04$	0.16	0.04

The combination at time $t_3 = 3$ as shown:

$$m_{t_1,t_2,t_3}\{h_1\} = 0.299$$

$$m_{t_1,t_2,t_3}\{h_2, h_3\} = 0.560$$

$$m_{t_1,t_2,t_3}\{h_1, h_2, h_3\} = 0.077$$

$$m_{t_1,t_2,t_3}\{\theta\} = 0.063$$

Here we had combine the sequence $\{<E_1,1>,<E_2,2>,<E_3,3>\}$ at $t_3 = 3$.

In Table 3 we compared the results of classic D-S approach and our temporal weighted combination method. Apparently, from row $t_3 = 3$ we can see that the temporal weighted approach is more sensitive to latest, up-to-date evidence, which yield a higher belief for hypothesis set $\{h_2, h_3\}$ in favor of the newly observed evidence $<E_3,3>$. Also we could infer from the line $t_2 = 2$ that when the latest evidence was similar to the former ones, the output beliefs of temporal weighted combination method is only slightly different to classic D-S approach.

Table 3. Comparison of temporal weighted combination method and classic D-S evidence combination

Time	BPA	Classic D-S	Temporal Weighted
$t_2 = 2$	$m_{t_1,t_2}\{h_1\}$	0.9	0.9
	$m_{t_1,t_2}\{h_1, h_2, h_3\}$	0.07	0.06
	$m_{t_1,t_2}\{\theta\}$	0.03	0.04
$t_3 = 3$	$m_{t_1,t_2,t_3}\{h_1\}$	0.511	0.299
	$m_{t_1,t_2,t_3}\{h_2, h_3\}$	0.227	0.560
	$m_{t_1,t_2,t_3}\{h_1, h_2, h_3\}$	0.034	0.077
	$m_{t_1,t_2,t_3}\{\theta\}$	0.227	0.063

4. CONCLUSION

The Dempster-Shafer Theory of evidence based model has been widely used to multi sensor fault detection and recognition. As an improvement to the DST, the temporal weighted evidence combination method could be beneficial to the balance of long term trend and abrupt fault recognition, especially for the online health management applications, compared with the classic DST combination method.

Our contribution could be summarized as follows: First, the problem of obsolete evidence of real time monitoring and diagnosis is analyzed. Then the temporal weighted evidence combination method is purposed. To make the method for efficiency, an improved framework that accumulates the past combination result is suggested. Furthermore, a case study was discussed to illustrate the temporal weighted D-S rule based assessment framework.

ACKNOWLEDGEMENTS

Funding for this paper was provided by BeiHang University. The authors gratefully acknowledge the sponsor of the School of Reliability and System Engineering, BeiHang University.

REFERENCES

Yang, B.S., Kim, K. J., (2006). Application of Dempster-Shafer theory in fault diagnosis of induction motors using vibration and current signals. *Mechanical Systems and Signal Processing*, 20:403–420.

Parikh,C.R., Pont, M.J., Jones, N.B., (2001). Application of Dempster-Shafer theory in condition monitoring systems: A case study, *Pattern Recognition Letters*, 22 (6-7): 777-785.

Fang, L., Wang, C., et.al, (2010). A Framework for Network Security Situation Awareness Based on Knowledge. *2nd International Conference on Computer Engineering and Technology*, Chengdu, China

Zomlot, L., Sundaramurthy. S., (2011). Prioritizing Intrusion Analysis Using Dempster-Shafer Theory, *Proceedings of the 4th ACM workshop on Security and artificial intelligence*, Chicago, Illinois, USA

McKeever, S., (2009). *Recognising Situations Using Extended Dempster-Shafer Theory*, Doctoral dissertation. National University of Ireland, Dublin

Beranek, L., Knizek, J., (2013).The Use of Contextual Information to Detection of Fraud on On-line Auctions. *Journal of Internet Banking and Commerce*, vol. 18, no.3

Dempster, A., (1968). A Generalization of Bayesian inference. *Journal of the Royal Statistical Society*, pp. 205–247.

Shafer, G., (1976). *A Mathematical Theory of Evidence*, New Jersey, Princeton University Press.

Sentz, K., (2002). Combination of Evidence in Dempster-Shafer Theory, Binghamton University, Binghamton, NY

Garvin, D., (1988). Managing quality. NY: Free Press.

Yu, D,. Frincke, D., (2005). Alert Confidence Fusion in Intrusion Detection Systems with Extended Dempster-Shafer Theory, *43rd ACM Southeast Conference*, Kennesaw, GA

Kay, R.U., (2007). Fundamentals of the Dempster-Shafer theory and its applications to system safety and reliability modeling. *Reliability: Theory and Applications*. vol 2(3-4), pp. 173-185

BIOGRAPHIES

Xiaoyun Wang Phd candidate of Aerospace Engineering, School of Reliability & System Engineering, BeiHang University, Beijing, China, telephone: (8610) 82317665, email:wangxys@gmail.com

Tingdi Zhao Professor of Engineering, School of Reliability & System Engineering, BeiHang University, Beijing, China, telephone:(8610)82316570, email:ztd@buaa.edu.cn