

An Optimal Rejuvenation Strategy for Increasing Service Reliability of a VOIP System with Multiple Components

Vandana Gupta[#], Rahul Kumar^{*}

Department of Operational Research
University of Delhi, Delhi, India

[#]khaitan.vandana@gmail.com; ^{*}rahulkumar886049@gmail.com

[#]Corresponding author

(Received October 10, 2016; Accepted December 30, 2016)

Abstract

In this day and age Voice over Internet Protocol (VoIP) has turned out to be a revolutionary technology in the field of telecommunication. Because of the escalating number of users of VoIP, it is essential to intend for a reliable VoIP system providing good quality of service. In this paper a stochastic model based on continuous time Markov chain is developed to analyze the reliability of a server based VoIP system comprising of significant components only. The concurrent breakdown of more than one component is taken into modeling concern. Redundancy at the component level is used with the purpose to boost the system reliability of VoIP. As a defensive course of action, software rejuvenation is implemented to avert or suspend software failures. And an optimal software rejuvenation strategy is proposed which leads to increased system reliability. Numerical results are presented for the quantitative examination of the suggested model.

Keywords- VoIP; System reliability, Continuous time Markov chain, Software rejuvenation, Optimal rejuvenation strategy.

1. Introduction

Voice over Internet Protocol (VoIP), also identified as Internet telephony, is the technology that allows people to exploit the Internet as the communication medium for voice communication. In this technology voice calls are made using a broad-band Internet connection in place of a regular (or analog) phone line. The major benefit of this technology over the conventional circuit switched communication is the lesser call cost mainly for overseas calls. With a major reduction in call cost, VoIP has become a well-accepted way of making calls amid the Internet users. Moreover, the recent advancements on hardware and software aspire to cut down the cost of their phone calls. It is very much evident that this technology has been budding relatively fast in the field of telecommunication in last few years (Karapantazis and Pavlidou, 2009; Koutras and Platis, 2007).

Nevertheless, as several other latest technologies, VoIP too experiences a number of shortcomings that telecommunication companies desire to get rid of with the intention to provide service comparable to the service provided by circuit switched communication companies. Consequently, loads of research has been done by the educational society and the telecommunication companies to overcome the shortcomings of this technology. And it is therefore indispensable for well-made VoIP networks to be dependable and to make available its services in a well-timed way.

A fair amount of research has been carried out on dependability analysis of VoIP systems, mainly on determining its availability and reliability. Due to growing number of calls in a VoIP system, resource exhaustion may occur which can result in system failure. In Koutras and Platis (2007),

the consequences of implementing software rejuvenation with the purpose of averting such failures in a VoIP system are checked. In Koutras et al. (2009), the authors modeled a VoIP system with essential components by means of a Markov chain, and utilized redundancies at component level and software rejuvenation to attain improved reliability and availability. A major shortcoming of the model suggested in this paper is that at a given time the breakdown of only one component is considered. This inspired us to include the simultaneous breakdown of two or more components in our proposed model. In addition to component redundancies, software rejuvenation is a technique used to rise above the problems of resource degradation and security breaches in a VoIP system (Trivedi et al., 2000; Koutras and Platis, 2006; Gupta and Dharmaraja, 2011). Software rejuvenation technique can be looked upon as a defensive continuance strategy to avert or suspend software breakdowns. This approach can be adopted from time to time to counter attack the observable fact of software aging (Koutras and Platis, 2006). The authors in, Koutras and Platis (2006), suggested the best possible rejuvenation strategy to improve the reliability of VoIP system in the occurrence of resource exhaustion. Following the same approach, we propose in this paper the most favorable rejuvenation strategy to improve the reliability of a VoIP system in the occurrence of simultaneous breakdown of two or more components.

In this paper, a Markov model that depicts a VoIP system with redundant components and software rejuvenation is presented. The Mean Time to Failure (MTTF) is obtained as a measure of system reliability. The purpose of this study is to derive an optimal rejuvenation strategy that will lead to an increased MTTF and hence would provide more reliable VoIP service. In order to obtain the optimal rejuvenation strategy, system MTTF is maximized subject to a constraint about the time that is spent at the rejuvenation mode. This paper is an extended version of our work published in Gupta and Dharmaraja (2013). In Gupta and Dharmaraja (2013), we incorporated component redundancies to increase system reliability, and also dealt with the simultaneous failure of more than one component. In this paper, we extend our previous work by proposing an optimal software rejuvenation strategy which further contributes to improved system reliability.

