

# New Evaluation Model on Web Site Credibility

Y.Y. Ma, F. Chen, Z.X. Sun

Key Lab of Broadband Wireless Communication and Sensor Network Technology  
 Ministry of Education  
 Nanjing University of Posts and Telecommunications  
 China

**Abstract**-In view of the existing web site credibility evaluation models failing to comprehensively consider the multiple attributes of trust,after learning the D - S evidence theory and fuzzy theory,a new web site credibility evaluation model is put forward in which the trust value is divided into historical trust value and the current trust value.The study shows that the model can track the change of user preferences in real-time as well as can meet the dynamic nature of universal environment,so the trust assessment is more reasonable and accurate.

**KeyWords**-evaluation model; d-s evidence theor; fuzzy theory

## I. RELATED TECHNOLOGY RESEARCH

Recently, many universities and IT companies at home and abroad have engaged in trust measurement research[1].The classic trust evaluation models[2,3]are Beth evaluation model, Josang model and Tang trust management model.On the basis of the three classic trust model, Literature[4] presents web site credibility evaluation model based on t-s fuzzy neural network by combining fuzzy theory with neural network theory,Literature[5] lists influence factors of web site credibility from the perspective of users, and then uses the method of multi-level fuzzy comprehensive evaluation model, Literature[6] proposes a trust model based on improved D - S evidence theory,in Literature[7],a dynamic model is set up for trust.Although there are a lot of trust evaluation models at present, common problems exist as follows:

- (1) Only consider one certain attribute of the trust without considering multiple attributes.
- (2) Be failure to realize that different types of web site should adopt different influence factors .

## II. COMPREHENSIVE EVALUATION MODEL ON WEB SITE CREDIBILITY

### A. Definition

Membership

If  $X=\{x_1,x_2, \dots,x_i\}$  is non-empty set,and if  $xz(z=1,2,\dots,t)$ is an element included in  $X$ ,and if the following mapping relationship is established for all  $xz$ :

$$\mu_A : X \rightarrow [0,1], x_z \mapsto \mu_A(x_z) \in [0,1]$$

,then  $T=\{(x_1/\mu_A(x_1),x_2/\mu_A(x_2),\dots,x_i/\mu_A(x_i))\}$  is fuzzy subset on  $X$ ,and  $\mu_A(x_z)$  reflects the degree to which  $xz$  belongs to  $T$ ,that is membership.

Interest group

Interest group is a "network" and people in the interest group have the same interest.Recommended information from

Interest group members is more valuable when users are evaluating the web site.

### B. Comprehensive Evaluation Model On Web Site Credibility

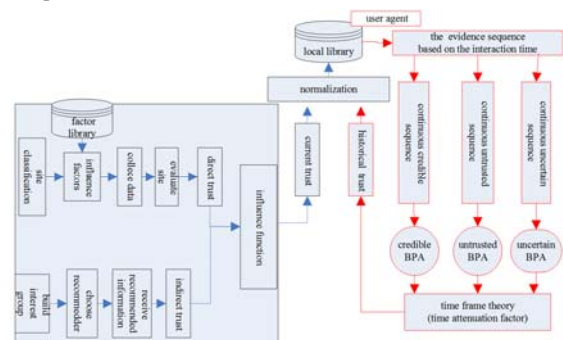


FIGURE I. COMPREHENSIVE EVALUATION MODEL ON WEB SITE CREDIBILITY.

The model presented in this paper is shown in fig.1.,D-S evidence theory and concept-tree theory are applied into the model,making the model meet subjective fuzziness of trust.At the same time,by the way of combining the historical trust with the current trust,the model can reflect trust changes dynamically over time.

### C. Computing Historical Trust Value Of Websites

Referencing literature [6] in this paper,D-S evidence theory is applied in calculating the historical trust value.As shown in the red part of fig.1.,the local library stored all users' assessing information on website j as long as the user interacts with website j and all assessing information is present in the form of evidence sequence[9],from which continuous credible service,continuous untrusted service and continuous uncertain service evidence subsequence are extracted using continuous sequence extraction algorithm.

Sequence extraction algorithm

Assuming that quality of the kth service when user i interacts with website j is  $S_{ijkt}$ ,and the length of the the kth continuous subsequence is  $length_k$ ,and the threshold of subsequence length is  $L, 0 < \alpha_i, \beta_i < 1$ .If  $\beta_i < S_{ijkt} < \alpha_i$ , the service is credible service  $\{T\}$ ;else if  $0 < S_{ijkt} < \alpha_i$ , the service is untrusted service  $\{-T\}$ ;else the service is uncertain service  $\{T, -T\}$ .

After extracting the subsequence, credible service set, untrusted service set and uncertain service set are formed,then some measures need to be taken: if  $length_k < L$ ,then the

subsequence should be removed from the set,else the subsequence is retained in the set.

