Elasticity Analysis of Web Browsing Behavior of Users

Diwakar Shukla
Department of Mathematics and Statistics, Sagar University, Sagar M. P. 470003, India.
Associate Member, DST-Centre for Interdisciplinary Mathematical Sciences, Banaras Hindu University, Varanasi, U.P.
Email: diwakarshukla@rediffmail.com
Sharad Gangele
Department of Computer Science, M.P. Bhoj (Open) University, Kolar Road, Bhopal, M. P., India.
Email: sharadgangele@gmail.com
Rahul Singhai
International Institute of Professional Studies, Devi Ahilya University. Indore, (M.P.), India.
Email: singhai_rahul@hotmail.com
Kapil Verma
Department of Computer Science, M.P. Bhoj (Open) University Kolar Road Bhopal, M.P., India.
Email: kapil_mca100@rediffmail.com

ABSTRACT

There are many browsers available in the market but only a few are liked and used by many. Every browser has a predefined level of popularity in the market. This paper considers the setup of only two browsers installed in a computer system and a user prefers to any one of them and if fails, switches to the other one. Elasticties of browser sharing are derived, index formation and graphical study are performed to compare the simulated output.

Keywords: Markov Chains (MC), Transition Probability Matrix (TPM), Quality of Service (QOS), Browser Failure (BF).

Date of Submission: June 15, 2011 Date of Acceptance: July 23, 2011

I. INTRODUCTION

For internet surfing many browsers are available in the market which can be used to connect a particular website as required by users. Every browser bears a failures rate related to connectivity. It occurs because of internal coding structure of browser, command execution pattern and speed of search engine. It is termed as browser failure (BF) and has some probability of occurrence every moment. Every browser serves to user and, therefore, has certain quality of service (QOS) also. Shukla and Singhai (2010) examined the web browsing behavior of user by a Markov chain model. This paper extends the same approach under elasticity property based approach and analysis.

The elasticity is a property defined as rate of change of one parameter with respect to another subject to condition when all other parameters are constant. This kind of study helps in decision making regarding continuity or stopping the process. Iso-curves are those showing the output when a level of input is fixed. In other words iso-curves are possible adjustments between two parameters. Index formation is a technique that shows the impact of relative change. This paper incorporates Elasticities analysis, iso-curve analysis and index formation on the contribution of Shukla and Singhai (2011).

II. A REVIEW

Catledge and Pitkow (1995) presented a contribution on characterization of browsing strategies. Pirolli. and pitkow (1996) suggested usable structure for web in light of many users. A similar study is performed due to pitkow (1997) regarding search of reliable usage data on www. Naldi (2002) presented Markov chain model based study in a multioperator environment. The detailed distribution of basics of Markov Chain Model is in Medhi (1991) and web browsing details are in Han and Kamber (2001). Shukla et al. (2007) discussed stochastic model for space decision switches for complete network. Shukla et al. (2007 a, b, c) suggested the use of Markov chain model in networking and operating system analysis. Shukla and Jain (2007) used Markov chain model for the analysis of multilevel Queue Scheduler in operating system. Shukla and Singhai (2010 a) discussed traffic analysis of massage flow in three crossbar architecture space division switches. Deshpande & Karypis (2004) discussed selective Markov chain model for predicting webpage access. Shukla et al. (2010 a, b, c, d, e, f, g, h,) discussed different aspects on Markov chain model in determining the system behavior. Shrivastava et al. (2000) presented a thought oriented contribution on the web page mining and application of

Volume: 03; Issue: 03; Pages:1162-1168 (2011)

usage patterns from web data. Shukla et al. (2011) is also a useful contribution.

Objective of this content are:

1. To examine Elasticities of browser blocking,
2. To study iso-curve of parameters,
3. To develop index for performance indications.