The remaining paper is structured as follows: In Section 2 the proposed Markov model is described and a reliability study in terms of MTTF is presented. In Section 3, an optimal rejuvenation strategy is proposed in order to increase systems MTTF. In Section 4, a numerical illustration of our study is presented. In the numerical example, an optimization problem is constructed and solved with the aim of deriving the most favorable rejuvenation strategy that improves systems reliability. Furthermore, the contribution of system's parameters to MTTF is examined. Finally, in Section 4 the conclusions of this study are presented.

2. Model Description

2.1 System Description

It is relatively complex to mold each and every feature of a VoIP system because of the intricacies involved. In this segment, a server based VoIP system which comprises of only the basic and important components that are necessary to make a VoIP call is taken into consideration. The most significant component of such a system which is accountable for the entire VoIP service is the server. A server is basically a type of resource which serves the VoIP calls. As large number of calls arrives, the server witnesses resource exhaustion and may ultimately stop working over a period of time. One more significant component of a VoIP system is the router. The key role of the VoIP router, as the name implies, is to route the different packets to their destinations on Internet. A router can stop working if there is overflow of packets. The

failure of the router can be prevented by a restart. Equivalently important is the function of Internet Service Provider (ISP). Majority of the VoIP companies do not possess their individual Internet networks, and thus they bank totally on the ISP. And last but not the least, one of the critical components of a VoIP system is the electrical power supply. If the power supply is broken up, it can have an adverse impact on VoIP's quality of service because of loss of calls. To counterbalance power supply discontinuance an Uninterrupted Power Supply (UPS) device is utilized.

A straightforward approach to attain improved reliability for any system is to make use of redundant components. In our model, we as well implement redundancy of components to get improved system reliability. At the outset, we take into consideration that the system comprises of two servers, one active and one backup. When the active server breaks down, the system control is switched mechanically to the backup server and the failed one is sent for repair.

In addition to redundancy of server, technique of software rejuvenation is applied to counter balance the consequences of resource exhaustion on quality of service of VoIP. Software rejuvenation is a defensive maintenance technique that proactively reboots a system or an application so as to avoid an unprepared failure owing to software aging (Koutras and Platis, 2006; Trivedi et al., 2000). This technique of defensive maintenance is employed only to the software component of the VoIP server. When the active server goes for rejuvenation, the control of the system is mechanically switched to the backup server. Besides we also consider that the system comprises of two routers in the similar condition as the servers, one active and one backup. In the situation of packet overcrowding on the active router, the routing of calls is instantly switched to the backup router. However, in contrast to the server and the router, there is a single ISP in the system with no backup. Lastly, redundancy for the power supply is provided through a UPS device.

Next, we introduce an analytical model for obtaining the service reliability of VoIP system with basic components only. In practical applications, nearly all of the failure and repair times follow distributions such as Weibull, Pareto and lognormal which have time-dependent failure rates. But analytical models with non-exponential distributions are generally not mathematically solvable. In literature (Osogami and Harchol-Balter, 2006) phase-type distribution is used for approximating many non-exponential distributions. Phase type distribution is basically convolution of many exponential phases. In view of the fact that exponential distribution is a particular case of phase-type distribution, in this paper we assume that the time to transit from one system state to another follows an exponential distribution, and hence develop the mathematically solvable Continuous Time Markov Chain (CTMC) model for the reliability analysis of the VoIP system.

2.2 Reliability Model

The state transition diagram of the underlying CTMC is depicted in Fig. 1. To begin with, the system is assumed to be in state Up . State Up is a completely functional state where each and every component of the system works as it should be. If the primary server breaks down at this state, then the system control is mechanically switched to the backup server and the system makes a transition from state Up to state S . Hence, in state S the system is controlled by the backup server and the failed primary server is sent for repair. The failure rate and the repair rate of the server are taken as λ_s and μ_s , respectively. While the system is in state Up and one of the router breaks down, then the system makes a transition from state Up to state R . In state R the routing of

