



# Performance analysis of P2P network with dynamic changes of servers based on M/M/c queuing model

Qiannan Si<sup>1</sup> · Zhanyou Ma<sup>1</sup> · Fengjiao Liu<sup>1</sup> · Rong Wang<sup>1</sup>

Accepted: 24 May 2021 / Published online: 4 June 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

Peer-to-Peer (P2P) network, which is distributed application architecture, distributes tasks and workloads among peers. One of most characteristics of P2P network is that peers play dual roles as servers and clients. To study the interaction process of nodes in P2P systems dynamically, the download delay of nodes, the energy consumption of the system and simulate the behavior of peers, analyze the performance of the P2P network, we develop an M/M/c queuing model with dynamic changes of servers, impatient customers and fault repairable. Using the quasi-birth-and-death process (QBD), the geometric solution of the matrix for the steady-state probability vector is derived. Since the matrix is quite complex, approximate solutions of the matrix equations are obtained using the Gauss–Seidel method, the performance indicators of the system are obtained. By employing Nash equilibrium and social optimal strategy, the optimal value of social benefits is obtained, which provides a theoretical basis and decision evaluation for scheduling of the nodes in the P2P system.

**Keywords** P2P network · M/M/c queue · Dynamic change · Work breakdown · Matrix-geometric solution method · Social optimization

## 1 Introduction

Driven by the research of Bit Torrent (BT), resource downloading, and live streaming, researchers are paying more and more attention to Peer-to-Peer (P2P) network technology, which has been developed and improved. It is a powerful advantage in distributed computing [1], teamwork, and multimedia [2, 3]. In the P2P system, the peer is regarded as the server, which downloading and transmission files, and the peer also is regarded as the consumer, which requesting service. With the progress of downloading and transmission, consumers are transformed into

servers that serve other peers, wherefore the service capability of the system increases exponentially [4].

Research on P2P performance analysis based on P2P structure is becoming more and more extensive. Li et al. [5] investigated an information pool protocol (IPP) based on grid architecture (IPBGA), which is a real hybrid P2P network and grid system. Their theoretical analysis showed that the IPP of IPBGA is more effective and has lower bandwidth and processing cost. A fast convergence iterative algorithm was proposed for the Cloud-P2P content distribution network system, and the optimal allocation of bandwidth between the cloud storage terminal and the P2P terminal were studied [6, 7]. Yang et al. [8] proposed a hybrid P2P network framework for interactive streaming media, called PPISM, which reduces the load of edge server. The framework adopted hierarchical structure, which integrated content distribution network (CDN), P2P network and tree grid structure. Zhou et al. [9] presented a hybrid strategy of Rarest First and Greedy. To ensure scalability, they designed an algorithm that adapts to the number of nodes dynamically [10], while retaining part of the work of the nodes to meet the requirements of instant playback and ensure low startup delay. In terms of network

---

✉ Zhanyou Ma  
mzhy55@ysu.edu.cn

Qiannan Si  
17862072593@163.com

Fengjiao Liu  
lfj456456@163.com

Rong Wang  
wangrong9327@163.com

<sup>1</sup> School of Science, Yanshan University,  
Qinhuangdao 066004, China

throughput, Ye et al. [11] proposed a distributed adaptive mixed media access control scheme and a throughput optimization framework. The aggregated data throughput is maximized by adjusting the optimal contention window size according to the change of voice and data service load. Aiming at the scalability problem of P2P, Edmundo et al. [12] developed a model, which calculated the achievable throughput under different nodes and block selection strategies. A new understanding of the behavior of P2P cluster system that provided a new mechanism for publishers and nodes to improve the overall performance of the system. Zhou et al. [13] proposed a resource search algorithm, which uses pull and push methods. In order to deal with the liquidity problem of nodes, the activities of nodes are often repeated on a day-to-day basis in the mobile P2P ad hoc networks (P2P MANETs). Through simulation, they studied that the resource search scheme can not only improve the search success rate, but also reduce the spread of information. Yang et al. [14] studied the performance of the P2P file-sharing system from the perspective of service capacity. They found that the well performance was showed in the P2P system as the load increased in both transient and steady-state conditions scalability. In the mobile P2P network environment, Lim et al. [15] researched an effective range query processing scheme, which includes two phases: query distribution phase and monitoring phase. Li et al. [16] addressed the multivirtual neural network embedding problem by a game-theoretic approach and designed a new priority admission control mechanism to handle overloaded service requests, reducing the amount of information exchange and computational complexity.

In recent years, it is gradually increasing that the problems of P2P network performance has been solving by queuing theory. Ge et al. [17] developed a multi-class closed queuing network in the P2P system, in which each class is composed of a fixed node group and studies the performance of the P2P file-sharing system through simulation of the model. Li et al. [18] modeled M/MMPP/1 queue in the BT system and solved the corresponding stable distribution using matrix geometric solution. The relevant results that performance analysis of BT system was given. Yin et al. [19] established a new dynamic model, which reflected the interaction among nodes in the Media Delivery Network (MDN) system. They proposed the node selection and bandwidth allocation strategy, which verified the rationality and validity of the model. To inculcate cooperation among the otherwise selfish communicating nodes, the pricing game between nodes as the Rubinstein Bargain model was formulated to enhance content distribution in the delay-tolerant networking (DTN) [20]. Ferragut et al. [21] and Hu et al. [22] analyzed the dynamic characteristics of that P2P file exchange swarm from the perspective of queue. The service rate of

