A Review of Vacation Queueing Models in Different Framework

A.P. Panta*, R.P. Ghimire², D. Panthi³ and S. R. Pant⁴

¹Department of Mathematics, Tri-Chandra Campus, Tribhuvan University
Kathmandu, Nepal
E-Mail: pantananda88@gmail.com

²Department of Mathematics, Kathmandu University, Kavre, Nepal
E-Mail: rpghimire@ku.edu.np

³Department of Mathematics, Valmeeki Campus, Nepal Sanskrit University
Kathmandu, Nepal
E-Mail: panthid06@gmail.com

⁴Central Department of Mathematics, Tribhuvan University, Kirtipur, Nepal
E-Mail: shankarpant21@gmail.com

*Corresponding author

Received 12 October 2021; accepted 22 November 2021

Abstract. This paper deals with the extensive survey of the vacation queueing systems from its birth in 1975 to till date. As the time passes on old-provisions cannot tackle new-problem so new concepts have to be developed. Some of the threshold policies that have been developed from time to time have been reported. Some of the prominent techniques of solution of vacation queueing system have also been cited. The challenges on tackling the problems of vacation queueing systems have been discussed.

Keywords: Vacation queue, working vacation, policy, batch arrivals, solution techniques

AMS Mathematics Subject Classification (2010): 60K25

1. Introduction
Vacation in the queueing system takes place in either absence of customer in the system or due to server breakdown. The benefit of the server vacation is that server can utilize its idle time for other purposes. Several authors contributed to the vacation queueing systems due to their numerous applications in the areas of computer telecommunication, manufacturing and other systems that can be represented by stochastic models. Some of the literatures on vacation models are highlighted in our studies. Levy and Yechiali [80] first introduced the concept of vacation in 1975, where the idle time of the server is utilized for additional work in a secondary system. Extensive surveys on the earlier works of vacation queueing
Vacation Markovian Queueing Systems

If the arrival rate follows a Poisson distribution and service rate follows an exponential distribution, then the queueing system is termed as the Markovian queueing system. Several researchers have contributed to the study of Markovian queueing systems with vacations. Levy and Yechiali [81] analyzed a multi-server M/M/s queue with servers’ vacations. Madan [97] studied an M/M/2 queueing system in which each server is equipped with a standby. He obtained the time dependent results giving probability generating functions for the number in the system under various states and derived the corresponding steady state results. Tian et al. [129] investigated the conditional stochastic decompositions of the stationary queue length and waiting time in a multi-server M/M/C queueing models with server vacations. Zhang and Tian [151] considered an M/M/c queue with a single vacation policy for some idle servers. They established conditional stochastic decomposition properties for the waiting time and the queue length when all servers are busy. Zhang and Tian [152] analyzed the multi-server M/M/c vacation queueing system as a quasi-birth and death process. They obtained the stationary distributions of queue length and waiting time, using matrix geometric method. Altman and Yechiali [4] considered customers’ impatience in queueing systems with server vacations where customers’ impatience is due to an absentee of servers upon arrival. They presented a comprehensive analysis of the M/M/1 and M/G/1 queues as well as M/M/c queue for both the single and multiple-vacation cases, and obtained various closed-form results. Ke et al. [61] developed the M/M/R machine repair problem with spares where the servers are characterized by single, multiple, vacations and hybrid single/multiple vacation. Perel and Yechiali [108] introduced and analyzed customers’ impatience that arises as a result of a slowdown in the servers’ service.
A Review of Vacation Queueing Models in Different Framework

rate. In order to analyze the model, they considered an M/M/c queue (c=1, 1< c < ∞, c=∞) operating in a 2-phase random environment. Yue et al. [148] analyzed customer’s impatience in an M/M/1 queueing system under server vacations, where they assumed that the ‘impatience timers’ of customers depend on the server’s states. Ibe and Isijola [49] dealt with an M/M/1 queueing system in which two types of vacations can be taken by the server and they obtained steady-state solution.

Recently, Ma et al. [95] threw light on a multiple vacation queueing system with complementary services. They investigated the equilibriums of the system in a time-based fee model under competition and monopoly cases respectively. Tian [132] studied the socially optimal strategies and optimal pricing strategies in unobservable M/M/1 queue with delayed multiple vacations. Very recently, Panta et al. [107] dealt with the optimization of M/M/s/N queueing model with reneging in a fuzzy environment and they obtained the various performance measures of the system.