Let $\{X_n, n \geq 0\}$ be a Markov chain on state space $\{C, Q, B_1, B_2, S\}$ see Shukla and Singhai (2011) where

State C: represents connecting state.
State Q: user quitting from the proce
State B_1: user attempts to surf through browser B_1.
State B_2: user attempts to surf through browser B_2.
State S: success for connectivity and surfing.

The $X^{(n)}$ denotes the position of random variable X in the state space at the n^{th} browser connectivity attempt made by a web user.

III. ASSUMPTIONS FOR USER BROWSING BEHAVIOR [see Shukla and Singhai (2011)]

(1) The user attempts for dial up connection to use Internet. If the connection is not established, user quits with the probability Pc.

(2) When connection is made user chooses any one of browsers B_1, B_2 with the probability p and 1–p respectively.

(3) User navigates to any one browser at a time when successfully opened.

(4) Browser B_i $(i=1, 2)$ failure occurs due to non-opening of any site through browser B_i .Then user either quits (with probability p_q) or switches to the next browser.

(5) Switching between browsers are on attempt by attempt basis *(n=1, 2, 3….)*.

(6) Initial preference for a browser is based on quality of services and variety of facility features are contained in both browsers.

(7) Failure probability of a browser B_1 is b_1 and of B_2 is b_2.

(8) Transition probability of surfing through B_1, being completed in a single attempt is $(1 – b_1)$.

(9) Absorbing state (transition from a state to itself) probability is 1. No further transition from this state occurs.

Under these assumptions user's browsing behavior is discussed in this paper by a Markov Chain Model (see fig.3.1) in which the transition probabilities are on the arcs connecting the circles and representing the chain states.

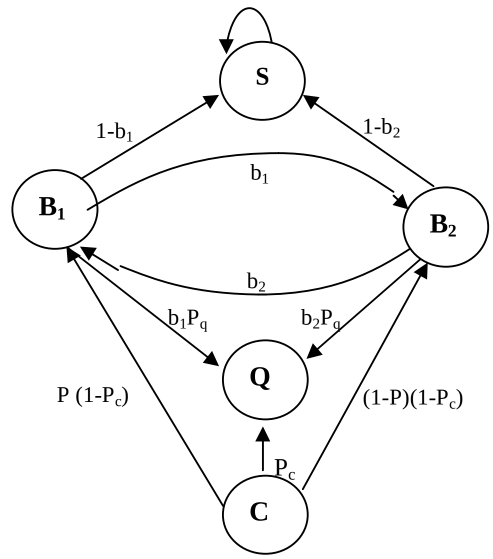


Fig. 3.1 [Transition Diagram of user browsing]

The initial conditions n=0, (state probability before the first surf attempt) are:

$$\left.\begin{aligned} P\left[X^{(0)} = C\right] &= 1 \\ P\left[X^{(0)} = B_1\right] &= 0 \\ P\left[X^{(0)} = B_2\right] &= 0 \\ P\left[X^{(0)} = S\right] &= 0 \\ P\left[X^{(0)} = Q\right] &= 0 \end{aligned}\right\} \quad \ldots (3.1)$$

The unit-step transition probability matrix is:

	B_1	B_2	S	Q	C
B_1	0	$b_1(1-b_1)$	$(1-b_1)$	b_1P_q	0
B_2	$b_2(1-b_2)$	0	$(1-b_2)$	b_2P_q	0
S	0	0	1	0	0
Q	0	0	0	1	0
C	$P(1-P_c)$	$(1-P)(1-P_c)$	0	P_c	0

Table: 3.1 [Transition Probability Matrix]

Using Shukla and Singhai (2011) we can write.

$$P\left[X^{(2n)} = B_1\right]$$
$$= b_1^{n-1} b_2^{n} (1-P)(1-P_C)(1-P_q)^{2n-1}; n > 0 \ldots.(3.2)$$

Volume: 03; Issue: 03; Pages:1162-1168 (2011)

$$P\left[X^{(2n+1)} = B_1\right]$$
$$= \left(b_1 b_2\right)^n P\left(1-P_C\right)\left(1-P_q\right)^{2n};\, n > 0 \ldots(3.3)$$

