Performance analysis of a queueing system based on working vacation with repairable fault in the P2P network

Shunzhi Wang
Yanshan University
Zhanyou Ma (mzhy55@ysu.edu.cn)
Yanshan University
Rong Wang
Yanshan University
Wenming Fang
Yanshan University

Research Article

Keywords: P2P network, Synchronous multiple working vacation, Repairable fault, Matrix-geometric solution method, Nash equilibrium

Posted Date: September 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1864515/v2

Performance analysis of a queueing system based on working vacation with repairable fault in the P2P network

Shunzhi Wang[1] · **Zhanyou Ma**[1] · **Wang Rong**[1] · **Wenming Fang**[1]

Received: date / Accepted: date

Abstract Nowadays, P2P (Peer-to-Peer) absolutely dominates the Internet as a traffic in content distribution and sharing service, the issue of energy consumption has emerged as a hot issue. This paper combines the basic theoretical knowledge of queueing theory with the currently more popular hybrid P2P network structure model. In order to address the problem of wasted energy consumption generated by unnecessary online behavior of nodes, the working sleep mechanism for the serving nodes is introduced, and the synchronous multiple working vacation strategy is also adopted to decrease the system energy consumption. In the meantime, negative customer and feedback strategies are introduced. Then a queueing model with working vacation and repairable fault is built, which used to simulate the resource transmission process of P2P and to study the performance of P2P sharing system. The steady-state probability vector is derived by employing quasi-birth-and-death process (QBD) and matrix-geometric solution method. And the expressions for system performance indicators, such as the average delay, are obtained by employing Gauss-Seidel iteration method. Through numerical experiments, the performance indicators of the P2P resource sharing system are studied, which provides a theoretical basis to decrease online energy consumption of nodes in the P2P network system. And through employing Nash equilibrium and social optimal strategy, the value of the social maximum benefit under the social optimization is obtained.

Keywords P2P network · Synchronous multiple working vacation · Repairable fault · Matrix-geometric solution method · Nash equilibrium

Shunzhi Wang
wdsysu@163.com

✉ Zhanyou Ma
mzhy55@ysu.edu.cn

Rong Wang
wangrong9327@163.com

Wenming Fang
fwm@ysu.edu.cn

1 School of Science, Yanshan University, Qinhuangdao 066004, China

1 Introduction

P2P (Peer-to-Peer) is a new network technology which enables people to communicate with each other directly over the Internet. In the P2P network, thousands of interconnected computers are in a peer-to-peer position. Each peer acts as both a server and a client, that is, providing resource services and sending service requests. The Internet application system structure and user behavior patterns have shifted from a simple centralized C/S structure to a variety of distributed P2P structures, the entire network no longer depends on a dedicated centralized server. It transforms communication on the network from centralized resource sharing to direct resource sharing among users, truly eliminating the middle links. With the increasing number of P2P applications over the past 20 years, a large number of users are uploading and downloading files through P2P services on a daily basis. In the Internet where P2P business traffic is absolutely dominant, the problem of energy consumption is gradually emerging as a vital obstacle to the rapid development of the Internet [1–3].

In recent years, there are many researches on the P2P network. Zhang et al. [4] described a multi-sender node streaming data distribution strategy based on a hybrid P2P and C/S architecture, then gave and compared the performance of related scheduling algorithms. Amit et al. [5] proposed a content distribution method for P2P applications based on content distribution and designed a strategy for balanced content replication in the P2P network. He et al. [6] introduced the powerful advantages of P2P network in distributed computing, collaborative work, multimedia, and the applications of P2P network in file sharing and distributed data storage, and analyzed the necessity of building secure peer-to-peer network. Yin et al. [7] established a P2P-based media distribution network that introduced a peer selection policy and a broadband allocation policy to improve the performance and efficiency of the system in the media content sharing process. Huo et al. [8] proposed a new dual-rate transmission energy conservation strategy in a study of cognitive radio network, established a payoff function based on the payoff expenditure structure, and studied the Nash equilibrium strategy and the socially optimal strategy.

Nowadays, queueing theory studies are also increasingly applied to the P2P network system.Ferragut and Paganini [9] analyzed the dynamic properties of P2P file exchange swarms from the queueing perspective, where the service rate received by nodes depends on the server and varied with the number of nodes. Yu et al. [10] used the repairable M/M/2 queueing model to study wireless communication network, and proposed a strategy to decrease energy consumption. Si et al. [11] developed an M/M/c queueing model with impatient customers and repairable fault for unstructured distributed P2P network, taking into account the node dynamic process and concluded that an appropriate increase in node conversion rate can avoid system congestion and improve energy savings. Singha and Singh [12] presented a new incentive mechanism to strengthen cooperation in wireless environments. Using Nash equilibrium, the simulation verified the effectiveness of these mechanisms in enhancing cooperation among users. Liu et al. [13] discussed the equilibrium strategy for the mixed overlay/underlay spectrum sharing approach in cognitive radio networks, divided into visible and invisible. Based on the social optimum and spectrum revenue maximization, a pricing mechanism was derived for cognitive users. Jin et al. [14] gave the expressions for performance indicator such as the system failure rate and the average data access time in the context of hybrid P2P network based on a delay repair strategy for fixed nodes. Xue et al. [15] established a system model based on hybrid queueing network and verified the effectiveness of the strategy and system model. Zhang and Yin [16] considered a P2P network system in which heterogeneous servers in processor sharing rules are operated, first using limited capacity M/M/1/N processor sharing, then the performance of the P2P network system is analyzed, and finally the computational results are verified by arithmetic examples and the performance of individual server is analyzed.

Due to the enormous number of nodes involved in the P2P network, considering the system performance and quality of service, the total energy consumption requires to be an important problem to be addressed. One of the strategies to decrease energy consumption is allowing the peer to cycle between normal mode and sleep mode. In recent years, there were several papers that also studied the P2P network from this aspect.Ma et al. [17] modeled the hybrid P2P network using a queueing model to address the ”free-riding” phenomenon in the P2P network, distinguished between requesting nodes of different reputation and introduced a work/sleep mechanism for the serving nodes. Trunfio [18] proposed a sleep-wake method, by allowing peer nodes to cycle between normal and sleep modes to save energy. In normal mode, the time a peer node passes depends on the number of files it provides. Jin et al. [19] proposed a task scheduling strategy with sleep delay timer and wake-up threshold. According to the random behavior of the strategy task, a synchronous vacation queueing model with vacation delay and N-policy is built. Based on this, the average stay time of the system at steady state and the energy saving level of the system are derived. Finally, the effect of the system parameters on the performance indicators is studied by numerical experiments.

The majority of papers mentioned above studied the P2P network in terms of dynamic changes of nodes, performance of network system, etc. In this paper, based on the actual situation that system nodes in the P2P network may have some faults due to overloaded tasks. This paper introduces node faults that can be repaired, and also considers the energy consumption generated by unnecessary online behavior when the nodes in the system are idle. The normal busy period and the working sleep period are abstracted as the online state and offline state respectively. The contributions of this paper are threefold.

