International Journal of Latest Trends in Software Engineering

Volume 1 No.1
June, 2011

Editorial Board

Table of Contents

An M/G/1 Two Phase Multi-Optional Retrial Queue with Bernoulli Feedback and Non-Persistent Customers

Kasturi Ramanath[1] and K.Lakshmi [2]

[1]*School of Mathematics, Madurai Kamaraj University,Madurai-21, Tamil Nadu, India.*

[2]*Swami Vivekananda Arts & Science College, Gingee Main Road,Orathur Cross Road, Lakshmipuram, Villupuram Dist 605601*

e-mail: 1kasturey@gmail.com, 2lakshmikrishna80@yahoo.com

Abstract: This paper has been motivated by the Interactive Voice Response System (IVRS).This system now become a common phenomenon in our everyday life. In this paper, we consider a Poisson arrival queueing system with a single server and two essential phases of heterogeneous service. The customer who completes the first phase has a choice of k options to choose for the second phase of service. The customer, who finds the server busy upon arrival, can either join the orbit or he/she can leave the system. After completion of both phases, the customer can decide to try again for service by joining the orbit or he/she can leave the system. By using the supplementary variables technique, we have obtained the steady state probability generating functions of the orbit size and the system size. We have obtained a necessary and sufficient condition for the existence of the steady state. We also obtain some useful performance measures of the system. We discussed some particular cases of the model.

Key words: Two essential phases of service, multi optional second phase, retrial queue, feedback, Stochastic decomposition.

1. Introduction

Queuing systems, in which arriving customers who find all servers and waiting positions occupied may retry for service after a period of time, are called retrial queues or queues with repeated calls. Retrial queues are useful in modeling many problems in telephone switching systems, computer and communication systems. A review of the literature on retrial queues can be found in Falin [8], Yang and Templeton [15], Artalejo and Gomez-correl [2].

Choi and Kulkarni [3] have investigated M/G/1 feedback retrial queues with no waiting position. The phenomenon of feedback in the retrial queueing systems occurs in many practical situations; for instance, multiple access telecommunication systems, where messages turned out as errors are sent again are modeled as retrial queues with feedbacks.

M.R.Salehirad and A.Badamchizadeh [13] have studied the multi-phase M/G/1 queueing system with random feedbacks. Yong Wan Lee and Young Ho Jang [16] discussed a retrial queueing system with a regular queue and an orbit. After completion of each service, the customer can decide to join the orbit or leave the system. Choudhry [5] has investigated an M/G/1 retrial queue with an additional phase of second service and general retrial times.

Madan [11] investigated an M/G/1 queue where the server first provides a regular service to all arriving customers, whereas, only some of them receive a second phase of optional service. The regular service follows a general distribution, but the second optional service is assumed to be exponentially distributed. Medhi [12] generalized the model by considering that the second optional service is also governed a by general distribution. Choudhury [4] investigated this model further in depth. However, Krishna Kumar et al. [10] studied an M/G/1 retrial queue with an additional phase of service, where at the first phase of service, the server may push out the customer who is receiving such service, to start the service of another priority customer. The interrupted customers joins a retrial group and the customer at the head of the queue is allowed to conduct a repeated attempt in order to start again his first phase of service after some random time. Artalejo and Choudhury [1] investigated a similar type of M/G/1 queue under a classical retrial policy (i.e.) retrial rate is ‘j ’ when the number of customer in the retrial group is ‘j’. This discipline is typical of telephone applications, when the retrials are made individually by each blocked customer following an exponential law of rate ‘ ’. The motivations for such types of models come from computer and communication networks, where messages are proceeds in two stages by a single server. Doshi [6] recognized its applications in a distributed system, where control of two phase execution is required by a central server. Jingting Wang [9] studied an M/G/1 queue with a second optional service and server breakdowns.

In this paper, we have investigated an M/G/1 retrial queueing system with Bernoulli feedback, non-persistent customers and a second phase of essential service. This second phase can be chosen from k-optional services which are available in the system. Our motivation for studying this queueing system comes from a study of the Interactive Voice

Response System (IVRS) available in banks, railway stations, telecommunication centers, call centers etc.

A customer who rings up such a system is offered several options from which he / she can choose the required type of service. However, if the system is busy the customer gets an engaged signal. The customer may persist and make repeated attempts until he / she succeeds in obtaining service or the customer may be a non-persistent customer who leaves the system.

We have therefore considered a single server Poisson input queueing system. The service time in the first phase of essential services as well as in the k-optional services available for the second phase of service is assumed to be generally distributed. We assume that the customers in the orbit make retrials independently of each other. After the completion of both phases of service, the customer either decides to make a feedback with probability r1 or he decides to leave the system with probability 1- r1. We have also assumed that whenever the customer finds the system busy He / She joins the orbit with probability p or decides to leave the system with probability 1-p.

The rest of the paper is organized as follows: In Sec-2, we present the mathematical model of the system. In Sec-3, we present the steady state analysis of the system, we derived the joint P.G.F.of the server state and orbit size and server state and system size. In sec-4, we derived the necessary and sufficient condition for the existence of the steady state. In sec-5 we consider some useful performance measures of the system. In sec-6, we derived some particular cases of the system. In sec-7, we obtained the stochastic decomposition of the system.

2. The Mathematical Model

We consider a single server retrial queueing system with a Poisson arrivals non-persistent customers and feedbacks. The arrival process is assumed to be a Poisson with parameter λ.Upon arrival If the customer finds the server busy, he/she either joins an orbit with probability p or leaves the system with probability 1-p.In the orbit the inter retrial times are assumed to be exponentially distributed with a parameter γ. From the orbit the customer keeps trying repeatedly for service until he succeeds in obtaining service. The retrial attempts made by a customer are assumed to be independent of the attempts made by other customers.

The service consists of two phases. The first phase service is generally distributed with a distribution function B(x),hazard rate function μ(x) and the Laplace Stieljes transform (LST) B*(S),the expected value β.

Upon completion of the first phase of service, the customer has a choice of k-services from which he can choose the second phase of service. We assume that the probability of choosing the j[th] option is αj ,j=1,2…k with $\sum_{j=1}^{k} \alpha_j$ =1.The service time in the j[th] optional service is assumed to have the distribution function $B_j(x)$,hazard rate function $\mu_j(x)$ and the Laplace Stieljes transform (LST) $B_j^*(s)$ and an expected value β_j.

Upon completion of both phases of service, if the customer so desires he/she may join the orbit to obtain service again. The customer may join the orbit with probability r_1 to obtain service again or leaves the system with probability 1- r_1.

We define N (t) to be the number of customers in the orbit at time t. C (t) to be the state of the server at time t, i.e.

$$C(t) = \begin{cases} 0, \text{ if the server is idle} \\ 1, \text{ if the server is performing the first phase of service} \\ 2, \text{ if the server is performing the second phase of service} \end{cases}$$

We introduce the following supplementary variable when C (t) =1 or 2 X (t) =the elapsed service time of the customer in service.

{C (t), N (t), X (t)} is a continuous time Markov process.

We define the following probability functions

$$P_{0,0}(t) = \Pr\{C(t)=0; N(t)=0\};\ n \geq 0$$
$$P_{1,n}(x,t)dx = \Pr\{C(t)=1; N(t)=n; x \leq X(t) \leq x+dx\}; n \geq 0$$
$$P_{2,n}^{(j)}(x,t)dx = \Pr\{C(t)=2; N(t)=n; x \leq X(t) \leq x+dx\}; n \geq 0$$

where X (t) is the elapsed service time in the j[th] optional service; j=1, 2… k

3. Steady State Analysis

Now, analysis of the queuing model can be performed with the help of the following Kolmogorov forward equations.

$$\frac{\partial}{\partial t}P_{0,n}(t)=-[\lambda+n\gamma]P_{0,n}(t)+(1-r_1)\sum_{j=1}^{k}\int_0^{\infty}P_{2,n}^{j}(x,t)\mu_j\ (x)\mathrm{dx}+\ (1-\delta_{0,n})\mathrm{r}_1\ \sum_{j=1}^{k}\int_0^{\infty}P_{2,n-1}^{j}(x,t)\mu_j(x)\mathrm{dx} \tag{1}$$

$$\frac{\partial}{\partial x}P_{1,n}(x,t)+\frac{\partial}{\partial t}P_{1,n}(x,t)=-\{p\lambda+\mu(x)\}P_{1,n}(x,t)+(1-\delta_{0n})p\lambda P_{1,n-1}(x,t) \tag{2}$$

$$\frac{\partial}{\partial x}P_{2,n}^{j}(x,t)+\frac{\partial}{\partial t}P_{2,n}^{j}(x,t)=-\left[p\lambda+\mu_j(x)\right]P_{2,n}^{j}(x,t)+p\lambda(1-\delta_{0,n})P_{2,n-1}^{j}(x,t),\mathrm{n}\geq 0,\mathrm{j{=}1\ to\ k} \tag{3}$$

The boundary conditions are as follows, for $n \geq 0$,

$$P_{1,\,\mathrm{n}}(0,t)=\ \lambda P_{0,n}(t)+(n+1)\gamma P_{0,n+1}(t) \tag{4}$$

$$P_{2,n}^{j}(0,t)=\ \alpha_j\int_0^{\infty}P_{1,\,\mathrm{n}}(x,t)\mu(x)dx\ ,\ \mathrm{j{=}1\ to\ k} \tag{5}$$

Assuming that the system reaches the steady state and taking the limits as $t\to\infty$ the equations (1) to (6) become

$$[\lambda+n\gamma]P_{0,n}=(1-r_1)\sum_{j=1}^{k}\int_0^{\infty}P_{2,n}^{j}(x)\mu_j\ (x)\mathrm{dx}+\ (1-\delta_{0,n})\mathrm{r}_1\ \sum_{j=1}^{k}\int_0^{\infty}P_{2,n-1}^{j}(x)\mu_j(x)\mathrm{dx} \tag{6}$$

$$\frac{d}{dx}P_{1,n}(x)=-\{p\lambda+\mu(x)\}P_{1,n}(x)+(1-\delta_{0n})p\lambda P_{1,n-1}(x) \tag{7}$$

$$\frac{d}{dx}P_{2,n}^{j}(x)=-\left[p\lambda+\mu_j(x)\right]P_{2,n}^{j}(x)+p\lambda(1-\delta_{0,n})P_{2,n-1}^{j}(x),\mathrm{j{=}1\ to\ k} \tag{8}$$

The boundary conditions are as follow j for, $n\geq 0$

$$P_{1,\,\mathrm{n}}(0)=\ \lambda P_{0,n}+(n+1)\gamma P_{0,n+1} \tag{9}$$

$$P_{2,n}^{j}(0)=\ \alpha_j\int_0^{\infty}P_{1,\,\mathrm{n}}(x)\mu(x)dx \tag{10}$$

We define the following partial probability generating functions j for $|z|\leq 1$

$$P_0(z)=\sum_{n=0}^{\infty}P_{0,n}z^n \tag{11}$$

$$P_1(x,z)=\ \sum_{\mathrm{n}=0}^{\infty}P_{1,\,\mathrm{n}}(x)\ \mathrm{z}^{\mathrm{n}} \tag{12}$$

$$P_2^{j}(x,z)=\ \sum_{\mathrm{n}=0}^{\infty}P_{2,n}^{j}(x)\ \mathrm{z}^{\mathrm{n}} \tag{13}$$

Let

$$P_1(z)=\int_0^{\infty}P_1(x,z)\mathrm{dx} \tag{14}$$

$$P_2^{j}(z)=\int_0^{\infty}P_2^{j}(x,z)\mathrm{dx} \tag{15}$$

Theorem: In the steady state, the joint distribution of the server state and the orbit size is given by

$$P(z)=e^{-\frac{\lambda}{\gamma}\int_0^z \frac{1-\phi(u)}{u-\phi(u)}du}\left[\frac{1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)}{e^{-\frac{\lambda}{\gamma}\int_0^1 \frac{1-\phi(u)}{u-\phi(u)}du}\left[1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)+\sum_{j=1}^{k}\alpha_j\lambda\left(\beta+\beta_j\right)\right]}\right]\left[\frac{p\left(z-\phi(z)\right)-1+\sum_{j=1}^{k}\alpha_j B^*\left(\lambda p(1-z)\right)B_j^*\left(\lambda p(1-z)\right)}{p\left(z-\phi(z)\right)}\right]$$

where $\phi(z)=\sum_{j=1}^{k}\alpha_j B^*\left(\lambda p(1-z)\right)B_j^*\left(\lambda p(1-z)\right)\left[(1-r_1)+r_1 z\right]$ and the joint distribution of the server state and system size is given by

$$K(z)=e^{-\frac{\lambda}{\gamma}\int_0^z \frac{1-\phi(u)}{u-\phi(u)}du}\left[\frac{1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)}{e^{-\frac{\lambda}{\gamma}\int_0^1 \frac{1-\phi(u)}{u-\phi(u)}du}\left[1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)+\sum_{j=1}^{k}\alpha_j\lambda\left(\beta+\beta_j\right)\right]}\right]\left[\frac{p\left(z-\phi(z)\right)-z++z\sum_{j=1}^{k}\alpha_j B^*\left(\lambda p(1-z)\right)B_j^*\left(\lambda p(1-z)\right)}{p\left(z-\phi(z)\right)}\right]$$