calls is instantly switched to the backup router and the failed router is sent for repair. The failure rate and the repair rate of the router are taken as λ_R and μ_R , respectively. While the system is in state Up and the electric power supply is broken up, then the system makes a transition from state Up to state P . In state P the system utilizes power supply through the UPS and the electric power supply is sent for repair. The failure rate and the repair rate of the electric power supply are taken as λ_P and μ_P , respectively. In addition to above mentioned scenarios, while the system is in state Up , then software rejuvenation can be triggered from time to time. Once rejuvenation is triggered, the system makes a transition from state Up to state Rj . At the onset of rejuvenation, the system control is mechanically switched to the backup server. Hence, in state Rj the system is managed by the backup server. The rate of rejuvenation and the rate of recovery after rejuvenation are taken as λ_{Rj} and μ_{Rj} , respectively. Furthermore, while the system is in state Up and the ISP gets failed, then the system makes a transition to state $Down$ which is an absolute down state. The failure rate of the ISP is taken as λ_I .

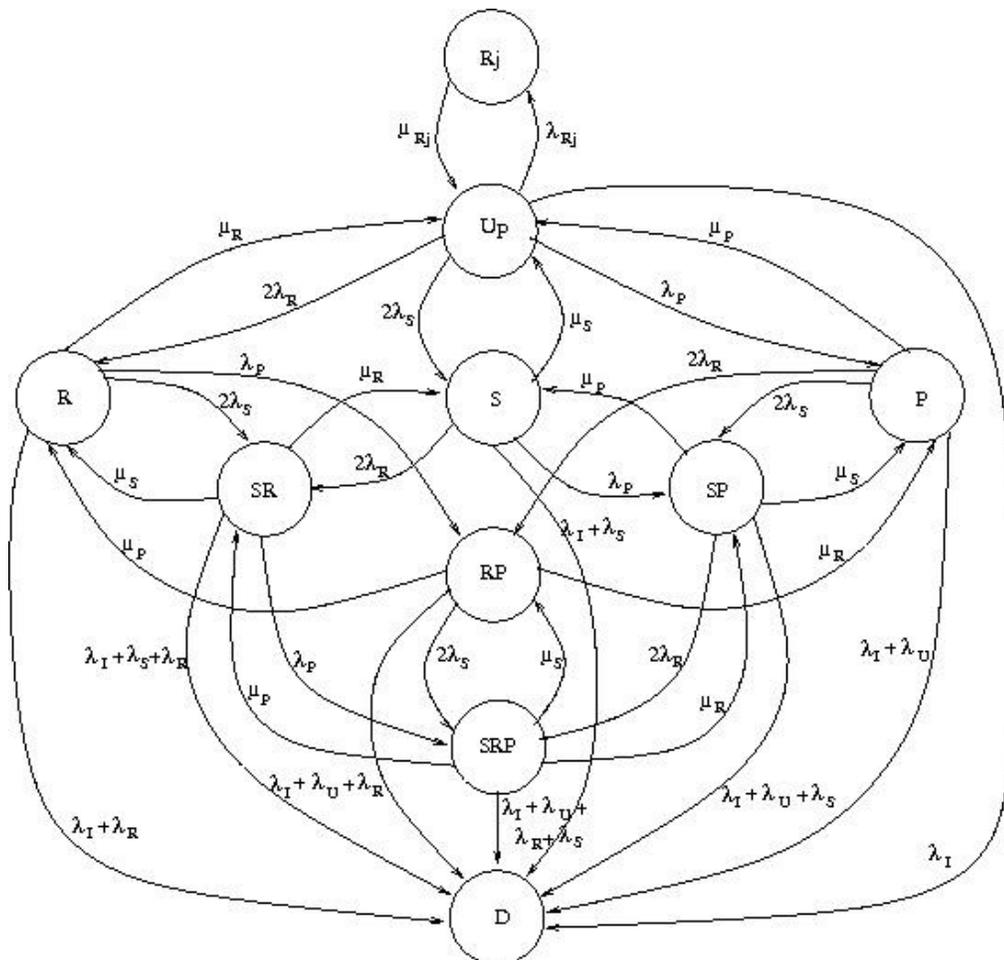


Fig. 1. Reliability model for a VoIP system with software rejuvenation

Next, we consider the simultaneous breakdown of more than one component in the system. While the system is in state S (where the primary server is undergoing repair and the system is managed

by the backup server), if one of the routers breaks down prior to the repair completion of the primary server, then the system makes a transition to state SR . In state SR the system is managed by the backup server and the routing of calls is done by the backup router. Furthermore, if the electric power supply is broken up prior to the repair completion of the primary server, then the system makes a transition to state SP . Hence, in state SP the system is managed by the backup server and utilizes the power supply through UPS. On the other hand, while the system is in state S and the backup server breaks down prior to the repair completion of the primary server, or the ISP gets failed, in these two conditions the system makes a transition to state $Down$.