(1) By using the M/M/c queueing model, the hybrid P2P network model is built. The synchronous working vacation strategy is employed to decrease system energy consumption, and then the energy consumption of the serving nodes in various state is quantified. In the meantime, the negative customer and feedback strategies are applied to the hybrid P2P network in order to make the model more compatible with the actual conditions of the hybrid P2P network. Compared with the traditional mode, this method can significantly reduce the energy consumption of P2P network.
(2) The expressions for performance indicators such as system energy consumption, are given by employing Gauss-Seidel iteration method. In order to improve the efficiency of resource request data in the P2P network, the system performance is studied by numerical experiments.
(3) The Nash equilibrium between the arrival rate and the individual benefit of the requesting nodes is studied to avoid the increase in system energy consumption due to an overload of the requesting nodes in the P2P network system.

The rest of this paper is organized as follows. The schematic diagram and the operating mechanism of the hybrid P2P network is introduced in Section 2. In Section 3, through using a three-dimensional Markov chain, the steady-state probability vector is derived. In Section 4, the performance indicators of the P2P resource sharing system are studied by numerical experiments, and the individual benefit and the optimization of the social benefit are analyzed by Nash equilibrium strategy and social optimal strategy. Conclusions are introduced in Section 5.

2 Hybrid P2P network and modeling

2.1 Hybrid P2P network

Hybrid P2P network presents a centralized architecture locally, which consists of different clusters. Each cluster contains a super node (SN) and several ordinary nodes (ON) with the same service resources, the details of hybrid P2P network can refer to [20]. The ordinary node stores its own resource information and area information, as a serving node receives resource requests from the requesting nodes, establishes peer-to-peer connections and then transmits the resource. The super node is responsible for responding to resource requests, transmitting resources and collecting resource information of ordinary nodes within their own region. In the meantime, the super node maintains and updates the information of the ordinary nodes in the region and establishes connections with the super nodes in other regions. The schematic diagram of the hybrid P2P network is shown in Figure 1. Specifically, when the P2P network system receives requests, the requesting nodes first query within the super node region. If the requested resource exists in the region, the requesting nodes receive the resource transmission in the region. If the requested resource does not exist in the region, the requesting nodes query across the region. By the resource transmission of the super nodes, the requesting nodes find information about the requested resource outside the region. The requesting nodes enter the resource transmission stage after querying the node with the requested resource. A requesting node that completes a resource transmission may continue to request additional resources, or it may simply leave the P2P network after completing the service.

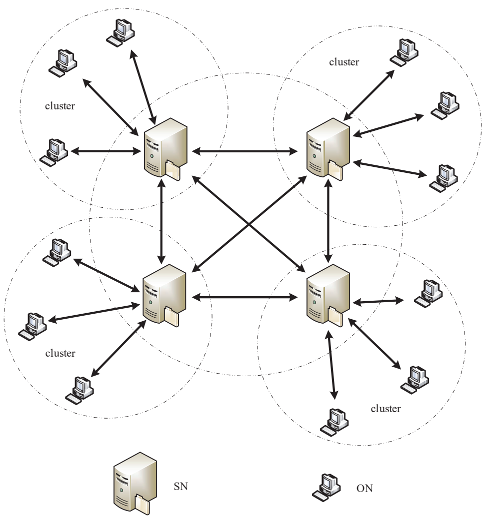

Fig. 1 Schematic diagram of hybrid P2P network.

2.2 Model description

In the P2P network, there are two types of nodes, the nodes that require to receive resource transmission are called the requesting nodes, and the nodes that provide resource request services are called the serving nodes. The so-called working sleep mechanism is to allow idle serving nodes to enter a low-energy working state to avoid unnecessary resource consumption. The resource transmission process of the P2P network system is abstracted as the service process, the working sleep process of the serving node is abstracted as the working vacation period, the normal resource transmission process is abstracted as the normal busy period. A new queueing system model with negative customers, feedback, working vacation and repairable fault is built in the P2P network. The model description and system parameters are as follows:

(1) The requesting nodes which need to receive resource transmission are regarded as positive customers. The arrival of the requesting nodes follow the Poisson process with the parameter λ_1. There is also a type of external interference signal in the system that is regarded as the negative customer. It is usually considered as a virus, a work disappearance signal, an operation error or system disaster, etc. The arrival of the external interference signal follows the Poisson process with the parameter λ_2. Under the Removal the Customers at the End (RCE) offset strategy, an arrival of external interference signal can offset a requesting node at the end of the queue in the infinite buffer. If there is no requesting node in the buffer, the external interference signal disappears automatically and is not served.

(2) In order to reduce energy consumption, the working sleep strategy of the serving nodes is being introduced into the system. When the serving nodes are in the normal busy period, the time consumed by each requesting node to complete the service follows the exponential distribution with parameter μ_1. When there is no requesting node in the system, the serving nodes enter a working sleep period with random length V. The working sleep time follows the exponential distribution with parameter ξ. During the working sleep period, the serving nodes remain at a lower service rate, the time consumed by each requesting node to complete the service follows the exponential distribution with parameter $\mu_2 (\mu_2 < \mu_1)$. If there is still no requesting node in the system after a working sleep period which has finished, the serving nodes continue an independent and identically distributed (i.i.d.) working sleep period. If there are requesting nodes in the system after the working sleep period has finished, the system enters into a normal busy period. After the normal busy period, the system again enters a working sleep period with random length V, then repeats the above process.

(3) If some requesting nodes that have completed the service want to obtain other resource, they are regarded as new customers and wait at the end of line after re-entering the system. Any requesting node re-enters the system with the probability $p(0 \leq p \leq 1)$, which is known as the feedback probability. If the requesting nodes that have completed service are idle, it will leave system with probability $\overline{p}(\overline{p} = 1 - p)$.

(4) The serving node may fault and the fault is repairable. All serving nodes may be overloaded with requests which causes faults such as the congestion and lag. The process of fault follows the Poisson process with parameter α. If the serving node faults, it can be repaired immediately, and the faulted node can only be repaired by one repairer. The repair time follows the exponential distribution with parameter β. Assume that the number of repairers is infinite.

(5) The service sequence is First-Come-First-Served (FCFS). It is assumed that the arrival interval, the service time, the working sleep time, the fault time and the repair time are mutually independent. The operation mechanism of the hybrid P2P network system is shown in Figure 2.

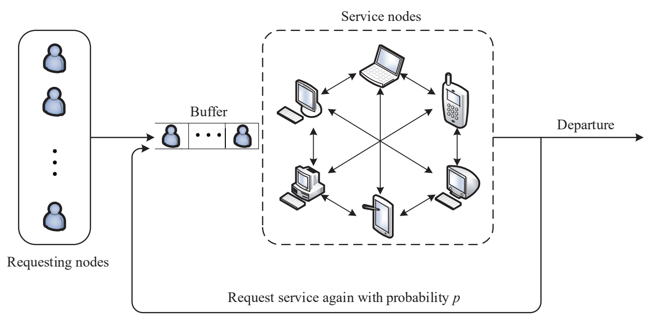

Fig. 2 The operation mechanism of the hybrid P2P network.

3 Modeling analysis of P2P system

3.1 State transition rate matrix

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

(1) $J(t)$=0 indicates that the serving nodes are in the working sleep period or the total fault state at time t,
(2) $J(t)$=1 indicates that the serving nodes are in the normal busy period at time t.