Similarly, for browser B_2

$$P\left[X^{(2n)} = B_2\right]$$
$$= b_1^n b_2^{n-1} P\left(1-P_C\right)\left(1-P_q\right)^{(2n-1)};\, n > 0 \ldots(3.4)$$

When n is odd

$$P\left[X^{(2n+1)} = B_2\right]$$
$$= \left(b_1 b_2\right)^n \left(1-P\right)\left(1-P_C\right)\left(1-P_q\right)^{2n};\, n > 0 \ldots(3.5)$$

IV. BROWSER SHARING

The probability that navigation successfully starts with browsers B_1 at the n^{th} attempt [as discussed by Shukla and Singhai (2010)]

$$P_1^{(n)} = P\left[X^{(n-1)} = B_1\right].P\left[X^{(n)} = S \middle| X^{(n-1)} = B_1\right]$$
$$n > 0 \ldots(4.1)$$

$$P_1^{(n)} = \left(1-b_1\right)\sum_{i=1}^{n} P\left[X^{(i-1)} = B_1\right] \quad \ldots\ldots.(4.2)$$

$$\text{Let} \quad b_1 b_2 \left(1-P_q\right)^2 = r$$

For n > 0; when n is even [Using Shukla and Singhai (2011)]

$$P^{(2n)} = \left(1-b_1\right)\begin{bmatrix} P\left(1-P_C\right)\left\{\dfrac{1-r^n}{1-r}\right\} \\ +\left(1-P\right)\left(1-P_C\right)\left(1-P_q\right)b_2\left\{\dfrac{1-r^{n-1}}{1-r}\right\}\end{bmatrix} \ldots..(4.3)$$

For n > 0; when n is odd

$$P^{(2n+1)} = \left(1-b_1\right)\left\{\frac{1-r^n}{1-r}\right\}\left[P\left(1-P_C\right)+\left(1-P\right)\left(1-P_C\right)\left(1-P_q\right)b_2\right] \quad \ldots\ldots\ldots.(4.4)$$

These two equations exhibit the overall probability of sharing two browsers by a user in a web environment. These two proportions also reflect how the sharing of two different browsers change with respect to the initial preference. The limiting cumulative probabilities of successful surfing through the two browsers can be obtained by taking the limit of expression over infinite attempts i.e. when the numbers of attempts are infinitely large.

When n is even

$$P_1 = \lim_{n\to\infty} \overline{P}_1^{(2n)}$$
$$= \left(1-b_1\right)\left(1-P_C\right)\left[\frac{P + (1-P)(1-P_q)b_2}{1 - b_1 b_2 (1-P_q)^2}\right] \ldots(4.5)$$

When n is odd

$$P_1 = \lim_{n\to\infty} \overline{P}_1^{(2n+1)}$$
$$= \left(1-b_1\right)\left(1-P_C\right)\left[\frac{P + (1-P)(1-P_q)b_2}{1 - b_1 b_2 (1-P_q)^2}\right] \ldots..(4.6)$$

These expressions are independent of parameter *n*. Browser sharing depends on p, p_q, b_1, b_2 and p_c.

V. ELASTICITIES OF BROWSER SHARING

Differentiate p_1 with respect to b_1

$$f_1(.) = \left(\frac{\partial p_1}{\partial b_1}\right)_{b_2, p_c, p_q, p} = \frac{\begin{array}{l}(1-p_c)\{p+(1-p)(1-pq)b_2\} \\ [\{(1-b_1)b_2(1-pq)^2\} \\ -\{1-b_1 b_2(1-pq)^2\}]\end{array}}{[1-b_1 b_2(1-pq)^2]^2} \ldots(5.1)$$