#### Application of time frame theory

Assuming that  $t_k$  represents the time when the  $k$ th subsequence happens ,  $t_{current}$  represents the current time and  $t_k' = t_{current} - t_k$ .In order to reward the reliable service and punish malicious service, the basic probability function is set up. In general, the  $t_k'$  is smaller,the users believe the subsequence more,and the length

of the subsequence is longer,the users believes the subsequence more.Therefore, when calculating trust values,time factor and the length of subsequence must be taken into account.

epn (1) is BPA function for continuous credible subsequence,in which time and length of the subsequence are mapped.

$$\left\{ \begin{array}{l} m(T)_{ij} = \frac{\sum_{k=1}^m f_{adek,ts} * length_{k,ts}}{N} \\ f_{adek,ts} = \rho \end{array} \right. \quad (1)$$

In epn(1), $f_{adek,ts}$  is time attenuation factor,and  $0 < \rho < 1$ .  $t_{\alpha + length - 1}$ , the time of the last service,can be viewed as the reference . $N$  is the total service times.

In order to punish the continuous untrusted service or continuous uncertain service,the basic probability function in literature [6] was improved,shown as epn (2).In epn (2), the reciprocal of time decay function and the length of the continuous untrusted service are viewed as penalty factor.

$$m(-T)_{ij} = \frac{\sum_{k=1}^z 1 / \sqrt{f_{adek,fs} * punish * length_{k,fs}}}{N} \quad (2)$$

Being different from the epn (1),  $t_{first}$ , the time of the first service,can be viewed as the reference.Besides that,another penalty factor exists.

$$punish = \varpi \tan(c_j - \varpi) \quad (3)$$

$c_j$  is the probability that service is transformed into untrusted service from trusted services.

$$c_j = \begin{cases} (m-1) / n, & \text{if } t_n = t_{\alpha m + length_m - 1}, \text{ and } length_m \geq 2 \\ m / n, & \text{otherwise} \end{cases} \quad (4)$$

$m$  represents for the number of continuous credible subsequence,  $n$  represents the length of the evidence sequence.

#### D. Computing The Current Trust Value Of Websites

Learning from the realistic society,in this paper,the current trust value can be divided into direct trust value and indirect trust value.In order to reflect that how the indirect trust effects the direct trust,the paper puts forward influence function. Direct trust based on the concept tree

Many factors can affect users' evaluation for the website.These factors can be classified and each class can be divided into many smaller classes, the same size classes are in the same layer,that is concept tree.After the weight of each influence factor is determined,calculating the trust value step by step from leaf nodes to root node,the final trust vector  $V$  can be got.

Before evaluating a site,the factors which can influence the site credibility must be known,in this paper,take a news site for example,the influencing factor are  $X = \{\text{service,technology,beauty,safety}\}$  and the trust is divided into five grades,they are  $T = \{\text{fully trust,trust,general trust,part of the trust,distrust}\}$ .

Given the concept of membership,the vector made up of many membership values reflecting the degree to which  $x_i$  (from the set  $X$ ) belongs to  $T$  is called trust vector.And the user can determine the weight of each factor according to the actual situation.Then comprehensive assessment expression is  $V = W.R$ .

#### Calculation of indirect trust value

When evaluating a website,the user can first ask people from interest group for the recommend information and then sort out these information as a reference.A calculation formula of influence function is given below:

$DTval(i,j)$ ,the direct trust reflecting how user  $i$  trusts the site  $j$ .

$RTval(k,j)$ ,the recommend trust value from user  $k$  reflecting how user  $k$  trusts site  $j$ .

$U_i$ ,the degree to which user  $i$  insists on direct trust value of site  $j$ .

$U_k$ ,the degree to which user  $k$  insists on direct trust value of site  $j$ .

The function  $f(i,k)$  reflecting how user  $k$  affect user  $i$  is as follows:

$$f(i,k) = \begin{cases} -U_k * (U_k - U_i), & \dots \dots U_k > U_i \text{ and } DTval(i,j) > RTval(k,j) \\ U_k * (U_k - U_i), & \dots \dots U_k > U_i \text{ and } RTval(k,j) > DTval(i,j) \\ 0, & \dots \dots U_k \leq U_i \text{ and } DTval(i,j) = RTval(k,j) \end{cases} \quad (5)$$

Given the credit factor  $p$ , the normalization formula is shown as below:

$$T = p * T_{current} + (1-p) * T_{historic} \quad 0 < p < 1 \quad (6)$$

The credit factor  $p$  can be given the appropriate value according to actual condition.

#### E. The Simulation

The simulation is carried out in Matlab and is carried out on two websites in order to verify the validity of the model by comparing with each other.

#### F. The Simulation Experiment On Historical Trust Value

When calculating the historical trust value,10 evidence sequences from the local library are got,they are: $S_1 = (0.4,0.3,0.4,0.8,0.2,0.1,0.2,0.7,0.4,0.3)$ ,

$$S_2=(0.6,0.5,0.8,0.7,0.3,0.7,0.8,0.9,0.7,0.9).$$

In experiment,  $\alpha = 0.4$ ,  $\beta = 0.7$ ,  $\rho = 0.8$ ,  $\varpi = 0.5$ . Trust values change respectively show as in fig.2.,the decline part represents continuous untrusted sequence, or continuous uncertain sequence and the curve part represents continuous trusted sequence.The result show that final credibility of website 1 is about 0.1, final credibility of website 2 is about 0.8.