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

$$\Omega = \Omega_1 \cup \Omega_2 \tag{1}$$

where $\Omega_1 = \{(i,l,j), i \geq 0, 0 \leq l \leq c, j = 0\}$, $\Omega_2 = \{(i,l,j), i \geq 1, 0 \leq l \leq c-1, j = 1\}$, the state set $(0,0,0),(0,1,0),(0,2,0),\cdots,(0,c-1,0),(0,c,0)$ is called level 0, and the state set $(i,0,0),(i,0,1),(i,1,0),(i,1,1),\cdots,(i,c-1,0),(i,c-1,1),(i,c,0)$ is called level i $(i \geq 1)$. When all serving nodes fail, the P2P network system enters the total fault state ($J(t) = 0$). Therefore, the state transition of the queueing model is shown in Figure 3.

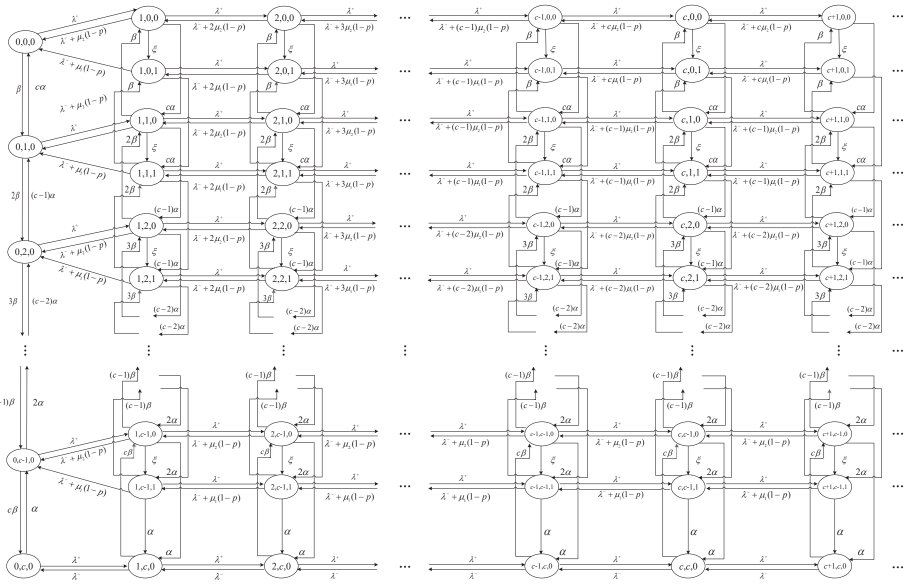

Fig. 3 State transition diagram of queueing model.

Arranging the states in lexicographic order, the state transition rate matrix of the system can be written as follows:

$$
\boldsymbol{Q} = \begin{bmatrix}
\boldsymbol{A}_0 & \boldsymbol{C}_0 & & & & & \\
\boldsymbol{B}_1 & \boldsymbol{A}_1 & \boldsymbol{C} & & & & \\
& \boldsymbol{B}_2 & \boldsymbol{A}_2 & \boldsymbol{C} & & & \\
& & \ddots & \ddots & \ddots & & \\
& & & \boldsymbol{B}_{c-1} & \boldsymbol{A}_{c-1} & \boldsymbol{C} & \\
& & & & \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{C} \\
& & & & & \ddots & \ddots & \ddots
\end{bmatrix} \tag{2}
$$

where $\boldsymbol{A}_0, \boldsymbol{C}_0, \boldsymbol{C}, \boldsymbol{A}_i \ (1 \le i \le c), \boldsymbol{B}_i \ (1 \le i \le c)$ are the state transition rate between corresponding levels, and let $\boldsymbol{A} = \boldsymbol{A}_c, \boldsymbol{B} = \boldsymbol{B}_c$. The specific matrix is as follows.
$(c+1) \times (c+1)$ dimensional matrix $\boldsymbol{A}_0$ is given as follows:

$$
\boldsymbol{A}_0 = \begin{bmatrix}
\eta_0 & c\alpha & & & & & \\
\beta & \eta_1 & (c-1)\alpha & & & & \\
& 2\beta & \eta_2 & (c-2)\alpha & & & \\
& & 3\beta & \eta_3 & (c-3)\alpha & & \\
& & & \ddots & \ddots & \ddots & \\
& & & & (c-1)\beta & \eta_{c-1} & \alpha \\
& & & & & c\beta & \eta_c
\end{bmatrix}
$$

where $\eta_i = -\lambda_1 - (c-i)\alpha - i\beta, \quad 1 \le i \le c$.

$(c+1)\times(2c+1)$ dimensional matrix $\boldsymbol{C}_0$ is given as follows:

$$\boldsymbol{C}_0 = \begin{bmatrix} \lambda_1 & 0 & & & & & & \\ & & \lambda_1 & 0 & & & & \\ & & & & \lambda_1 & 0 & & \\ & & & & & \ddots & \ddots & \\ & & & & & & \lambda_1 & 0 \\ & & & & & & & \lambda_1 \end{bmatrix}.$$

$(2c+1)\times(c+1)$ dimensional matrix $\boldsymbol{B}_1$ is given as follows:

$$\boldsymbol{B}_1 = \begin{bmatrix} \omega & 0 & & & \\ \delta & 0 & & & \\ & \omega & 0 & & \\ & \delta & 0 & & \\ & & \ddots & \ddots & \\ & & & \omega & 0 \\ & & & \delta & 0 \\ & & & & \lambda_2 \end{bmatrix}$$

where $\omega = \lambda_2 + \mu_2(1-p), \quad \delta = \lambda_2 + \mu_1(1-p)$.

$(2c+1)\times(2c+1)$ dimensional matrix $\boldsymbol{C}$ is given as follows:

$$\boldsymbol{C} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_1 \end{bmatrix}.$$

In order for the convenience of writing, when $1 \le k \le c$, the symbols are defined as follows:

$$a_{k,i} = \begin{cases} -\lambda_1 - \lambda_2 - k\mu_1(1-p) - (c-i)\alpha - i\beta, & 0 \le i \le c-k, \\ -\lambda_1 - \lambda_2 - (c-i)\mu_1(1-p) - (c-i)\alpha - i\beta, & c-k < i \le c, \end{cases}$$

$$a^*_{k,i} = \begin{cases} -\lambda_1 - \lambda_2 - k\mu_2(1-p) - (c-i)\alpha - i\beta - \xi, & 0 \le i \le c-k, \\ -\lambda_1 - \lambda_2 - (c-i)\mu_2(1-p) - (c-i)\alpha - i\beta - \xi, & c-k < i \le c. \end{cases}$$

Then $(2c+1)\times(2c+1)$ dimensional matrix $\boldsymbol{A}_k$ is given as follows:

$$\boldsymbol{A}_k = \begin{bmatrix} a^*_{k,0} & \xi & c\alpha & & & & \\ 0 & a_{k,0} & 0 & & & & \\ \beta & 0 & a^*_{k,1} & \xi & (c-1)\alpha & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & (c-1)\beta & 0 & a^*_{k,c-1} & \xi & \alpha \\ & & & (c-1)\beta & 0 & a_{k,c-1} & \alpha \\ & & & & c\beta & 0 & a^*_{k,c} \end{bmatrix}.$$