2(ii). Working Vacations Markovian Queueing

In the vacation queueing system, the server completely stops service when he is on vacation whereas in the working vacation scheme the server works at a different rate instead of being completely idle during the vacation period. Servi and Finn [112] first introduced the concept of working vacations in 2002. Markovian queueing systems with working vacations have been studied extensively due to their wide applications in computer systems, communication networks, production management and so forth. Servi and Finn [112] first studied M/M/1 queues with working vacations and they generalized classical single server vacation model to consider a server which works at a different rate rather than completely stops during the vacation period. Liu et al. [93] demonstrated stochastic decomposition structures of the queue length and waiting time in an M/M/1 queue with working vacations and obtained the distributions of the additional queue length and additional delay. Li and Tian [86] analyzed the M/M/1 queueing system with working vacations and vacation interruptions. In terms of the quasi birth and death process and matrix-geometric solution method, they obtained the distributions and the stochastic decomposition structures for the number of customers and the waiting time and provided some indices of systems. Xu et al. [142] analyzed the M/M/1 queue with single working vacation and set-up times by using quasi birth and death process and matrix-geometric solution method. They derived the distributions for the stationary queue length and waiting time of a customer in the system, and obtained stochastic decomposition structures of stationary indices. Lin and Ke [91] analyzed multi-server M/M/R queueing system with single working vacation, and developed the explicit formula for the probability distributions of queue length and other system characteristics by using Neuts’ matrix-geometric approach. Jain and Jain [53] studied a single server working vacation queueing model with multiple types of server breakdowns and proposed a matrix geometric approach for computing the stationary queue length distribution. Wu et al. [137] studied a M/M/s queueing system with multiple
working vacations. They applied the direct search method and Quasi-Newton method to get an approximate solution to the constrained optimization problem. Maurya [100] demonstrated a mathematical modeling for analyzing a Markovian queueing system with two heterogeneous servers and working vacation and obtained various performance measures of the system with varying parameters under steady state using matrix geometric method. Yue et al. [147] considered an M/M/1 queueing system with working vacations and impatient customers, where the customers’ impatient is due to slow service rate in a working vacation. Zhang et al. [149] investigated the equilibrium balking behavior of customers in an M/M/1 queue with working vacations. Tian et al. [133] analyzed the customer’s equilibrium and socially optimal joining-balking behavior in M/M/1 queues with multiple working vacations and vacation interruptions.

Recently, Bouchentouf and Yahiaoui [18] analyzed a single server Markovian feedback queueing system with reneging and retention of reneged customers, multiple working vacations and Bernoulli schedule vacation interruption. Jain et al. [51] considered a M/M/1 queue subject to sudden halt and functioning vacation policy by including the realistic assumption of customer’s intolerance. They solved the steady state probabilities for the system states in terms of rate matrix equations by using the matrix geometric approach. Liu and Hlynka [92] considered both working vacations and regular vacations and they compared system with vacations to the regular M/M/1 system via mean service rates and expected number of customers by using matrix-analytic methods. They showed that the system with vacations performs better than the regular M/M/1 system under certain conditions. Laxmi et al. [65] dealt with an infinite capacity single server Markovian queue with a single working vacation and reneging of customers due to working vacation. Kalyanaraman and Sundaramoorthy [55] analyzed M/M/1 queue with multiple working vacation and with partial breakdown using matrix geometric method. Very recently, Jain et al. [54] investigated Markovian queue with working vacation, retrial and impatient customers by including the feature of imperfect service during working vacation.

3(i). Poisson Arrival General Service Queueing System with Vacations

Many researchers have contributed to the study of M/G/1 type vacation models. Levy and Yechiali [80] first studied an M/G/1 queueing model where the idle time of the server is utilized for additional work in a secondary system. They derived Laplace Stieltjes transforms of the occupation period, vacation period and waiting time for the model. Fuhrmann [37] considered an M/G/1 queueing model with server vacations and provided a simple and elegant method of solution. Fuhrmann and Cooper [38] studied the M/G/1 queue with generalized vacations and they demonstrated that the M/G/1 decomposition property holds for a very general class M/G/1 queueing models. Teghem [124] analyzed single server M/G/1 type queues with vacation periods and investigated the transient and steady-state behavior of the waiting time process for the FIFO discipline. Levy and Kleinrock [79] analyzed a delay of a queue with starter and a queue with vacations by decomposition.
A Review of Vacation Queueing Models in Different Framework