Proof: Multiplying (7) & (8) by z^n and summing for n from 0 to ∞, the solutions of the resulting equation are

$$\mathrm{P}_1(x,z)=\mathrm{P}_1(0,z)\ \mathrm{e}^{-[\lambda p(1-z)]x}\ \left[1-B(x)\right] \tag{16}$$

$$P_2^j(x,z)=\ e^{-\lambda p(1-z)x}\ \left[1-B_j^*(x)\right]\ P_2^j(0,z),\ \mathrm{j}=1\text{to k} \tag{17}$$

similarly multiplying eqn(6) by z^n *and* summing *for n from* 0 *to* ∞
and by using eqn(17),we get

$$\lambda \mathrm{P}_0(z)+z\gamma \mathrm{P}_0^{'}(z)=\left[(1-r_1)+r_1 z\right]\sum_{j=1}^{k}B_j^*\left(\lambda p(1-z)\right)P_2^j(0,z) \tag{18}$$

A similar procedure with eqn(10)

$$P_2^j(0,z)=\sum_{j=1}^{k}\alpha_j\int_0^\infty P_1(x,z)\mu(x)dx$$

$$P_2^j(0,z)=\sum_{j=1}^{k}\alpha_j B^*\left(\lambda p(1-z)\right)P_1(0,z) \tag{19}$$

$$P_1(0,z) = \lambda P_0(z) + \gamma P_0'(z) \quad (20)$$

from (18) & (20) we get

$$\frac{P_0'(z)}{P_0(z)} = -\frac{\lambda}{\gamma}\frac{1-\phi(z)}{z-\phi(z)}$$

where $\phi(z) = \sum_{j=1}^{k} \alpha_j B^*(\lambda p(1-z)) B_j^*(\lambda p(1-z))\left[(1-r_1)+r_1 z\right]$

$$P_0'(z) = -\frac{\lambda}{\gamma}\ \frac{1-\phi(z)}{z-\phi(z)} e^{-\frac{\lambda}{\gamma}\int_0^z \frac{1-\phi(u)}{u-\phi(u)}du}\ P_{0,0} \quad (21)$$

$$P_0(z) = e^{-\frac{\lambda}{\gamma}\int_0^z \frac{1-\phi(u)}{u-\phi(u)}du}\ P_{0,0} \quad (22)$$

$$P_1(0,z) = \lambda e^{-\frac{\lambda}{\gamma}\int_0^z \frac{1-\phi(u)}{u-\phi(u)}du} \left[\frac{z-1}{z-\phi(z)}\right] P_{0,0} \quad (23)$$

$$P_2^j(0,z) = \lambda e^{-\frac{\lambda}{\gamma}\int_0^z \frac{1-\phi(u)}{u-\phi(u)}du} \sum_{j=1}^{k} \alpha_j \left[\frac{z-1}{z-\phi(z)}\right] B^*(\lambda p(1-z)) P_{0,0} \quad (24)$$

$$P_1(z) = \left[\frac{1-B^*(\lambda p(1-z))}{p(1-z)}\right]\left[\frac{z-1}{z-\phi(z)}\right] e^{-\frac{\lambda}{\gamma}\int_0^z \frac{1-\phi(u)}{u-\phi(u)}du} P_{0,0} \quad (25)$$

$$P_2^j(0,z) = \sum_{j=1}^{k} \alpha_j \left[\frac{z-1}{z-\phi(z)}\right] e^{-\frac{\lambda}{\gamma}\int_0^z \frac{1-\phi(u)}{u-\phi(u)}du} \left[\frac{\left[1-B_j^*(\lambda p(1-z))\right] B^*(\lambda p(1-z))}{p(1-z)}\right] P_{0,0} \quad (26)$$

The orbit size P.G.F

$$P(z) = P_0(z) + P_1(z) + \sum_{j=1}^{k} P_2^j(z)$$

$$P(z) = e^{-\frac{\lambda}{\gamma}\int_0^z \frac{1-\phi(u)}{u-\phi(u)}} \left[\frac{p(z-\phi(z)) - 1 + +\sum_{j=1}^{k} \alpha_j B^*(\lambda p(1-z)) B_j^*(\lambda p(1-z))}{p(z-\phi(z))}\right] P_{0,0} \quad (27)$$

To obtain the value of $P_{0,0}$ we use the normalizing condition as $z \to 1$, $P(1) = 1$ and applying L'Hospital's rule in an appropriate place we get

$$P_{0,0} = \frac{1 - r_1 - \sum_{j=1}^{k} \alpha_j \lambda p(\beta+\beta_j)}{e^{-\frac{\lambda}{\gamma}\int_0^1 \frac{1-\phi(u)}{u-\phi(u)}du} \left[1 - r_1 - \sum_{j=1}^{k} \alpha_j \lambda p(\beta+\beta_j) + \sum_{j=1}^{k} \alpha_j \lambda(\beta+\beta_j)\right]} \quad (28)$$

Substituting (28) in (27) we get the orbit size P.G.F

$$P(z)=e^{-\frac{\lambda}{\gamma}\int_0^z\frac{1-\phi(u)}{u-\phi(u)}du}\left[\frac{1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)}{e^{-\frac{\lambda}{\gamma}\int_0^1\frac{1-\phi(u)}{u-\phi(u)}du}\left[1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)+\sum_{j=1}^{k}\alpha_j\lambda\left(\beta+\beta_j\right)\right]}\right]$$
$$\left[\frac{p\left(z-\phi(z)\right)-1++\sum_{j=1}^{k}\alpha_j B^*\left(\lambda p\left(1-z\right)\right)B_j^*\left(\lambda p\left(1-z\right)\right)}{p\left(z-\phi(z)\right)}\right] \qquad (29)$$

The K (z) be the system size P.G.F

$$K(z)=e^{-\frac{\lambda}{\gamma}\int_0^z\frac{1-\phi(u)}{u-\phi(u)}du}\left[\frac{1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)}{e^{-\frac{\lambda}{\gamma}\int_0^1\frac{1-\phi(u)}{u-\phi(u)}du}\left[1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)+\sum_{j=1}^{k}\alpha_j\lambda\left(\beta+\beta_j\right)\right]}\right]$$
$$\left[\frac{p\left(z-\phi(z)\right)-z++z\sum_{j=1}^{k}\alpha_j B^*\left(\lambda p\left(1-z\right)\right)B_j^*\left(\lambda p\left(1-z\right)\right)}{p\left(z-\phi(z)\right)}\right] \qquad (30)$$

4. The Embedded Markov Chain

In this section, we establish the necessary and sufficient condition for the ergodicity of the given system considered in the previous section. We consider the ergodicity of corresponding embedded Markov chain of the process. We use the Foster's criterion for sufficiency.

Foster's criterion: For an irreducible and aperiodic Markov chain ξ_j with state space S, a sufficient condition for ergodocity is the existence of a nonnegative function f(s), $s\in S$ and $\varepsilon>0$ such that the mean drift $x_s=E\left(f\left(\xi_{i+1}\right)-f\left(\xi_i\right)/\xi_i=s\right)$ is finite for all $s\in S$ and $x_s\le-\varepsilon$ for all $s\in S$ except perhaps a finite number.

Theorem: The M/G/1 retrial queuing system with feedbacks considered in the previous section is ergodic iff

$$r_1+\sum_{j=1}^{k}\lambda p\alpha_j\beta_j+\lambda p\beta<1$$

Proof: We use the foster's criterion to prove the sufficiency of the condition.

Let $X_n=N\left(t_n+\right)$ denote the number of customers in the orbit after the n[th] serviced customer. Then

$$q_{m,n}=\frac{m\gamma}{\lambda+m\gamma}\left[\left(1-r_1\right)K_{n-m+1}+r_1K_{n-m}\right]+\frac{\lambda}{\lambda+m\gamma}\left[\left(1-\delta_{m,n}\right)r_1K_{n-m-1}+\left(1-r_1\right)K_{n-m}\right]$$

$$q_{0,0}=K_0\left(1-r_1\right)$$

For $n\ge1$, $q_{0,n}=\left(1-r_1\right)K_n+r_1K_{n-1}$

Where K_n =Pr {n arrivals during the service time of one customer}

$$K_n = \sum_{r=0}^{n} K_{1r} K_{2,n-r}$$

where K_{1r} = *prob.of r arrivals during the* 1st phase of service

$$= \int_0^\infty e^{-\lambda px} \frac{(px)^r}{r!} dB(x)$$

The expected number of arrivals in the first phase of service is given by

$$-\lambda p\, B^{*'}(0) = \lambda p \beta$$

$K_{2,n-r}$ = *probability* of n-r arrivals in the second phase of service

$$= \sum_{j=1}^{k} \alpha_j \int_0^\infty e^{-\lambda px} \frac{(px)^{n-r}}{(n-r)!} dB_j(x)$$

The expected number of arrivals during the second phase of service is given by

$$\lambda p \sum_{j=1}^{k} \alpha_j \beta_j$$

The expected number of arrivals during the service of a customer is given by

$$\sum_{n=0}^{\infty} n\, K_n = \lambda p \beta + \lambda p \sum_{j=1}^{k} \alpha_j \beta_j$$

To use Foster's criterion, we use the test function f (n) =n, $\forall$ n ≥ 1.

For m ≥ 1,

Let

$$\xi_m = E\left[f(X_{n+1}) - f(X_n) / X_n = m \right]$$
$$= E\left[X_{n+1} - X_n / X_n = m \right]$$
$$= \sum_{n=m-1}^{\infty} (n-m)\, q_{m,n}$$

$$= \frac{m\gamma}{\lambda + m\gamma}(1-r_1)\left[\lambda p\beta + \sum_{j=1}^{k} \alpha_j\beta_j - 1 \right] + \frac{m\gamma}{\lambda + m\gamma} r_1 \left[\lambda p\beta + \lambda p \sum_{j=1}^{k} \alpha_j\beta_j \right]$$
$$+ \frac{\lambda}{\lambda + m\gamma} r_1 \left[\lambda p\beta + \lambda p \sum_{j=1}^{k} \alpha_j\beta_j + 1 \right] + \frac{\lambda}{\lambda + m\gamma}(1-r_1)\left[\lambda p\beta + \lambda p \sum_{j=1}^{k} \alpha_j\beta_j \right]$$

$$= \lambda p \left[\beta + \sum_{j=1}^{k} \alpha_j\beta_j \right] - (1-r_1)\frac{m\gamma}{\lambda + m\gamma} + \frac{r_1\lambda}{\lambda + m\gamma}$$

For m ≥ 1,

$$\xi_m = \lambda p \left[\beta + \sum_{j=1}^{k} \alpha_j\beta_j \right] + r_1 - \frac{m\gamma}{\lambda + m\gamma}$$

As m $\to \infty$, $\frac{m\gamma}{\lambda + m\gamma} \to 1$

If

$$r_1 + \lambda p \left[\beta + \sum_{j=1}^{k} \alpha_j\beta_j \right] < 1,$$

choose $\varepsilon < 1 - r_1 - \lambda p \left[\beta + \sum_{j=1}^{k} \alpha_j\beta_j \right]$

For this ε there exists M such that

$$1 - \varepsilon < \frac{m\gamma}{\lambda + m\gamma} < 1, \ \forall\ m \geq M$$

$$\xi_m = r_1 + \lambda p \left[\beta + \sum_{j=1}^{k} \alpha_j\beta_j \right] - \frac{m\gamma}{\lambda + m\gamma}$$

$$< r_1 + \lambda p \left[\beta + \sum_{j=1}^{k} \alpha_j\beta_j \right] - (1-\varepsilon)$$

$$= r_1 + \lambda p \left[\beta + \sum_{j=1}^{k} \alpha_j\beta_j \right] - 1 + \varepsilon < 0, \text{by the choice of } \varepsilon.$$

Therefore ξ_m is finite for all m ≥ 0 and $\xi_m < 0$ for all m $\geq$ M. By Foster's criterion, the given M/G/1queueing system is ergodic. Therefore, the condition stated the theorem is sufficient for the steady state to exist.

