Malaysian Journal of Computer Science, Vol. 13 No. 1, June 2000, pp. 84-89

EVIDENCE SETS APPROACH FOR WEB SERVICE FAULT DIAGNOSIS

Wen Cheong Chin and V. Ramachandran
Faculty of Information Technology
Multimedia University
63100 Cyberjaya, Selangor, Malaysia
Tel.: 603-83125250
Fax: 603-83125264
email: wcchin@mmu.edu.my

Chee Way Chong
Faculty of Engineering
Multimedia University
63100 Cyberjaya, Selangor, Malaysia
Tel.: 603-83125250
Fax: 603-83125264

ABSTRACT

The main objective of this paper is to diagnose the fault distribution of the web service process using expert subjective opinions. In order to capture the fuzziness, non-specificity and conflictive of human judgement, evidence set proposed by L. M. Rocha is utilized to represent the expert opinions linguistically. The fault propagation is illustrated via coherent fault tree, and then the minimal cut sets (MCS) are derived. The highest occurrence of MCS is determined with acceptable conflictive measure. Proper web performances' improvement can be done based on the criticality of each MCS.

Keywords: ***Fault Tree Analysis, Evidence Sets, Dempster Shafer Theory***

1.0 INTRODUCTION

The WWW performance at present is not very convincing. Dangling referential link, floating and time-varying access rate, slow server response, lengthy surfing and complex web designs are some of the problems experienced by the web users. Certain web-sites have fixed and short life spans due to limitation of server bandwidth availability. The user may gain or lose the access rights to a particular link or set of links in the web-sites at random. Further, it is common for the web users to wait for long time while retrieving the target page and yet fail to access the page due to access time out and long request queue.

Thus, diagnosis of web service degradation and faults becomes increasingly significant in order to pinpoint various key failures within the web process. The faults may include inconsistency of hardware operation, network congestion, limitation of client and server bandwidth, random user surfing behaviour, protocol conflict etc. All these possibilities or combination of possibilities constitute the basic events leading to the downgrading of web service within a particular web process.

This paper aims to illustrate the inter-relations among these basic events symbolically and to indicate the physical connections in highly time-varying web process. Coherent fault tree is constructed to pictorially represent the propagation of all the combinations of basic events leading to web service failure and minimal cut sets are derived. However, the identification of web service failures involves many uncertainties and randomness since their occurrence frequently relying on imprecise or vague input data namely. Michael A. S. Guth [1] has proposed to use Dempster-Shafer theory (DST) [2] of evidence in order to model this vagueness in a probabilistic manner. He has utilised DST to assign upper and lower bounds for the probability on elements of the state space. The imprecise data are represented via a 3-valued logic derived from DST probability assignment. Although the method proposed above is simple and well programmed on the computer, it is still impractical to apply in real web process since it is difficult to estimate the uncommitted states for all the web service failures.

Meanwhile, D. Singer [3] has argued that the conventional fault tree does not concern the tolerances of the probability values of hazards. The causes of inaccuracy of relative frequencies are general non-stationary and non-ergodicity of natural phenomena especially in man-made systems. Thus, he proposed to use fuzzy numbers to represent the relative frequencies of the basic events. He has shown the use of n-ary possibilistic AND, OR, and NEG operators to construct possibilistic fault tree. Nevertheless, the method proposed has not fully concerned that the non-specificity and conflictive of human involvement in judging the occurrence of the basic events. Thus, it is not easy to determine the probability as well as its tolerance.

The proposals of Guth and Singer seems just to take into account the fuzziness or vagueness of input data. However, the non-specificity and conflictive features in expert evaluation for occurrence of basic events have not been fully considered. In this paper, evidence set proposed by L. M. Rocha [4] is utilised to fully model the uncertainties pertaining to input data of the fault tree analysis of web service process.

2.0 EVIDENCE SETS

Evidence set [4] offers a way to model the fuzziness, non-specificity and conflictive of human evaluation. It offers both an independent characterisation of membership and a formalisation of judgements imposed on the membership. However, conventional set structures (crisp, fuzzy or interval valued fuzzy set) alone offer only an independent degree of membership while evidence theory by itself offers a formalisation of belief which constraints the elements of a universal set with a probability restriction.