Next, we suppose that the system is in state R where one of the routers is in failed condition and is undergoing repair. While the system is in this state and the primary server breaks down prior to the repair completion of the failed router, then the control of the system is mechanically switched to the backup server. The system then makes a transition from state R to state SR . Likewise, if the electric power supply gets broken up prior to the repair completion of the failed router, then the system makes a transition from state R to RP . Therefore, in state RP calls are routed by the backup router and the system utilizes power supply through the UPS. But, while the system is in state R and the backup router breaks down prior to the repair completion of the failed one, or else the ISP gets failed, in these two situations the system makes a transition to state $Down$.

Next, we consider the situation when the system is in state P where the electric power supply is broken up and is undergoing repair. While the system is in this state and the primary server breaks down prior to the repair completion of the failed electric power supply, then the control of the system is mechanically switched to the backup server, and the system makes a transition from state P to state SP . In the same way, if one of the routers breaks down prior to the repair completion of the failed electric power supply, then the system makes a transition from state R to state RP . On the other hand, while the system is in state P and the UPS breaks down prior to the repair completion of the electric power supply, or the ISP gets failed, in these two situations the system makes a transition to state $Down$. The failure rate of UPS is taken as λ_U .

Now suppose that the system is in state SR where the primary server and one of the routers are in failed condition and both are undergoing repair. At this stage, if the electric power supply gets broken up prior to the repair completion of any of the server or the router, then the system makes a transition from state SR to state SRP . In this state SRP , the system is managed by the backup server, routing of calls is done by the backup router, and the system functions on the power supply through the UPS. Nevertheless, while the system is in state SR and if any one of the backup server or router (backup) breaks down prior to the repair completion of any one of the respective failed ones, or the ISP gets failed, in all these circumstances the system makes a transition from state SR to state $Down$.

Next, we consider the situation when the system is in state SP where the primary server as well as the electric power supply are in failed condition and are undergoing repair. If at this stage one of the routers breaks down prior to the repair completion of any of the server or the power supply, then the system makes a transition from state SP to state SRP . On the other hand, while the system is in state SP and if any one of the backup server or the UPS breaks down prior to the repair completion of the primary server or the electric power supply, or the ISP gets failed, in all these conditions the system makes a transition from state SP to state $Down$.

Next, suppose that the system is in state RP where one of the routers is in failed condition and the electric power supply too is broken up, and both are undergoing repair. At this stage, if the primary server breaks down and the control of the system is mechanically switched to the backup server, then the system makes a transition from state RP to state SRP . It is to be noted that while the system is in state RP and any of the backup router or the UPS breaks down prior to the repair completion of any of the failed router or the electric power supply, or the ISP gets failed, in all these state of affairs the system makes a transition from state RP to $Down$.

Lastly, consider the situation when the system is in state SRP where the primary server, one of the two routers and the electric power supply all are in failed condition, and are undergoing repair. If at this stage, any one of the backup components (i.e., backup server, router and UPS) breaks down prior to the repair completion of any one of the respective primary components or the ISP gets failed, then the system makes a transition to state $Down$.

2.3 Reliability Study

After describing the reliability model of the VoIP system with software rejuvenation action, our objective now is to obtain service reliability of VoIP with regard to its MTTF. With this purpose, we identify a random process $\{X(t), t > 0\}$, which corresponds to the evolution of the system with time. The state space of $\{X(t), t > 0\}$ is defined as $\{Up, S, R, P, SR, RP, SP, SRP, Rj, Down\}$. As we have considered all failure times and repair times to be exponentially distributed in the reliability model, $X(t)$ is a Continuous Time Markov Chain (CTMC). Further, to find out the MTTF, we need the time that the process invests at each and every state. From Fig. 1, we can determine the rate matrix Q for the CTMC $X(t)$. In the rate matrix Q the rows and the columns represent the states $\{Up, S, R, P, SR, RP, SP, SRP, Rj, Down\}$ correspondingly, and the elements of the rate matrix Q represent the transition rates between the above states. Hence, elements $q_{(ij)}$, $i, j \in \{Up, S, R, P, SR, RP, SP, SRP, Rj, Down\}$ of the transition matrix denote the rate of transition from i^{th} state to j^{th} state.