Also

$$P_{0,0} = \frac{1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)}{e^{-\frac{\lambda}{\gamma}\int_0^1\frac{1-\phi(u)}{u-\phi(u)}du}\left[1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)+\sum_{j=1}^{k}\alpha_j\lambda\left(\beta+\beta_j\right)\right]}$$

For the steady state to exist, it is necessary that $P_{0,0} > 0$.This implies that the numerator of $P_{0,0}$ must be positive.

$\therefore\ r_1+\lambda p\left[\beta+\sum_{j=1}^{k}\alpha_j\beta_j\right]<1$ is a necessary condition for the system to be ergodic.

5. Performance Measures

In this section, we derive the analytical expression for some useful performance measures of the system

(a) The expected number of customers in the system is given by

$$E(L)=\lim_{z\to 1}\frac{d}{dz}K(z)$$

$$=\frac{\left[1-\phi'(1)\right]\left\{\lambda^2p^2\left[\left(\beta_j^{(2)}-\beta_j\right)+\left(\beta^{(2)}-\beta\right)+2\beta\beta_j\right]+2\lambda p\beta_j-\lambda p(1-\beta)-p\phi''(1)\right\}+\phi''(1)\left[p\left(1-\phi'(1)\right)-2+\lambda p\left(\beta+\beta_j\right)\right]}{2p\left[1-\phi'(1)\right]\left[1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)+\sum_{j=1}^{k}\alpha_j\lambda\left(\beta+\beta_j\right)\right]}$$

$$where\ \phi'(1)=\sum_{j=1}^{k}\alpha_j\left[r_1+\lambda p\left(\beta_j+\beta\right)\right]$$

$$\phi''(1)=\sum_{j=1}^{k}\alpha_j\left[\lambda^2p^2\left[\left(\beta_j^{(2)}-\beta_j\right)+2\beta\beta_j\right]+2\lambda pr_1\left(\beta+\beta_j\right)\right]$$

(b) The expected number of customers in the orbit

$$E\left(L_q\right)=P'(1)=\frac{\left[1-\phi'(1)\right]\left\{\lambda^2p^2\left[\left(\beta_j^{(2)}-\beta_j\right)+\left(\beta^{(2)}-\beta\right)+2\beta\beta_j\right]-p\phi''(1)\right\}+\phi''(1)\left[p\left(1-\phi'(1)\right)+\lambda p\left(\beta+\beta_j\right)\right]}{2p\left[1-\phi'(1)\right]\left[1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)+\sum_{j=1}^{k}\alpha_j\lambda\left(\beta+\beta_j\right)\right]}$$

$$where\ \phi'(1)=\sum_{j=1}^{k}\alpha_j\left[r_1+\lambda p\left(\beta_j+\beta\right)\right]$$

$$\phi''(1)=\sum_{j=1}^{k}\alpha_j\left[\lambda^2p^2\left[\left(\beta_j^{(2)}-\beta_j\right)+2\beta\beta_j\right]+2\lambda pr_1\left(\beta+\beta_j\right)\right]$$

Where $\beta^{(2)} and\ \beta_j^{(2)}$ represent the second moments of the service times in the first and second phase of service.

The steady state distribution of the server state is given by

Pr {server is idle} =

$$P_0(1)=\frac{1-\rho}{1-\rho+\left(1-r_1\right)\lambda\sum_{j=1}^{k}\alpha_j\left[\beta+\beta_j\right]}$$

Pr {server is performing 1st phase of service} = $P_1(1)$

$$=\frac{\lambda\beta}{\left(1-r_1\right)\left(1-\rho\right)+\lambda\sum_{j=1}^{k}\alpha_j\left(\beta+\beta_j\right)}$$

Pr {server is performing 2nd phase of service} =

$$\sum_{j=1}^{k}P_2^j(1)=\frac{\lambda\sum_{j=1}^{k}\alpha_j\beta_j}{\left(1-r_1\right)\left(1-\rho\right)+\lambda\sum_{j=1}^{k}\alpha_j\left(\beta+\beta_j\right)}$$

Where $\rho=\sum_{j=1}^{k}\alpha_j\left[\lambda p\left(\beta+\beta_j\right)\right]$

6. Particular cases

In this section, we consider two particular cases of our model,

(i) We first assume that there is no second phase of service and there are no feedbacks (i.e.) we assume that each = 0 and B*j(s) =1 for j=1 to k, r1=0

The value of P0, 0 becomes

This is the same as the results obtained by Falin [5] on page no: 208 for a single sever model with impatient subscribers p is replaced by H_1

(ii) If we assume that there are no retrials (i.e.) retrial rate μ = ∞, we then obtain

$$P(z)=\left[\frac{1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)}{\left[1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)+\sum_{j=1}^{k}\alpha_j\lambda\left(\beta+\beta_j\right)\right]}\right]$$
$$\left[\frac{p\left(z-\phi(z)\right)-1++\sum_{j=1}^{k}\alpha_j B^*\left(\lambda p(1-z)\right)B_j^*\left(\lambda p(1-z)\right)}{p\left(z-\phi(z)\right)}\right]$$

This is the P.G.F of an M/G/1 queue with non-persistent customers, feedbacks and a multi optional second phase of service.

In this expression p=1, β_j = 0, $B^*_j(s)$ =1; $B^*(s)=\dfrac{\mu}{s+\mu}$ then P (z) becomes

$$P(z)=1+\rho\sum_{n=0}^{\infty}\rho^n z^n \text{ where } \rho=\frac{\lambda}{\mu(1-r_1)}$$

equating the coefficient of z^n we obtain

$$P_n=(1-\rho)\,\rho^{n+1}$$

= Probability of (n+1) customers in the system, n ≥ 1.This expression is the same as the expression given by Thangaraj and Vanitha [14].

7. Stochastic Decomposition

P (z) = Q (z). R (z)

$$R(z)=\left[\frac{1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)}{\left[1-r_1-\sum_{j=1}^{k}\alpha_j\lambda p\left(\beta+\beta_j\right)+\sum_{j=1}^{k}\alpha_j\lambda\left(\beta+\beta_j\right)\right]}\right]$$
$$\left[\frac{p\left(z-\phi(z)\right)-1++\sum_{j=1}^{k}\alpha_j B^*\left(\lambda p(1-z)\right)B_j^*\left(\lambda p(1-z)\right)}{p\left(z-\phi(z)\right)}\right]$$

$$\therefore \mathrm{N}(t)=N_1(t)+N_2(t)$$

The random variable $N_2(t)$ represents the number of customers in the queue. Queue with feedback non-persistent customers without retrials. N_2 (t) is the increase in the queue (orbit) size due to the the presence of the retrial customers in the system. Q (z) is the pgf of N_1 (t), R (z) is the pgf of the N_2 (t) in the steady state.

References

1. Artalejo, J.R and Choudhry, G., (2004), *steady state analysis of an M/G/1 queue with repeated attempts and two phase service*, Quality Technology and quantitative Management,Vol.1,PP.189-199.
2. Artalejo J.R. and Gomez-Corral, (2008), *Retrial queueing systems*, Springer.
3. Choi, B.D., and Kulkarni, V.G., (1992), *feedback retrial queueing system, Stochastic model and related model*, 93-105.
4. Choudhry, G., (2003), *some aspects of M/G/1 queueing system with optional second service*, Top, Vol.11, pp.141-150.
5. Choudhry, G., (2009), *An M/G/1 retrial queue with an additional phase of second service and general retrial times*, Information and Management sciences, 20, 1-14.
6. Doshi, B.T, (1991), *Analysis of a two phase queueing system with general service times*, operations research Letters, Vol.10 pp.265-272.
7. Falin, G.I. and Templeton, C.J.G. (1997), *Retrial Queues.* Chapman Hall, London.
8. Falin, G.I., (1990), *A survey of retrial queues*, Queueing Systems, 7, 127- 167.
9. Jinting Wang, (2004), *An M/G/1 queue with second optional service and server breakdowns*, An international journal of computer and Mathematics with applications 47, pp.1713 - 1723.
10. Krishna Kumar, B., Arivudainambi, D and Vijayakumar, A, (2002), *An M/G/1 queueing system with two phase service and preemptive resume*, Annuals of Operations Research, vol.113, pp.61-79.
11. Madan.K.C, (2000), *An M/G/1 queue with second optional service*, Queueing Systems, vol.31, pp-37-36.
12. Medhi.J, (2003),*A single server Poisson input queue with a second optional Channel*, queueing systems, vol.11, pp.141-150.
13. Salchirad.M.R and A.Badamchizadeh, (2009), *On the multi-phase M/G/1 queueing system with random feedback*, CEJOR, 17:131-139.
14. Thangaraj.V and Vanitha.S, (2010), *An M/M/1 queue with feedback a continued fraction approach*, International Journal of Computational and applied Mathematics, vol.5 No.2, pp-129-139.
15. Yang and Templeton.J.G.C, (1987), *A survey on retrial* queueing systems 2, 102-133,
16. Yang Wan Lee and Young Ho Jang, (2009), *the M/G/1 feedback retrial queue with Bernoulli schedule*, Journal of applied Mathematics and Informatics, vol.27, No.1, pp.259-266

Vacation Queues With Impatient Customers and a Waiting Server

R.Padmavathy [1], K.Kalidass[2], Kasturi Ramanath[3]

School of Mathematics, Madurai Kamaraj University, Madurai-625 021. Tamil Nadu. India

[1] vetriselvi.divya@gmail.com

[2] dassmaths@gmail.com

[3]kasturey@gmail.com

***Abstract*— In this paper, we study single server queueing models with impatient customers, server vacations and a waiting server. The arrival process is assumed to be Poisson. The server is allowed to take a vacation whenever the system is empty after waiting for a random period of time. Whenever the server is on vacation, each customer in the queue independently of the other customers sets up an impatience timer. If the server does not return from the vacation before the expiry of the customer impatience time, the customer abandons the system forever. We first take up the case where the service times, customer impatience times, waiting times of the server in the empty system and the duration of the server vacations are all exponentially distributed. We then generalize our results to the case where the service times and the server waiting times have general distributions. We derive the steady state probability generating functions (PGF) of the system size and the queue size. We obtain some useful performance measures for the systems. We discuss some particular cases.**

***Keywords*—** Server vacations, a waiting server, impatient customers, probability generating functions, steady state analysis.

1. Introduction

Queueing systems with server vacations have been extensively studied from the mid 1970's onwards. Various types of vacation schemes have been considered in the literature , the single vacation scheme, the multiple vacation scheme, gated vacation, a limited service discipline, exhaustive service discipline etc. The comprehensive list of results and applications of vacation models is available in the survey paper of Doshi [6] and the monographs of Takagi [11] and Tian and Zhang [12]. A more recent reference is the survey paper of Ke et al. [8]. The concept of a server vacation with a waiting server was first introduced by Boxma et al. [4]. In this vacation scheme, upon finding the system empty, the server waits for a random period of time before proceeding on a vacation. This option of the server wait reflects many real life queueing systems, particularly when dealing with human behaviour.

Customer impatience has been dealt with in the queueing literature mainly in the context of customers abandoning the queue due to either a long wait already experienced, or a long wait anticipated upon arrival. However, there are situations where customer impatience is due to an absentee of servers upon arrival. This situation is encountered, in particular, when observing human behaviour in service systems: if an arriving customer sees no server present in the system, he/she may abandon the queue if no server shows up within sometime.

Altman and Yechiali [2] have studied this phenomenon extensively in their paper. They have considered situations where the server takes multiple vacations as well as situations where the server takes a single vacation. Subsequently Yechiali [13] studied queues with disasters and customer impatience. Adan et al. [1] studied queueing systems with vacations and synchronized reneging. Bae and Kim [3] studied an G/M/1 queue with impatient customers. Chakravarthy [5] studied a disaster queue with Markovian arrivals and impatient customers. Kapodistria [7] studied the M/M/1 queue with synchronized abandonments.

In this paper we have combined these concepts of customer impatience during a server vacation and a server vacation system with a waiting server.

The remaining part of the paper is organised as follows. In section 2, we consider the case when all the random variables involved are exponential. We derive the probabilty generating functions(PGF) of the system size in the steady state. We obtain expressions for the expected numbers in the system when the server is busy and when the server is on vacation. We also obtain the steady state probabilities of the server state. In section 3 , we take up the case where the service times and the waiting times of the server are generally distributed. We use the method of supplementary variables to derive the expressions for the PGF's and the performance measures of the system in the steady state.

