A FUZZY FORECASTING MODEL FOR NETWORK TRAFFIC

Onianwa, U. Charles;' Imianvan, Anthony; and Enaholo, Emmanuel

Abstract

A useful model for network accessibility and planning is x-rayed. The fuzzy inference system consequent on this paper, incorporates numerical data to derive server traffic index based on four workload parameters. The fuzzy model proposed illustrates (he state transitions of server resource utilization based on expert linguistic evaluation of stationary transition probability. Fuzzy rules are established by taking into account different combinations of fuzzified traffic values with regard to predefined membership functions. A steady state algorithm is applied to explore the convergence of server traffic after n transition period(s).

Keywords: Network Traffic, Fuzzy Logic, Forecasting Models.

Introduction

Traffic distribution in homogeneous and heterogeneous network servers is growing in leaps and bounds. To provide highly reliable service, workload handling and forecasting is paramount. Additionally, traffic burstiness in peak hours at certain time scales tends to crawl servers services. In this paper, a fuzzy model is formed to represent the transitions of the server resource utilization and to forecast the future server workload stale based on expert linguistic evaluation of state transition matrix.

Server utilization index is obtained **through** a fuzzy inference system that employs four servers' performance metrics: server's latency (millisecond), service rate (connection per second), throughput (megabits per second) as well as the occurrence of error (error per second). Further, the fuzzy rules are established by taking into account the different combinations of fuzzified workload metrics with regard to pre-defined membership functions. A fuzzy algorithm is utilized to explore the steady states of the server traffic after n transition period.

Section 2, provides a review of previous research relevant to our study. In section 3, we derive a fuzzy forecasting model for servers' traffic whilst section 4 provides numerical estimates of server resource utilization. Section 5 draws our study to a climax.

Literature Review

The failure of multiserver system to provide scalability and flexibility of network load traffic [4, 5] as well as the existence of traffic burstiness at peak time scale [2,3] tends to crawl server services. Against this backdrop, network server traffic handling and forecasting is vital.

There abounds lots of statistical and artificial intelligence forecasting systems. The statistical approach comprises moving average, exponential smoothing, lime series regression and economic modeling **[1J.** Artificial Intelligence forecasting approach on the other hand, include use knowledge engineering, expert system, fuzzy logic, neural network and genetic algorithm. In both, the numerical data collected are vague due to fluctuation of server workload distribution. Therefore, fuzzy control appears to be a better candidate. Lots of fuzzy forecasting models abound in the horizon. Nie [6] provides a fuzzy scheme for time series prediction whereas Wang and Mendel (7) provide a guide to developing fuzzy forecasting models.

Fuzzy Modeling of Server Traffic

For the purpose of this paper, let server utilization index (I) be governed by a 4-tuple performance metrics such as:

$I = \{M,, M_2, M_3, M_4\}$

where:

Mi = server's service rate (connection per second).

Mj= server's throughput (megabits per second).

M_3 = server's latency (milliseconds).

M_4 = server's error frequency (error per millisecond).

Also, let A,, A_2, A_3, A_4 represent the fuzzy sets for M,, M_2, M_3, as well as M_4 over time window Tc, where c is the integer number.

The fuzzifications of the four medics values are based on predefined membership functions (Table 1). At a specific instant of time, all the measured data for M,, M_2, M,, and M_4 are fuzzified by membership function to their associate fuzzy sets A,, A_2, A,, A_4.

For example, the workload metrics are fuzzified into four linguistic spaces as follows:

A, = {Low, Medium, High}

A_2 = {Small, Fair, Large}

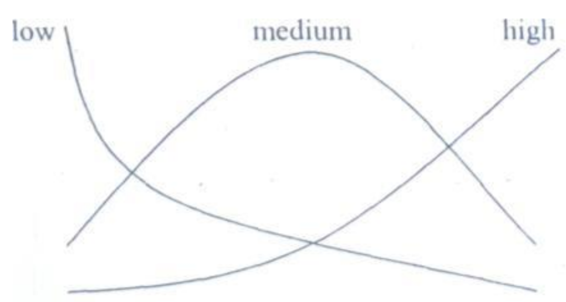
A_3 = {Slow, Moderate, Fast}

A_4 = {Not frequent, Moderate, Very frequent}

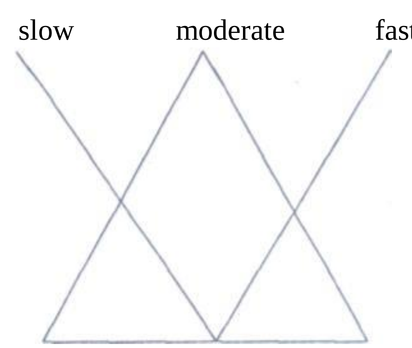
Table 1 also contains the rules that decide the utilization values. Further, fuzzy inference is used as a max-min composition in the calculation of server's utilization value. Fuzzy inference with predefined rules will integrate all the 4-tuples intensity that finalizes a server's utilization states with four states such as *Very Critical*, *Critical*, *Normal* and *Insignificant.* Figure 1 illustrates the linguistic input and output of the model system.

IF							THEN
Al	O	A2	O	A3	()	A4	Utilization State
Low	And	Small	And	Slow	And	None	Not Significant
Low	And	Small	And	Moderate	And	Not Frequent	Not Significant
Low	And	Small	And	Moderate	And	Moderate	Not Significant
Low	And	Small	And	Moderate	And	Very Frequent	Normal
IF							THEN
Al	O	A2	O	A3	O	A4	Utilization State
High	And	Large	And	Moderate	And	Moderate	Extremely Critical
High	And	Large	And	Moderate	And	Very Frequent	Extremely Critical
High	And	Large	And	Fast	And	Not Frequent	Extremely Critical
High	And	Large	And	Fast	And	Moderate	Extremely Critical
High	And	Large	And	Fast	And	Very Frequent	Extremely Critical

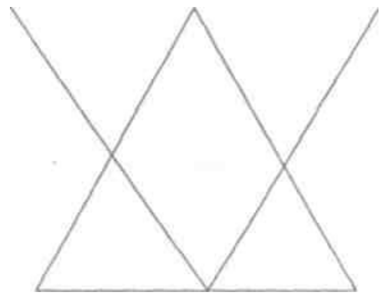
Table 1: Fuzzy Rules



(a) Input variable: "throughput".

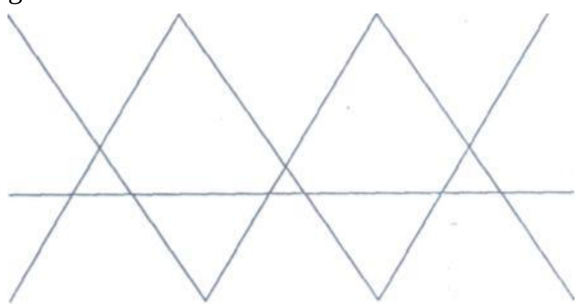


(b) Input variable: "service rate".



(c) Input variable: "latency".

(d) Input variable: "error"



(e) Output variable "Utilization Stale"

Figure 1: Parameters and Associate Memberships Function.

Let a discrete state space (S_b S_2, S_3,...,S,,) represent a finite process in discrete time. A finite fuzzy process similar to a stochastic process can be established based on the following condition:

a. The transition probabilities in a finite square n x n matrix P with the following structure, which is known as fuzzy state transition matrix of a fuzzy process:

M = (my)

$$
= \begin{array}{c c}
 & \begin{array}{cccc} S_1 & S_2 & & S_3 \end{array} \\
\begin{array}{c} S_1 \\ S_2 \\ \\ S_n \end{array} & \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ m_{21} & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & & \vdots \\ m_{n1} & m_{2n} & \dots & m_{nn} \end{bmatrix}
\end{array}
$$

Where niy represents the grade of membership of the transition going from state I to state j which confines in 0 < m,, < 1. b. Generally, the fuzzy state $U_X^{(M}$ at each time interval is a row vector:

$$U_x^{(k)} = (V_{i1}^{(k)}, V_{i2}^{(k)}, \dots, V_{in}^{(k)})$$

Where k denotes the instant of time. V denotes the grade of membership of sets with respec :o fuzzy set X.