Differentiate p_1 with respect to b_2

$$f_2(.) = \left(\frac{\partial p_1}{\partial b_2}\right)_{b_1, p_c, p_q, p} = \frac{(1-b_1)(1-p_c)1-pq\begin{bmatrix} b_1(1-pq).\{p+(1-p) \\ (1-pq)b_2\}+(1-p). \\ \{1-b_1 b_2(1-pq)^2\}\end{bmatrix}}{[1-b_1 b_2(1-pq)^2]^2} \ldots(5.2)$$

Differentiate p_1 with respect to p_c

$$f_3(.)\left(\frac{\partial p_1}{\partial P_c}\right)_{b_1, b_2, p_q, p} = (1-b_1)\left[\frac{p+(1-p)(1-pq)b_2}{b_1 b_2 (1-pq)^2 - 1}\right] \ldots(5.3)$$

VI. ISO-QUIETING CURVE

$$p_c = \frac{\{p+(1-p)(1-pq)b_2\} \atop (1-b_1)-p_1\left\{1-b_1b_2\,(1-pq)^2\right\}}{\{p+(1-p)(1-pq)b_2\}(1-b_1)} \quad \ldots\ldots\ldots\ldots(6.1)$$

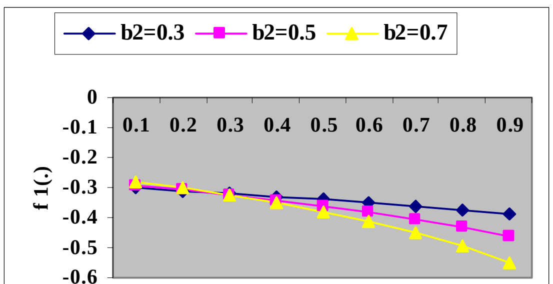
VII. Browser Popularity Index (BPI)

We define BPI as

$$I(B_1) = \frac{P_1}{P_1+P_2}$$

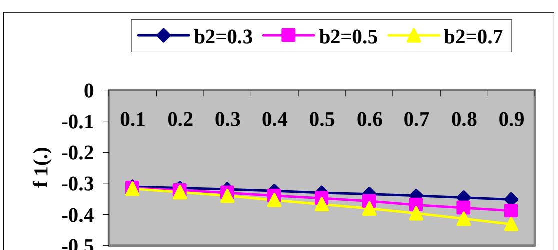
$$I(B_1) = \frac{\left[(1-b_1)(1-p_c)\left[p+(1-p)(1-pq)b_2\right]\right]}{\left[\begin{array}{l}(1-b_1)(1-p_c)\{p+(1-p)(1-pq)b_2\}\\ +(1-b_2)(1-p_c)(1-p)+p(1-pq)b_1\end{array}\right]} \ldots(7.1)$$

VIII. SIMULATION STUDY



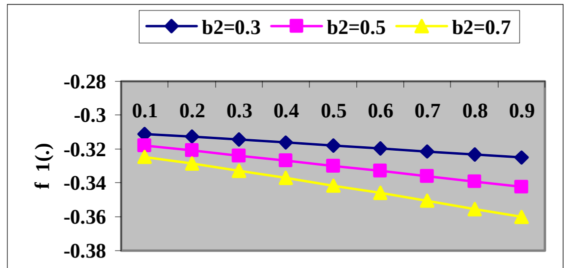
Failure probability of (b1)
fig. 8.1 pc=0.5,p=0.6,pq=0.3

In view of figure 8.1 if browser blocking is high the elasticities of browser share decrease with constant rate slowing adverse effect on browser sharing. If b_2 is high then this effect is more serious, specially at while $p_q = 0.3$



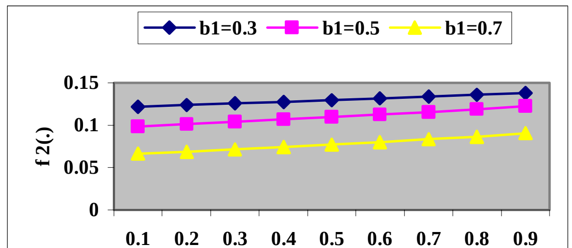
Failure probability of (b1)
fig. 8.2 (pc=0.5, p=0.6, pq=0.5)

In figure 8.2, when p_q is 0.5 (higher a earlier), the elasticities are more negative than earlier, but having similar pattern.