Evidence set captures fuzziness, non-specificity, and conflict in their membership degrees. Fuzzy set captures solely fuzziness, and interval valued fuzzy set (IVFS) capture fuzziness and non-specificity [4]. The membership degrees are defined by the belief measures of the Dempster-Shafer theory of evidence.

Evidence set A of X is defined generally by a membership function of the form:

$$A(x) : X \rightarrow \mathcal{B}[0, 1]$$

where $\mathcal{B}[0, 1]$ is the set of all possible bodies of evidence (F^x, m^x) on the set of intervals of the unit interval $\zeta[0, 1]$, and F^x and m^x are IVFS and basic probability assignment for x respectively.

For instance, the membership function of an evidence set A of X is given, for each x by n intervals weighted by a basic probability assignment m^x:

$$A(x) = \{< I_1^x, m_1^x >, < I_2^x, m_2^x >, \ldots\ldots, < I_n^x, m_n^x >\}$$

where

$$I^x = [I_{inf}^x, I_{sup}^x] \subseteq [0,1]$$

The complementary, intersection and union operations of evidence sets are shown as follows:

2.1 Complementation

The complement of A(x) is:

$$A^c(x) = \{< (I_1^x)^c, m_1^x >, < (I_2^x)^c, m_2^x >, \ldots\ldots, < (I_n^x)^c, m_n^x >\}$$

where,

$$(I^x)^c = [1 - I_{sup}^x, 1 - I_{inf}^x]$$

2.2 Intersection and Union

The intersection and union of two IVFS are defined as the minimum and maximum of their respective lower and upper bounds of their membership intervals. Given two intervals of [0,1]:

$$I = [I_L, I_U] \subseteq [0,1]$$

$$J = [J_L, J_U]$$

The minimum of both intervals is an interval:

$$K = Min(I, J) = [Min(I_L, J_L), Min(I_U, J_U)]$$

The maximum of both intervals is an interval:

$$K = Max(I, J) = [Max(I_L, J_L), Max(I_U, J_U)]$$

Given two evidence sets A and B defined for each x of X by:

$$A(x) = \{< I_i^x, m_A^x(I_i) >\} \qquad i = 1, \ldots\ldots\ldots, n$$

$$B(x) = \{< J_j^x, m_B^x(J_j) >\} \qquad j = 1, \ldots\ldots\ldots, m$$

where I_i and J_j are intervals of [0,1]. Their intersection and union is an evidence set C(x) = A(x)B(x), whose intervals of membership K_k and respective basic probability assignment $m_C(K_k)$ are defined by ;

For intersection:

$$m_C^x(K_k^x) = \sum_{MIN(I_i^x, J_j^x) = K_k^x} m_A^x(I_i^x).m_B^x(J_j^x)$$

For union:

$$m_C^x(K_k^x) = \sum_{MAX(I_i^x, J_j^x) = K_k^x} m_A^x(I_i^x).m_B^x(J_j^x)$$

3.0 FAULT TREE ILLUSTRATION OF WEB SERVICE FAILURES

In this section, a typical HTTP request/response process will be considered and all possibilities of the process parameters (basic events) such as server capacity, client bandwidth, server bandwidth, queue length etc. contributing to the failure of web service (top event) are investigated. The details of how the effect of these basic events affect the top event are illustrated by the fault tree as follows:

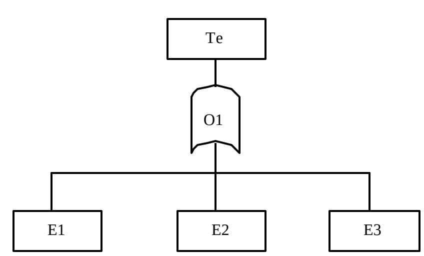
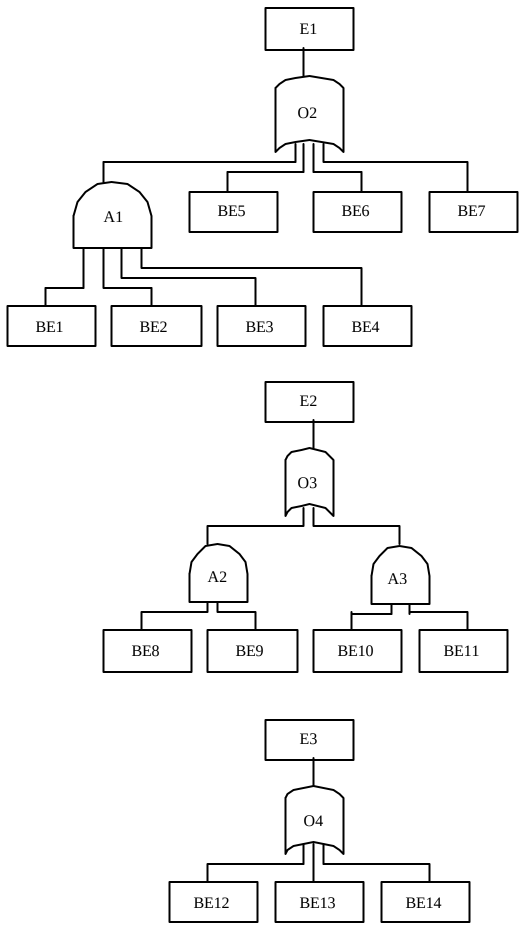


Fig 1: Classification of Web Service Failures

Basically, the failures or downgrading of web service (Te) can be classified into three main categories:

(a) Server Side Failures (E1)
(b) Client Side Failures (E2)
(c) Communication Failures (E3)

The details of intermediate events {E1, E2, E3} can be further decomposed as shown in Fig. 2.



BE1	≡	Long Queue
BE2	≡	Non-Intelligent Server Management
BE3	≡	Slow Server Processing & Bandwidth
BE4	≡	Insufficient Network Bandwidth
BE5	≡	Virus Implantation
BE6	≡	Incompatible Software
BE7	≡	Configuration and Installation Error
BE8	≡	Incompatible Browser
BE9	≡	No External Plug-ins
BE10	≡	BE4
BE11	≡	Low Client Processing Capacity
BE12	≡	Protocol Conflict
BE13	≡	Hardware Failures
BE14	≡	Network Congestion

Fig. 2: Web Service Fault Tree

The algorithms consist of two criteria:

a. if the top event gate is an OR gate(O) then each input to the OR gate represents an entry for each row of the list matrix.

b. for AND gate(A), each input to the AND gate represents an entry for each column of the list matrix.

The algorithm starts at O_1 directly under $\mathbf{T_e}$. In this case, the O_1 is an "OR" gate and hence the inputs to O_1 are placed in separate rows as follows:

Step 1:
E_1
E_2
E_3

Since each of the combination of the three inputs will lead to the top event, each of them constitutes a cut set. The following analysis is stated as follows:

Step 2:
O_2
O_3
O_4

Step3:
A_1
BE_5
BE_6
BE_7

A_2
A_3

BE_{12}
BE_{13}
BE_{14}

Step 4:
BE_1
BE_2 BE_3 BE_4
BE_5
BE_6
BE_7

BE_8 BE_9
BE_{10} BE_{11}

BE_{12}
BE_{13}
BE_{14}

Each gate in the Fig. 2 is then replaced with its inputs until one has gone through the whole tree and left with only the subsets of links. When this algorithm is fully completed, nine minimal cut sets in this example are obtained as below:

Order 1 MCS(s):
{BE5}, {BE6}, {BE7}, {BE12}, {BE13}, {BE14}

Order 2 MCS(s):
{BE8, BE9}, {BE10, BE11}

Order 4 MCS(s):
{BE1, BE2, BE3, BE4}

4.0 NUMERICAL ESTIMATES OF FAULT TREE ANALYSIS

Each and every basic event contributing to the MCS of the fault tree will be evaluated based on the historical technical information collected. Four experienced web administrators will be responsible to subjectively judge the relative frequencies of occurrence for each event for a specific time window using the following linguistic terms that determine by the intuition [7] of the four experts:

Low (L) : 1.0/1 + 0.8/2 + 0.6/3 + 0.4/4 + 0.2/5

Fair (F) : 0.6/1 + 1.0/2 + 0.8/3 + 0.4/4 + 0.3/5

Above Average (Av) : 0.5/1 + 0.7/2 + 1.0/3 + 0.5/4 + 0.4/5

High (H) : 0.4/1 + 0.5/2 + 0.6/3 + 1.0/4 + 0.7/5

Extremely High (EH) : 0.3/1 + 0.5/2 + 0.7/3 + 0.9/4 + 1.0/5

Table 1 illustrates the evaluation of first minimal cut set by all the web administrators (WA). The interval valued membership degrees contained in the evidence sets are represented linguistically along with their belief (basic probability assignment).