In order for the convenience of writing, when $2 \leq k \leq c$, the symbols are defined as follows:

$$b_{k,i} = \begin{cases} \lambda_2 + k\mu_1(1-p), & 0 \leq i \leq c-k, \\ \lambda_2 + (c-i)\mu_1(1-p), & c-k < i \leq c, \end{cases}$$

$$b^*_{k,i} = \begin{cases} \lambda_2 + k\mu_2(1-p), & 0 \leq i \leq c-k, \\ \lambda_2 + (c-i)\mu_2(1-p), & c-k < i \leq c. \end{cases}$$

Then $(2c+1) \times (2c+1)$ dimensional matrix $\boldsymbol{B}_k$ is given as follows:

$$\boldsymbol{B}_k = \begin{bmatrix} b^*_{k,0} & & & & & & & \\ & b_{k,0} & & & & & & \\ & & b^*_{k,1} & & & & & \\ & & & b_{k,1} & & & & \\ & & & & \ddots & & & \\ & & & & & b^*_{k,c-1} & & \\ & & & & & & b_{k,c-1} & \\ & & & & & & & b^*_{k,c} \end{bmatrix}.$$

3.2 Steady-state analysis

The structure of the matrix $\boldsymbol{Q}$ indicates that the Markov process $\{(L(t),Y(t),J(t)),t \geq 0\}$ is a quasi-birth-and-death process (QBD). When the Markov process is positive recurrent, the steady-state distribution is defined as follows:

$$\pi_{i,l,j} = \lim_{t\to\infty} P\{X(t) = i, Y(t) = l, J(t) = j\}, \quad (i,l,j) \in \Omega,$$

$$\pi_0 = (\pi_{0,0,0}, \pi_{0,1,0}, \pi_{0,2,0}, \cdots, \pi_{0,c-1,0}, \pi_{0,c,0}),$$

$$\pi_i = (\pi_{i,0,0}, \pi_{i,0,1}, \pi_{i,1,0}, \pi_{i,1,1} \cdots, \pi_{i,c-1,1}, \pi_{i,c,0}), \quad i \geq 1,$$

$$\Pi = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \cdots) \tag{3}$$

The sufficient and necessary conditions that QBD $\{(L(t),Y(t),J(t)),t \geq 0\}$ is positive recurrent is that the matrix qua-
dratic equation:

$$\boldsymbol{R}^2\boldsymbol{B} + \boldsymbol{R}\boldsymbol{A} + \boldsymbol{C} = \boldsymbol{0} \tag{4}$$

has a minimum non-negative solution, and the spectral radius $SP(\boldsymbol{R}) < 1$. Therefore, $2c^2 + 2c + 1$ dimensional stochastic matrix:

$$B[\boldsymbol{R}] = \begin{bmatrix} \boldsymbol{A}_0 & \boldsymbol{C}_0 & & & & \\ \boldsymbol{B}_1 & \boldsymbol{A}_1 & \boldsymbol{C} & & & \\ & \boldsymbol{B}_2 & \boldsymbol{A}_2 & \boldsymbol{C} & & \\ & & \ddots & \ddots & \ddots & \\ & & & \boldsymbol{B}_{c-1} & \boldsymbol{A}_{c-1} & \boldsymbol{C} \\ & & & & \boldsymbol{B} & \boldsymbol{B}\boldsymbol{R} + \boldsymbol{A} \end{bmatrix} \tag{5}$$

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

$$\begin{cases} (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \cdots, \boldsymbol{\pi}_c)B[\boldsymbol{R}] = 0, \\ \boldsymbol{\pi}_0 \boldsymbol{e}_0 + \sum_{i=1}^{c-1} \boldsymbol{\pi}_i \boldsymbol{e} + \boldsymbol{\pi}_c (\boldsymbol{I} - \boldsymbol{R})^{-1} \boldsymbol{e} = 1, \\ \boldsymbol{\pi}_i = \boldsymbol{\pi}_c \boldsymbol{R}^{i-c}, \quad i \geq c \end{cases} \tag{6}$$

where $\boldsymbol{e}_0$ is a $c+1$ dimensional column vector which all elements are 1, $\boldsymbol{e}$ is a $2c+1$ dimensional column vector which all elements are 1, $\boldsymbol{I}$ is a $2c+1$ dimensional unit matrix.

The analytical process of the above conclusion employs matrix-geometric solution method, and the specific proof process can be referred to [21]. Due to the relative complexity of the matrix $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, it is difficult to obtain an explicit expression for the minimum non-negative solution $\boldsymbol{R}$ of the above matrix equation $\boldsymbol{R}^2\boldsymbol{B} + \boldsymbol{R}\boldsymbol{A} + \boldsymbol{C} = \boldsymbol{0}$. Therefore, the Gauss-Seidel iterative method is employed to address the difficulties mentioned above. The numerical results of $\pi_{i,l,j}$ are obtained. The specific algorithm is shown in Table 1.

Table 1 Iterative algorithm of rate matrix $\boldsymbol{R}$.

Step	Operation
Step1	Error accuracy $\varepsilon(0 < \varepsilon < 1)$, Initialize $c, \lambda_1, \lambda_2, \alpha, \beta, \xi, p, \mu_1, \mu_2$, Rate matrix $\boldsymbol{R}_0 = \boldsymbol{0}$, $\boldsymbol{R} = \boldsymbol{R}_0$
Step2	Input $\boldsymbol{A}$,$\boldsymbol{B}$,$\boldsymbol{C}$,$n = 1$
Step3	$\boldsymbol{R}_n = \boldsymbol{R}$ $n = n+1$ $\boldsymbol{R}_n = -(\boldsymbol{R}_{n-1}{}^2\boldsymbol{B} + \boldsymbol{C})\boldsymbol{A}^{-1}$
Step4	If $\|\boldsymbol{R}_n - \boldsymbol{R}_{n-1}\|_\infty > \varepsilon$ Go to Step3; else Go to Step5;
Step5	$\boldsymbol{R} = \boldsymbol{R}_n$

3.3 Performance indicators

According to the above analysis, the performance indicators of the P2P network system are obtained as follows:

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

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

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

$$E(W) = \frac{1}{\lambda_1} \sum_{i=c+1}^{\infty} (i-c) \sum_{l=0}^{c-1} \sum_{j=0}^{1} \pi_{i,l,j} + \frac{1}{\lambda_1} \sum_{i=c+1}^{\infty} (i-c)\pi_{i,c,0}.$$

(3) The probability that the serving node is in the working sleep period is given by:

$$P_1 = \sum_{i=0}^{\infty}\sum_{l=0}^{c-1}\pi_{i,l,0}.$$

(4) The probability that the serving node is in the normal busy period is given by:

$$P_2 = \sum_{i=1}^{\infty}\sum_{l=0}^{c-1}\pi_{i,l,1}.$$

(5) The probability that all serving nodes are in the total fault state is given by:

$$P_3 = \sum_{i=0}^{\infty}\pi_{i,c,0}.$$

(6) The average number of faulted serving nodes in the system is given by:

$$E(Z) = \sum_{l=1}^{c} l\pi_{0,l,0} + \sum_{i=1}^{\infty}\sum_{l=1}^{c-1} l(\pi_{i,l,0}+\pi_{i,l,1}) + \sum_{i=1}^{\infty} c\pi_{i,c,0}.$$