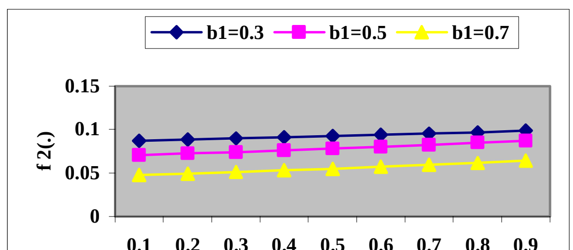


failure probability of (b 1)
fig. 8.3 (pc=0.5, p=0.6, pq=0.7)

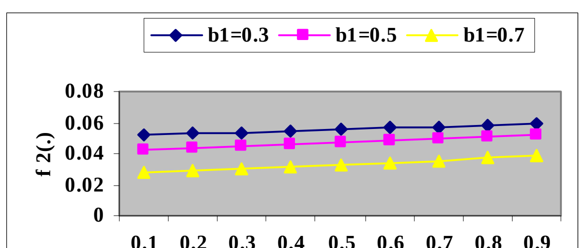
In figure 8.3 contains high value of $p_q = 0.7$ and due to this the stability in elasticity appears better. It seems if p_q is high the elasticity factor f_1 (.) shall be stable.



failure probability b 2
fig. 8.4 (p=0.6, pq=0.5, pc=0.3)



failure probability b 2
fig. 8.5 (p=0.6, pq=0.5, pc=0.5)



failure probability b 2
fig 8.6 (p=0.6, pq=0.5, pc=0.7)

When to examine elasticities with respect to b_2 as obtain in f_2 (.), the output in figure 8.4-8.6 are showing reverse pattern. With the increase of b_2, f_2 (.) has increasing pattern. On the p_c probability if high contains a stability factors in Elasticities variation. The elasticity reduces with the combined increment in b_1 and p_c both.

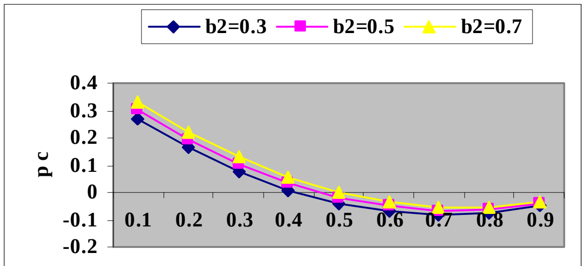


fig.8.7 (p1=0.4, p=0.6, pq=0.5)

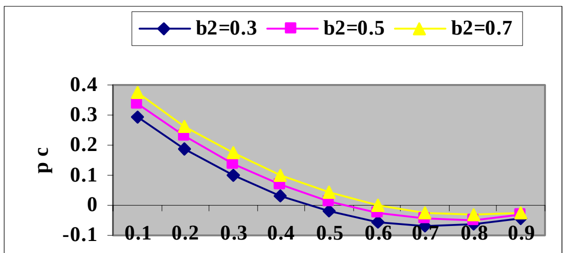


fig. 8.8 (p1=0.4, p=0.6, pq=0.3)

Figure 8.7, 8.8 show downward trend for increasing value of b_1. When b_2 is high better gain appears to probability p_c. The iso-queiting curves to b1 are of exponential nature.