Table 1: First MCS and Web Administrator Evaluation

	BE5
WA1	{<[L, F], 0.3>, <[F, Av], 0.7>}
WA2	{<[F, H], 0.5>, <[L, Av], 0.5>}
WA3	{<[L, Av], 0.8>, <[F, H], 0.2>}
WA4	{<[F, F], 0.6>, <[L, F], 0.4>}

To obtain the overall opinion of all web administrators, the intersection (minimum) of the individual opinion (evidence sets) from WA1 to WA4 is taken as follows:

WA1(BE5)∩WA2(BE5)
= {<[(L ∩ F), (F ∩ H)], 0.15>,
<[(L ∩ L), (F ∩ Av)], 0.15>,
<[(F ∩ F), (Av ∩ H)], 0.35>,
<[(F ∩ L), (Av ∩ Av)], 0.35>}

WA3(BE5)∩WA4(BE5)
={<[(L ∩ F), (Av ∩ F)], 0.48>,
<[(L ∩ L), (Av ∩ F)], 0.32>,
<[(F ∩ F), (F ∩ H)], 0.12>,
<[(F ∩ L), (H ∩ F)], 0.08>}

WA1(BE5)∩WA2(BE5)∩WA3(BE5)∩WA4(BE5) =

{<([0.4, 0.6]/1+[0.5, 0.8]/2+[0.6, 0.6]/3+[0.4, 0.4]/4+ [0.2, 0.3]/5), 0.218>,
<([0.5, 1.0]/1+[0.7, 0.8]/2+[0.6, 0.8]/3+[0.4, 0.4]/4+ [0.2, 0.3]/5), 0.120>,
<([0.4, 0.6]/1+[0.5, 0.8]/2+[0.6, 0.8]/3+[0.4, 0.4]/4+ [0.2, 0.3]/5), 0.340>,
<([0.5, 0.6]/1+[0.7, 0.8]/2+[0.6, 0.8]/3+[0.4, 0.4]/4+ [0.2, 0.3]/5), 0.280>}
<([0.4, 0.6]/1+[0.5, 1.0]/2+[0.6, 0.8]/3+[0.4, 0.4]/4+ [0.3, 0.3]/5), 0.042>}

The aggregate values are obtained by averaging the interval valued fuzzy sets and are given as follows:

WA1(BE5)∩WA2(BE5)∩WA3(BE5)∩WA4(BE5) =

{<(0.50/1+0.65/2+0.60/3+0.40/4+0.25/5), 0.218>,
<(0.75/1+0.75/2+0.70/3+0.40/4+0.25/5), 0.120>,
<(0.50/1+0.65/2+0.70/3+0.40/4+0.25/5), 0.340>,
<(0.55/1+0.75/2+0.70/3+0.40/4+0.25/5), 0.280>,
<(0.50/1+0.75/2+0.70/3+0.40/4+0.30/5), 0.042>}

Similarly, the above procedures are applied to obtain the overall opinions of all MCS(s) from the web administrators. Each minimal cut set will constitute final evidence set, and the highest occurrence MCS can be obtained by comparing the centre of mass of the evidence sets. Further, the conflict for each MCS evidence set will be estimated as follows [6]:

$$S(m) = -\sum_{A\in F} m(A)\log_2 \sum_{B\in F} m(B).SUB(A,B)$$

for all A, B ∈ P(X), power set of X, and a body of evidence (F,m) defined on X, and SUB(A, B) denotes the subsethood of set A in set B of X. For this specific discrete domain, its value is given by the ratio $|A \cap B| / |A|$.

The centre of mass for MCS BE5 is obtained as follows:
Centre of Mass (CoM) BE5:

{0.54/1+0.69/2+0.68/3+0.40/4+0.25/5}

and the respective CoM and S(m) for each MCS is illustrated in Table 2.