(7) The total energy consumption of the P2P network system includes the energy consumption of the serving nodes in the normal busy period and the working sleep period. Assuming that g_1 is the energy consumption of a single serving node in the working sleep period, g_2 is the energy consumption of a single serving node in the normal busy period, and the faulted serving node has no energy consumption. Therefore, the average energy consumption G_1 of the serving node in the working sleep period is given by:

$$G_1 = g_1 E(L_1) = g_1 \sum_{i=0}^{\infty}\sum_{l=0}^{c-1}(c-l)\pi_{i,l,0}.$$

(8) The average energy consumption G_2 of the serving node in the normal busy period is given by:

$$G_2 = g_2 E(L_2) = g_2 \sum_{i=1}^{\infty}\sum_{l=0}^{c-1}(c-l)\pi_{i,l,1}.$$

(9) The total energy consumption of the P2P network system is given by:

$$G_y = g_1 \sum_{i=0}^{\infty}\sum_{l=0}^{c-1}(c-l)\pi_{i,l,0} + g_2 \sum_{i=0}^{\infty}\sum_{l=0}^{c-1}(c-l)\pi_{i,l,1}.$$

4 Numerical experiments of the P2P network

The minimum non-negative solution of the matrix quadratic equation (4) is obtained by the Gauss-Seidel iterative method. Then, the numerical results of $\pi_{i,l,j}$ are obtained by solving the system of steady-state distribution equations, and the performance indicators of the P2P system are obtained. Programming by computer software, the relationship graph that performance indicators changed with parameters is obtained, and then analyzing the effect of parameter variations on the performance indicators, which provides the basis for P2P network system to decrease energy consumption.

4.1 Effect of parameter on performance indicators

By numerical experiments, the effect of the system parameters on the performance indicators of the system is analyzed. Assuming that $c = 10, \lambda_2 = 1, \mu_2 = 1, \alpha = 0.8, \beta = 1.0$, Fig. 4 reflects the effect of the feedback probability p, the arrival rate λ_1, the working sleep parameter ξ and the service rate μ_1 on the average queue length $E(L)$. When the other three parameters are fixed, $E(L)$ increases with the increase of p, $E(L)$ increases with the increase of λ_1, $E(L)$ decreases with the increase of ξ, $E(L)$ decreases with the increase of μ_1. The main reason is that when the feedback probability increases, it means that the number of the requesting nodes re-entering the system to continue queueing for service increases, thus the average queue length increases. When the arrival rate increases, it means that the number of the requesting nodes arriving at the system per unit time increases, thus the average queue length increases. When the working sleep parameter increases, it means that the time when they are in the normal busy period increases. Because of the higher service rate in the normal busy period, thus the average queue length decreases. When the service rate increases, it means that the time consumed by the requesting node to complete the service decreases, thus the average queue length decreases. It can be seen that in order to reduce the average queue length, it is required to decrease the feedback probability and the arrival rate, and increase the working sleep parameter and the service rate to some extent.

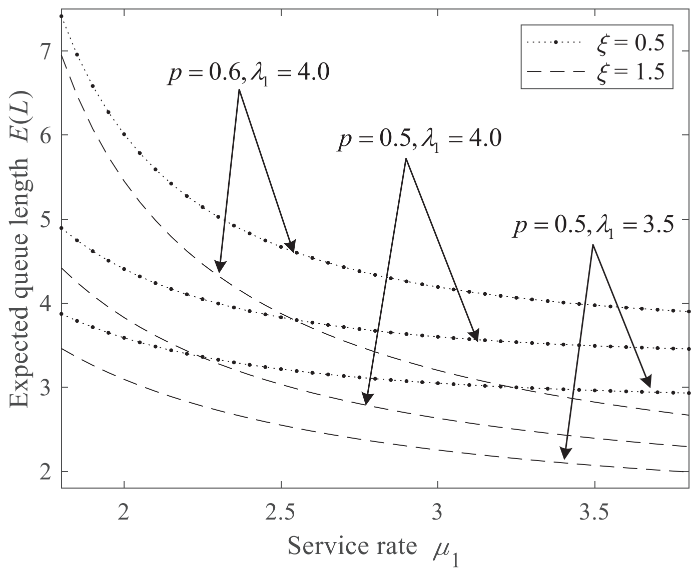

Fig. 4 The relationship between $E(L)$ and p, λ_1, ξ, μ_1.

Assuming that $c = 10, \lambda_2 = 1, \mu_1 = 3, \mu_2 = 1, p = 0.6, \alpha = 0.8, \beta = 1.0$, Fig. 5 reflects the effect of the working sleep parameter ξ, the service rate μ_1, the fault rate α and the arrival rate λ_1 on the average delay $E(W)$. When the other three parameters are fixed, $E(W)$ decreases with the increase of ξ, $E(W)$ decreases with the increase of μ_1, $E(W)$ increases with the increase of α, $E(W)$ increases with the increase of λ_1. The main reason is that when the working sleep parameter increases, it means that the time when they are in the normal busy period increases. Because of the higher service rate in the normal busy period, thus the average delay decreases. When the service rate increases, it means that the time consumed by the requesting node to complete the service decreases, thus the average delay decreases. When the fault rate increases, the number of faulted serving node in the system increases, it means that the number of the available serving nodes decreases with a certain number of requesting nodes, thus the average delay increases. When the arrival rate increases, it means that the number of the requesting nodes arriving at the system per unit time increases, thus the average delay increases. It can be seen that in order to reduce the average delay, it is required to decrease the fault rate and the arrival rate, and increase the working sleep parameter and the service rate to some extent.

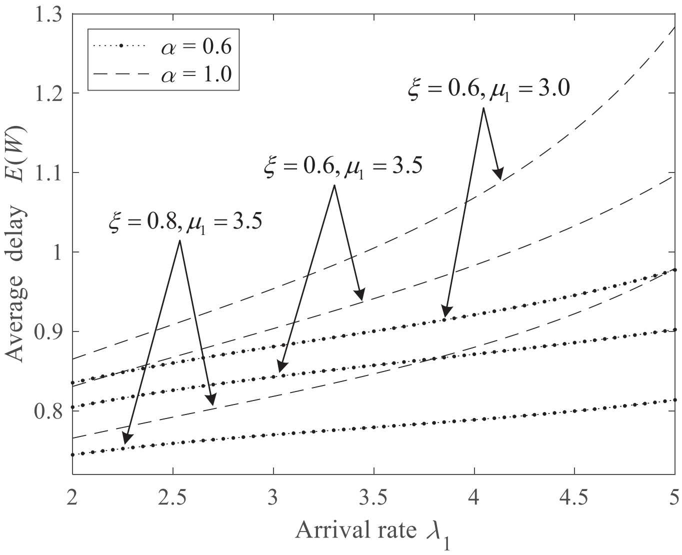

Fig. 5 The relationship between $E(W)$ and $\xi, \mu_1, \alpha, \lambda_1$.