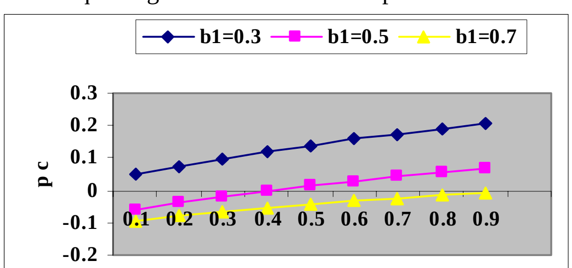


fig.8.9 (p1=0.4, p=0.6, pq=0.3)

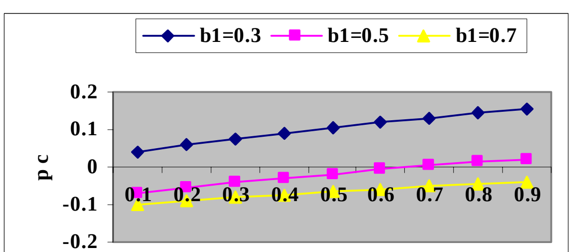


fig. 8.10 (p1=0.4, p=0.6, pq=0.5)

Figure 8.9, 8.10 are showing the linear trend when b_2 is increasing constantly. The trend line goes down while b_1 is high. These iso-queiting curves are positive in nature as discussed earlier.

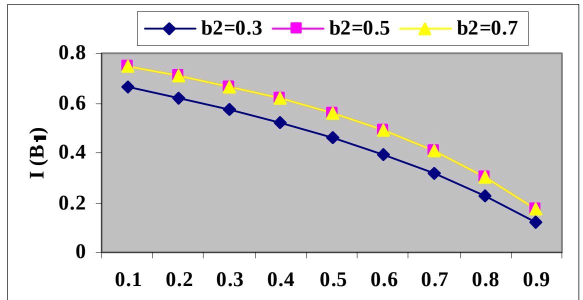


fig 8.11(pc=0.5, pq=0.5, p=0.6)

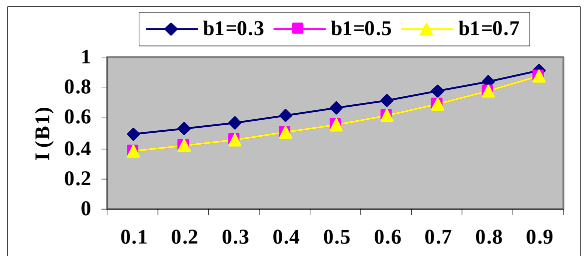


fig.8.12 (pc=0.5, pq=0.5, p=0.6)

Figure 8.11, 8.12 relate to browser share index. This index depends on b_1, b_2 probabilities. The increment in b_1 reduces the index level where as increase in b_2 increases the same. This index shows the relative browser sharing chance in light of browser blocking probability.

XI. CONCLUDING REMARKS:

The elasticity pattern found negative with respect to b_1 for first browser but opposite in nature for second browser. This indicates that a browser has to keep lesser blocking level comparing to other in order to maintain the same popularly level. The quieting probability is a function of browser blocking probability therefore a fair balance must be maintained between these two. The relative increment in browsing index shows that input parameters must be present in a particular interval to maximize the index values.

ACKNOWLEDGEMENT

One of authors, Dr. D. Shukla is thankful to Coordinator, DST-Center for Interdisciplinary Mathematical Science, Banaras Hindu University, Varanasi, U.P. for providing Associate Membership.

REFERENCES

[1] Catledge, L. D. and Pitkow J. E. (1995): "Characterizing browsing strategies in the World-Wide Web", Computer Networks and ISDN Systems,26 (6),pp 1065.

[2] Deshpande, M. and Karypis G.(2004): "Selective Markov Models for Predicting Web-Page Accesses", ACM Transactions on Internet Technology, Vol 4, No.2, May 2004,pp.168-184.

[3] Han, J. and Kamber, M.(2001): "Data Mining-Concept & Techniques", Academic Press, San Diego..

[4] Medhi, J. (1991): "Stochastic processes", Ed 4, Wiley Eastern Limited (Fourth reprint), New Delhi.

[5] Naldi, M. (2002): 'Internet access traffic sharing in a multi operator environment", Computer Network, vol. 38, pp. 809-824.