However, the web administrators' expectations on each MCS may vary. For instance, the reduction of server side failures is much more necessary than client side failures since client side failures are not fully under control and uncertain. Thus, the threshold for each MCS occurrence is varied. The pre-defined threshold [7] of the CoM of BE5 is given as:

AR_CoM(BE5) = 0.80/1+0.60/2+0.50/3+0.40/4+0.1/5

Table 2: Centre of Mass and Conflict of BE5

	Centre of Mass	Conflict
{BE5}	{0.54/1+0.69/2+0.68/3+0.40/4+0.25/5}	{0.40/1+0.30/2+0.25/3+0.00/4+0.25/5}
{BE6}	{0.70/1+0.60/2+0.40/3+0.40/4+0.30/5}	{0.40/1+0.30/2+0.25/3+0.40/4+0.30/5}
{BE7}	{0.90/1+0.70/2+0.50/3+0.40/4+0.20/5}	{0.40/1+0.40/2+0.40/3+0.40/4+0.30/5}
{BE12}	{0.50/1+0.90/2+0.70/3+0.60/4+0.40/5}	{0.30/1+0.50/2+0.30/3+0.30/4+0.40/5}
{BE13}	{0.60/1+0.80/2+0.90/3+0.60/4+0.50/5}	{0.40/1+0.30/2+0.35/3+0.50/4+0.40/5}
{BE14}	{0.60/1+0.70/2+0.75/3+0.80/4+0.90/5}	{0.25/1+0.30/2+0.50/3+0.30/4+0.30/5}
{BE8, BE9}	{0.40/1+0.40/2+0.50/3+0.70/4+0.60/5}	{0.25/1+0.50/2+0.25/3+0.30/4+0.25/5}
{BE10, BE11}	{0.60/1+0.80/2+0.75/3+0.60/4+0.40/5}	{0.30/1+0.40/2+0.50/3+0.40/4+0.25/5}
{BE1, BE2, BE3, BE4}	{0.50/1+0.70/2+0.80/3+0.85/4+0.90/5}	{0.40/1+0.40/2+0.40/3+0.30/4+0.25/5}

Similarly, the conflict of overall evaluation should be maintained within a specific tolerance. Pre-defined threshold set of the conflict for all MCS(s) is given as follows and the threshold distance to the AR_Conflict is assigned as $<\pm$ N(where N is an real number and defined as 0.60):

AR_Conflict= 0.25/1 + 0.30/2 + 0.25/3 + 0.30/4 + 0.25/5

The most critical MCS, fuzzy Hamming distances (FHD) [7] of BE5 with respect to pre-defined threshold set for CoM and conflict are estimated as follows:

$$FHD = \sum_{i=1}^{r} | \beta^{i}_{CoM} - \beta^{i}_{AR} |$$

where

r $\equiv$ number of discrete elements in fuzzy set AR

β^i $\equiv$ grades of membership of i^{th} element in fuzzy set AR

For CoM of BE5:

FHD (BE5) = | 0.54-0.80 | + | 0.69-0.60 | + | 0.68 - 0.50 |+ | 0.40-0.40 | + | 0.25-0.10 |
= 0.68

For conflict of BE5:

FHD (BE5) =| 0.40-0.25 | + | 0.30-0.30 | + | 0.25 - 0.25 |+ | 0.0-0.30 | + | 0.25-0.25 |
= 0.45

Similarly, the fuzzy Hamming distance of CoM and conflicts for all the MCS(s) are obtained and illustrated in Table 3.

Table 3: CoM and Conflict Fuzzy Hamming Distance

	FHD of CoM	FHD of Conflict
{BE5}	0.68	0.45
{BE6}	0.50	0.30
{BE7}	0.40	0.55
{BE12}	0.70	0.45
{BE13}	0.90	0.60
{BE14}	0.85	0.30
{BE8, BE9}	0.80	0.20
{BE10, BE11}	0.65	0.50
{BE1, BE2, BE3, BE4}	1.05	0.40

The largest Hamming distance of CoM originates from MCS {BE1, BE2, BE3, BE4}, which concludes that the most critical MCS is {BE1, BE2, BE3, BE4} namely the combination of long queue and inefficiency of server and bandwidth management. Thus, a proper intelligent server management system should be figured out to optimise the web service, and this in turn will increase its reliability.