Assuming that $c = 10, \lambda_1 = 2, \lambda_2 = 1, \mu_1 = 3, \mu_2 = 1, p = 0.6, \xi = 0.8$, Tab. 2 reflects the effect of the fault rate α and the repair rate β on the average number $E(Z)$ of the faulted serving nodes in the system. When α is a fixed value, $E(Z)$ decreases with the increase of β. When β is a fixed value, $E(Z)$ increases with the increase of α. The main reason is that when the fault rate increases, the serving nodes in the system are more probable to fail, thus the average number of faulted serving nodes increases. When the repair rate increases, the time to repair the faulted node decreases, thus the average number of the faulted serving nodes decreases.

Assuming that c=10,$\lambda_2 = 1.0, \mu_2 = 1.0, p = 0.2, \alpha = 0.8, \beta = 0.8, g_1 = 5.0$, Fig. 6 reflects the effect of the working sleep parameter ξ, the service rate μ_1 and the arrival rate λ_1 on the average energy consumption G_1 in the working sleep period. When λ_1 and μ_1 are fixed value, G_1 increase with the increase of ξ. When λ_1 and ξ are fixed value, G_1 increase with the increase of μ_1. When μ_1 and ξ are fixed value, G_1 decreases with the increase of λ_1. The main reason is that when the working sleep parameter increases, the average working sleep time of the serving nodes decreases, thus the average energy consumption in the working sleep period decreases. When the arrival rate increases, the number of the

Table 2 The relationship between $E(Z)$ and α, β.

	$E(Z)$					
	α=0.5	0.7	0.9	1.1	1.3	1.5
$\beta = 0.5$	5.000	5.833	6.429	6.875	7.222	7.500
$\beta = 0.9$	3.571	4.375	5.000	5.500	5.909	6.250
$\beta = 1.4$	2.632	3.333	3.913	4.400	4.815	5.172
$\beta = 1.8$	2.174	2.800	3.333	3.793	4.194	4.546

requesting nodes increases, and the time that the serving nodes are in the normal busy period increases, and they are less likely to enter the working sleep period, thus the average energy consumption in the working sleep period decreases. When the service rate increases, the rate at which the system enters the working sleep period after all services are completed increases, the average time in the working sleep period increases, thus the average energy consumption in the working sleep period increases. It can be seen that in order to decrease the average energy consumption in the working sleep period, it is required to decrease the service rate and increase the working sleep parameter and the arrival rate to some extent.

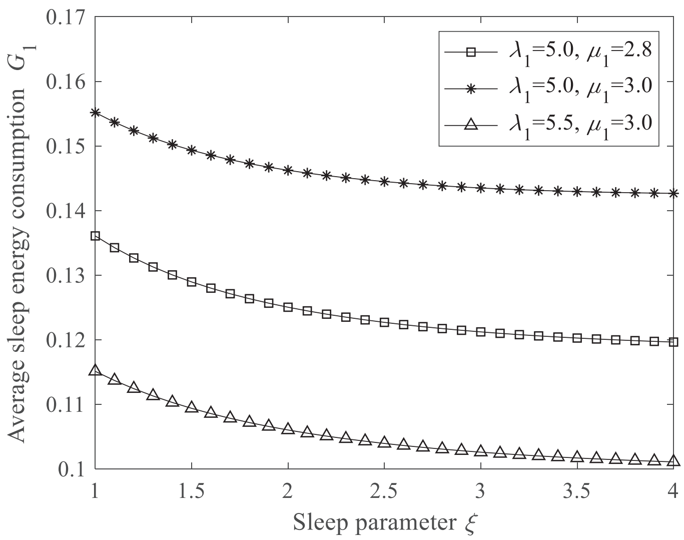

Fig. 6 The relationship between G_1 and ξ, λ_1, μ_1.

Assuming that c=15, $\lambda_2 = 1, \mu_2 = 5, p = 0.2, \alpha = 1.0, \beta = 2.0, g_1 = 3, g_2 = 10, \xi = 0.8$, Fig. 7 reflects the effect of the arrival rate λ_1 and the service rate μ_1 on the energy consumption G_1 in the working sleep period and the energy consumption G_2 in the normal busy period. When μ_1 is a fixed value, G_1 decreases with the increase of λ_1, G_2 increases with the increase of λ_1. When λ_1 is a fixed value, G_1 increases with the increase of μ_1, G_2 decreases with the increase of μ_1. The main reason is that when the arrival rate increases, the average time when the system is in the working sleep period decreases while the average time when it is in the normal busy period increases, thus the energy consumption in the working sleep period decreases while the energy consumption in the normal busy period increases. When the service rate increases, the rate at which the system enters the working sleep period after completing all services increases, thus the energy consumption in the working sleep period increases while the energy consumption in the normal busy period decreases. It can be seen that in order to decrease the energy consumption in the normal busy period, it is required to increase the service rate, and decrease the arrival rate to some extent.

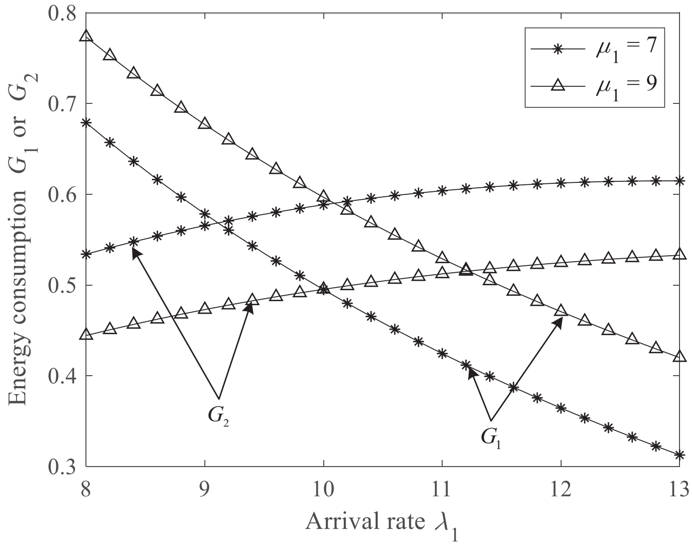

Fig. 7 The relationship between G_1, G_2 and λ_1, μ_1.

Assuming that $c = 15, \lambda_2 = 1, \mu_2 = 2, p = 0.2, \beta = 1.5, g_1 = 2, g_2 = 10$, Fig. 8 reflects the effect of the arrival rate λ_1, the fault rate α, the working sleep parameter ξ and the service rate μ_1 on the total energy consumption G_y. When the other three parameters are fixed, G_y decreases with the increase of μ_1, G_y increases with the increase of λ_1, G_y increases with the increase of ξ, G_y increases with the increase of α. The main reason is that the energy consumption of a single serving node in the normal busy period is higher than that in the working sleep period. When the service rate increases, it means that the system is more easily able to enter the working sleep period after serving all the requesting nodes, and the time in the normal busy period decreases, thus the total energy consumption decreases. When the working sleep parameter increases, it means that the time when they are in the normal busy period increases, thus the total energy consumption increases. When the arrival rate increases, it means that the number of the requesting nodes arriving at the system per unit time increases, the time when the serving nodes are in the normal busy period increases, thus the total energy consumption increases. When the fault rate increases, the number of the faulted serving node in the system increases, it means that the number of available serving nodes decreases with a certain number of the requesting nodes, and the time when the serving nodes are in the normal busy period increases, thus the total energy consumption increases. It can be seen that in order to decrease the total energy consumption, it is required to decrease the fault rate, the arrival rate and the working sleep parameter, and increase the service rate to some extent.