Let the state space of the process be defined by E , which is decomposed into subsets U and D such that $U \cup D = E$ and $U \neq \emptyset$; $U \neq E$. Subset U contains the Up states and correspondingly D contains the Down states. In the model presented above subset D contains only the absorbing state D . The transition rate matrix is given now by

$$\begin{bmatrix} Q^U & Q^{UD} \\ Q^{DU} & Q^D \end{bmatrix}$$

Consequently, Q^U is the sub matrix corresponding to the working states. Q^D is the transition rate matrix inside the non-working subset D and Q^{UD} , Q^{DU} are the transition rate matrices from subset U to subset D , and from subset D to subset U correspondingly (Platis et al., 1998).

Let $\pi^U(0) = Pr(X(0) = i, i \in \{Up, S, R, P, SR, RP, SP, SRP, Rj, Down\})$, subvector of the initial distribution vector $\pi(0)$, be the initial distribution restricted in the working states (upstates). Further, to obtain the time τ_i^U invested at each state i , the subsequent system of equations has to be solved:

$$\tau^U Q^U = -\pi^U(0) \tag{1}$$

where $\tau^U = [\tau_{Up}^U \ \tau_S^U \ \tau_R^U \ \tau_P^U \ \tau_{SR}^U \ \tau_{RP}^U \ \tau_{SR}^U \ \tau_{SRP}^U \ \tau_{Rj}^U \ \tau_{Down}^U]$ is the vector giving the time invested at each state (Trivedi, 2001). In addition to this, it is presumed that the initial distribution subvector is:

$$\pi^U(0) = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0] \quad (2)$$

The MTTF can now be computed (Trivedi, 2001) using the following equation:

$$MTTF = \tau_{Up}^U + \tau_S^U + \tau_R^U + \tau_P^U + \tau_{SR}^U + \tau_{RP}^U + \tau_{SR}^U + \tau_{SRP}^U + \tau_{Rj}^U + \tau_{Down}^U \quad (3)$$

which represents the total time that the process invests at each and every state before going into state *Down* which is the absorbing failure state.

In fact, the problem to be solved consists of determining the optimal rejuvenation strategy which maximizes the MTTF subject to a constraint about the time invested at the rejuvenation state. Deriving the optimal rejuvenation strategy consists of determining the value of the rejuvenation rate that corresponds to the maximum MTTF value with respect to the constraint mentioned above.

As far as the constraint is concerned, the time that the process invests at the rejuvenation state cannot exceed the time invested at the other operational and check states. Hence, the constraint that is derived by this assumption is given by:

$$\tau_{Rj}^U \leq \tau_{Up}^U + \tau_S^U + \tau_R^U + \tau_P^U + \tau_{SR}^U + \tau_{RP}^U + \tau_{SR}^U + \tau_{SRP}^U + \tau_{Down}^U \quad (4)$$

Summarizing the previous assumptions, the problem that finally has to be solved in order to achieve the maximum value of MTTF and hence increased VoIP service reliability is:

$$\begin{aligned} \text{Max } MTTF &= \tau_{Up}^U + \tau_S^U + \tau_R^U + \tau_P^U + \tau_{SR}^U + \tau_{RP}^U + \tau_{SR}^U + \tau_{SRP}^U + \tau_{Rj}^U + \tau_{Down}^U \\ \text{subject to } \tau_{Rj}^U &\leq \tau_{Up}^U + \tau_S^U + \tau_R^U + \tau_P^U + \tau_{SR}^U + \tau_{RP}^U + \tau_{SR}^U + \tau_{SRP}^U + \tau_{Down}^U \end{aligned} \quad (5)$$

The objective function in the above optimization problem, MTTF, is a function of the rate of recovery μ_{Rj} after rejuvenation. For this reason, the optimization is performed with respect to μ_{Rj} . Furthermore, the recovery rate is involved into the constraint as the time that the process $X(t)$ invests at the rejuvenation state is a function of μ_{Rj} . Determining vector τ^U by solving equation (1), indicates that the time invested at each state except the rejuvenation state Rj is constant. Only the time invested at the rejuvenation state depends on μ_{Rj} . Solving the optimization problem of equation (5) leads to a solution that provides the maximum value of MTTF and the corresponding μ_{Rj} value. Hence, in order to derive the optimal rejuvenation strategy, rejuvenation action has to be performed as often or as rarely as the solution of the problem indicates.