Keilson and Servi [63] discussed the dynamic aspects of the M/G/1 vacation queueing system and obtained the relaxation time of the system. Shanthikumar [114] provided an analytic proof for the stochastic decomposition theorem in an M/G/1 queue with a class of more general vacation policies. Lee [68] investigated the M/G/1 queue with exceptional first vacation and derived the server idle probability and the system size distribution of the system. Shanthikumar [115] illustrated the power and simplicity of level crossing analysis and presented a conservation identity for M/G/1 priority queues with server vacations. Doshi [28] investigated conditional and unconditional distributions for M/G/1 type queues with server vacations. Takagi [118] analyzed the time-dependent process in M/G/1 vacation models with exhaustive service and demonstrated the decomposition property in the queue size and the remaining service time. Madan [96] considered a single server M/G/1 queueing system with compulsory server vacations. Takine and Hasegawa [122] dealt with the M/G/1 queueing system with server vacations and showed that the decomposition property is valid for the joint probability distribution of the queue length and the forward recurrence service time. Miyazawa [104] considered single server vacation models with general vacation policies and obtained a unified approach for deriving decomposition formulas for these vacation models by using the rate conservation law. Madan [98] studied the steady state behavior of an M/G/1 queueing system with deterministic server vacations by using the supplementary variable technique. Sikadar and Gupta [116] dealt with an M/G/1 batch service queue with single vacation and obtained the probability generating functions of the queue length distributions at various epochs by using the supplementary variable method. Wu et al. [139] studied a BMAP/G/1 G-queues with second optional service and multiple vacations. Singh et al. [117] considered an M/G/1 queueing model with state dependent arrival rates and vacation.

3(ii). Poisson Arrival General Service Time Queueing System with Working Vacations

Some researchers have studied and proposed mathematical models of the type M/G/1 with working vacation as well. Wu and Takagi [138] dealt with an M/G/1 queue under multiple working vacations and exhaustive service discipline. They derived the distributions for the queue size and the system time for an arbitrary customer in the steady state. Li et al. [89] analyzed the M/G/1 queueing system with exponentially working vacations and obtained the distribution for the stationary queue length at departure epochs, using the matrix analytic approach. Also, they derived a conditional stochastic decomposition result, based on the classical vacation decomposition in the M/G/1 queue. Zhang and Hou [150] dealt with an M/G/1 queue with exponential working vacations and vacation interruption. They obtained the queue length distribution and service status at an arbitrary epoch under the steady state conditions, using supplementary variable and the matrix analytic methods. Gao and Liu [39] treated an M/G/1 queueing system with single working vacation and vacation
A.P. Panta, R.P. Ghimire, D. Panthi and S. R. Pant

interruption under Bernoulli schedule. They developed a variety of stationary performance measures for the system and gave a conditional stochastic decomposition result.

4(i). General Arrival Exponential Service Time Queueing Systems with vacations

Many researchers have contributed to the study of G/M/1 type queueing systems with vacations. Tian et al. [130] analyzed the GI/M/1 queue with exhaustive service and multiple exponential vacations. They expressed the transition matrix of the imbedded Markov chain as a block-Jacobi form and gave a matrix geometric solution. Tian and Zhang [125] considered a GI/M/1 queue with phase-type setup times or vacations and proved stochastic decompositions properties of the queue length and waiting time in this system. Tian and Zhang [126] studied GI/M/c queueing system with phase-type vacations and proved conditional stochastic decomposition properties for the queue length and the waiting time when all servers are busy. Li and Liu [84] analyzed a GI/M/1 queue with vacations and multiple service phases. They obtained the distributions of the stationary system size at both arrival and arbitrary epochs by using the matrix geometric solution method and semi-Markov process.