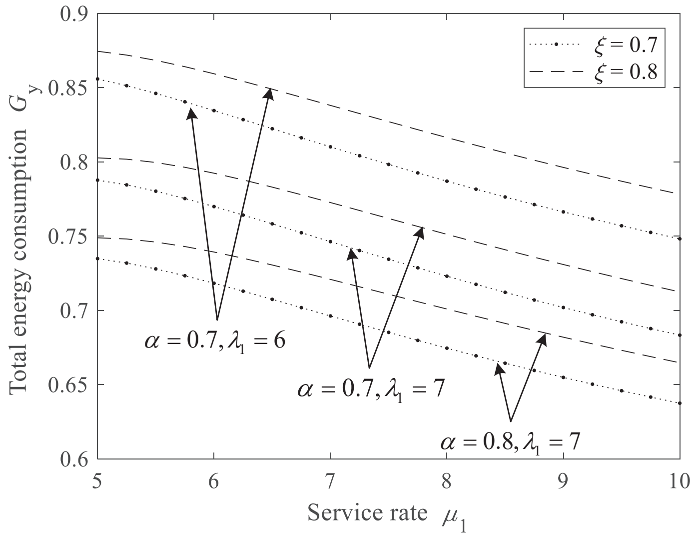

Fig. 8 The relationship between G_y and $\lambda_1, \alpha, \xi, \mu_1$.

4.2 Nash equilibrium strategy

In hybrid P2P network, to avoid an overwhelming number of requesting nodes to send the non-essential service requests to the system which increase the unnecessary energy consumption of the system, this section studies the Nash equilibrium between the arrival rate and the individual benefit of a single requesting node. Assuming that h_1 indicates the benefit obtained after the service is completed, C_1 indicates that the average delay cost of the node unit time, h_2 indicates the fee to be paid for each requesting node that enters the system. Therefore, the individual benefit function of each requesting node is defined as follows:

$$U_L = h_1 - C_1 E(W) - h_2.$$

Assuming that the requesting node is not aware of the state of the serving nodes in the system, nor is it aware of the number of the serving nodes that have failed. Assuming that Λ indicates the potential arrival rate of the requesting node, and Γ indicates the response probability that the requesting node is under the Nash equilibrium strategy. Therefore, the Nash equilibrium strategy of the requesting node is given by:

$$\Gamma = \begin{cases} 0, & h_1 C_1 E(W)|_{\lambda=0} + h_2, \\ \frac{\lambda^e}{\Lambda}, & C_1 E(W)|_{\lambda=0} + h_2 \le h_1 \le C_1 E(W)|_{\lambda=\Lambda} + h_2, \\ 1, & h_1 C_1 E(W)|_{\lambda=\Lambda} + h_2 \end{cases}$$

where λ^e is the solution of $U_L = 0$.

Assuming that $c = 15, \alpha = 0.8, \beta = 0.8, \mu_1 = 3, \mu_2 = 1, p = 0.6, h_1 = 17, h_2 = 2, C_1 = 3$, Fig. 9 reflects the effect of the arrival rate λ_1 and the working sleep parameter ξ on the individual benefit U_L of each requesting node. When ξ is a fixed value, U_L decreases with the increase of λ_1. When λ_1 is a fixed value, U_L increases with the increase of ξ. The main reason is that when the arrival rate increases, the number of the requesting nodes increases, and the average delay increases, thus the individual benefit decreases. When the working sleep parameter increases, the time in the normal busy period increases. The service rate in the normal busy period is higher than the service rate in the working sleep period. The average delay decreases, thus the individual benefit increases. As shown in Figure 9,

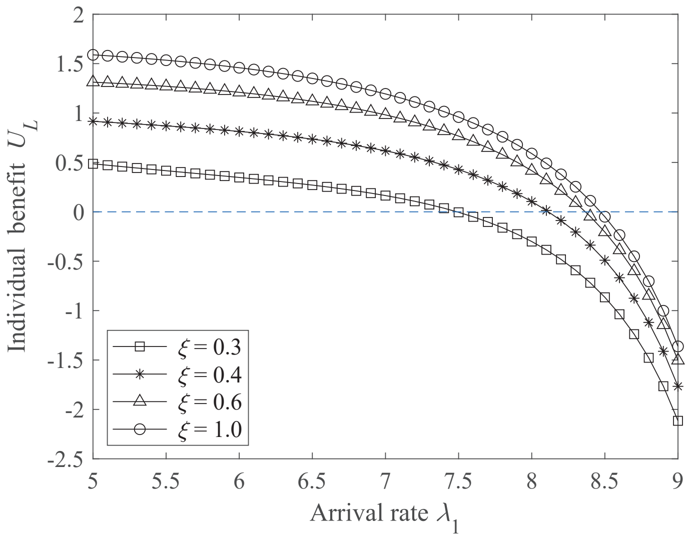

Fig. 9 The relationship between U_L and λ_1, ξ.

whichever value ξ is taken, $\lambda_1 = \lambda_1{}^{\min}, U_L \geq 0, \lambda_1 = \lambda_1{}^{\max}, U_L \leq 0$. It indicates that only a part of the requesting nodes can obtain the positive individual benefit, while other requesting nodes obtain the negative individual benefit due to the long waiting time in the buffer. Nash equilibrium is reached when $U_L = 0$, the optimal arrival rate $\lambda_1{}^{\min} \leq \lambda_1{}^{e} \leq \lambda_1{}^{\max}$. Tab. 3 reflects the relationship between the individual benefit and the arrival rate λ_1 when ξ=0.3. It can be seen that the individual benefit U_L decreases with the increase of λ_1. When $\lambda_1 = 7.5$, $U_L = 0$. The optimal arrival rate to reach the Nash equilibrium state under this condition is $\lambda_1{}^{e} = 7.5$.

In the case of different working sleep parameters of the serving nodes, the numerical results of the individual benefit and the arrival rate are shown in Table 4.

4.3 Social optimal strategy

In order to discuss the social optimal strategy, assuming that the individual benefit function of all single resource request nodes is the same and the individual benefit of all single requesting nodes are additive. Therefore, the social benefit function is defined as follows:

$$U_S = \lambda_1(h_1 - C_1 E(W) - h_2).$$

Table 3 The relationship between U_L and λ_1

λ_1	U_L
5.0	0.5
5.8	0.4
6.6	0.3
7.0	0.2
7.5	0.0
8.2	-0.5
9.0	-2.1

Table 4 The effect of λ_1 and ξ on individual benefit U_L

	U_L					
	$\lambda_1 = 5.0$	5.8	6.6	7.4	8.2	9.2
ξ=0.3	0.486	0.375	0.252	0.037	-0.481	-2.116
ξ=0.4	0.916	0.838	0.716	0.473	-0.088	-1.767
ξ=0.6	1.312	1.237	1.096	0.818	0.213	-1.504
ξ=1.0	1.589	1.492	1.323	1.014	0.379	-1.362

Assuming that $c = 15, \alpha = 0.8, \beta = 0.8, \mu_1 = 3, \mu_2 = 1, p = 0.6, h_1 = 17, C_1 = 3, h_2 = 2$, Fig. 10 reflects the effect of the arrival rate λ_1 and the working sleep parameter ξ on the social benefit U_S. When λ_1 is a fixed value, U_S increases with the increase of ξ. When ξ is a fixed value, U_S first increases and then decreases with the increase of λ_1. Therefore, the actual arrival rate and the value of the social maximum benefit under the social optimization exist. When $\xi = 1.0, \lambda_1 = 8.3$, the social benefit exists $maxU_S = 93.4041$. The main reason is that when the working sleep parameter increases, the time in the normal busy period increases. The average delay decreases, thus the social benefit increases. In the meantime, before the system is congested, that is, before the serving nodes are fully utilized, when the arrival rate increases, the utilization of the serving nodes increases, thus the social benefit increases. When the arrival rate exceeds the actual arrival rate under the social optimum, more and more nodes request the download service, causing congestion in the system, and the average delay increases, thus the social benefit decreases. In conclusion, the social benefit shows a trend of first increasing and then decreasing with the increase of λ_1, and the value of the social maximum benefit under the social optimization exists.