2. The M/M/1 queue with customer impatience, server vacations and a waiting server

In this section, we consider a single server queueing system with Poisson arrivals with an intensity λ. The service times are exponentially distributed i.i.d random variables with mean service time $\frac{1}{\mu}$. The server proceeds on a vacation whenever he finds the system empty after a random period of time. This random period of time is exponentially distributed with a parameter η. The duration of the server vacation is assumed to be exponentially distributed with a parameter γ. When the server is on a vacation, each customer sets up an impatience timer independently of the other customers in the system, which is again assumed to be exponentially distributed with a parameter ξ. The customer impatience times, the server waiting times, the inter arrival times, service times and the duration of the server vacations are all assumed to be independent of each other.

2.2 Steady state analysis

Let J denote the number of servers in the system (J=0 implies that the server is on vacation) while L denotes the total number of customers in the system. Then the pair (J,L) defines a continuous time Markov process.

Let $P_{j,n} = Prob\{J = j, L = n\}$ (j=0, 1 ; n=0,1, $\cdots$) denote the steady state system probabilities. The set of balance equations is given as follows:

$$j = 0\begin{cases} n = 0, & (\lambda+\gamma)P_{0,0} = \xi P_{0,1} + \eta P_{1,0}, \\ n \geq 1, & (\lambda + n\xi + \gamma)P_{0,n} = \lambda P_{0,n-1} \\ & + (n+1)\xi P_{0,n+1}. \end{cases} \tag{1}$$

$$j = 1\begin{cases} n = 0, & (\lambda+\eta)P_{1,0} = \gamma P_{0,0} + \mu P_{1,1}, \\ n \geq 1, & (\lambda + \mu)P_{1,n} = \mu P_{1n+1} + \gamma P_{0,n} \\ & + \lambda P_{1,n-1}. \end{cases} \tag{2}$$

Define the following partial generating functions

$$G_0(z) = \sum_{n=0}^{\infty} P_{0,n} z^n,$$

$$G_1(z) = \sum_{n=0}^{\infty} P_{1,n} z^n.$$

Using the above functions in (1) and (2),

$$\xi(1-z)G_{0'}(z) - (\lambda(1-z)+\gamma)G_0(z) = -\eta P_{1,0},$$

$$(\lambda z-\mu)(1-z)G_1(z) - \gamma z G_0(z) = z(\mu P_{1,0} - \eta P_{1,0}) - \mu P_{1,0}. \tag{4}$$

From (3),

$$G_{0'}(z) - \frac{\lambda(1-z)+\gamma}{\xi(1-z)} G_0(z) = \frac{-\eta P_{1,0}}{\xi(1-z)}.$$

Hence,

$$G_0(z) = \frac{-\eta P_{1,0}}{\xi(1-z)^{\frac{\gamma}{\xi}}} e^{\frac{\lambda z}{\xi}} \int_0^z e^{\frac{-\lambda s}{\xi}} (1-s)^{\frac{\gamma}{\xi}-1} ds + \frac{G_0(0)}{(1-z)^{\frac{\gamma}{\xi}}} e^{\frac{\lambda z}{\xi}}. \tag{5}$$

Taking limits as $z \to 1$, we get,

$$G_0(1) = \frac{e^{\frac{\lambda}{\xi}}\left(G_0(0) - \frac{\eta P_{1,0}}{\xi} \int_0^1 e^{\frac{-\lambda s}{\xi}} (1-s)^{\frac{\gamma}{\xi}-1} ds \right)}{\lim_{z\to 1} (1-z)^{\frac{\gamma}{\xi}}}.$$

Since $G_0(1) = \sum_{n=0}^{\infty} P_{0,n} = P_{0*} > 0$ and $\lim_{z\to 1} (1-z)^{\frac{\gamma}{\xi}} = 0$, we must have,

$$G_0(0) = \frac{\eta P_{1,0}}{\xi} \int_0^1 e^{\frac{-\lambda s}{\xi}} (1-s)^{\frac{\gamma}{\xi}-1} ds.$$

Define $K = \int_0^1 e^{\frac{-\lambda s}{\xi}} (1-s)^{\frac{\gamma}{\xi}-1} ds$,

$$G_0(0) = P_{0,0} = \frac{\eta P_{1,0}}{\xi} K.$$ (8)

Hence,

$$G_0(z) = G_0(0)\frac{e^{\frac{\lambda z}{\xi}}}{(1-z)^{\frac{\gamma}{\xi}}}\left[1-\frac{\int_0^z e^{\frac{-\lambda s}{\xi}}(1-s)^{\frac{\gamma}{\xi}-1}ds}{\int_0^1 e^{\frac{-\lambda s}{\xi}}(1-s)^{\frac{\gamma}{\xi}-1}ds}\right].$$ (9)

By applying L'Hospitals rule,

$$G_0(1) = P_{0*} = G_0(0)\frac{\xi}{\gamma K} = \frac{\xi P_{0,0}}{\gamma K},$$ (10)

$$G_0(1) = P_{0*} = \frac{\eta P_{1,0}}{\gamma} \quad (\,from\ (8)),$$ (11)

$$P_{1*} = \sum_{n=0}^{\infty} P_{1,n} = 1 - P_{0*} = 1 - \frac{\xi P_{0,0}}{\gamma K}.$$ (12)

2.2 Derivation of $P_{0*}, P_{1*}, P_{0,0}, E(L_0), E(L_1)$

From (4),

$$G_1(z) = \frac{\gamma z G_0(z) + z(\mu P_{1,0} - \eta P_{1,0}) - \mu P_{1,0}}{(\lambda z - \mu)(1-z)}.$$ (13)

By applying L'Hospitals rule and using (11),

$$G_1(1) = \frac{\gamma G_{0'}(1) + \mu P_{1,0}}{\mu - \lambda} = \frac{\gamma E(L_0) + \mu P_{1,0}}{\mu - \lambda},$$ (14)

$$or \quad P_{1*} = \frac{\gamma E(L_0) + \mu P_{1,0}}{\mu - \lambda},$$

$$E(L_0) = \frac{(\mu - \lambda)P_{1*} - \mu P_{1,0}}{\gamma}.$$ (15)

From (3),

$$G_{0'}(z) = \frac{-\eta P_{1,0} + (\lambda(1-z) + \gamma)G_0(z)}{\xi(1-z)}.$$

By applying L'Hospitals rule,

$$G_{0'}(1) = \frac{-\lambda P_{0*} + \gamma G_{0'}(1)}{-\xi},$$

$$E(L_0) = \frac{\lambda P_{0*}}{\gamma + \xi}.$$ (16)

Equating (15) and (16) and using $P_{0*} + P_{1*} = 1$,

$$P_{0*} = \frac{(\gamma + \xi)(\mu(1 - P_{10}) - \lambda)}{\mu\gamma + \xi(\mu - \lambda)},$$

$$P_{0*} = \frac{(\gamma + \xi)(\mu(\eta K - \xi P_{0,0}) - \lambda\eta K)}{(\mu\gamma + \xi(\mu - \lambda))\eta K} \quad (\,from\ (8)),$$ (17)

$$P_{1*} = 1 - P_{0*} = \frac{P_{0,0}\xi\mu(\gamma + \xi) + \lambda\gamma\eta K}{(\mu\gamma + \xi(\mu - \lambda))\eta K}.$$ (18)

Equating P_{0*} from (10) and (17),

$$\frac{\xi P_{0,0}}{\gamma K} = \frac{(\gamma + \xi)(\mu(\eta K - \xi P_{0,0}) - \lambda\eta K)}{(\mu\gamma + \xi(\mu - \lambda))\eta K},$$

$$P_{0,0} = \frac{\gamma(\gamma + \xi)\eta K(\mu - \lambda)}{(\mu\gamma\eta + \xi\mu\eta - \xi\lambda\eta + \gamma^2\mu + \xi\gamma\mu)\xi}.$$ (19)

From (8) and (19),

$$P_{1,0} = \frac{\gamma(\gamma+\xi)(\mu-\lambda)}{\mu\gamma\eta + \xi\mu\eta - \xi\lambda\eta + \gamma^2\mu + \xi\gamma\mu}. \qquad (20)$$

Substituting (17) in (16),

$$E(L_0) = \frac{\lambda(\mu(\eta K - \xi P_{0,0}) - \lambda\eta K)}{(\mu\gamma + \xi(\mu-\lambda))\eta K}. \qquad (21)$$

Now to find $E(L_1)$,

From (4)

$$G_1(z) = \frac{\gamma z G_0(z) + z(\mu P_{1,0} - \eta P_{1,0}) - \mu P_{1,0}}{(\lambda z - \mu)(1-z)}.$$

By applying L'Hospitals rule,

$$G_{1'}(1) = \frac{2\mu\gamma G_{0'}(1) - (\lambda-\mu)\gamma G_0''(1) + 2\lambda\mu P_{1,0}}{2(\lambda-\mu)^2},$$

$$G_{1'}(1) = \frac{2\mu\gamma E(L_0) - (\lambda-\mu)\gamma E(L_0(L_0-1)) + 2\lambda\mu P_{1,0}}{2(\lambda-\mu)^2},$$

$$E(L_1) = \frac{2\mu\gamma E(L_0) - (\lambda-\mu)\gamma E(L_0(L_0-1)) + 2\lambda\mu P_{1,0}}{2(\lambda-\mu)^2}.$$

where E(L $_0$) is given by equation (21), $P_{1,0}$ is given by equation (20) and

$$E(L_0(L_0-1)) = \frac{2\lambda^3\eta(\mu-\lambda)}{(2\xi+\gamma)[\eta(\lambda\mu(\xi+\gamma)-\lambda^2\xi)+\mu\lambda\gamma(\xi+\gamma)]}$$

The PGF of the system size is given by,

$$G(z) = G_0(z) + G_1(z),$$

$$= G_0(z)\left[1 + \frac{\gamma z}{(\lambda z-\mu)(1-z)}\right] + \frac{z(\mu P_{1,0} - \eta P_{1,0}) - \mu P_{1,0}}{(\lambda z-\mu)(1-z)}. \qquad (22)$$

The PGF of the queue size is given by,

$$H(z) = G_0(z) + P_{1,0} + \frac{1}{z}[G_1(z) - P_{1,0}],$$

$$= G_0(z)\left[1 + \frac{\gamma}{(\lambda z-\mu)(1-z)}\right] + P_{1,0}\left[1 + \frac{\lambda z - \lambda - \eta}{(\lambda z-\mu)(1-z)}\right]. \qquad (23)$$

where $G_0(z)$ is given by (9).

2.3 Particular cases

Taking limits as η tends to ∞ in (19), (9) and (13),

$$P_{0,0} = \frac{\gamma K}{\xi}\frac{(\gamma+\xi)(\mu-\lambda)}{\mu\gamma+\xi(\mu-\lambda)},$$

$$G_0(z) = \frac{G_0(0)e^{\frac{\lambda}{\xi}z}}{(1-z)^{\frac{\gamma}{\xi}}}\left[1 - \frac{\int_0^z (1-s)^{\frac{\gamma}{\xi}-1} e^{-\frac{\lambda}{\xi}s} ds}{\int_0^1 (1-s)^{\frac{\gamma}{\xi}-1} e^{-\frac{\lambda}{\xi}s} ds}\right],$$

$$G_1(z) = \frac{[\gamma G_0(z) - (\mu P_{1,1} + \gamma P_{0,0})]z}{(\lambda z-\mu)(1-z)}.$$

These values coincide with (2.18), (2.12) and (2.3) of [2]. In this case, the system reduces to the multiplication vacation scheme.

3. The M/G/1 queue with customer impatience, server vacations and a waiting server

In this section, we consider a M/G/1 queueing system with server vacations and a waiting server. Whenever the server finds the system empty, he waits for a random period of time before proceeding on a vacation. This random period of time is assumed to be generally distributed with a distribution function V(x), a hazard rate function $\eta(x)$ and a Laplace Stieltjes transform (LST) $V^*(s)$ respectively. The duration of the server vacation is assumed to be exponentially distributed with a parameter γ. The service time distribution is taken as B(x) with a hazard rate function r(x) and a Laplace Stieltjes transform $B^*(s)$. When the server is away on a vacation, each customer sets up an impatience timer independently of the other customers. If the duration of the timer is completed before the server returns from the vacation, the customer abandons the queue.

The server waiting times, the customer impatience times, interarrival times, service times and the duration of the server vacations are all assumed to be independent of each other. Let C(t) denote the state of the server at time t i.e.,

$$C(t) = 0, \textit{if the server is on vacation},$$

$$C(t) = 1, \textit{if the server is in the system.}$$

Let N(t) denote the number of customers in the system at time t.