3. Numerical Analysis

In order to illustrate the above study, it is necessary to provide a numerical example based on experimental data. The aim of this example is to present the optimal rejuvenation strategy consisting of the value of the rejuvenation rate μ_{Rj} , which maximizes the MTTF. For the sake of numerical illustration, the parameter values of various components of network elements are given in Table 1 (Koutras et al., 2009).

Parameters	Values in years ⁻¹	Parameters	Values in years ⁻¹
λ_S	4	μ_S	525600
λ_R	0.2	μ_R	17520
λ_P	15	μ_P	26280
λ_{Rj}	6	λ_U	17520
λ_I	0.00066	μ_P	0.01

Table 1. Model parameters

Fig. 2 shows the behavior of system MTTF for different values of server failure rate (λ_S). The graph also depicts the optimal values of the rejuvenation rate μ_{Rj} which maximizes the MTTF. It is also observed that MTTF increases with decrease in the server failure rate (λ_S) as anticipated.

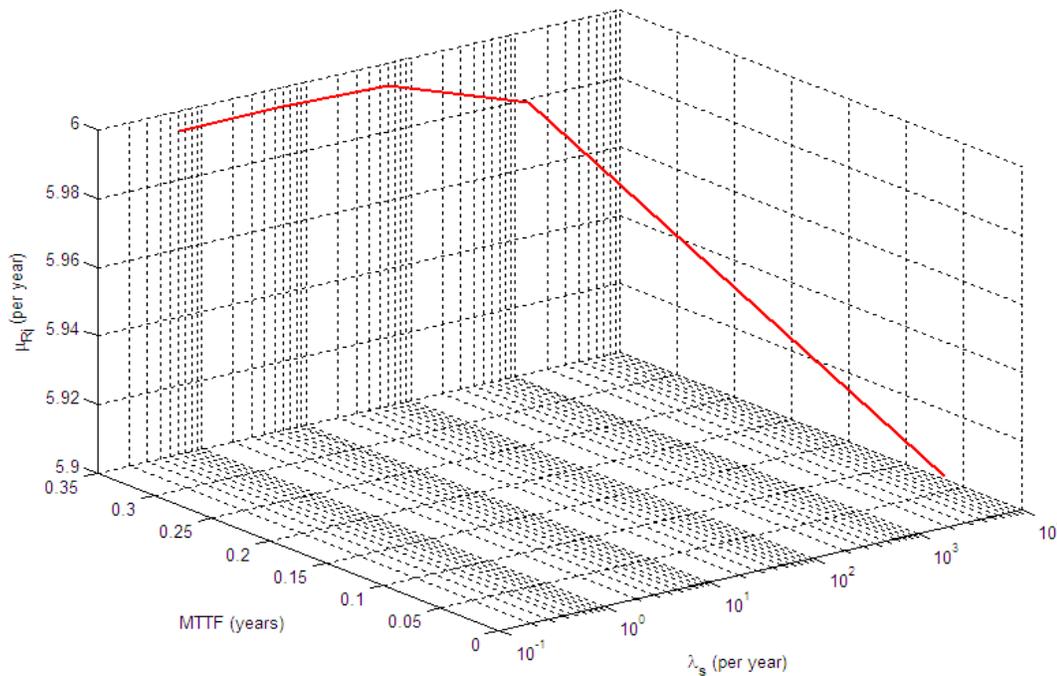


Fig. 2. MTTF vs λ_S

Fig. 3 shows the behavior of system MTTF for different values of router failure rate (λ_R). The graph also illustrates the optimal rejuvenation strategy to maximize the MTTF. Also, it can be examined that MTTF increases with decrease in the router failure rate (λ_R) as predictable.