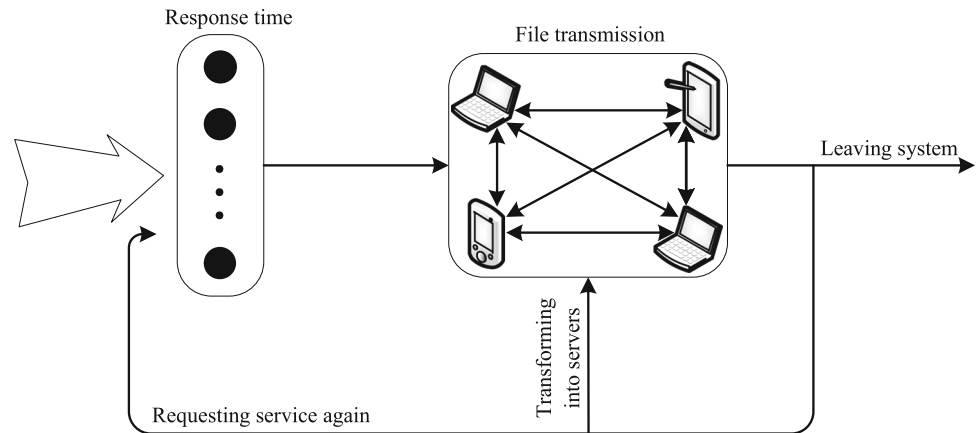
receiving peer depends on a main fixed central component (servers or seeders) and the number of peers. Zhao et al. [23] presented the M/M/m(t) queuing model, in which servers are variable based on the migration of virtual machines in the cloud system. The stationary distribution and performance indicators are obtained by using analytic theory of matrices. Under the constraints of service level agreement (SLA), the least number of virtual machines in open and the impact of different parameters on performance are analyzed. The above-mentioned papers have studied from the perspectives of service capacity, broadband allocation, node selection strategy, network performance, and so on. However, there is little literature that analyzed the service process of multiple servers dynamically by the quasi birth and death process (QBD) and the method of matrix geometric solution.

Ma et al. [24] formulated the M/M/1 queuing system with multiple vacations and repairable faults by the method of generating functions. This method provides a useful ideological reference for being assumed that there are repairable peers in the P2P system. Wu et al. [25] studied a queuing system that limited capacity with a change in the number of servers. Therefore, it can be assumed that the dynamic changes of peers in the file transfer process of the P2P system. If the resource package required by the requesting node is less, it is not easy to query and will cause the larger query traffic. If there are too many nodes requesting the same resource, the conversion probability of the node into a service node is low, which leads to too long delay of the requesting node and abandoning the request.

In this paper, considering dynamic changes of the loads during the process of the nodes scheduling, the performance of the P2P system based on the classical M/M/c queuing model with the strategies of repairable work failure and impatient customers is studied. And we propose a reasonable, effective and applicable peer scheduling strategy, which can reduce energy consumption and improve the service performance of P2P network systems. The contributions of this paper are threefold.

- (1) To guarantee the stability of system operation and the efficiency of transmission between nodes, the M/M/c queuing model is proposed with the strategies of dynamic changes of servers, repairable work failure and impatient customers. The main purpose of this paper is to investigate how the service efficiency and the conversion efficiency of the nodes can lead to a low energy consumption of the system and a smaller average delay of the nodes.
- (2) A three-dimensional continuous time Markov stochastic model is established, the expressions of average delay of nodes and throughput rate of system

**Fig. 1** The operation mechanism of the P2P system



are obtained by using matrix-geometric solution method and Gauss–Seidel iterative method.

- (3) The social utility function is defined, the effects of parameter variations on the performance metrics of the P2P network system are derived through numerical analysis. The Nash equilibrium and the social optimization strategies are analyzed in detail, and the conversion rate of nodes is increased appropriately to reduce system delay and congestion, enhance throughput of system and social utility.

The rest of this paper is organized as follows. In Sect. 2, the operating mechanism of the P2P system is introduced, and the relevant parameters are assumed. In Sect. 3, the three-dimensional Markov chain is built, and the steady-state distribution is given. In Sect. 4, the performance indicators are evaluated by using some numerical results and the Nash equilibrium and social optimal strategy are analyzed. Conclusions are introduced in Sect. 5.

## 2 Model description

Unstructured distributed P2P networks, in which nodes are accessing the whole network by communicating with neighboring nodes. And nodes are connected irregularly without stable relationships. Nodes search for needed resources by employing the mechanism of query packets. Specifically, the node sends a query packet containing a query to its neighboring nodes, and the query packet is transmitted through an incremental manner in the network. At first a large number of requesting nodes enter the system to wait for service with the first come first served (FCFS) rules, which is called response time. And requesting nodes enter the file transfer phase after querying the nodes that has the desired resource package. Nodes that finish downloading have three options, including: continue to request downloads in order to obtain other resources, transform into servers to serve other requesting nodes, and

leave the system after completing the service. The operation mechanism of the P2P network system is shown in Fig. 1.

The process of file transmission is abstracted as the process of service in P2P network system. The requesting node is abstracted as customer. The service node is abstracted as sever. A new queuing model is established in the P2P system. The strategies of model that including impatient customers, repairable work failure, and dynamic changes of nodes are introduced. Meanwhile, the relevant parameters are assumed as follows.

- (1) The requesting node follows the Poisson process with the parameter  $\lambda$ . The model is assumed that the capacity of the requesting node is infinite in P2P system. The system allows to be idle when the server is not requested. In the idle state, it may be attacked by unsafe factors, but it does not cause congestion of system.
- (2) The time consumed by each node to complete the service follows the exponential distribution with parameter  $\mu_1$ . Assuming that there are  $n$  requesting nodes, there are  $m$  seeds (nodes that have a complete copy after the downloading) in the system, and the service rate are  $m\mu_1$  ( $m < n$ ),  $n\mu_1$  ( $n \leq m$ ) respectively. The minimum number of nodes that is online is 1, and the number of all nodes available is  $c$ .
- (3) The nodes that has completed service can remain in the system with probability  $\alpha$  ( $0 < \alpha < 1$ ). In the system the remaining nodes have two choices: (i) if some nodes want to download other files again, they are regarded as new customers and wait at the end of line after re-entering the system. For any requesting node, re-entering the system with the probability  $\beta$  ( $0 < \beta < \alpha$ ), which is known as the feedback probability; (ii) the node that has finished the downloading can transform into seeds with probability  $\gamma$  ( $\gamma = \alpha - \beta$ ) in the system. If the node that