Let X(t) denote the elapsed service time of the customer in service when C(t)=1, N(t)=n ≥ 1. Let Y(t) denote the elapsed waiting time of the server when C(t)=1, N(t)=0. Then $\{C(t), N(t), X(t), Y(t) : t \geq 0\}$ is a continuous time Markov chain. We define

$$P_{0,n}(t) = Pr\{C(t) = 0, N(t) = n\},$$

$$P_{1,0}(x,t)dx = Pr\{C(t) = 1, N(t) = 0, x \leq Y(t) < x + dx\},$$

$$P_{1,n}(x,t)dx = Pr\{C(t) = 1, N(t) = n, x \leq X(t) < x + dx\}.$$

The Kolmogorov forward equations for the system are as follows:

$$P_{0,0}'(t) = -(\lambda + \gamma)P_{0,0}(t) + \int_0^\infty P_{1,0}(x,t)\eta(x)dx + \xi P_{0,1}(t), \tag{24}$$

$$P_{0,n}'(t) = -(\gamma + \lambda + n\xi)P_{0,n}(t) + \lambda P_{0,n-1}(t) + (n+1)\xi P_{0,n+1}(t), \quad n \geq 1, \tag{25}$$

$$\frac{\partial}{\partial t}P_{1,0}(x,t) + \frac{\partial}{\partial x}P_{1,0}(x,t) = -(\eta(x) + \lambda)P_{1,0}(x,t), \tag{26}$$

$$\frac{\partial}{\partial t}P_{1,1}(x,t) + \frac{\partial}{\partial x}P_{1,1}(x,t) = -(\lambda + r(x))P_{1,1}(x,t), \tag{27}$$

$$\frac{\partial}{\partial t}P_{1,n}(x,t) + \frac{\partial}{\partial x}P_{1,n}(x,t) = -(\lambda + r(x))P_{1,n}(x,t) + \lambda P_{1,n-1}(x,t), \quad n \geq 2, \tag{28}$$

$$P_{1,0}(0,t) = \gamma P_{0,0}(t) + \int_0^\infty P_{1,1}(x,t)r(x)dx, \tag{29}$$

$$P_{1,1}(0,t) = \gamma P_{0,1}(t) + \int_0^\infty P_{1,2}(x,t)r(x)dx + \int_0^\infty \lambda P_{1,0}(x,t)dx, \tag{30}$$

$$P_{1,n}(0,t) = \gamma P_{0,n}(t)$$

$$+\int_0^\infty r(x)P_{1,n+1}(x,t)dx, \quad n \geq 2. \qquad (31)$$

Assume that the steady state is attained.

Taking limits as $t \to \infty$, we obtain,

$$(\lambda+\gamma)P_{0,0} = \int_0^\infty P_{1,0}(x)\eta(x)dx + \xi P_{0,1}, \qquad (32)$$

$$(\gamma+\lambda+n\xi)P_{0,n} = \lambda P_{0,n-1} + (n+1)\xi P_{0,n+1}, \quad n \geq 1, \qquad (33)$$

$$\frac{d}{dx}P_{1,0}(x) = -(\eta(x)+\lambda)P_{1,0}(x), \qquad (34)$$

$$\frac{d}{dx}P_{1,1}(x) = -(\lambda + r(x))P_{1,1}(x), \qquad (35)$$

$$\frac{d}{dx}P_{1,n}(x) = -(\lambda + r(x))P_{1,n}(x)$$

$$+\lambda P_{1,n-1}(x), \quad n \geq 2 \qquad (36)$$

$$P_{1,0}(0) = \gamma P_{0,0} + \int_0^\infty r(x)P_{1,1}(x)dx, \qquad (37)$$

$$P_{1,1}(0) = \gamma P_{0,1} + \int_0^\infty r(x)P_{1,2}(x)dx + \int_0^\infty \lambda P_{1,0}(x)dx,$$

$$P_{1,n}(0) = \gamma P_{0,n} + \int_0^\infty r(x)P_{1,n+1}(x)dx, \quad n \geq 2.$$

(39)

Define

$$G_0(z) = \sum_{n=0}^\infty P_{0,n}z^n,$$

$$G_1(x,z) = \sum_{n=0}^\infty P_{1,n}(x)z^n,$$

$$P_1(x,z) = \sum_{n=1}^\infty P_{1,n}(x)z^n.$$

From (35) and (36), after multiplying by z^n, and summing over all values of n, we obtain the following equation,

$$P_1(x,z) = P_1(0,z)(1-B(x)e^{-\lambda(1-z)x}. \qquad (40)$$

Similarly, from (32), (33),

$$[\lambda(1-z)+\gamma]G_0(z) - \xi(1-z)G_{0'}(z) = \int_0^\infty P_{1,0}(x)\eta(x)dx. \qquad (41)$$

Solving (34),

$$P_{1,0}(x) = P_{1,0}(0)(1-V(x))e^{-\lambda x}. \qquad (42)$$

Solving (35),

$$P_{1,1}(x) = P_{1,1}(0)(1-B(x))e^{-\lambda x}. \qquad (43)$$

From (38),(39) after multiplying by z^n, and summing over all values of n, we obtain the following equation,

$$P_1(0,z) = \gamma G_0(z) + \frac{1}{z}\int_0^\infty r(x)P_1(x,z)dx - P_{1,0}(0) \qquad (38)$$

$$+\int_0^\infty \lambda P_{1,0}(x)zdx. \qquad (44)$$

Substituting (42) in (41) and solving,

$$G_{0'}(z) - \frac{\lambda(1-z)+\gamma}{\xi(1-z)}G_0(z) = -\frac{P_{1,0}(0)V^*(\lambda)}{\xi(1-z)}.$$

Solving (45),

$$G_0(z) = (1-z)^{-\frac{\gamma}{\xi}} e^{\frac{\lambda z}{\xi}} \left(G_0(0) - \frac{P_{1,0}(0)}{\xi} V^*(\lambda) \times \int_0^z (1-s)^{\frac{\gamma}{\xi}-1} e^{\frac{-\lambda s}{\xi}} ds \right). \quad (46)$$

$$Now,\ G_0(1) = \frac{e^{\frac{\lambda}{\xi}} \left(G_0(0) - \frac{P_{10}(0)}{\xi} V^*(\lambda) \times \int_0^1 (1-s)^{\frac{\gamma}{\xi}-1} e^{\frac{-\lambda s}{\xi}} ds \right)}{\lim_{z\to 1} (1-z)^{\frac{\gamma}{\xi}}}.$$

Since $G_0(1) = \sum_{n=0}^{\infty} P_{0n} > 0$ and $\lim_{z\to 1}(1-z)^{\frac{\gamma}{\xi}} = 0,$

we have,

$$G_0(0) = \frac{P_{1,0}(0)}{\xi} V^*(\lambda) K,\ where$$
$$K = \int_0^1 (1-s)^{\frac{\gamma}{\xi}-1} e^{-\frac{\lambda s}{\xi}} ds. \quad (47)$$

By substituting (47) in (46),

$$G_0(z) = (1-z)^{-\frac{\gamma}{\xi}} e^{\frac{\lambda z}{\xi}} \frac{P_{1,0}(0)}{\xi} V^*(\lambda) \times \left[K - \int_0^z (1-s)^{\frac{\gamma}{\xi}-1} e^{-\frac{\lambda s}{\xi}} ds \right]. \quad (48)$$

Substituting (42) in (44),

$$P_1(0,z) = \gamma G_0(z) + \frac{1}{z}\int_0^\infty r(x) P_1(x,z) dx - P_{1,0}(0)(1-z) - zV^*(\lambda) P_{1,0}(0). \quad (49)$$

Substituting (40) in (49),

$$P_1(0,z) = \frac{z[\gamma G_0(z) - P_{1,0}(0)(1-z) - zV^*(\lambda)P_{1,0}(0)]}{z - B^*(\lambda(1-z))}. \quad (50)$$

Substituting (50) in (40),

$$P_1(x,z) = \frac{z[\gamma G_0(z) - P_{1,0}(0)(1-z) - zV^*(\lambda)P_{1,0}(0)]}{z - B^*(\lambda(1-z))} \times (1-B(x)) e^{-\lambda(1-z)x}. \quad (51)$$

Consider $G_1(z) = \int_0^\infty G_1(x,z) dx$,

$$= \frac{z[\gamma G_0(z) - P_{1,0}(0)(1-z) - zV^*(\lambda)P_{1,0}(0)][1 - B^*(\lambda(1-z))]}{[z - B^*(\lambda(1-z))]\lambda(1-z)} + P_{1,0}(0)\frac{1-V^*(\lambda)}{\lambda}. \quad (52)$$

From (48), by applying L'Hospitals rule,

$$G_0(1) = \frac{P_{1,0}(0) V^*(\lambda)}{\gamma}. \quad (53)$$

From (52), by applying L'Hospitals rule,

$$G_1(1) = P_{1,0}(0)\frac{1-V^*(\lambda)}{\lambda} + \frac{[\gamma G_{0'}(1) + P_{1,0}(0)(1-V^*(\lambda))]E(B)}{1-\rho}. \quad (54)$$

where $\rho = \lambda E(B) = -\lambda B^{*'}(0)$.

From (45), by applying L'Hospitals rule,

$$G_{0'}(1) = \frac{\lambda G_0(1)}{\xi+\gamma}$$

$$= \frac{\lambda}{\xi+\gamma}\frac{P_{1,0}(0)V^*(\lambda)}{\gamma} \quad (\textit{from}\ (52)). \tag{55}$$

Substituting (55) in (54),

$$G_1(1) = P_{1,0}(0)\frac{V^*(\lambda)[\lambda\rho-\xi-\gamma]+(\xi+\gamma)}{\lambda(1-\rho)(\xi+\gamma)}. \tag{56}$$

Using $G_0(1)+G_1(1)=1$,

$$P_{1,0}(0) = \frac{\gamma\lambda(1-\rho)(\xi+\gamma)}{V^*(\lambda)[\lambda(\xi+\gamma)-\lambda\rho\xi-(\xi+\gamma)\gamma]+(\xi+\gamma)\gamma}. \tag{57}$$

From (56) and (57),

$$G_1(1) = \frac{\gamma[V^*(\lambda)(\lambda\rho-\xi-\gamma)+(\xi+\gamma)]}{V^*(\lambda)[\lambda(\xi+\gamma)-\lambda\rho\xi-(\xi+\gamma)\gamma]+(\xi+\gamma)\gamma}. \tag{58}$$

System size PGF

$$S(z) = G_0(z)+G_1(z)$$

$$= P_{1,0}(0)\left\{\begin{aligned}&\frac{e^{\frac{\lambda z}{\xi}}V^*(\lambda)}{(1-z)^{\frac{\gamma}{\xi}}\xi}\left[K-\int_0^z(1-s)^{\frac{\gamma}{\xi}-1}e^{-\frac{\lambda s}{\xi}}ds\right]\\&\times\left[1+\frac{\gamma z(1-B^*(\lambda(1-z)))}{[z-B^*(\lambda(1-z))]\lambda(1-z)}\right]\end{aligned}\right.$$

$$\left.-\left[\frac{[1-z+zV^*(\lambda)]z[1-B^*(\lambda(1-z))]}{[z-B^*(\lambda(1-z))]\lambda(1-z)}\right]+\frac{1-V^*(\lambda)}{\lambda}\right\}. \tag{59}$$

Queue size PGF

$$H(z) = G_0(z)+\frac{1}{z}[G_1(z)-P_{1,0}]+P_{1,0}$$

$$= P_{1,0}(0)\left\{\begin{aligned}&\frac{e^{\frac{\lambda z}{\xi}}V^*(\lambda)}{(1-z)^{\frac{\gamma}{\xi}}\xi}\left[K-\int_0^z(1-s)^{\frac{\gamma}{\xi}-1}e^{-\frac{\lambda s}{\xi}}ds\right]\\&\times\left[1+\frac{\gamma(1-B^*(\lambda(1-z)))}{[z-B^*(\lambda(1-z))]\lambda(1-z)}\right]\end{aligned}\right.$$

$$\left.+\frac{1-V^*(\lambda)}{\lambda}-\frac{[1-z+zV^*(\lambda)][1-B^*(\lambda(1-z))]}{[z-B^*(\lambda(1-z))]\lambda(1-z)}\right\}. \tag{60}$$

3.1 Performance measure

Pr{server is idle} = $P_{1,0} = \int_0^\infty P_{1,0}(x)dx,$

$$= P_{1,0}(0)\frac{[1-V^*(\lambda)]}{\lambda}, \tag{61}$$

$$= \frac{\gamma(1-\rho)(\xi+\gamma)(1-V^*(\lambda))}{V^*(\lambda)\big[\lambda(\xi+\gamma)-\lambda\rho\xi-(\xi+\gamma)\gamma\big]+(\xi+\gamma)\gamma}. \tag{62}$$