Fig. 4 shows the behavior of system MTTF for different values of power supply failure rate (λ_P). The graph also represents the best possible values of rejuvenation rate μ_{Rj} to maximize the MTTF. In addition, it is also observed that MTTF increases with decrease in the power supply failure rate (λ_P) as expected.

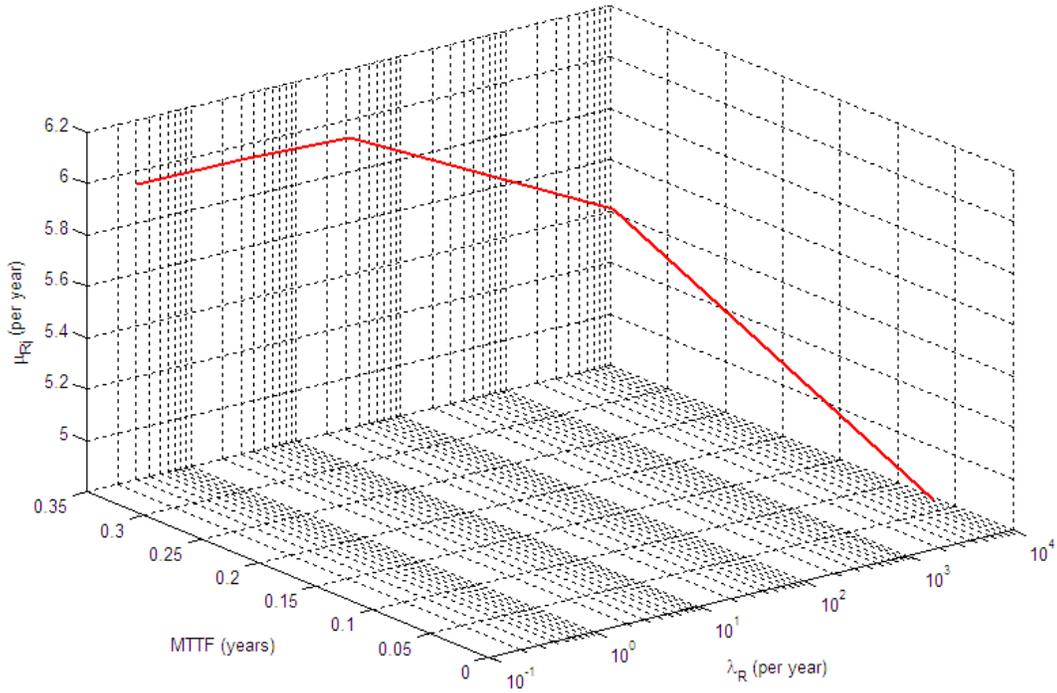


Fig. 3. MTTF vs λ_R

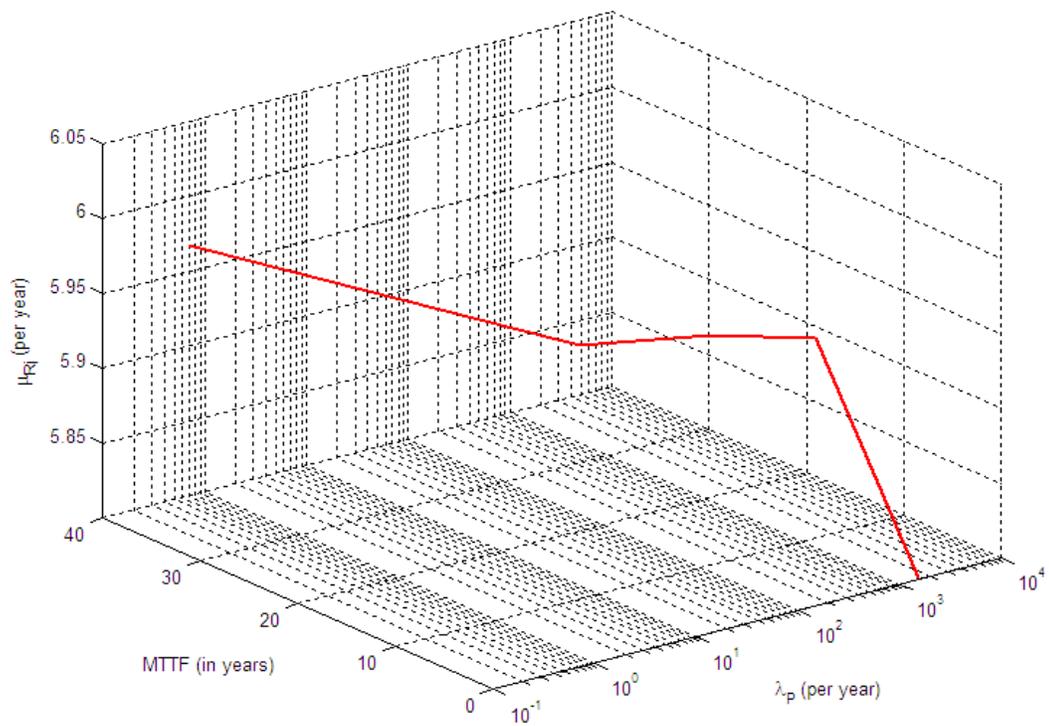


Fig. 4. MTTF vs λ_P

4. Conclusion

In this paper a VoIP system is studied and a new approach for attaining improved VoIP service reliability is suggested. The VoIP system under this study comprises of merely the significant components. Redundancies at component level and the technique of software rejuvenation are implemented to attain improved system reliability. In addition, simultaneous breakdown of more than one component is taken into modeling concern. Furthermore, it is assumed that the time for breakdown and time to repair for all components follow exponential distribution. As a consequence, a reliability model for the VoIP system is developed using a CTMC.

From the CTMC, MTTF is obtained as a reliability measure. An optimization problem maximizing the MTTF is then formulated with respect to a constraint about the time invested at the rejuvenation state of the model. The solution of the optimization problem provides the optimal value of the rejuvenation rate which maximizes the MTTF for the system. An optimal rejuvenation strategy is then proposed concerning on how often rejuvenation should be performed for achieving the maximum MTTF with respect to the other parameters involved. The rejuvenation strategy gives the optimal rejuvenation time interval that maximizes MTTF which consequently improves the quality of service of the VoIP system.

Acknowledgement

This research work is supported by the grant sanctioned under Research and Development Scheme of the University of Delhi, Delhi.

References