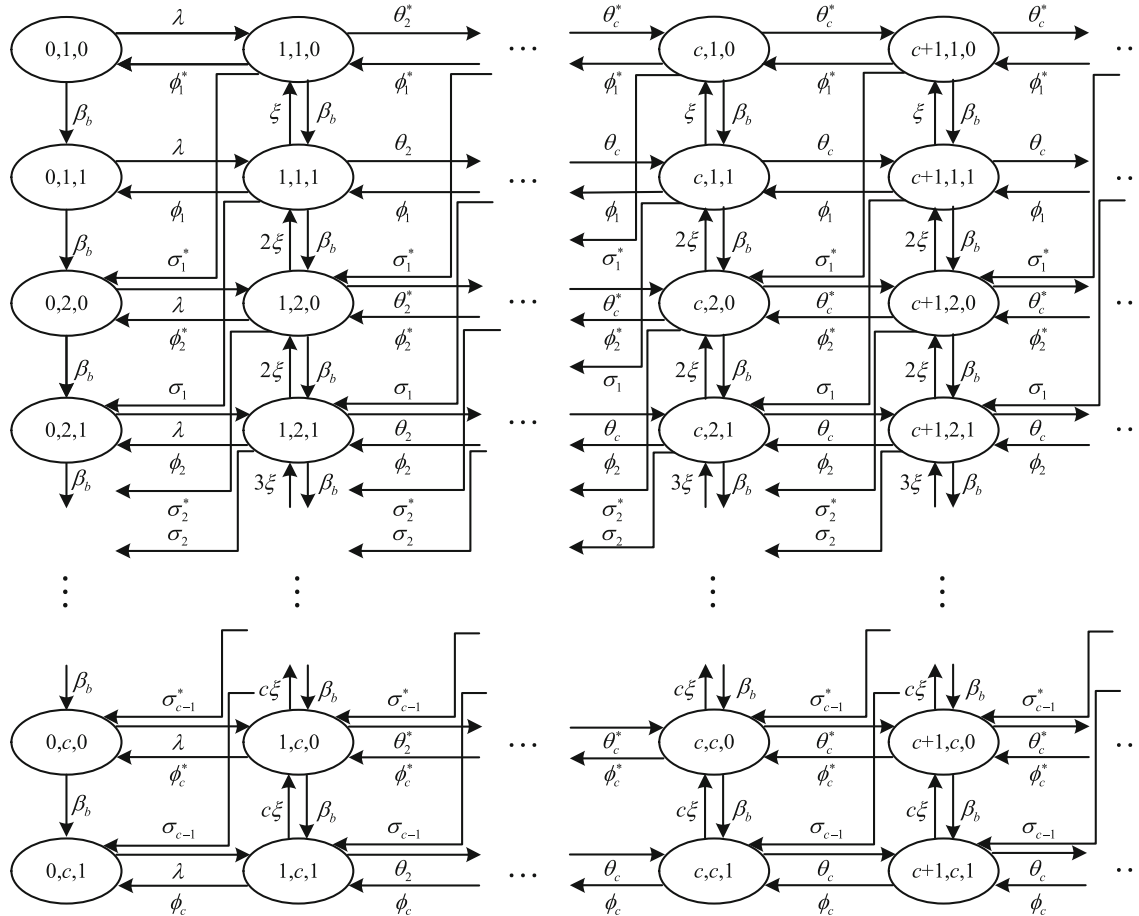


Fig. 2 State transition diagram of queuing model

completed service is idle, it will leave system with probability  $\bar{\alpha}$  ( $\bar{\alpha}=1 - \alpha$ ).

- (4) Failed nodes that may congest or delay because the downloading is too heavy can be repaired. The failed node does not allow completely down, and can serve others at a lower service rate. The process of failure follows the Poisson process with parameter  $\xi$ . The failed node that provides service with the lower service rate  $\mu_2$ . Due to the relatively low service rate during failures, some nodes are impatient. If the requesting node is not received within time  $T$ , it may become an impatient node and leave system. And the time  $T$  follows the exponential distribution with parameter  $\eta$ . If the working node fails, it can be repaired immediately by waiting in line and the same nodes can only be repaired by one repairman. The repair time follows the exponential distribution with parameter  $\beta_b$ .
- (5) The service sequence is FCFS. It is assumed that the arrival time, the service time, and the repair time are mutually independent.

### 3 Model analysis

#### 3.1 State transition rate matrix

Let  $L(t)$  denotes the number of requesting nodes at time  $t$ ,  $Y(t)$  denotes the number of service nodes at time  $t$ ,  $J(t)$  denotes the state of the nodes at time  $t$ , and  $J(t)$  is defined as follows:

- (1)  $J(t) = 0$  indicates that the nodes is in a fault state or idle state at time  $t$ .
- (2)  $J(t) = 1$  indicates that the nodes is in a work state at time  $t$ .

Then  $\{(L(t), Y(t), J(t)), t \geq 0\}$  indicate a three-dimensional Markov process, which has the state space as follows:

$$\Omega = \{(i, l, j), i \geq 0, 1 \leq l \leq c, j = 0, 1\}$$

where the state set  $(i, 1, 0), (i, 1, 1), \dots, (i, c, 0), (i, c, 1)$  is called level  $i$  ( $i \geq 0$ ). The level 0 denotes the idle state of system, which the number of requesting nodes is zero.

When servers are in a fault state (i.e.  $j = 0$ ), and  $1 \leq l \leq c$ , we can obtain



radius  $SP(\mathbf{R}) < 1$ . Simultaneously, the  $2c^2 + 2c$  dimensional stochastic matrix is as follows:

$$B[\mathbf{R}] = \begin{bmatrix} \mathbf{A}_0 & \mathbf{C}_0 & & & & & \\ \mathbf{B}_1 & \mathbf{A}_1 & \mathbf{C}_1 & & & & \\ & \mathbf{B}_2 & \mathbf{A}_2 & \mathbf{C}_2 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & & & \mathbf{B}_{c-1} & \mathbf{A}_{c-1} & \mathbf{C}_{c-1} \\ & & & & & & \mathbf{B} & \mathbf{BR} + \mathbf{A} \end{bmatrix}$$

has left-zero vector, when QBD is positive recurrent, the steady-state distribution satisfies the following equations:

$$\begin{cases} (\pi_0, \pi_1, \dots, \pi_c)B[\mathbf{R}] = \mathbf{0}, \\ \sum_{i=0}^{c-1} \pi_i \mathbf{e} + \pi_c (\mathbf{I} - \mathbf{R})^{-1} \mathbf{e} = \mathbf{1}, \\ \pi_i = \pi_c \mathbf{R}^{i-c}, \quad i \geq c \end{cases} \tag{3}$$

where  $\mathbf{e}$  denotes a  $2c$  dimensional column vector, in which all elements are 1.