Pr{server is busy}= $G_1(1)-P_{1,0}$,

$$= \frac{\gamma\rho\big[V^*(\lambda)(\lambda-(\xi+\gamma))+(\xi+\gamma)\big]}{V^*(\lambda)\big[\lambda(\xi+\gamma)-\lambda\rho\xi-(\xi+\gamma)\gamma\big]+(\xi+\gamma)\gamma}. \tag{63}$$

Pr{server is on vacation}= $G_0(1)$,

$$= \frac{V^*(\lambda)\lambda(1-\rho)(\xi+\gamma)}{V^*(\lambda)\big[\lambda(\xi+\gamma)-\lambda\rho\xi-(\xi+\gamma)\gamma\big]+(\xi+\gamma)\gamma}. \tag{64}$$

Expected number of customers in system

$S'(1) = G_{0'}(1) + G_{1'}(1),$

$$= \frac{E(B)\gamma G_0''(1)}{2(1-\rho)} + P_{1,0}(0)$$

$$\times\left\{\begin{array}{l}\dfrac{\lambda V^*(\lambda)}{\gamma(\xi+\gamma)} \\ \times\left[1+\dfrac{\gamma[\lambda\rho B^{*''}(0)+(\lambda B^{*''}(0)-2B^{*'}(0)) \times(1-\rho)]}{2(1-\rho)^2}\right]\end{array}\right.$$

$$\left.+\frac{[\lambda\rho B^{*''}(0)+(\lambda B^{*''}(0)-2B^{*'}(0))(1-\rho)] \times(1-V^*(\lambda))}{2(1-\rho)^2}\right\}.$$

Expected number of customers in queue

$$H'(1) = G_{0'}(1) + G_{1'}(1) - G_1(1) + P_{1,0},$$

$$= \frac{E(B)\gamma G_0''(1)}{2(1-\rho)} + P_{1,0}(0)$$

$$\times\left\{\frac{\lambda V^*(\lambda)}{\gamma(\xi+\gamma)}\left[1+\frac{\gamma[\lambda\rho B^{*''}(0)+(\lambda B^{*''}(0)-2B^{*'}(0)) \times(1-\rho)]}{2(1-\rho)^2}\right]\right.$$

$$+\frac{[\lambda\rho B^{*''}(0)+(\lambda B^{*''}(0)-2B^{*'}(0))(1-\rho)](1-V^*(\lambda))}{2(1-\rho)^2}$$

$$\left.-\frac{V^*(\lambda)[\lambda\rho-(\xi+\gamma)]+(\xi+\gamma)}{\lambda(1-\rho)(\xi+\gamma)}+\frac{1-V^*(\lambda)}{\lambda}\right\},$$

$$\text{where } G_0''(1) = \lim_{z\to 1} G_0''(z) = \frac{2\lambda^2 V^*(\lambda)P_{1,0}(0)}{\gamma(2\xi+\gamma)(\xi+\gamma)}.$$

3.2 Particular cases

Case 1: Taking service time to be exponential with parameter μ and waiting time of the server to be exponential with a parameter η ,

$$G_0(z) = \frac{G_0(0)e^{\frac{\lambda z}{\xi}}}{(1-z)^{\frac{\gamma}{\xi}}}\left[1-\frac{\int_0^z e^{-\frac{\lambda s}{\xi}}(1-s)^{\frac{\gamma}{\xi}-1}ds}{\int_0^1 e^{-\frac{\lambda s}{\xi}}(1-s)^{\frac{\gamma}{\xi}-1}ds}\right],$$

$$G_1(z) = \frac{\gamma z G_0(z) + z(\mu P_{1,0} - \eta P_{1,0}) - \mu P_{1,0}}{(\lambda z-\mu)(1-z)}.$$

These expressions coincide with (9) and (13).

Case 2: Assume $V^*(\lambda)=1$, then the queueing system reduces to the system with multiple vacation,

$$P_{0,0} = G_0(0) = \frac{\gamma K}{\xi}\frac{(\xi+\gamma)(\mu-\lambda)}{\gamma\lambda+(\xi+\gamma)(\mu-\lambda)}. \quad (65)$$

The value of $P_{0,0}$ in (65) coincide with the $P_{0,0}$ on page number 268 of [2].

References

[1] Adan, I., Economou, A. and Kapodistria, S., *Synchronized reneging in queueing systems with vacations*, Queueing Systems , 62: 1–33, 2009.

[2] Altman, E. and Uri Yechiali., *Analysis of customers impatience in queues with server vacation*, 52, (261-279), 2006.

[3] Bae, J. and Kim, S., *The stationary workload of the G/M/1 queue with impatient customers.* Queueing systems, 64, 253-265, 2010.

[4 Boxma O.J., Schlegel S. and Yechiali U., *A note on an M/G/1 queue with a waiting server timer and vacations,* American Mathematical society Translations, series 2, 207, (25-35), 2002.

[5] Charavarthy, S. R., *A disaster queue with Markovian arrivals and impatient customers*, Applied Mathematics and Computation 214) 48–59. 2009.

[6] Doshi, B. Queueing systems with vacations-a survey. Queueing Systems. 1, 29-66, 1986.

[7] Kapodistria, S., *The M/M/1 queue with synchronized abandonments*, Queueing systems, 68, 79-109. 2011.

[8] Ke, J.C., Wu, C.H. and Zhang, Z. G. *Recent developments in vacations queueing models: A short survey*, International Journal of Operations Research , 7 (4), 3-8 2010.

[9] Medhi, J., *Stochastic models in Queueing Theory, Second Edition*, Academic Press

[10] Medhi, J., *Stochastic Processes, Second Edition*, Wiley Eastern Ltd.

[11] Takagi, H.: Queueing analysis, Vol. 1, Vacation and Priority systems. North-Holland Elsevier, Amsterdam, 1991.

[12] Tian, N. and Zhang, Z. Vacation queueing models-theory and applications. Springer, New York, 2006

[13] Yechiali, U., *Queues with system disasters and impatient customers when system is down*, Queueing systems, 56, 195-202. 2007

Fault Detection in SRAM Cell Using Wavelet Transform Based Transient Current Testing Method

N.M.Sivamangai[#1], K.Gunavathi[*2]

[#]Assistant Professor, [*]Professor

Department of Electronics and Communication Emgineering, PSG College of Technology, Peelamedu

Coimbatore-641004, TamilNadu, India.

[1]nmsivam@gmail.com

Abstract— **This paper proposes a novel transient power supply current (IDDT) testing method to detect faults in static random access memories (SRAM) cell, using wavelet transform based IDDT waveform analysis. IDDT provides a window of observability into the internal switching characteristics of a SRAM cell. Wavelet transform has the potential to resolve a signal in both time and frequency domain simultaneously. Hence, by monitoring a dynamic current pulse during a transition write operation and analyzing the waveform through wavelet coefficients decomposition, faults are detected. The time complexity of the proposed technique is 2n (n is the number of bits) as compared to the conventional March tests with minimum time complexity of 4n. Hence 50% reduction of test time is obtained by the proposed technique. Moreover, the test vector generation in the proposed technique is independent to the type of fault present in the SRAM cell resulting in higher fault coverage. For detecting the fault, the complete set of wavelet coefficients for the IDDT waveform is used as a signature to compare a faulty cell with a fault-free one and to make pass/fail decisions. The proposed technique has been validated by SPICE simulation and Matlab wavelet tool box.**

Keywords— *SRAM, transient power supply current, Wavelet transform, March tests, fault coverage.*

1. Introduction

Memory testing is one of the toughest issues in the area of testing due to the large amount of memory cells required by the applications and the limited number of chip access pins available on a single chip [1]. Testing SRAM is different from testing logic circuits since SRAM are mostly mixed-signal devices whose faulty behaviors are often analog in nature [2]. Thus, fault coverage provides a better estimate for the overall test quality. Test engineers typically seek to maximize fault coverage by applying different test algorithms such as March tests [3]. March algorithm has a time complexity of O(n) where n is the number of bits. Hence, as the density of memory increases the test time for March algorithm also increases. In addition, the very long test lengths are due to the algorithms extensive read operations [3]. Many algorithms have been proposed to reduce the test length [4]. But the fault coverage of shorter test length algorithm is minimum.

SRAM cell faults can also be tested by measuring quiescent power supply current (IDDQ) testing [5]. Normal IDDQ levels in a SRAM cell are mostly due to current leakages. Elevated IDDQ levels could indicate SRAM faults. However, if a fault such as Pattern Sensitive Faults (PSF) does not result in a measurable extra IDDQ in the memory circuit, then this fault remains undetected. In addition, IDDQ test measurements require the circuit to stabilize in the steady state resulting in additional delay, thus increasing the test time. While many solutions have been proposed to deal with the background leakage elevation in IDDQ testing, IDDT testing has emerged as an alternative and/or supplementary testing method. The IDDT waveform analysis is an effective technique to detect many of the faults that can occur in SRAM cells, including resistive opens, which may not be detected by conventional IDDQ testing methods. The IDDT pulse peak level which is appreciably greater than IDDQ is due to two components – direct transient path between power and ground, path between the output of the switching gate and either power or ground depending on whether the cell is switching from a high state to a low state or from a low state to a high state. Defects such as opens may prevent the gate from having a normal conducting path between power and ground during switching which can appreciably change the level of IDDT.

Many of the researchers worked on detecting faults using IDDT pulse. In [6], authors have evaluated a technique that uses power supply charge as the test observable and shown that the method is efficient to detect those faults that prevent current elevation. In [7], authors have investigated the potential of transient current testing in faulty chip detection with silicon devices. A method of testing dynamic CMOS circuits using the transient power supply current is proposed in [8]. In [9], an analysis of transient current testing technique that measures and computes the charge delivered to the circuit during the transient operation was made and the results indicate that the Charge Based Testing (CBT) can successfully test submicron ICs.

In [10], authors have developed and analyzed a hierarchical power-grid equivalent circuit model for transient current testing evaluation. A new dynamic power supply current testing method to detect open defects in CMOS SRAM cells by monitoring the dynamic current pulse during a transition write operation or read operation is proposed in [11]. In [12] the test efficiency of SRAM has been improved by using, on chip dynamic power supply current sensors. While there have been many investigations on fault detection by monitoring the magnitude of the transient current pulse directly, some faults may remain undetected due to lack of analysis of the current pulse. Hence in this work, a deeper analysis of the transient current signature is made using wavelet transform for better fault coverage and minimum test time.

2. Transient Current Testing

A basic assumption for synchronous circuits is that the primary input signals can change only when all internal node signals in the circuit are in stable state. In a stable state, IDDT is very small. But once the primary input signals change, it takes some time for all internal node signals to stabilize, resulting in IDDT. IDDT testing is a test method by means of measuring the average transient current of the V_{DD} power supply during the time interval from the input vector change to corresponding stable state [13]. When some inputs change from logic 1(0) to logic 0(1), the power supply current instantaneously varies with current pulses before all gate outputs are stabilized. Its time average is taken in the transition duration termed average transient current, IDDT. All gate signals are modeled as logical signals, and the time interval from the first input change to the last output change is called the transition duration. It is noticed that a multiple logical transition at some gate output can occur due to its multiple input change at different times. It is just those multiple logical transitions that result in IDDT pulses.

The amount of IDDT is influenced by many factors, for instance, power supply voltage, threshold voltage, IC technological parameters etc. Some defects in IC may significantly influence IDDT. Suppose IDDT increases or decreases significantly due to a defect in the IC under test, and its variation can be large enough to be observable. Then, those faults are considered to be IDDT testable, and the input vector pair resulting in this large IDDT variation is the IDDT test pair for the fault. Hence an IDDT testable fault needs to be identified and IDDT test pair generated.

For analysis, a CMOS six-transistor SRAM cell was designed at 0.13µm CMOS technology as shown in Fig. 1.

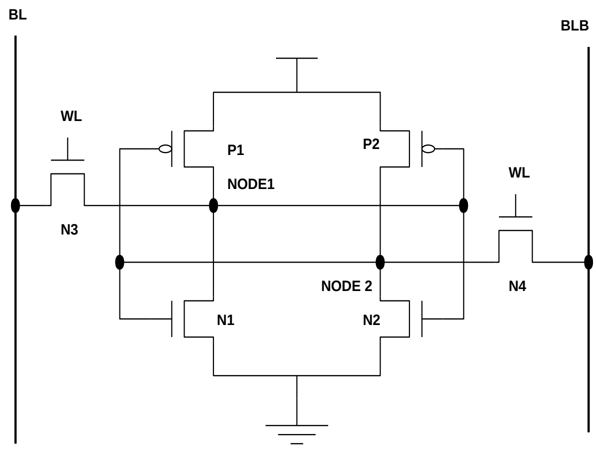

Figure 1. Conventional fault free 6T SRAM cell

'WL' indicates word line, 'BL' and 'BLB' represent bit line and bit line bar respectively. Whenever the cell switches its state, a measurable transient current pulse is established. Table I shows the peak value of the transient current pulse, observed in the fault free SRAM cell during its operations. The peak value of the transient current pulse gives useful information about the switching behavior of an SRAM cell [14].