;. For a row vector with initial time, k = 0 is known as the initial state designator of X with th following form:

$$U_x^{(0)} = (V_{i1}^{(0)}, V_{i2}^{(0)}, \dots, V_{in}^{(0)})$$

Hence, the state designator of X at time T = n can be obtained by the following equation:

$$U_x^{(n)} = U_x^{(0)} \bullet M^{(n)} = U_x^{(0)} \bullet \mid m_{ij}^{(n)} \mid)$$

Where, the multiplication between the initial designator row vector and the fuzzy transition trix involve and intersection basics operations.

Numerical Estimates of Network Traffic

The existence of imprecision and vagueness in a server's resources utilization is represented by assigning the linguistic variable as the utilization index. Consideration of a network server that delivers few types of items such as images, software, audio and video clips, as well as some formatted document as a monitoring target analyzes the four predefined metrics. By implementing the Fuzzy Inference System (FIS), network administrator compromised the server's utilization status as illustrated in Table 2.

States	Server's **utilization State**
Extremely Critical (1)	> 0.600
Critical (2)	0.400 - 0.600
Normal (3)	0.200 0.400
Insignificant (4)	< 0.200

Table 2: Utilization States References

Table 3 shows the initiated predefined traffic parameters. From Table 2, the occurrence of each state within the time period is counted and their normalized values are determined as illustrated in Table 4.

Average **Time period in an hour (3 minutes)**	C/s	M/s	**t/ms**	e/ms	**FIS output**	**Server utilization state.**
Time 1	5.5	2.0	3	0.001	0.28	Normal (3)
Time 2	12.5	3.5	4	0.005	0.38	Normal (3)
Time 3	18.3	4.0	16	0.002	0.57	Critical (2)
Time 18	15.0	3.0	15.0	0.0 06	0.51	Critical (2)
Time 19	1.0	5.0	5.0	0.0 05	0.33	Normal (2)
Time 20	4.0	3.0	5.0	0.0 . 01	0.31	Normal (1)

Table 3: Network Server's Traffic States Determination

States	**Number of occurrence**	**Normalized value**
Extremely critical (1)	**6**	0.75
Critical (2)	**8**	1.00
Normal (3)	**4**	0.50
Not significant (4)	**2**	0.33

Table 4: Normalized Value of State Occurrence

Further, the state transition matrix is determined by experts' subjective evaluation by referring to Table 4 on the distribution and changes of server utilization index. We assume that the state transition matrix and initial designator vector for this specific case are given as follows:

11	VII	M	L
V	EH	L	VI
M	L	M	L
L	VL	M	L

With discrete state space (1,2, 3,4), and$U_x^{(>)}$ = [VII Ell M L]

The state space diagram is illustrated in Figure 2.

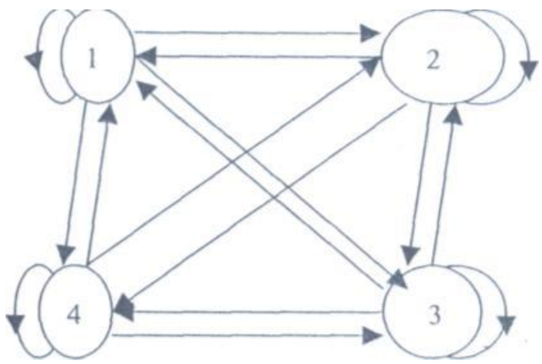


Figure 2: State Space Diagram

The server resource utilization may occur due to the users implosive demands or variability of network condition. The experts' subjective judgments, the fuzzy grades of membership of linguistic variable Low, Moderate, High and Extremely High are defined by the fuzzy set below:

Low (L) = 0.9/0.1 + 0.7/0.2 + 0.4/0.3 + 0.2/0.4 + 0.1/0.5

Moderate (M) = 0.3/0.2 l 0.5/0.3 + 1.0/0.4 + 0.7/0.5 + 0.2/0.6

High (H) = 0.2/0.5 + 0.7/0.6 + 1.0/0.7 + 0.5/0.8 + 0.2/0.9

Extremely High (EH) = 0.6/0.6 + 0.7/0.7 + 0.8/0.8 + 0.9/0.9 + 1.0/1.0

According to Zadeh[8], the fuzzy sets for very low, very medium, very high and very extremely high are interpreted as:

Very Low (VL)= low x low = 0.81/0.1 + 0.49/0.2 + 0.16/0.3 + 0.04/0.4 + 0.01/0.5

Very Medium (VM) = 0.09/0.2 + 0.25/0.3 + 1.0/0.4 + 0.49/0.5 + 0.04/0.6

Very High (VH) = 0.04/0.5 + 0.49/0.6 + 1.0/0.7 + 0.25/0.8 + 0.04/0.9

Very Extremely High (VEH) = 0.36/0.6 + 0.49/0.7 + 0.64/0.8 + 0.81/0.9 + 1.00/1.0

Using Zadeh[8] decision principle, every slate transition is selected for I' for time T = 2. The multiplication of the fuzzy state transition matrix I'^2 is illustrated in the tree diagram in Figure 3. The tree diagram will demonstrate the status after 2 steps given that the fuzzy process started in state 1, state 2 and state 3 respectively. Consequent on Figure 3, the I'^z is defined as follows:

$$\begin{vmatrix} H & VII & M & L \\ VM & EH & M & VL \\ M & M & M & L \\ M & L & M & L \end{vmatrix}$$

Therefore, for time T = 2, the state designator $U_x^{<2)}$ is obtained as:

$u,^w$- $W^{0)}$ • P^2

$$[VH \quad EH \quad M \quad L] \begin{Vmatrix} H & & M & L \\ VM & EH & M & VL \\ M & M & M & L \\ M & L & M & L \end{Vmatrix}$$

And, based on basic union and intersection operation, the state designator $U_x^{,2)}$ is obtained as:

$U_x^{2)}$ =[u,,,u,2,u13,uM] = [II VII M L]

where:

U,, = (VII $_A$H) $_v$ (EH $_A$ VM) $_v$ (M$_A$M) $_v$ (L$_A$M) =11 $_v$ VM $_v$ M $_v$L =11

$u_{12} \;=\; (VH_{\wedge} VH)\ _{v}\ (EH_{\wedge} EH)\ _{v}\ (M_{\wedge} M)\ _{v}\ (L_{\wedge} L) = VH$
$u_{13} \;=\; (VH_{\wedge} M)\ _{v}\ (EH_{\wedge} M)\ _{v}\ (M_{\wedge} M)\ _{v}\ (L_{\wedge} M) = M$
$u_{14} \;=\; (VH_{\wedge} L)\ _{v}\ (EH_{\wedge} VL)\ _{v}\ (M_{\wedge} L)\ _{v}\ (L_{\wedge} L) = L$

For quantitative measurement of the fuzzy linguistic term, rough estimate of the $U_x^{(2)}$ can be obtained by selecting the optimal grade of memberships. This optimal grade indicates the highest possibility of server utilization in state 1 and 2 after two transitions. The quantitative calculation becomes:

$$U_x^{(2)} = [\,H \quad VH \quad M \quad L\,]$$

$$= [$$

0.2/0.5 + 0.7/0.6 + **1.0/0.7** + 0.5/0.8 + 0.2/0.9
0.04/0.5 + 0.49/0.6 + **1.0/0.7** + 0.25/0.8 + 0.04/0.9
0.3/0.2 + 0.5/0.3 + **1.0/0.4** + 0.7/0.5 + 0.2/0.6
0.9/0.1 + 0.7/0.2 + 0.4/0.3 + 0.2/0.4 + 0.1/0.5