- Gupta, V., & Dharmaraja, S. (2011). Semi-Markov modeling of dependability of VoIP network in the presence of resource degradation and security attacks. *Reliability Engineering & System Safety*, 96(12), 1627-1636.
- Gupta, V., & Dharmaraja, S. (2013). Reliability and performance modelling of VoIP system with multiple component failures. *International Journal of Reliability and Safety*, 7(1), 58-74.
- Karapantazis, S., & Pavlidou, F. N. (2009). VoIP: A comprehensive survey on a promising technology. *Computer Networks*, 53(12), 2050-2090.
- Koutras, V. P., & Platis, A. N. (2006). Applying software rejuvenation in a two node cluster system for high availability. In proceedings *International Conference on Dependability of Computer Systems (DEPCOS-RELCOMEX 06)*, 2285-2290.
- Koutras, V. P., & Platis, A. N. (2006). Optimal rejuvenation policy for increasing VoIP service reliability. In proceedings *European Safety and Reliability Conference*. 2285-2290.
- Koutras, V. P., & Platis, A. N. (2007). VoIP availability and service reliability through software rejuvenation policies. In proceedings *International Conference on Dependability of Computer Systems*, 262-269.
- Koutras, V. P., Salagaras, C. P. S., & Platis, A. N. (2009). Software rejuvenation for higher levels of VoIP availability and mean time to failure. In proceedings *International Conference on Dependability of Computer Systems*, 99-106.
- Osoyami, T., & Harchol-Balter, M. (2006). Closed form solutions for mapping general distributions to quasi-minimal PH distributions. *Performance Evaluation*, 63(6), 524-552.
- Platis, A. N., Limnios N., & Le Du, M. (1998). Dependability analysis of systems modeled by non-homogenous Markov chains. *Reliability Engineering and System Safety*, 61, 235-249.

Trivedi, K. S. (2001), *Probability and statistics with reliability, queueing, and computer science applications*, 2nd Ed. New York: Wiley.

Trivedi, K. S., Vaidyanathan, K., & Goseva-Postojanova, K. (2000). Modeling and analysis of software aging and rejuvenation. In proceedings *IEEE Annual Simulation Symposium*, 270-279.