Table 1 shows that a transition write operation i.e., changing the contents of a cell from 0 to 1 or from 1 to 0 establishes a higher transient current than other operations. Hence the test sequence Write '0' Write '1' forms the test vector for the proposed technique. The current values are measured at the time instant when the cell contents change. Thus for fault detection, the peak value of the transient current pulse of a cell is sensed during a transition write operation and then compared it with the fault free condition. If the transient current pulse peak value of the observed cell is different from that of the fault free cell, it is conclude that the cell is faulty.

Table 1. Transient current pulse values for fault free SRAM cell

SRAM cell operation	Peak value of Transient current pulse
Write ‘0’ read ‘0’	5.3 µA
Write ‘1’ read ‘1’	5.3 µA
Write ‘0’ Write ‘1’	112.53 µA
Write ‘1’ Write ‘0’	112.53 µA

2.1 SRAM fault models

To identify the peak value of the transient current pulse due to different faults, a memory cell designed at 0.13µm CMOS technology with six different faults were injected into the SRAM cell as shown in Fig. 2. It depicts the scheme of a conventional 6T SRAM cell, where the defects are injected into the cell by adding resistances of appropriate values at appropriate locations. They are placed on the interconnections, where the probability of occurrences is higher [16]. The defects are not injected into all possible locations because of the symmetry of the core-cell, the chosen six locations allow an exhaustive analysis of the resistive-open and resistive-short defects in the core cell.

Table 2 shows a summary of the fault models identified for each injected resistive-open and resistive-short defects, according to the conditions which maximize the fault detection, i.e., the minimum detectable resistance value. The first column gives the defect names. The second column gives the minimum resistance value that induces a faulty behaviour and the third column shows the related fault models.

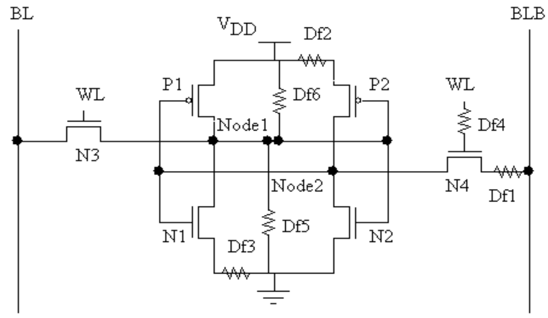

Figure 2. Faults injected 6T SRAM cell

Table 2. Summary of Fault models

Defects	Resistance value	Fault model
Df1	1GΩ	TF
Df2	1GΩ	DRF
Df3	1GΩ	RDF
Df4	1GΩ	SOpF
Df5	1Ω	SA0
Df6	1Ω	SA1

The definitions of the fault models reported in Table 2 are the following ones:

Transition Fault (TF): A cell is said to have a TF if it fails to undergo a transition (0 → 1 or 1 → 0) when it is written.

Data Retention Fault (DRF): A cell in the presence of a data retention fault can write and memorize the input data, but it fails to retain the logic value after some time.

Read Destructive Fault (RDF): A cell is said to have an RDF if a read operation performed on the cell changes the data in the cell and returns an incorrect value on the output.

Stuck-Open Fault (SOpF): A cell is said to be stuck-open if it is not possible to access the cell by any action on the cell.

Stuck-At-0 Fault (SAF0): A cell is said to be stuck-at-0 if the logic value of the cell remains 0 and cannot be changed by any action on the cell or by influences from other cells.

Stuck-At-1 Fault (SAF1): A cell is said to be stuck-at-1 if the logic value of the cell remains 1 and cannot be changed by any action on the cell or by influences from other cells.

Simulation is carried out by injecting each fault in the conventional 6T SRAM cell, to detect the peak value of the transient current pulse. Table 3 shows the peak value of the transient current pulse for a faulty SRAM cell during a transition write operation.

Table 3 shows that increase in peak value of transient current pulse for SOpF (124.21 µA) is very small when compared to fault free SRAM cell (112.53 µA). This may cause the fault to be undetected by only measuring the peak value of the current pulse. Moreover, the transient currents of the cells are of very short duration and hence are difficult to be measured. They are pulses of few µA existing only for a few µs. Hence detection of these values and using them for comparisons between faulty and free cells is tedious. Hence the need of a transform which widens this time pulse for detection and comparison of faults without losing the information about the timing instants of these pulses is necessary. Therefore, wavelet transform is used in this paper to analyze the current pulses, due to its ability to resolve signals in both time and frequency domain.

Table 3. Transient current pulse values for faulty SRAM cell

Fault model	Peak value of Transient current pulse
TF	190.45 µA
DRF	64.55 µA
RDF	64.21 µA
SOpF	124.21 µA
SA0	406.15 µA
SA1	658.82 µA

3. Wavelet Transform

The signals when analyzed using Fourier has a serious drawback since it transforms signal in frequency domain losing all information on how the signal is spatially distributed. Wavelet Transform (WT) of a signal, on the other hand, decomposes signal in both time and frequency domain, which turns out to be very useful in fault detection [15]. The Wavelet transform is a tool that divides up data, functions or operators into different frequency components and then studies each component with a resolution matched to its scale. It helps in archiving the localization in frequency and time, and is able to focus on short-time intervals for high frequency components and long intervals for low frequency components, making it a well suited tool for analyzing high frequency transients in the presence of low frequency components.

3.1.1 Continuous Wavelet Transform (CWT)

The continuous wavelet transform (CWT) of a continuous signal x (t) is defined as

$$WT\left(a,b\right)=\int_{\infty}^{-\infty} X(t)\psi_{a,b}^{*}dt \qquad (1)$$

$$\psi_{a,b}^{*}=\frac{1}{\sqrt{a}}\psi\,\frac{(t-b)}{a} \qquad (2)$$

where ψ (t) is the basis function or mother wavelet, a, b are real and * indicates complex conjugate. Wavelet analysis attempts to express the signal x(t) in terms of a series of shifted and scaled prototype functions or wavelets $\psi_{ab}(t)$, where 'a' determines the amount of time-scaling or dilation and the variable 'b' represents time shift or translation.

3.1.2 Discrete Wavelet Transform (DWT)

To avoid generating redundant information, the base functions are generated discretely by selecting

$$a=a_0^m \qquad b=nb_0a_0^m$$

The discrete wavelet transform (DWT) is defined as

$$DWT\left(m,n\right)=2^{-m/2}\sum_m\sum_n X(n)\psi^{*}\left(\frac{t-n2^m}{2^m}\right) \qquad (3)$$

where, the discretized mother wavelet becomes

$$\psi_{m,n}^{*}(t)=\frac{1}{\sqrt{a_0^m}}\psi\left(\frac{t-nb_0a_0^m}{a_0^m}\right) \qquad (4)$$

a_0, b_0 are fixed constants with $a_0 > 1$, $b_0 > 1$. m, n ∑ Z; where Z is the set of integers.

The DWT is easier to implement than CWT. CWT is computed by changing the scale of the analysis window, shifting the window in time, multiplying the signal and integrating over all times. In discrete case, filters of different cut-off frequencies are used to analyze the signal at different scales. The signal is passed through a series of high-pass filter (HPF) to analyze the high frequency components and it is passed through a series of low-pass filter (LPF) to analyze the low frequency components. The wavelet decomposition results of a signal are called DWT coefficients.

3.1.3 Multi Resolution Analysis (MRA)

MRA allows the decomposition of signal into various resolution levels. The level with coarse resolution contains approximate information about low frequency components and retains the main features of the original signal. The level with finer resolution retains detailed information about the high frequency components. This is an elegant technique in which a signal is decomposed into scales with different time and frequency resolutions, and can be efficiently implemented by using only two filters: one HPF and one LPF. The results are then down sampled by a factor of two and thus same two filters are applied to the output of LPF from the previous stage. The HPF is derived from the wavelet function (mother wavelet) and measures the details in a certain input. The LPF, on other hand, delivers a smooth version of input signal and this filter is derived from a scaling function, associated to the mother wavelet.

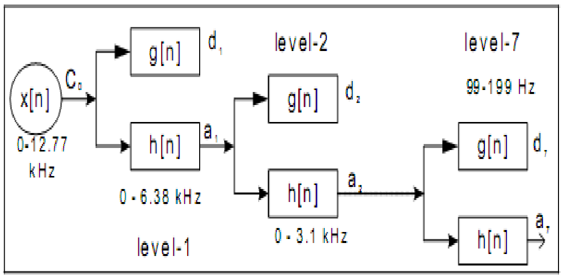

Figure 3. Wavelet decomposition using MRA

For a recorded digitized time signal $C_0(n)$ which is a sampled copy of x(t) as shown in Fig. 3, the smoothened version (called the Approximation) $a_1(n)$ and the detailed version $d_1(n)$ after a first-scale decomposition are given by

$$a_1(n) = \sum_k h(k-2n)C_0(k) \qquad (5)$$

$$d_1(n) = \sum_k g(k-2n)C_0(k) \qquad (6)$$

where h(n) has a low-pass filter response and g(n) has a high pass filter response. The coefficients of the filters h(n) and g(n) are associated with the selected mother wavelet and a unique filter is defined for each. The next higher scale decomposition is based on a_1 (n) instead of C_0 (n). At each scale, the number of the DWT coefficients of the resulting signals (e.g., a_1 (n) and d_1 (n)) is half of the decomposed signal (e.g., C_0 (n)).

4. Transient current testing analysis using Wavelets

The technique behind the fault detection using wavelets is based on comparing current signature of the SRAM cell with the signature of the golden (fault-free) SRAM cell. Input test stimulus is chosen randomly in our detection process, so as to excite the fault. After applying the input stimulus, the wavelet coefficients of the transient current are computed and are compared with those for the golden SRAM cell with the same input. The comparison in this method is made by calculating the Mean Square Error (MSE) between the two sets of wavelet coefficients. Mean Square Error is chosen for comparison because it is simple metric that can effectively detect faults. The pass/fail criterion can be decided by comparing the value of the MSE with a pre-selected test margin. Since the transient current signature is based on wavelet coefficients, both time and frequency components are taken into account simultaneously. This gives a better sensitivity in fault detection than methods based on only spectral or only time-domain components.

4.1.1 Mother wavelet selection

Choice of mother wavelet is an important issue. One of the advantages of wavelet transform is that it is adaptive i.e. we can select a mother wavelet which can best approximate the input waveform. Upon experimenting with number of mother wavelets, db8 and coif4 were found to have high correlation with the input signals.

4.1.2 Effects of sampling frequency

The sampling rate at which the IDDT waveform should be monitored is important because it affects the measurement noise and applicability of the method in real time. Ideally the IDDT waveform should be sampled at above the Nyquist rate (twice the maximum frequency) to keep all the frequency components in the sampled data. However for detecting a fault, it has been observed that a high sampling frequency is not needed. In the proposed method, the current signatures were sampled at ten times the input pulse frequency.

5. Simulation Results

5.1.1 Fault detection

The output responses of the memory cell array in terms of current signatures were sampled for both faulty and fault-free conditions. Matlab wavelet tool box was used to perform wavelet decomposition on the current signatures. Mother wavelet used in wavelet transform is db8. The wavelet coefficients obtained from the faulty cell are then compared with the fault free cell by calculating the mean square error (MSE) given as

$$MSE = \frac{\sum_{i=1}^{N}\left(X_i - Y_i\right)^2}{N} \qquad (7)$$

where X is the wavelet coefficient for fault-free output, Y is the wavelet coefficient for faulty output and N is length (x) = length(y).

Table 4. Fault detection results for db8 and Coif4 wavelets

Injected faults	Mean Square Error (MSE) for db8	Mean Square Error (MSE) for Coif4
TF	1.675	0.260
DRF	0.931	0.169
RDF	0.862	0.138
SOpF	1.594	0.247
SA0	1.925	0.311
SA1	1.967	0.315

The impact of process variation has to be taken into account in determining test margin for fault detection. If the measured MSE value lies outside this test margin, then the particular fault is found to be detected. Table 4 describes the fault detection sensitivity of db8 mother wavelet and coif4 mother wavelet. The comparison of the fault detection sensitivity of both mother wavelets shows that the performance of db8 is better than coif4. Hence in this work db8 is used as the mother wavelet.