Because expressions of the matrices  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  are relatively complex, It is difficult to obtain the exact expression of the minimum non-negative solution  $\mathbf{R}$  of the matrix  $\mathbf{R}^2 \mathbf{B} + \mathbf{R} \mathbf{A} + \mathbf{C} = \mathbf{0}$ . Therefore, the Gauss–Seidel iterative method is used to solve the problems mentioned above and the numerical results of  $\pi_{i,l,j}$  are obtained. The specific algorithm is as follows:

---

**Algorithm:** Gauss-Seidel iterative algorithm

---

```

Input: Error accuracy  $\epsilon$  (greater than 0, far less than 1),
Output: Rate matrix  $\mathbf{R}$ .
Initialization:  $\mathbf{R}_0 \leftarrow \mathbf{0}$ ,  $\mathbf{R} \leftarrow \mathbf{R}_0$ ;
Begin
   $n \leftarrow 1$ ;
   $\mathbf{R}_1 \leftarrow -(\mathbf{R}_0^2 \mathbf{B} + \mathbf{C}) \mathbf{A}^{-1}$ ;
  While  $\max\|\mathbf{R}_n - \mathbf{R}_{n-1}\| > \epsilon$ 
     $n \leftarrow n + 1$ ;
     $\mathbf{R}_n \leftarrow -(\mathbf{R}_{n-1}^2 \mathbf{B} + \mathbf{C}) \mathbf{A}^{-1}$ ;
  End
   $\mathbf{R} \leftarrow \mathbf{R}_n$ 
End
```

---

The process of the above analysis uses the matrix geometry solution method. The relevant specific proof process can refer to [26].

### 3.3 Performance indicators

According to the above analysis, the performance indicators of requesting nodes are given by using the Little’s law.

(1) The average length of the requesting node is given by

$$E(L) = \sum_{i=0}^{\infty} iP(L = i) = \sum_{i=0}^{\infty} i \left( \sum_{l=1}^c \sum_{j=0}^1 \pi_{i,l,j} \right).$$

(2) The average delay of the requesting node is given by

$$E(W) = \frac{1}{\lambda} \sum_{i=c}^{\infty} (i - c) \left( \sum_{l=1}^c \sum_{j=0}^1 \pi_{i,l,j} \right).$$

(3) The departure rate of the node in the P2P system is given by

$$D = \eta P(J = 0) = \eta \sum_{i=0}^{\infty} \sum_{l=1}^c \pi_{i,l,0}.$$

(4) The total energy consumption of the system includes the energy consumption of the node in two states of work and failure. Assuming that  $g_1$  denotes the energy consumption of the nodes in the failed state, and  $g_2$  denotes the energy consumption of the nodes in the working state. Therefore, the total energy consumption of the system is given by

$$W_y = g_1 \sum_{i=0}^{\infty} \sum_{l=1}^c l \pi_{i,l,0} + g_2 \sum_{i=0}^{\infty} \sum_{l=1}^c l \pi_{i,l,1}.$$

(5) The maximum throughput  $\Theta$  of the system is the maximum amount of data requests that all nodes can accept and serve per unit time. Hence, the maximum throughput of the system is given by

$$\Theta = \mu_2 \sum_{i=0}^{\infty} i \sum_{l=1}^c \pi_{i,l,0} + \mu_1 \sum_{i=0}^{\infty} i \sum_{l=1}^c \pi_{i,l,1}.$$

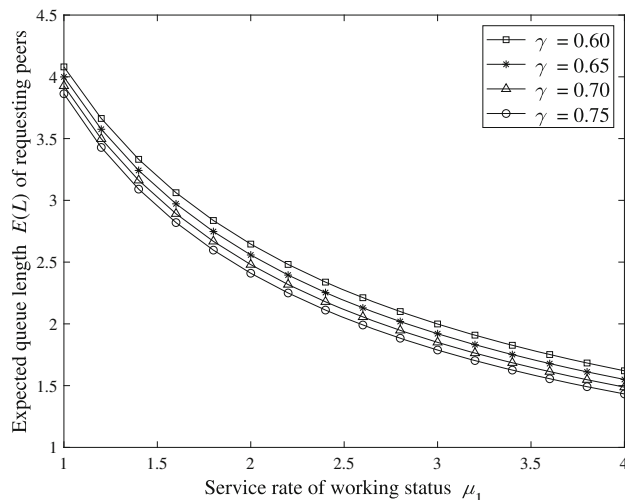


Fig. 3 The relationship between  $E(L)$  and  $\mu_1, \gamma$

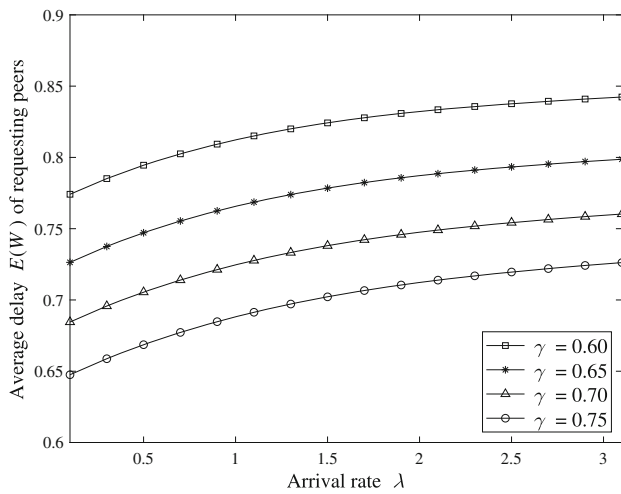


Fig. 4 The relationship between  $E(W)$  and  $\lambda, \gamma$

### 4 Numerical examples

#### 4.1 Impact of parameter variation on P2P systematic performance indicators

The minimum non-negative solution of the matrix quadratic equation Eq. (2) is obtained by Gauss–Seidel iterative method. The numerical results and various performance indicators of the P2P system can be obtained by solving equations of the steady-state distribution. By using MATLAB drawing, the relationship graph that performance index changes with parameters is obtained. According to the image, the influence of parameter changes can be analyzed in terms of performance indicators, which provide a basis for the scheduling strategy of the peers in P2P system.