4(ii). General Arrival Exponential Service Time Queueing Systems with Working Vacations

Several researchers have contributed to the study of G/M/1 type models with working vacations. On the GI/M/1-type systems with the general arrival rate, Baba [9] first presented a GI/M/1 queue with multiple working vacations using the matrix analytic method. Banik et al. [14] analyzed the finite buffer GI/M/1/N queue with exhaustive service discipline and multiple working vacations. Li et al. [88] analyzed the single server GI/M/1 queue with working vacations and vacation interruption. Chen et al. [24] considered a GI/M/1 queueing system with phase-type working vacations and vacation interruption. Chae et al. [21] provided the comparison analysis between the GI/M/1 and the GI/Geo/1 queues with single working vacation. Tian et al. [128] provided a survey on working vacation queues and matrix analytic method. Li and Tian [87] discussed a GI/M/1 queue with single working vacation and obtained the steady-state distributions of the number of customers in the system at both arrival and arbitrary epochs by using the matrix-geometric approach. They also presented the stochastic decomposition structures for the queue length and waiting time but such property was not used in Baba [9]. Recently, Ye and Liu [146] studied a new class of the GI/M/1 queue with single working vacation and vacations, and derived the stationary distribution of the system size at arrival epochs by using the matrix geometric solution method. They obtained the stochastic decompositions of the system size and conditional system size given that the server is in the regular service period.
5(i). General Arrival, General Service Time Queueing System with Vacation
The G/G/1 (general arrival and general service, one server) can be considered in an analytical queueing model to address congested links through router. This model is chosen when we have to use results of the statistical analyses pointed to non-exponential distributions, both for the arrival process and the service process in the network that follow general distributions. Doshi [29] proved some stochastic decomposition results for variations of the GI/G/1 queue. Keilson and Servi [64] dealt with the oscillating random walk models for GI/G/1 vacation systems with Bernoulli schedules. They showed that the decomposition results for exhaustive service can be extended to the more general class of Bernoulli schedules. Li and Niu [85] investigated a generalization of the GI/G/1 queue in which the server is turned off at the end of each busy period and is reactivated only when the sum of the service times of all waiting customers exceeds a given threshold of size D.

5(ii). Geometric Arrival, General Service Time Queueing Systems and General Identical Arrival, geometric Service Time Queueing Systems with Vacations
Queues with disasters have been intensively studied in the past by various researchers due to their great applications in complex modern communication systems, networks and manufacturing systems. Since the discrete-time queue is more suitable for describing the telecommunication network, digital communication systems and other related areas, there is a growing interest in the analysis of the discrete-time queues with disasters. These are studied under the queueing models of the type Geo/G/1 wherein the geometrical inter-arrival time distribution takes place. The time axis is divided into equal time intervals and is marked with 0,1,2,...,n,... All the arrivals only happen at boundary epochs of time slots in discrete-time and GI/Geo/1 wherein the departures only happen at boundary epochs of time slots in discrete-time. Meisling [103] studied discrete-time queueing theory and analyzed a single-server queueing system treating time as a discrete variable. Li et al. [90] studied the discrete-time GI/Geo/1 queueing system with multiple working vacations. Using the matrix-geometric solution method, they obtained the steady-state distribution of the number of customers in the system and presented the stochastic decomposition property of the queue length. Li [83] investigated a discrete-time Geo/G/1 queueing system with multiple working vacations and analyzed the model using matrix analytic approach and the stochastic decomposition theory. Recently, Gao and Wang [40] analyzed a discrete time Geo/G/1 retrial queue with server vacation and two waiting buffers based on ATM networks by using the supplementary variable technique and the generating function approach.

5(iii). Various Queueing Systems with Vacations
Servi [111] analyzed a D/G/1 queue with vacations and investigated the waiting time probability distribution at the primary queue in terms of the primary task arrival rate, the service time distribution for the primary tasks, the probability distribution of the vacation duration, and the processor’s service schedule. Jain and Agrawal [52] investigated a state
dependent $M/E_k/1$ queueing system with server breakdown and working vacation. They derived several performance indices in explicit form by applying generating function approach. Baba [10] considered the $M/PH/1$ queueing system with phase type working vacations and vacation interruption. He obtained the conditional stochastic decomposition structures of queue length and waiting time when the service distribution in the regular busy period is exponential. Alfa [3] dealt with the $MAP/PH/1$ vacation queueing system at arbitrary times using the matrix-analytic method and obtained decomposition results for the $R$ and $G$ matrices. Goswami and Selvaraju [42] analyzed the non-homogeneous quasi-birth-death model of a $PH/M/C$ queue with impatient customers and multiple working vacations. They used the finite truncation method to determine the stationary distribution. Recently, Chakravarthy et al. [22] studied a $MAP/PH/1$ type queueing model with server vacations, breakdowns, repairs and backup server. They performed a qualitative study of the model in steady-state through a number of system performance measures including the tail probabilities of the sojourn time for various scenarios. Very recently, Dudin et al. [33] analyzed a single server $MAP/PH/1$- type queueing system with server vacations which is useful for the analysis of multiple access systems with pooling discipline without transmission interruption. Sakuma et al. [110] considered an $M/PH/1$ queue with workload-dependent processing speed and vacations. For this system, they obtained the steady-state workload distribution and its moments of any order. Panta et al. [106] presented the development trend of queueing systems, their forms and applications in modern real life situations.