[6] Pirolli, P., J. Pitkow, and R. Rao (1996): "Silk from a sow's ear: Extracting usable structures from the web", Conference on Human Factors in Computing Systems, CHI '96, Vancouver, Canada.

[7] Pitkow, J. E. (1997): "In Search of Reliable Usage Data on the WWW", Sixth International World Wide Web Conference, Santa Clara, CA.

[8] Shukla, D. and Jain, S. (2007): "A Markov chain model for multi-level queue scheduler in operating system", Proceedings of the International Conference on Mathematics and Computer Science, ICMCS-07, pp. 522-526,.

[9] Shukla, D. and Singhai, R. (2010 a): "Traffic analysis of message flow in three crossbar archicture space-division switches", Journal of Computer Science (JCS), Coimbatore, Vol. 4, No.3, pp. 1560.

[10] Shukla, D., Gadewar, S. and Pathak, R.K. (2007): "A Stochastic model for space-division switches in computer networks", Applied Mathematics and Computation (Elsevier Journal), Vol. 184, Issue 2, pp. 235-269.

[11] Shukla, D. and Singhai, Rahul(2011): Analysis of user web browsing behavior using markov chain model, International journal of Advanced networking and Application, Vol. 02, issue 05,pp 824-830

[12] Shukla, D., Jain, A. and Choudhary, A. (2010 b): "Estimation of ready queue processing time under SL- Scheduling scheme in multiprocessors environment", International Journal of Computer Science and Security (IJCSS), Vol. 4, No. 1, pp. 74-81.

[13] Shukla, D., jain, A. and Choudhary, A. (2010 c): "Estimation of ready queue processing time under usual group lottery Scheduling (GLS) in multiprocessors environment ", Journal of Computer Applications, Vol. 8, No. 4, pp. 39-45.

[14] Shukla, D., Jain, S. and Ojha, S. (2010 d): "Effect of data model approach in state probability analysis of multilevel queue scheduling", Journal of Advanced Networking & Applications (IJANA), Vol. 02, No. 1, pp. 419-427.

[15] Shukla, D., Jain, S., Singhai, R. and Agarwal, R. K.(2009 a): "A Markov chain model for the analysis of round-robin scheduling scheme", journal of Advanced Networking & Applications (IJANA), Vol. 01, No. 1, pp. 1-7.

[16] Shukla, D., Jain, S. and Ojha, S. (2010 e): "Data Model based analysis of multi level queue scheduling using Markov Chain Model, International Journal of Applied Computer Science and Mathematics, Vol. 8, No. 4, pp. 50-56.

[17] Shukla, D., Ojha, S. and Jain, S. (2010 f): "Performance evalution of a general class of multi-level queue scheduling scheme, Computer Science and Telecommunications (GEST), Vol. 3, No. 26, pp. 99-121.

[18] Shukla, D., Singhai, R., Gadewar, S.K., Jain, S., Ojha, S. and Mishra, P.P. (2010 g): "A Stochastic model approach for reaching probabilities of message flow in Space Division Switches", International Journal of Computer Network (IJCN), Vol. 2, No. 2, pp.140-150.

[19] Shukla, D., Tiwari, V., and Kareem, A. (2009 b): "All comparision analysis of internet traffic sharing using Markov Chain Model in Computer Network", Georgian Electronic Scientific Journal: Computer Science and Telecommunications (GESJ), Vol. 06, No. 23, pp. 108-115.

[20] Shukla, D., Tiwari, V., Thakur, S. and Deshmukh, A.K. (2009 c): "Share loss analysis of Internet traffic distribution in computer network", International Journal of Computer Science and Security (IJCSS), Vol. 3, No. 4, pp. 414-427.

[21] Shukla, D., Tiwari, V., Thakur, S. and Deshmukh, A.K. (2010 h): " Two call based analysis of internet traffic sharing", International Journal of Computer and Engineering, Vol. 1, No. 1, pp. 14-24.