According to the above analysis, the performance indicators of the system are obtained, and numerical experiments are used to verify the influence of parameters on performance indicators in P2P system. Assuming that  $\lambda = 2, \mu_2 = 0.5, c = 15, \zeta = 0.1, \beta_b = 1.5, \eta = 0.3$ . Figure 3 describes the relationship between the average length  $E(L)$  of requesting node and the service rate  $\mu_1$  under

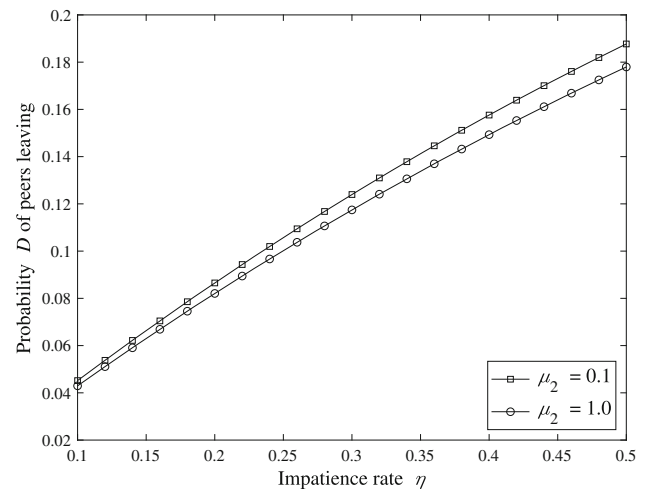


Fig. 5 The relationship between  $D$  and  $\eta, \mu_2$

normal operation of the node and the probability  $\gamma$  of nodes converting into seeds. When  $\gamma$  is a fixed value,  $E(L)$  decreases with the increase of  $\mu_1$ . When  $\mu_1$  is a fixed value,  $E(L)$  decreases with the increase of  $\gamma$ . The main reason is that the service rate of the node increases, the downloading speed of the nodes is the higher, so the possibility of congestion is the lower in the system. At the same time, the node is more likely to be converted into the seed, the average downloads increases per unit time. Therefore, the average queue length of the requesting node is smaller.

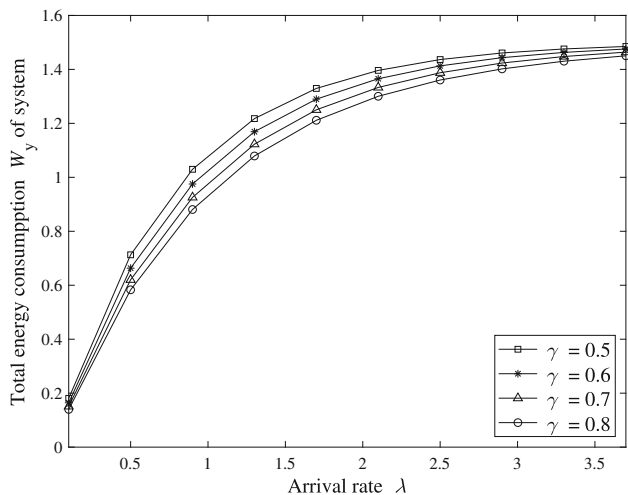
Assuming  $\mu_1 = 3, \mu_2 = 0.5, c = 15, \zeta = 0.1, \beta_b = 2.5, \eta = 0.3$ . Figure 4 describes the relationship between the average delay  $E(W)$  of the requesting node and the arrival rate  $\lambda$  and the probability  $\gamma$  of nodes that have finished service transforming into seeds. When  $\gamma$  is a fixed value,  $E(W)$  increases with the increase of  $\lambda$ . When  $\lambda$  is a fixed value,  $E(W)$  decreases with the increase of  $\gamma$ . The main reason is that the average delay of system is increasing with the arrival rate of requesting nodes increase. With the increase of the arrival rate, more and more nodes are converted into seeds, which increases the service efficiency of the system, so that the growth of average delay is flattening out. When the transformation rate of a node

Table 1 The relationship between  $E(W)$  and  $\lambda, \gamma$

	$E(W)$								
	$\lambda = 0.1$	$\lambda = 0.5$	$\lambda = 0.9$	$\lambda = 1.3$	$\lambda = 1.7$	$\lambda = 2.1$	$\lambda = 2.5$	$\lambda = 2.9$	...
$\gamma = 0.60$	0.7741	0.7945	0.8093	0.8200	0.8278	0.8335	0.8377	0.8409	...
$\gamma = 0.65$	0.7264	0.7471	0.7625	0.7739	0.7823	0.7886	0.7933	0.7971	...
$\gamma = 0.70$	0.6845	0.7055	0.7213	0.7332	0.7422	0.7490	0.7543	0.7585	...
$\gamma = 0.75$	0.6475	0.6686	0.6848	0.6971	0.7066	0.7139	0.7196	0.7242	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

**Table 2** The relationship between  $D$  and  $\eta, \mu_2$

	$D$								
	$\eta = 0.10$	$\eta = 0.16$	$\eta = 0.22$	$\eta = 0.28$	$\eta = 0.34$	$\eta = 0.40$	$\eta = 0.46$	$\eta = 0.52$	...
$\mu_2 = 0.1$	0.0452	0.0705	0.0943	0.1168	0.1379	0.1576	0.1761	0.1933	...
$\mu_2 = 1.0$	0.0430	0.0669	0.0894	0.1107	0.1306	0.1493	0.1669	0.1833	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮



**Fig. 6** The relationship between  $W_y$  and  $\lambda, \gamma$

increases by 25%, the average delay of the system decreases by about 15%. When the arrival rate of nodes is increased by 20 times, the average delay of the system increases by 8–11%. In the case of the different transformation rates of the nodes, the numerical results of the average delay and arrival rate are shown in Table 1.