6(i). N-policy Imposed Queueing System

The N-policy is to turn on the server when the queue size reaches the number $N$ and turn him off when the system is empty. The concept of N-policy was introduced by Yadin and Naor [143] in 1963 under the investigation of a single-server queueing system with constant Poisson input and a removable service station. Some of researchers have been attracted towards the N-policy queueing systems. Hayman [45] discussed the optimal operating policies for $M/G/1$ queueing systems and established the recursion relation to find the optimal (non-stationary) policy for finite horizon problems. Balachandran [12] considered two parametric control policies (N-policy and D-policy) for $M/G/1$ queueing system and compared the optimum policies for both the policies when the costs assumed were linear. Shanthikumar [113] studied an $M/G/1$ queueing system with non-preemptive priority service discipline and N-control policy. Borthakur et al. [16] dealt with a Poisson input single server queueing system with startup time and under control operating policy. Tian et al. [131] obtained various transient and steady-state results for the queue size, the delay times and the waiting times for the $M/G/1$ queue with controllable vacations. Medhi and Templeton [102] examined the steady state behavior of an $M/G/1$ queue under N-policy and with a general start up time. They generalized the results obtained by Borthakur et al. [16]. Muh [105] studied a single-server queueing system with a compound Poisson
input stream and generally distributed service times, N-control operating policy, bi-level controlled service discipline and start-up time. Takagi [120] analyzed M/G/1/K queueing system with N-policy and setup times. He obtained the queue length distributions and the mean waiting times for the exhaustive service system, the gated service system, the E-limited service system and the G-limited service system. Lee and Park [70] studied an M/G/1 queueing system with early set-up and N-policy. They applied the results of the decomposition property to the system with (m, N) policy and derived the queue length distribution. Artalejo [5] studied the M/G/1 queue with N-policy. He showed some applications of the stochastic decomposition property for the queue size and observed a new stochastic decomposition property for the waiting time. Hur and Paik [48] studied an M/G/1 queue with general server setup time under an N-control policy. They considered the case when the arrival rate varies according to the server’s status: idle, setup and busy states. Lee et al. [71] established an explicit form of matrix decompositions of the queue length distributions in the MAP/G/1 queues under multiple and single vacations with N-policy. Ke [57] investigated the control policy of the N policy M/G/1 queue with server vacations, startup and breakdowns. Ke and Pearn [59] considered the optimal management policy for M/M/1 queueing system with heterogeneous arrivals under N policy, in which the server was characterized by vacations and break downs. Liu et al. [94] investigated an M/G/1 retrial G-queue with preemptive resume and feedback under N-policy vacation subject to the server breakdowns and repairs. Kannadasan et al. [56] analyzed an F/M/FM/1 queue with multiple working vacation with N-policy. Very recently, Gao and Zhang [41] analyzed the performance of a hospital service system by modelling it as a continuous time M/G/1 queue with retrial customers due to service vacation and derived various performance measures by using supplementary variable technique and transform methods.

6(ii). D-policy, T-policy and F-policy Imposed Queueing Systems
Sometimes we encountered such a complex-problem in which solution of the problem is only possible when we are confined to some rules in constructing the queueing models, these rules are well defined which we term as D-policy, T-policy and F-policy. The D-policy is defined as a control policy which turns the server on when the total work to be done reaches the value D. The concept of D-policy was introduced by Balachandran [88] in 1973. Balachandran and Tijms [13] considered the D-policy for the M/G/1 queue and it was shown to be superior to the Heyman, Yadin-Naor N-policy for the service time distributions distributed as exponential, with decreasing failure rates and for some cases with increasing failure rates. In the T-policy, the server is turned off for a fixed time interval T at the end of each busy period and then either resumes the queue service or stays idle depending on whether or not there are waiting customers at the end of T. The concept of T-policy was introduced by Heyman [46] in 1977 under the development of M/G/1 queue which activates the server T time units after the end of the last busy period. Teghem [123]
A.P. Panta, R.P. Ghimire, D. Panthi and S. R. Pant