The Web administrators' assessments consistency can be obtained by comparing the real number N. For instance, according to another predefined real number $<\pm$N (say, N=0.5), the assessments of MCS BE7 and BE13 are not consistent and must be reassigned to ensure the non conflict judgement amongst the Web administrators.

4.1 Comparison with Standard Fuzzy Sets Approach

To evaluate Web administrators' judgements using the fuzzy set approach, the IVFS of the previous results is examined. Select the highest probability IVFS, and evaluate the maximum and minimum fuzzy sets:-

For minimum case:
WA1(BE5) = F
WA2(BE5) = F
WA3(BE5) = L
WA4(BE5) = F

The overall of the WA is obtained as follow:
$WA1 \cap WA2 \cap WA3 \cap WA4$
$= \{ F \cap F \cap L \cap F\}$
= 0.6/1 + 0.8/2 + 0.6/3 + 0.4/4 + 0.2/5

For maximum case:
WA1(BE5) = Av
WA2(BE5) = H
WA3(BE5) = Av
WA4(BE5) = F

The overall of the WA is obtained:
$WA1 \cap WA2 \cap WA3 \cap WA4$
$= \{ Av \cap H \cap Av \cap F\}$
= 0.4/1 + 0.5/2 + 0.6/3 + 0.4/4 + 0.3/5

The mean value of the minimum and maximum fuzzy sets is obtained as follow:
$WA1 \cap WA2 \cap WA3 \cap WA4$
= 0.5/1 + 0.65/2 + 0.6/3 + 0.4/4 + 0.25/5

Then, compare this fuzzy set with the threshold AR_CoM(BE5) and the distance is obtained.
FHD(BE5) = 0.6

From the above illustration, the conflictive measure of the evaluation is not known using the fuzzy set approach. The inconsistencies and incorrect assessments of the Web administrators are hardly to be identified despite the fuzziness of human judgement can be well modelled.

5.0 CONCLUSION

In this paper, web service failure analysis is done using a fault tree approach. Evidence set is used to capture the fuzziness, non-specificity and conflictive of human categorisation on basic event evaluation. The subjective evaluation of fault diagnosis for HTTP process is vital since it may provide another alternative reference guide to the web administrators about the MCS leading to the degradation of web service. In this specific case, it may eventually boost up intelligent request/response control as well as load minimisation in the future since the quality of web service is mainly downgraded by inefficient server management. However, the typicality of evaluation can be further improved if the conflict of all MCS(s) can be significantly reduced.

REFERENCES

[1] M. A. S. Guth, "A Probabilistic Foundation for Vagueness and Imprecision in Fault Tree Analysis", *IEEE Transactions on Reliability*, Vol. 40, No. 5, December 1991, pp. 563-571.

[2] Ronald R. Yager (Editor), *Advances in the Dempster-Shafer Theory of Evidence*. John Wiley & Sons, 1994.

[3] D. Singer, "A Fuzzy Set Approach to Fault Tree and Reliability Analysis", *Fuzzy Sets and Systems*, 34 (1990), pp. 145-155.

[4] L. M. Rocha, "Interval Based Evidence Sets", In, *B. Ayyub (Ed.). Proceedings of the ISUMA-NAFIPS'95*, pp. 624-629.

[5] Timothy J. Ross, Fuzzy Logic With Engineering Applications. McGraw-Hill International Edition, 1995.

[6] G. J. Klir and A. Ramer, "Uncertainty in the Dempster-Shafer Theory: A Critical Re-examination", *International Journal of General System*, 18, pp. 155-166.

[7] V. Ramachandran, N. P. Padhy, and S. R. Paranjothi, "Fuzzy Decision System for Recruiting Candidates", *National Conference in Industrial and Applied Mathematics*, India, 1994.

BIOGRAPHY

W. C. Chin is a research student in Multimedia University. His current research interests include fuzzy logic, fuzzy modelling application and Internet and web application.

C. W. Chong is a lecturer in FET Multimedia University. His current research interests are web security and auditing, image analysis and evidence sets.

V. Ramachandran was an associate professor in Multimedia University. His research areas include fuzzy logic application in power system, Internet and web application and networking.