Assuming  $\mu_1 = 2, \lambda = 1, c = 15, \xi = 0.1, \beta_b = 1, \alpha = 0.9, \beta = 0.1, \gamma = 0.8$ . Figure 5 describes the relationship between the departure rate  $D$  of the system, the impatience rate  $\eta$  and the service rate  $\mu_2$  of the server in the fault state. When  $\mu_2$  is a fixed value,  $D$  increases with increase of  $\eta$ . When  $\eta$  is a fixed value,  $D$  increases with decrease of  $\mu_2$ . The main reason is that the requesting node

has a lower service rate when it fails. The node may leave the system due to the waiting time is too long. If the impatience rate in the system is higher and the service rate of the failed node is lower, the more nodes may leave the system. In the case of the different impatience rates of the nodes, the numerical results of the departure rate and service rate of the server in the fault state are shown in Table 2.

Assuming  $\mu_1 = 2, \mu_2 = 0.5, c = 20, \xi = 0.1, \beta_b = 2, \eta = 0.3, g_1 = 1, g_2 = 2$ . Figure 6 describes the relationship between the total energy consumption  $W_y$  of the system and the arrival rate  $\lambda$  and the probability  $\gamma$  of nodes that have finished service transforming into seeds. When  $\gamma$  is a fixed value,  $W_y$  increases with increase of  $\lambda$ . When  $\lambda$  is a fixed value,  $W_y$  increases with decrease of  $\gamma$ . The main reason is that initially the number of nodes is very small. As the arrival rate gradually increases, the system energy consumption is higher. The probability of the node that finished service converting into seeds is the smaller, so the number of servers is the less. It is more likely to cause congestion and the server fails, so the total energy consumption of the system is also greater. In the case of the different arrival rates of the nodes, the numerical results of the total energy consumption and transformation rate of the nodes are shown in Table 3.

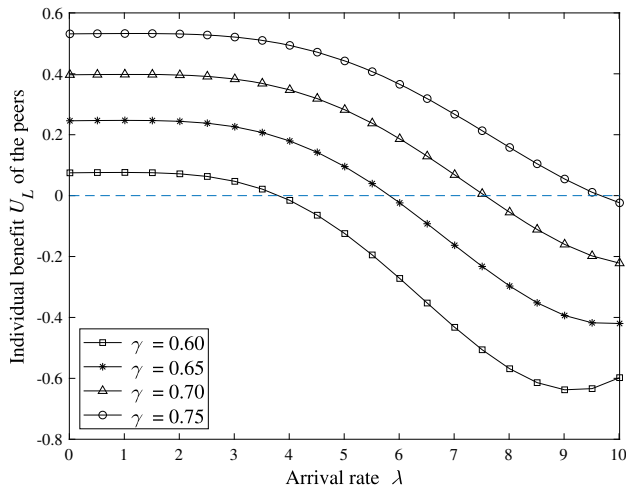
### 4.2 Nash equilibrium and social optimal

In order to analyze the social optimal strategy, the personal utility function and the social utility function are constructed. Supposing  $f_1$  represents the fee received after the

**Table 3** The relationship between  $W_y$  and  $\lambda, \gamma$

	$W_y$								
	$\lambda = 0.1$	$\lambda = 0.5$	$\lambda = 0.9$	$\lambda = 1.3$	$\lambda = 1.7$	$\lambda = 2.1$	$\lambda = 2.5$	$\lambda = 2.9$	...
$\gamma = 0.50$	0.1080	0.7128	1.0297	1.2177	1.3295	1.3963	1.4365	1.4609	...
$\gamma = 0.60$	0.1644	0.6632	0.9752	1.1691	1.2898	1.3654	1.4131	1.4435	...
$\gamma = 0.70$	0.1510	0.6204	0.9257	1.1228	1.2502	1.3331	1.3876	1.4236	...
$\gamma = 0.80$	0.1398	0.5830	0.8808	1.0791	1.2114	1.3004	1.3606	1.4019	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮





**Fig. 7** The relationship between  $U_L$  and  $\lambda, \gamma$

service is completed,  $G_1$  indicates that the average waiting cost of the node unit time due to delay, and  $f_2$  indicates the fee paid by each node entering the system. Then the personal utility function of the node is defined as

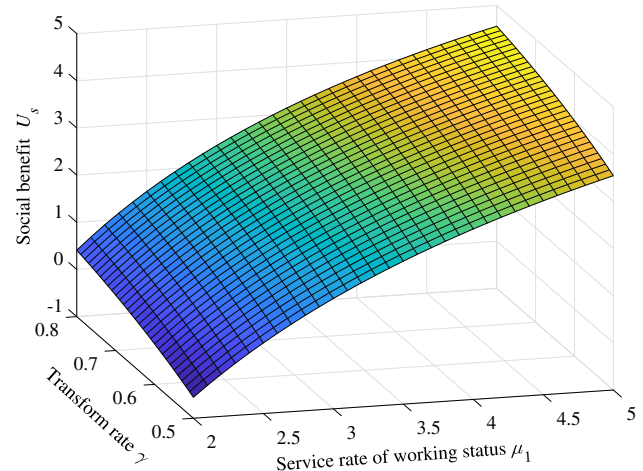
$$U_L = f_1 - G_1 E(W) - f_2.$$

Supposing that the arriving node, which face a completely invisible queuing situation, that is, the requesting node neither knows the throughput rate of the system nor understands the status of the nodes. Assuming  $\Phi$  that represents the potential arrival rate of the requesting node, and  $h^c$  represents the probability of response that the requesting node under the equilibrium strategy. The equilibrium strategy of the requesting node is given by

$$H(t) = \begin{cases} 0, & \text{when } f_1 < G_1 E(W)|_{\lambda=0} + f_2, \\ \frac{\lambda^c}{\Phi}, & \text{when } G_1 E(W)|_{\lambda=0} + f_2 \leq f_1 \leq G_1 E(W)|_{\lambda=\Phi} + f_2, \\ 1, & \text{when } f_1 > G_1 E(W)|_{\lambda=\Phi} + f_2 \end{cases}$$

where  $\lambda^c$  is the solution of  $U_L = 0$ .