Fig. 4 to Fig. 9 shows the db8, scale 6, wavelet coefficients of the current signatures of SRAM cell under fault-free and faulty conditions. Simulation results clearly indicate an increase in the peak value of the current signatures under faulty conditions as compared to the peak value of the current signatures of the fault free conditions. This enables fault detection and maximizes the fault coverage without increase in test time.

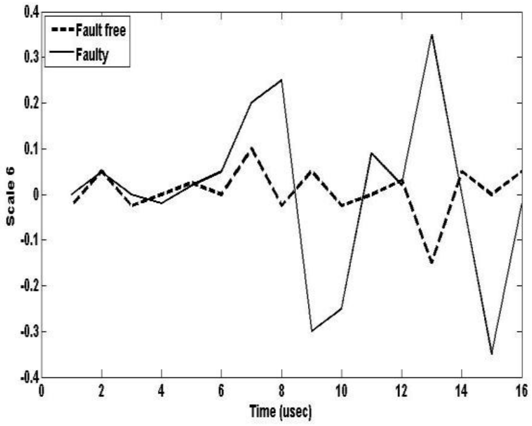

Figure 4. Wavelet coefficient of current signatures for TF under fault free and faulty condition

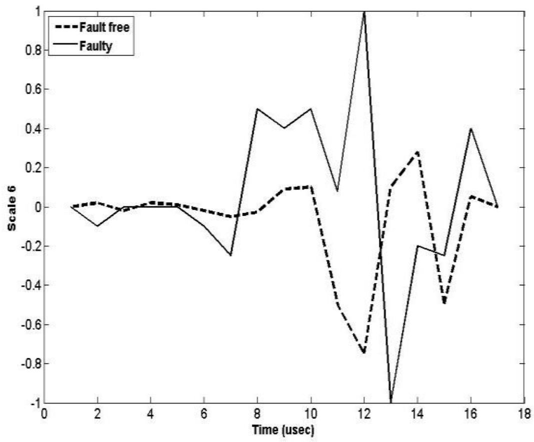

Figure 5. Wavelet coefficient of current signatures for DRF under fault free and faulty condition

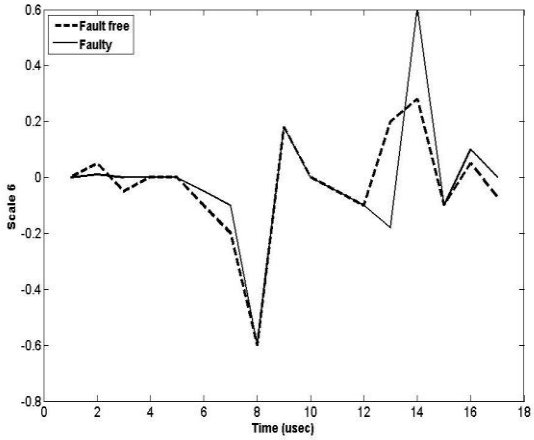

Figure 6. Wavelet coefficient of current signatures for RDF under fault free and faulty condition

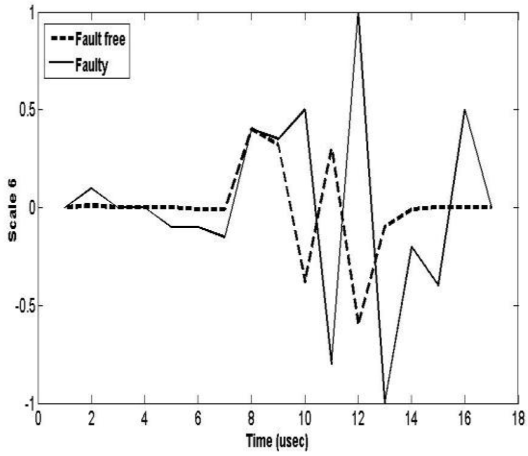

Figure 7. Wavelet coefficient of current signatures for SOpF under fault free and faulty condition

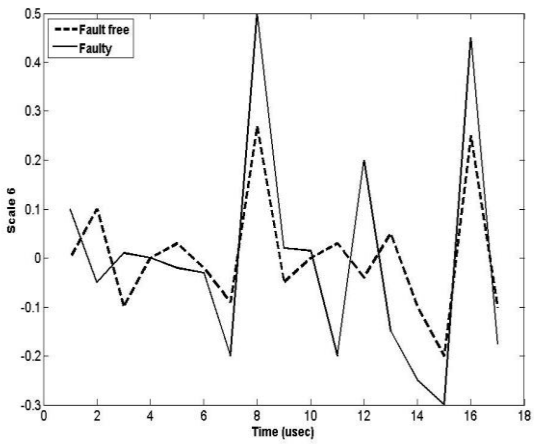

Figure 8. Wavelet coefficient of current signatures for s-a-0 under fault free and faulty condition

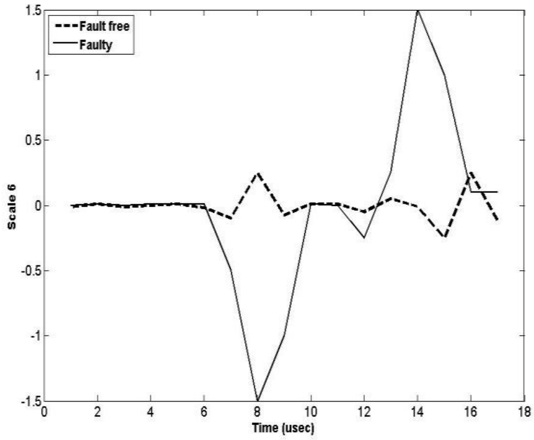

Figure 9. Wavelet coefficient of current signatures for s-a-1 under fault free and faulty condition

5.1.2 Time complexity analysis

An efficient and economical memory test should provide best fault coverage in minimum test time. The proposed technique has the advantage that it has minimum time complexity as compared to the conventional March tests.

Table 5 shows the comparison of number of test vectors required for testing two different faults using conventional March test and transient current testing methods. Conventional March test requires a time complexity of 4n, where n is the size of the array, while the transient current testing technique requires a time complexity of 2n. Thus the number of test vectors to detect faults using transient current testing method is reduced by twice as compared to conventional March test, resulting in 50% reduction of test time. Moreover the sequence of operations is fault specific in March test, but in the proposed technique the sequence of operations remain the same to detect any types of faults in SRAM cells

Table 5. Comparison of March test and Transient current testing

Faults	Test vectors for March test	Test vectors for transient current detection
TF	write ‘0’ read ‘0’ write ‘1’ read ‘1’	Write ‘0’ write ‘1’
s-a-1	write ‘1’ read ‘1’ write ‘0’ read ‘0’	Write ‘0’ write ‘1’

Thus the proposed method of measuring the peak value of the transient current pulse and analyzing using wavelet transform identifies any type of fault present in the SRAM cell providing better fault coverage as compared to conventional March test.

6. Conclusion

In this paper, wavelet transform based transient current testing for fault detection in CMOS SRAM cell has been presented. From the simulation results, the peak value of the transient current pulse of the faulty cell is found to be different from that of the fault free cell. But if the difference is less, the fault may be undetected. Moreover, transient current pulses are in the order of few µA existing only for few µs. Therefore, wavelet transform is used in this paper to widen the time pulse for detection and comparison of faults without losing the information about the timing instants. Wavelet transform has the ability to resolve signals in both time and frequency domain. This makes wavelet a more suitable candidate for fault detection in CMOS SRAM cell, than pure time domain and pure frequency domain methods. The time complexity analysis of transient current testing indicates a test time saving of about 50% as compared to conventional March test. Moreover, in conventional March test method, the test vector applied for fault detection is not capable of detecting all types of faults, since it is fault dependent. But in the proposed method the test vector for fault detection is independent of the nature of the fault present in SRAM cell. Hence, it provides better fault coverage than conventional March test method.

References

[1] R.D.Adams, *High Performance Memory Testing*, Springer-Verlag, 2003.

[2] Andrei Pavlov and Manoj Sachdev, *CMOS SRAM Circuit Design and Parametric Test in Nano-Scaled Technologies*, Springer, 2008.

[3] Qikai Chen, Hamid Mahmoodi, Swarup Bhunia and Kaushik Roy, "*Efficient Testing of SRAM with optimized March sequences and a novel DFT technique for emerging failures due to process variations*", IEEE Trans. on VLSI systems, vol. 13, No. 11, pp. 1286-1295, Nov. 2005.

[4] Balwinder Singh, Sukhleen Bindra Narang and Arun Khosla, *Modeling and Simulation of Efficient March Algorithm for Memory Testing*, Springer, 2010.

[5] V.H.Champac, Jose Castillejos and Joan Figueras, "*IDDQ Testing of Opens in CMOS SRAMs*", Journal of Electronic Testing: Theory and Applications, vol. 15, pp. 53-62, Aug.1999.

[6] De Paul, R. Picos, J. L. Rossello, M. Roca, E. Isern, J. Segura and C. Hawkins, "*Transient Current testing based on Current (Charge) Integration*", in Proc. IEEE Int. Workshop on IDDQ Testing, California, pp. 26–30, 1998.

[7] Manoj Sachdev, Peter Janssen and Victor Zieren, "*Defect Detection with Transient Current Testing and its potential for Deep Sub- micron CMOS ICs*", in Proc. Int. Test Conf., Washington, pp. 204–209,1998.

[8] Ali Chehab, Ayman Kayssi, Anis Nazer and Najwa Aaraj, "*Transient Current Testing of Dynamic CMOS Circuits in the presence of Leakage and Process Variation*", in Proc. 16th Int. Conf. on Microelectronics, Tunisia, pp. 381–387, 2004.

[9] B.Alorda, M.Rosales, J.Soden, C.Hawkins and J. Segura, "*Charge Based Transient Current Testing (CBT) for Submicron CMOS SRAMs*", in Proc. Int. Test Conf., Washington, pp. 947–953, 2002.

[10] B.Alorda, V.Canals, and J.Segura, "*A Two-Level Power-Grid Model for Transient Current Testing Evaluation*", Journal of Electronic Testing: Theory and Applications, vol. 20, No. 3, pp. 543-552, Oct. 2004.

[11] Doe-Hyun Yoon, Hong-Sik Kim and Sungho Kang, "*Dynamic Power Supply Current Testing for Open Defects in CMOS SRAMS* ", ETRI Journal, vol. 23, No. 2, pp. 77-84, Jun. 2001.

[12] J. Liu and R. Z. Makki, "*SRAM Test Using On-Chip Dynamic Power Supply Current Sensor*", in Proc. IEEE Int. Workshop on Memory Technology, Design and Testing, California, pp. 57–62, 1998.

[13] Nilesh P Chitre, *Fault modeling for transient power supply current (IDDT) testing at very deep sub-micron (0.25[mu]m) level.* Charlotte, University of North Carolina, 2000.

[14] S.Su and R. Z. Makki, "*Dynamic Power Supply Current Testing of CMOS SRAMs*", Journal of Electronic Testing: Theory and applications, vol. 16, No.5, pp. 499-511, Oct. 2000

[15] Swarup Bhunia and Kaushik Roy, "*Dynamic Supply Current Testing of Analog Circuits Using Wavelet Transform*", in Proc. 20th IEEE VLSI Test Symposium, California, pp. 302-307, 2002.

[16] Simone Borri and Megali Hage-Hassan, "*Analysis of Dynamic Faults in Embedded-SRAMs: Implications for Memory Test*", Journal of Electronic Testing: Theory and Applications, vol. 21, No.2, pp. 169-179, Apr. 2005.

International Journal of Latest Trends in

Software Engineering

Call for Paper

The journal welcomes high quality original research papers, survey paper, review paper, tutorial, technical notes as well as the discussion papers.

- Software Process / Process Model
- Software Engineering Practice
- Web Engineering
- Quality Management
- Managing Software Projects
- Advanced Topics in Software Engineering
- Multimedia and Visual Software Engineering
- Software Maintenance and Testing
- Languages and Formal Methods
- Web-based Education Systems and Learning Applications
- Software Engineering Decision Making
- Knowledge-based Systems and Formal Methods
- Search Engines and Information Retrieval
- Aspect-oriented programming
- Functional decomposition

- Aspect-oriented Programming
- Functional Decomposition
- Object-Oriented Programming
- Post-object Programming
- Structured Programming
- Rule-Based Programming
- Project Management
- Patterns, Anti-Patterns
- Agile Software Development
- Extreme Programming
- Pair Programming
- Refactoring
- Quality Assurance
- Aspect-oriented programming
- Functional decomposition