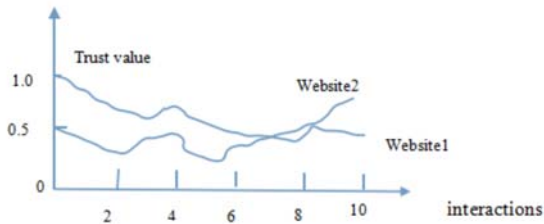


FIGURE II. THE HISTORICAL TRUST OF WEBSITE.

### G. The Simulation Experiment On Current Trust Value

In fig.3.,the inputs are influence factors and the outputs are website evaluation result.In fig.4.,set membership functions for each variable, five membership functions curve respectively represent "distrust", "part of trust", "trust", "very trust", "totally trust" by adjusting the parameters of the membership function.

Normalize the historical trust value and the current trust value and the final value is gotten:T1=0.43,T2=0.556.Here,p=0.6.Also,p can be other value according to the actual situation.

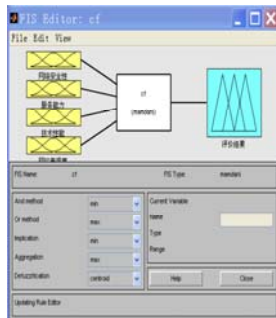


FIGURE III. SET THE INPUT AND OUTPUT.



FIGURE IV. SET MEMBERSHIP FOR VARIABLES.

Normalize the historical trust value and the current trust value and the final value is gotten:T1=0.43,T2=0.556.Here,p=0.6.Also,p can be other value according to the actual situation.

From the above simulation, current trust value of website 1 is higher,but the final value decreased after combining with the historic trust value and website 2 is on the contrary.By comparing the model with the models in literature [4] and literature [6],we can see that the models in literature [4] and literature [6] are one-sided because they only consider the historical trust or only consider the current trust. But the model here integrated fusion historical trust and the current trust, the evaluation results are more comprehensive.

### III. CONCLUSIONS

Fuzzy evaluation is a kind of comprehensive evaluation,and it is a kind of fuzzy evaluation conclusion combining many people's opinion by a suitable algorithm.The trust evaluation model established in this paper well solved the problem how to get site trust in the pervasive computing,providing great convenience for the Internet users.

### REFERENCE

- [1] P.Liu &X.Liu.,Trust management research.Computer engineer and application.(32)pp,39-42,2004.
- [2] Balder Ten Cate,Enrico Franconi & Inanc Seylan.Beth definability in expressive description logics. Proceedings of the Twenty-Second international joint conference on Artificial Intelligence.Jul,2011.
- [3] J.J.Chin,S.Y.Tan & S.H.Heng. On the security of a modified Beth identitybased identification scheme.Information Processing Letters(113)pp,14-16
- [4] Y.M.Hou & C.S.Wang.The research on software reliability prediction model.
- [5] D. Zhao & N Zhou. Research on enterprise web site credibility from the network consumer.School of Information Management,Wuhan University,2009.
- [6] L.Zhang&J.W.Liu.Trust evaluation model based on the improved D - S evidence theory. Journal of communication, (7),2013.
- [7] Q.Liu &Y.Zhong,Study on China's Internet News ,May,2011.
- [8] Aliyena,N&Ismayilov,E.The analysis of how the choice of membership functions influences the quality of recognition system.2012International Symposium on Innovations in Intelligent Systems and Applications.pp,1-3,2012.
- [9] Edwards,G.J&Taylor,C.J.Improving identification performance by intergrating evidence from sequences.1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.(1),1999.
- [10] Y.H.Hu &F. Wu,Mining multi-level time-interval sequential patterns insequence databases.2010 2th International Conference on Software Engineering and Data Mining.(2)pp,416-421,2010.
- [11] Lulic,D&Marceta,Z.Real-timeframe-freezingdetectionsystemevaluation .Tele-communications Forum,(19)pp,1143-1146,2011.
- [12] J. Liu &X.YDeng.Multi-attribute decision-making method based on intervalvalued intuitionistic fuzzy sets and D-S theory of evidence.Control and Decision Conference,(24)pp,2651-2654,2012.
- [13] J.Wan &T.M.Jiang.A comprehensive assessment method for reliability enhacement testing based on D-S theory of evidence.2010 IEEE International Conference on Industrial Engineering and Engineering Management.pp,2357-2378,2010.
- [14] C.L. Xia&X.H.Jiang.Notice of Retraction Dynamic item-based recommendation algorithm with time decay.2010 6th International Conference on Natural Computation.(1)pp,242-247,2010.
- [15] Y.J.Weng&L.Shi.Price spike forecasting using concept-tree approach based on cloud model.ISECS International Colloquium on CCCM,(2)pp,352-355,2009.

- [16] H.Z.Liu&B.Hong.A novel vector space model for tree based concept similar- ity measurement.2010 2th IEEE International Conference on Information Management and Engineering.(2)pp,144-148,2010.