Assuming  $\mu_1 = 2, \mu_2 = 0.5, c = 10, \xi = 0.05, \beta_b = 3, \eta = 0.3, f_1 = 8, G_1 = 2.5, f_2 = 3$ . Figure 7 describes the relationship between the personal utility  $U_L$  and the arrival



**Fig. 8** The relationship between  $U_S$  and  $\mu_1, \gamma$

rate  $\lambda$  and the probability  $\gamma$  of nodes that have finished service are transformed into seeds. When  $\gamma$  is a fixed value,  $U_L$  decreases with the increase of  $\lambda$ . When  $\lambda$  is a fixed value,  $U_L$  increases with the increase of  $\gamma$ . The main reason is that there are relatively few nodes that finished service to convert into seeds, the average delay of the node gradually increases, which leads to decrease in personal utility. When seeds increase continuously, the average delay of the node decreases, and the individual benefits is increasing. The influence of the transformation rate and the arrival rate on individual utility and the zero point is shown in Fig. 7, the equilibrium arrival rate exists. In order to improve the energy efficiency of P2P networks, the transformation rate and service rate can be increased appropriately. In the case of the different arrival rates of the nodes, the numerical results of the personal utility and transformation rate of the nodes are shown in Table 4.

To discuss the social optimal strategy,  $h^*$  is assumed that the optimal probability of response after the requesting node entering the system.  $\lambda^*$  is the actual request rate of the node under the social optimum. Meanwhile,  $\lambda^* = h^* \Phi$ . Then the social utility function is defined as

$$U_S = \lambda(f_1 - G_1 E(W) - f_2) - G_2 W_y.$$

**Table 4** Influence of  $\gamma$  and  $\lambda$  on individual utility  $U_L$

	$U_L$						...
	$\lambda = 0.01$	$\lambda = 2.01$	$\lambda = 4.01$	$\lambda = 6.01$	$\lambda = 8.01$	$\lambda = 10.01$	
$\gamma = 0.60$	0.1243	0.0954	- 0.0339	- 0.2349	- 0.4093	- 0.4562	...
$\gamma = 0.65$	0.2274	0.2062	0.1071	- 0.0575	- 0.2225	- 0.3108	...
$\gamma = 0.70$	0.3159	0.2998	0.2227	0.0886	- 0.0601	- 0.1657	...
$\gamma = 0.75$	0.3925	0.3799	0.3190	0.2100	0.0795	- 0.0299	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Assuming  $c = 15$ ,  $\beta = 0.1$ ,  $\mu_2 = 0.5$ ,  $\xi = 0.1$ ,  $\beta_b = 1$ ,  $\eta = 0.3$ ,  $f_1 = 10$ ,  $G_1 = 2.5$ ,  $f_2 = 3$ ,  $G_2 = 3$ ,  $g_1 = 2$ ,  $g_2 = 3$ . Figure 8 describes the relationship of  $U_S$ ,  $\mu_1$  and  $\gamma$ . When  $\gamma$  is given different values, the increase of  $U_S$  is relatively large at first and then tends to be flat with the increase of  $\mu_1$ . When  $\mu_1 = 5$ ,  $\gamma = 0.8$ , there is a maximum value of social utility  $U_S = 4.5594$  under the conditions of this paper. The main reason is that the transformation rate of nodes increases, the number of nodes that have finished downloading translated into seeds increases, the service speed is the faster. System congestion and failure may be less, and the social benefit is the greater.

## 5 Conclusions

For unstructured distributed P2P networks, the M/M/c queuing model with impatient customers and fault repairable is established, after considering the process of node dynamics, combining the presence of factors such as environmental disturbances and physical damage. A steady-state analysis of the system model is performed by using QBD and the matrix-geometric solution method, and expressions for performance metrics such as average delay of nodes and throughput rate are given. The effects of parameter variations on the performance metrics of the P2P network system are derived through numerical analysis. Throughput is a measure of system transmission effectiveness and social efficiency is a measure of system benefits. The optimal parameters are derived by considering the throughput rate and social benefits, etc. By observing the results of the numerical analysis, it is known that an appropriate increase in the conversion rate can reduce the average delay in the system, avoid system congestion, improve energy saving, and increase personal and social benefits.

**Acknowledgements** This work was supported in part by the National Natural Science Foundation of China under Grant Nos. 61973261, 61872311, Natural Science Foundation of Hebei Province under Grant Nos. A2020203010, A2018203088.