[22] Shrivastava J., Cooley R., Deshpande M. and Tan P. (2000): "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data", ACM SIGKDD, Jan. 2000.

[23] Shukla, D., Singhai, Rahul and Thakur, N.S. (2010): A New Imputation Method for Missing Values in Data Mining, Journal of Applied Computer Science and Mathematics, Vol. 10, pp.28-38.

[24] Shukla, D., Gangale sharad, Verma, kapil and Singh, Pankaj (2011): "Elasticity of Internet Traffic

Volume: 03; Issue: 03; Pages:1162-1168 (2011)

Distribution in Computer Network in Two Market Environment", Journal of Global Research in Computer Science, Vol. 2, No. 6, pp. 6-12.

[25] Shukla, D., Gangele Sharad, Verma, Kapil and Singh, Pankaja (2011): "Elasticities and Index Analysis of Usual Internet Traffic Share Problem", International Journal of Advanced Research in Computer Science (IJARCS),Vol. 02, No. 04, pp.473-478.

Authors Biography

Dr. Diwakar Shukla is working as an Associate Professor in the Department of Mathematics and Statistics, Sagar University, Sagar, M.P. and having over 21 years experience of teaching to U.G. and P.G. classes. He obtained M.Sc.(Stat.) Ph.D.(Stat.), degrees from Banaras Hindu University, Varanasi and served the Devi Ahilya University, Indore, M.P. as a Lecturer over nine years and obtained the degree of M. Tech. (Computer Science) from there. During Ph.D., he was junior and senior research fellow of CSIR, New Delhi qualifying Fellowship Examination (NET) of 1983. Till now, he has published more than 90 research papers in national and international journals and participated in more than 35 seminars / conferences at national level. He also worked as a selected Professor at the Lucknow University, Lucknow, U.P., for one year and visited abroad to Sydney (Australia) and Shanghai (China) for conference participation. He has supervised thirteen Ph.D. theses in Statistics and Computer Science both; and five students are presently enrolled for their doctoral degree under his supervision. He has authored/ Co-authored two books. He is member of 10 learned bodies of Statistics and Computer Science both at national level. The area of research he works for are Sampling Theory, Graph Theory, Stochastic Modeling, Inventory Modeling, Computer Network and Operating Systems.

Sharad Gangele has completed M.Sc. (Statistics) in 2005 from Devi Ahilya University, Indore (M.P.), P.G.D.C.A. in 2006 and M.Sc. (Computer Science) in 2008 both from M.C.R.P.V., Bhopal, (M.P). He is presently working as an Assistant Professor (Statistics) in the Department of Mathematics in M.B.Khalsa College Indore, (M.P). He has published three research papers. He is pursuing research in the field of stochastic modeling of computer network system. He presented six research papers in National/International conferences. His current research interest is stochastic modeling of computer network system and internet traffic sharing analysis.

Dr. Rahul Singhai has obtained M.C.A. degree from H.S. Gour University, Sagar, MP, in 2001 and obtained M.Phil degree in Computer Science from Madurai Kamaraj University, Madurai, Tamilnadu in 2008.The Ph.D degree in Computer Science was awarded in 2011 by Sagar university. He is a permanent faculty member in IIPS, Devi Ahilya University, Indore. His research interest includes Computer Network, Data mining, operating System & Software Testing. He has authored and co-authored 12 research papers in international/national journals and proceeding. Currently, he is working on to develop new probability based methods for data preprocessing in data mining.

Kapil Verma has completed M.C.A degree from Dr. H. S. Gour Central University, Sagar [M.P.]. He is presently working as a guest Lecturer in the Department of Computer Science & Applications in Govt. Autonomous Girls P. G. Excellence College Sagar, M.P. He is pursuing research in the field of Probability Based Mathematical Modeling of Internet Traffic in Computer Network. He has authored and co-authored three research papers. He presented two research papers in National/International conferences. His current research interest is to study the internet traffic sharing under various traffic variants.