dealt with the determination of the optimal operating rule for the behavior of a removable server and examined the case of an individual service process in $M/G/1$ queueing system with $N$-policy, $D$-policy & $T$-policy. Dshalalow [31] dealt with the queueing process in a class of $D$-policy models with Poisson bulk input, general service time and a multiple vacation, single vacation and idle server. Artalejo [6] studied some aspects of the $M/G/1$ queue with $D$-policy in more depth. Feinberg and Kella [35] studied the optimality of $D$-policies in an $M/G/1$ queue with a removable server and they proved an old conjecture that $D$-policies are optimal for the average cost per unit time criterion. Agrawal and Dshalalow [1] studied a variant of the $M/G/1$ queueing model with bulk input, $D$-policy and multiple vacations. They explored a new fluctuation technique of multivariate marked counting processes and derived the Laplace-Stieltjes transform of the busy period. Ke [58] studied the modified $T$ vacation policy of an $M/G/1$ queueing system with a un-reliable server and startup, and he derived the LSTs of various system characteristics by using the analytical results. Lee et al. [72] investigated the steady-state queue length and waiting time of the $M/G/1$ queueing system under the $D$-policy and multiple vacations. Xu and Zhang [140] studied a multi-server $M/M/c$ queueing system with a single vacation $(e, d)$-policy. They formulated the system as a quasi-birth-and-death process and developed the various stationary performance measures. Also, they proved several conditional stochastic decomposition properties.

In the queueing model with $F$-policy, when the number of customers in the system reaches its capacity $K$ ($K < \infty$), no further customers are allowed to enter into the system until a certain number of customers, who are already in the system, have been served in order to make the number of customers in the system decreases to a predetermined threshold $F(0 \leq F \leq K - 1)$. Gupta [43] first introduced the concept of $F$-policy in 1995 and established the interrelationship between $F$-policy problem and the truncated $N$-policy problem. Wang et al. [136] investigated the optimal management problem in an $M/G/1/K$ queueing system with combine $F$ policy and an exponential startup time. They established a cost model to determine the optimal management $F$-policy at minimum cost. Yang et al. [145] studied the $F$-policy $M/M/1/K$ queueing system with single working vacation and an exponential startup time, and developed the steady-state probability vectors in matrix form by the matrix-analytic approach. They also carried out an optimization numerical analysis by using the direct search method and Quasi-Newton method.

7. Finite Capacity Vacation Queueing Systems

Tremendous number of literatures can be found in the study of queueing system in which the infinite capacity has been provisioned but in the real life situations it is not always relevant. Queueing models with finite capacity have been studied by several authors in various frameworks with their field of applications. Some such models are worthwhile to mention here. Courtois [27] analyzed a finite capacity $M/G/1$ queue with delays in the steady-state and obtained closed form expressions for the distribution of the queue length.
A Review of Vacation Queueing Models in Different Framework

and the first two moments of the queueing time distribution. Lee [77] studied a finite capacity M/G/1/N queueing system with server vacation and exhaustive service discipline. Lee [78] considered an M/G/1/N queueing model with vacation time and limited service discipline. Using a combination of the supplementary variables and sample biasing techniques, he derived the waiting time distribution, blocking probability and general queue length distribution. Takagi [121] considered in detail an M/G/1/N queueing system with the server’s multiple vacations and exhaustive service. He explicitly obtained the distributions of the unfinished work, the virtual waiting time and the real waiting time in the steady state. Frey and Takahashi [36] dealt with a finite capacity M/GI/1/N queue with vacation time and exhaustive service discipline. Focusing only on the service completion epochs, they presented a simple analysis for the queue length distribution at an arbitrary time as well as for the waiting time distribution. Li and Zhu [82] derived a computationally tractable scheme for obtaining the exact stationary queue length distribution and then some other important performance characteristics in M (n) /G/1/N queues with generalized vacations and exhaustive service.

Hui and Wei [47] studied the two-phases - service M/M/1/N queueing system with the server breakdown and multiple vacations. They derived the equations of steady state probability by applying the Markov process theory and they obtained matrix form solution of steady - state probability by using blocked matrix method. Yang and Wu [144] investigated the transient behavior of M/M/1/N queue with working breakdowns and multiple vacations. They applied a numerical technique based on the fourth-order Runge-Kutta method to compute the transient state probabilities. Agrawal et al. [2] dealt single server finite Markovian queue with breakdown, repair and state dependent arrival rate under N- policy by using recursive method. They obtained steady state results for different states by using the probability generating function technique. Recently, Bouchentouf et al. [19] analyzed a finite capacity M/M/1 feedback queueing system with vacation and impatient customers and they obtained measures of effectiveness of the model by using the stationary distribution.