## Declarations

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

## References

- Pan, M., & Lin, Y. (2018). Efficient data dissemination for Wi-Fi peer-to-peer networks by unicasting among Wi-Fi P2P groups. *Wireless Networks*, 24(8), 3063–3081.
- Wang, X. L., & Zhang, J. (2010). Summary of research on P2P key technology. *Application Research of Computers*, 27(3), 801–805 (in Chinese).
- Liu, Y., Mao, J. P., Shen, C. X., & Cui, Y. L. (2008). Performance analysis of P2P network resource service. *Computer Engineering and Applications*, 44(36), 7–10 (in Chinese).
- He, W. H., Liu, H., & He, J. S. (2019). Research on the status and development of P2P network. *Software Engineering*, 22(4), 1–5 (in Chinese).
- Li, D., Chen, Z. G., Liu, H., Vasilakos, A. V., & Pen, Y. (2008). IPBGA: A hybrid P2P based grid architecture by using information pool protocol. *The Journal of Supercomputing*, 49(2), 159–189.
- Li, Z.H., Zhang, T., Huang, Y., & Dai, Y. (2012). Maximizing the bandwidth multiplier effect for hybrid cloud P2P content distribution. In *2012 IEEE 20th International Workshop on Quality of Service*, Coimbra, pp. 1–9.
- Jinlong, E., Li, Z. H., Cui, Y., & Ruan, M. K. (2020). HyCloud: Tweaking hybrid cloud storage services for cost-efficient filesystem hosting. *IEEE/ACM Transactions on Networking*, 28(6), 2629–2642.
- Yang, H., Liu, M. D., & Li, B. C. (2019). A P2P network framework for interactive streaming media. In *2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, Hangzhou, pp. 288–292.
- Zhou, Y. P., Chiu, D. M., & John, C. S. (2016). A simple model for analyzing P2P streaming protocols. In *IEEE International Conference on Network Protocols*, Beijing, pp. 226–235.
- Liu, C., & Lai, C. (2018). A group-based data-driven approach for data synchronization in unstructured mobile P2P systems. *Wireless Networks*, 24(7), 2465–2482.
- Ye, Q., & Zhang, W. H. (2017). Distributed and adaptive medium access control for internet-of-things-enabled mobile networks. *IEEE Internet of Things Journal*, 4(2), 446–460.
- Edmundo, D. S. E. S., Leao, R. M. M., Menasche, D. S., & Towsley, D. (2019). On the scalability of P2P swarming systems. *Computer Networks*, 151, 93–113.
- Zhou, S., & Meng, X. F. (2020). A location and time-aware resource searching scheme in mobile P2P ad hoc networks. *The Journal of Supercomputing*, 76(9), 6809–6833.
- Yang, X. Y., & Veciana, G. (2016). Performance of peer-to-peer networks: Service capacity and role of resource sharing policies. *Performance Evaluation*, 63(3), 175–194.
- Lim, J., Bok, K., & Yoo, J. (2017). An efficient continuous range query processing scheme in mobile P2P networks. *The Journal of Supercomputing*, 76(10), 7805–7819.
- Li, J. L., Shi, W. S., Ye, Q., Zhang, N., Zhang, W. H., & Shen, X. M. (2021). Multi-service function chain embedding with delay-guarantee: A game-theoretical approach. *IEEE Internet of Things Journal*. <https://doi.org/10.1109/JIOT.2021.3051905>.
- Ge, Z. H., Figueiredo, D., Jaiswal, S., & Towsley, D. (2003). Modeling peer-peer file sharing systems. In *22th Annual Joint Conference of the IEEE Computer and Communications Societies*, San Francisco, pp. 2188–2198.
- Li, M., & Tang, J. S. (2009). Performance analysis of Bit Torrent P2P system based on M/MMPP/1. *Journal of Xi'an University of Posts and Telecommunications*, 13(3), 32–37 (in Chinese).
- Yin, B. Q., Guo, D., Huang, J., & Wu, X. M. (2012). Modeling and analysis for the P2P-based media delivery network. *Mathematical and Computer Modeling*, 55(3–4), 1529–1539.
- Haq, A., & Faheem, Y. (2020). A peer-to-peer communication based content distribution protocol for incentive-aware delay tolerant networks. *Wireless Networks*, 26(1), 583–601.
- Ferragut, A., & Paganini, F. (2015). Queueing analysis of Peer-to-Peer swarms: Stationary distributions and their scaling limits. *Performance Evaluation*, 93, 47–62.

22. Hu, C. C., Lai, C. F., Hou, J. G., & Huang, Y. M. (2017). Timely scheduling algorithm for P2P streaming over MANETs. *Computer Networks*, 127, 56–67.
23. Zhao, G. X., Jin, S. F., Ma, C. B., Cao, J. Y., & Xu, L. J. (2019). Analysis of cloud service system with variable number of servers. *Journal of Beijing University of Posts and Telecommunications*, 42(4), 114–120. (in Chinese).
24. Ma, Z. Y., Cao, J., Yu, X. R., & Guo, S. S. (2019). Multiple vacation queuing system with impatient customers and work failure. *Journal of Chongqing Normal University (Natural Science Edition)*, 36(4), 7–13. (in Chinese).
25. Wu, D., Liu, Y., & Keith, W. (2010). Queuing network models for multi-channel P2P live streaming systems. In: *IEEE International Conference on Computer and Information Technology*, pp. 1849–1859.
26. Tian, N. S., & Yue, D. Q. (2002). *The quasi birth and death process and matrix-geometric solution*. Beijing: Science Press (in Chinese).

systems with vacations and performance evaluation models in communication networks



**Fengjiao Liu** received his Bachelor's degree in Mathematics and Applied Mathematics from Hebei Agricultural University, Baoding, China. Currently, he is a postgraduate at the School of Science, Yanshan University, Qinhuangdao, China. His research interests include P2P networks, queueing systems with vacations and performance evaluation models in communication networks

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Qiannan Si** received her Bachelor's degree in Mathematics and Applied Mathematics from Binzhou University, Binzhou, China. Currently, she is a postgraduate at the School of Science, Yanshan University, Qinhuangdao, China. Her research interests include P2P networks, queueing systems with vacations and performance evaluation models in communication networks.



**Rong Wang** received her Bachelor's degree in Mathematics and Applied Mathematics from Hengshui University, Hengshui, China. Currently, she is a postgraduate at the School of Science, Yanshan University, Qinhuangdao, China. Her research interests include P2P networks, queueing systems with vacations and performance evaluation models in communication networks.



**Zhanyou Ma** received his Bachelor's degree in Mathematics from Jilin Normal University, Siping, China, Master's degree in Operations Research and PhD in Management Science and Engineering from Yanshan University, Qinhuangdao, China. Currently, he is a Professor at the School of Science, Yanshan University, Qinhuangdao, China. His research interests include cognitive radio networks, wireless sensor networks, P2P networks, queueing