]

$$= [\,0.7 \quad 0.7 \quad 0.4 \quad 0.1\,]$$

Furthermore, if the fuzzy transition matrix M is converging to a limit when T = n, where

$$n \rightarrow \infty, \text{ all the rows of } \lim_{n \rightarrow \infty} [U_x^{(n)}]$$

are equivalent to the steady-state fuzzy designator vector $U^{\infty} = [u_1, u_2, u_3, u_4]$

The fuzzy steady-state designator vector $U^{\infty} = [u_1, u_2, u_3, u_4]$ can be determined using:

$$\bigcup_{n} (u_{i \wedge} m_{ij}) = u_i$$

where i and j =1,2,…,n (states).

Thus, by considering the four states vector, the four components of the fuzzy steady-state row vector becomes:

$u_1 = (u_{1\wedge} m_{11})\ _{v}\ (u_{2\wedge} m_{21})\ _{v}\ (u_{3\wedge} m_{31})\ _{v}\ (u_{4\wedge} m_{41})$
$u_2 = (u_{1\wedge} m_{12})\ _{v}\ (u_{2\wedge} m_{22})\ _{v}\ (u_{3\wedge} m_{32})\ _{v}\ (u_{4\wedge} m_{42})$
$u_3 = (u_{1\wedge} m_{13})\ _{v}\ (u_{2\wedge} m_{23})\ _{v}\ (u_{3\wedge} m_{33})\ _{v}\ (u_{4\wedge} m_{43})$
$u_4 = (u_{1\wedge} m_{14})\ _{v}\ (u_{2\wedge} m_{24})\ _{v}\ (u_{3\wedge} m_{34})\ _{v}\ (u_{4\wedge} m_{44})$

Finally, to solve the arising row vector, employ the following algorithm:

Step 1: Initialize the four components u_1, u_2, u_3, u_4.

Step 2: Let σ be the threshold limit. We calculate each component of the vector using: $/ R_x - L_x / < \sigma$

where x = number of state.

R and L are the right-hand-side and the left-hand-side of the four equations respectively.

Step 3: Terminate the computation if the desired σ has been fulfilled. Else, randomly generate the values of u_1, u_2, u_3, u_4. and goto step 2.

T = 0	T = 1	T = 2		$\min(m_{ij})$	$\max\min(m_{ij})$
		(H)	1	$m_{11} = H$	
	(H) 1	(VH)	2	$m_{12} = H$	
		(M)	3	$m_{13} = M$	
		(L)	4	$m_{14} = L$	$m_{11}^{(2)} = H$
		(VM)	1	$m_{11} = VM$	
	(VH) 2	(EH)	2	$m_{12} = EH$	
		(L)	3	$m_{13} = L$	
		(VL)	4	$m_{14} = VL$	$m_{12}^{(2)} = VH$
1	(M) 3	(M)	1	$m_{11} = M$	
		(L)	2	$m_{12} = L$	
		(M)	3	$m_{13} = M$	
		(L)	4	$m_{14} = L$	$m_{13}^{(2)} = M$
	(L) 4	(L)	1	$m_{11} = L$	
		(VL)	2	$m_{12} = VL$	
		(M)	3	$m_{13} = L$	
		(L)	4	$m_{14} = L$	$m_{14}^{(2)} = L$
		(H)	1	$m_{21} = H$	
	(VM) 1	(VH)	2	$m_{22} = EM$	
		(M)	3	$m_{23} = L$	
		(L)	4	$m_{24} = VL$	$m_{21}^{(2)} = VM$
		(VM)	1	$m_{21} = H$	
	(EH) 2	(EH)	2	$m_{22} = EH$	
		(L)	3	$m_{23} = M$	
		(VL)	4	$m_{24} = L$	$m_{22}^{(2)} = EH$
2	(L) 3	(M)	1	$m_{21} = M$	
		(L)	2	$m_{22} = M$	
		(M)	3	$m_{23} = M$	
		(L)	4	$m_{24} = L$	$m_{23}^{(2)} = M$
	(VL) 4	(L)	1	$m_{21} = L$	
		(VL)	2	$m_{22} = L$	
		(M)	3	$m_{23} = L$	
		(L)	4	$m_{24} = L$	$m_{24}^{(2)} = VL$