8(i). Vacation Queueing Models with Batch Arrival

Instead of single arrival and providing service for a single customer, sometimes arrivals and the services take place in a fixed or random group size which is known as the batch or bulk arrival/service such as production/manufacturing systems, communication systems, transportation systems, computer networks etc. The study of bulk queues originated with the pioneering paper by Bailey in 1954. Choudhary and Templeton [25] and Medhi [101] discussed the batch arrival queueing models at great length.

Many researchers have contributed to the study of Markovian type vacation queueing models with batch arrival in different frame-works. Wang et al. [135] analyzed the M[K]/M/1 queueing system with multiple vacations and server breakdowns and they developed the approximate formulae for the probability distributions of the number of
customers in the system using the maximum entropy principle. Xu et al. [141] studied a bulk input $M^{[x]}/M/1$ queue with single working vacation and they derived the probability generating function of the stationary queue length distribution by matrix analysis method. Baba [11] treated a batch arrival $M^{X}/M/1$ queue with multiple working vacation. Using a quasi-upper triangular transition probability matrix of two dimensional Markov chain and matrix analytic method, he obtained the probability generating function of the stationary system length distribution and obtained the stochastic decomposition structure of the system length. Mary et al. [99] studied the batch arrival $M^{X}/M/1$ queueing system along with server breakdowns and multiple working vacation under exponential distribution. They obtained the PGF of the system size through the Chapman-Kolmogorov balanced equations satisfied by the steady state system size probabilities and verified the stochastic decomposition property. Uma and Manikandan [134] studied single server bulk queueing system with three stage heterogeneous service, compulsory vacation and balking. They computed transient solution, steady state results, the average number of customers in the queue and the system. Very recently, Bouchentouf and Guendouzi [17] analyzed an $M^{X}/M/c$ Bernoulli feedback queueing system with variant of multiple working vacations and reneging which depend on the states of the servers and retention of reneged customers.

Many researchers have studied $M/G/1$ type vacation models with batch arrivals. Baba [7] studied the $M^{X}/G/1$ queueing system with vacation time and derived the general queue length distribution, the waiting time and busy period distributions. Baba [8] considered the $M^{X}/G/1$ queue with and without vacation time under non-preemptive last-come first-served discipline. He found the relationship between the second moments of the waiting time under first-come first-served and last-come first-served disciplines. Lee [69] studied the $M/G/1$ bulk arrival queues with server vacations and derived the server idle probabilities, waiting time distributions and the mean system sizes. Rosenberg and Yechiali [109] considered the $M^{X}/G/1$ queue without server vacations and with multiple vacations under the LIFO service regime. To each model, they derived explicit formulae for the Laplace-Stieltjes transform, mean and second moment of the waiting time of an arbitrary customer and extended the range of Fuhrmann’s general result. Lee [66] studied a finite capacity $M^{X}/G/1/K$ queueing system with multiple server vacations and $(\gamma, N)$ policy. Under a classical cost structure, he characterized an optimal policy and developed an algorithm to find an optimal policy which minimizes the expected cost per unit time. Choudhury [26] investigated an $M^{X}/G/1$ queueing system with a vacation period and a random setup period. Chang and Takine [23] considered the $BMAP/G^{b}/1$ queue with generalized vacations and derived a new stochastic decomposition formula for the stationary queue length distribution. Begum and Jose [15] analyzed an $M^{X}/G/1$ queueing system in which server takes at most $J$-working vacations during the idle period and they derived the PGF of the system using supplementary variable technique. Haridass and Nithya [44] analyzed the server break down with interrupted vacation in a $M^{X}/G(a,b)/1$
A Review of Vacation Queueing Models in Different Framework

queueing system. They obtained PGF of the steady state queue size distribution at an arbitrary time epoch and expressions for various performances.

8(ii). Batch arrival vacation queueing models with N-policy
Variable batch size in the study of queueing systems plays a paramount role to tackle tactical and strategic problems on which some of the researchers are interested in working such queueing models with the N-policy. Lee and Srinivasan [67] first studied the $M^X/G/1$ queueing system with N-policy. They derived the mean waiting time of an arbitrary customer and developed the procedure to find the stationary optimal policy under a linear cost structure. Federgruen and So [34] considered an $M^X/G/1$ queueing system with server vacations. They obtained the optimality of the no-vacation policy or of that of a threshold or N-policy and various generalizations and comparison results. Lee et al. [75] analyzed the $M^X/G/1$ queueing system with N-policy and multiple vacations. Lee et al. [73] discussed the operational characteristics of the $M^X/G/1$ queue with N-policy and derived and gave appropriate interpretations for system size distribution, the waiting time distribution and other performance measures. Lee et al. [76] dealt with a batch arrival $M^X/G/1$ queueing system under N-policy and single vacation. They derived the system size distribution and queue waiting time distribution of an arbitrary customer and presented a procedure to find the optimal stationary under a linear cost structure. Chae and Lee [20] presented heuristic interpretations of the mean queue waiting times of the $M^X/G/1$ vacation (single and multiple) models with N-policy. Lee et al. [74] analyzed the fixed size batch service $M/G^X/1$ queue with single and multiple server vacations. They derived the decompositions of the queue length distributions and provided relevant interpretations. Jacob and Madhusoodanan [50] dealt with the transient solution for a finite capacity $M/G^{a,b}/1$ queue with server vacations. Dshalalow [32] analyzed the queueing process in stochastic systems with bulk input, batch state dependent service, server vacations and three post vacation disciples: waiting, or leaving on multiple vacation trips with or without emergency. Ke et al. [60] derived the probability generating function of the number of customers in the various server’s state for the $M^{[K]}/G/1$ queueing system with N-policy and at most J vacations by using the supplementary variable technique.

9. Challenges on using vacation concept in queueing models
Most of the contributions made to date are concerning only with single stage queueing systems vacations but there are several real life situations where multi stage queueing systems can exist which have not been investigated yet. Vacation models with heterogeneous arrivals and service systems are felt dreaded even today due to multi-dimensional characteristics. Multi-threshold polies create another challenge in the queue. Accounting for the coefficients of correlation, variation of inter-arrival times and coefficient of vacation of service time under PH-distribution have rarely been dealt yet.
10. Techniques used in vacation queueing models
Various techniques have been introduced by many researchers to analyze the queueing models with vacation. Some of their techniques include the matrix analytic method (matrix geometric solution method), supplementary variable method, generating function method, recursive method, iterative method, Runge-Kutta method, direct search method, Quasi-Newton method, transform method, finite truncation method, Markov process, semi-Markov process, birth-death process, Quasi-birth-and-death process, theory of regenerative process, semi-regenerative process, censoring technique, embedded Markov chain technique, sample biasing technique, enhances fluctuation technique, maximum entropy principle, GIL algorithm, rate conservation law, control policy, stochastic decomposition theory, Maple computer program. Most of the parameters we take into account under the study of vacation queueing systems exhibit stochastic process due to which the variables we denote are random variables such as arrival rate, service rate, reneging rate, balking rate, failure rate, repair rate, resource sharing rate.

11. Conclusion
A survey in vacation queueing systems has been made so as to give insight into its pioneering to its advancement. Some of the vacation queueing policies which have been used in various situations have been illustrated with their pioneers. The up to date development of the vacation queueing system reported in the survey will enable the reader to get knowledge in a bird’s eye-view.

Acknowledgements. The authors would like to thank the reviewers for their valuable suggestions and comments for improving this paper.

Conflict of interest. The authors declare that they have no conflict of interest.

Authors’ Contributions. All the authors contribute equally to this work.

REFERENCES
A Review of Vacation Queueing Models in Different Framework

A.P. Panta, R.P. Ghimire, D. Panthi and S. R. Pant


A Review of Vacation Queueing Models in Different Framework


47. Z. Hui and G. Wei, The two-phases service M/M/1/N queueing system with the server breakdown and multiple vacations, *ICICA*, 2011 (2011) 200-207.


A Review of Vacation Queueing Models in Different Framework


A Review of Vacation Queueing Models in Different Framework


112. L.D. Servi and S.G. Finn, $M/M/1$ queues with working vacations ($M/M/1/WV$), *Performance Evaluation*, 50(1) (2002) 41-52.


A Review of Vacation Queueing Models in Different Framework