T = 0	T = 1	T = 2	min(m_{ij})	maxmin(m_{ij})
	(VH) 1	(H) 1	m_{31} = H	
		(VH) 2	m_{32} = EM	
		(M) 3	m_{33} = L	
		(L) 4	m_{34} = VL	$m_{31}^{(2)}$ = M
	(EH) 2	(VM) 1	m_{31} = H	
		(EH) 2	m_{32} = EH	
		(L) 3	m_{33} = M	
		(VL) 4	m_{34} = L	$m_{32}^{(2)}$ = M
3	(EM) 3	(M) 1	m_{31} = EM	
		(L) 2	m_{32} = EM	
		(M) 3	m_{33} = EM	
		(L) 4	m_{34} = L	$m_{33}^{(2)}$ = M
	(L) 4	(L) 1	m_{31} = L	
		(VL) 2	m_{32} = L	
		(M) 3	m_{33} = L	
		(L) 4	m_{34} = L	$m_{34}^{(2)}$ = L
	(VH) 1	(M)	m_{41} = H	
		(L)	m_{42} = EM	
		(M)	m_{43} = L	
		(L)	m_{44} = VL	$M_{41}^{(2)}$ = M
	(VEH)) 2	(VH) 1	m_{41} = H	
		(EH) 2	m_{42} = EM	
		(L) 3	m_{43} = M	
		(VL) 4	m_{44} = L	$m_{42}^{(2)}$ = L
4	(H) 3	(M) 1	m_{41} = H	
		(L) 2	m_{42} = H	
		(M) 3	m_{43} = EM	
		(L) 4	m_{44} = L	$m_{43}^{(2)}$ = M
	(L) 4	(L) 1	m_{41} = L	
		(VL) 2	m_{42} = L	
		(M) 3	m_{43} = L	
		(L) 4	m_{44} = L	$m_{44}^{(2)}$ = L

Conclusion

In this paper, a fuzzy system has been presented for network server traffic index. Our fuzzy forecasting scheme, integrate human experience and judgment with numerical data collected to forecast incoming server traffic. A steady state algorithm is applied to explore the convergence of server resource utilization after n transition period(s). The forecasting scheme obtained provides a useful reference for future server traffic.

References

Chiraphadhanakul, S.; Dangprasert, P.; and Avatchanakom, V. (1997). Genetic Forecasting Algorithm with Financial Applications. *Intelligent Information Systems, ISS'97 Proceedings,* Pp. (174-178).

Crovella, M. and Bestavros. A. (1995), *Explaining World Wide Web Traffic Self-Similarity.* Tech Rep. BUCS-TR95F-015, Boston University, Boston MA 02215.

Leiand, W. and Taqqu, M., (1993). On the Self-similar Nature of Ethernet Traffic. In *Proceedings of SIGCOMM'93.*

Michele Colajanni and Philip, S. Yu, (1997) Scheduling Algorithms for Distributed Web Servers, *Proc. ICDCS'97,* Baltimore, MD, Pp. (169-176).

Michele Colajanni, (1998). Dynamic Load Balancing in Geographically Distributed Heteiogeneousiy Web Servers, In *Proceedings of IS" International Conference on Distributed Computing Systems,* Pp. (295-302).

Nie, J. (1994). A Fuzzy-Neural Approach To Time-Series Prediction, In *Proceedings of IEEE International Conference on Neural Network.* New Jersey: (Pp. 3164-3169).

Wang, L. X. and Mende,l J. M., (1992) Generating Fuzzy Rules By Learning From Example, IEEE Transaction Systems. *Man and Cybernetics,* 22, Pp (1414-1427).

Zadeh, L. A. (1975). Linguistic Approach and Its Applications in Decision Analysis. *Directions in Large Scale Systems.* London: Plenum Press (Pp. 339-357).