5 Conclusion

In order to address the problem of wasted energy consumption generated by unnecessary online behavior of the nodes, the working sleep mechanism for the serving nodes is introduced, and the synchronous multiple working vacation strategy is adopted to decrease system energy consumption. Based on this, combining factors such as node feedback and environmental disturbances, an M/M/*c* queueing model with negative customer, feedback, working vacation and repairable fault is built. By employing the quasi-birth-and-death process and matrix-geometric solution method, the steady-state probability vector is derived. Further, the expressions for performance indicators, such as the average delay and the total

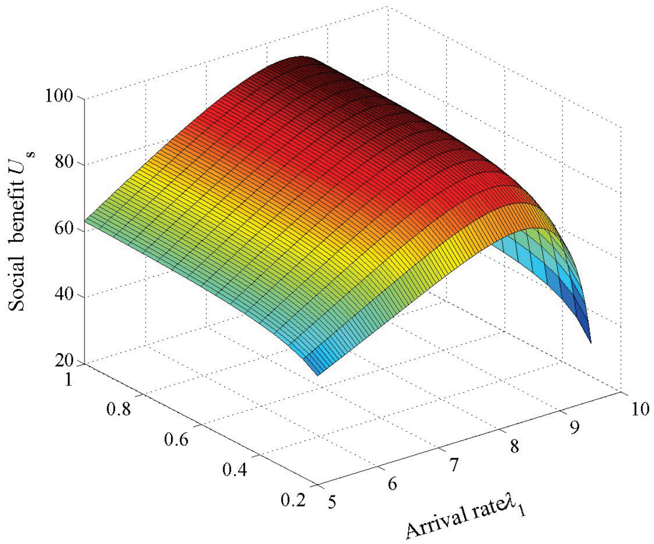

Fig. 10 The relationship between U_S and λ_1, ξ.

energy consumption, are given. Through the numerical analysis, the energy consumption of the serving nodes in various states is quantified and analyzed. Finally, the Nash equilibrium between the arrival rate and the working sleep parameter and the individual benefit of a single requesting node is analyzed, and then the optimization of the social benefit is studied. In conclusion, it can be seen that an appropriate decrease in the fault rate can avoid system congestion, introducing the working sleep mechanism and an appropriate increase in the working sleep parameter can decrease the average delay and the total energy consumption in the P2P network system, besides, the individual benefit and social benefit also can be increased.

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 No. A2020203010, A2018203088, Project of Hebei Key Laboratory of Software Engineering, No. 22567637H.

Declarations

- **Conflict of interest** The authors declare that they have no conflict of interest.
- **Ethics approval** Not applicable.
- **Consent to participate** Not applicable.
- **Consent for publication** All authors have reviewed the study and consented to its publication.
- **Authors' contributions** Conceptualization: all authors; Model building: Shunzhi Wang; Method guidance: Zhanyou Ma; Programming: Shunzhi Wang; Process supervision: Zhanyou Ma; Numerical experiments: Shunzhi Wang; Literature search: Rong Wang; Graphic beautification: Wenming Fang; Error checking: all authors.
- **Availability of data and materials** Not applicable.

References

1. Pan MS, Lin YP (2018) Efficient Data Dissemination for Wi-Fi Peer-to-Peer Network by Unicasting Among Wi-Fi P2P Groups. Wireless Network 24(8):3063-3081
2. Brienza S, Cebeci SE, Masoumzadeh SS, Hlavacs H, Anastasi G (2015) A Survey on Energy Efficiency in P2P Systems: File Distribution, Content Streaming, and Epidemics. Acm Computing Surveys 48(3):1-37
3. Azzedin F (2010) Trust-based Taxonomy for Free Riders in Distributed Multimedia Systems. International Conference on High Performance Computing & Simulation, pp 362-369
4. Zhang XR, Yang TB, Tan YC, Zhu GH (2007) Research on Multi-sender Data Dispatching Scheme based on P2P. Computer Systems & Applications 30(6):46-49 (in Chinese)
5. Mondal A, Trestian I, Zhen Q, Kuzmanovic A (2012) P2P as a CDN: A New Service Model for File Sharing. Computer Network 56(14):3233-3246
6. He WH, Liu H, He JS (2019) Research on the Current Situation and Development of P2P Network. Software Engineering 22(4):1-5 (in Chinese)
7. Yin BQ, Dong G, Jing H, Wu XM (2012) Modeling and Analysis for the P2P-based Media Delivery Network. Mathematical and Computer Modeling 55(3-4):1529-1539
8. Huo ZQ, Li XW, Jin SF, Wang ZH (2017) Nash equilibrium of an energy saving strategy with dual rate transmission in wireless regional area network. Wireless Communications and Mobile Computing 2017:1-10
9. Ferragut A, Paganini F (2015) Queueing Analysis of Peer-to-Peer Swarms: Stationary Distributions and their Scaling Limits. Performance Evaluation 93:47-62
10. Yu XR, Ma ZY, Guo SS, Chen L (2020) Performance Analysis of Wireless Communication Network with Threshold Activation Process and Interference Signals. International Journal of Communication Network and Distributed Systems 25(1):78-94
11. Si QN, Ma ZY, Liu FJ, Wang R (2021) Performance Analysis of P2P Network with Dynamic Changes of Servers based on M/M/*c* Queueing Model. Wireless Network 27(5):3287-3297
12. Singha N, Singh Y (2021) New incentive mechanism to enhance cooperation in wireless P2P networks. Peer-to-Peer Networking and Applications 14(3):1218-1228
13. Liu JP, Jin SF, Wang BS (2017) Study on optimization strategy for hybrid underlay/overlay spectrum sharing based on queuing model. Journal on Communications 38(9):55-64 (in Chinese)
14. Jin SF, Wang CF, Yin CX, Huo ZQ (2014) Performance Analysis for the Lazy Repair Strategy of the Fixed Nodes in a Hybrid P2P Network. Journal of Chinese Mini-Micro Computer Systems 35(6):1315-1319 (in Chinese)
15. Xue YZ, Jin SF, Wang XS (2018) A Task Scheduling Strategy in Cloud Computing with Service Differentiation. KSII Transactions on Internet and Information Systems 12(11):5269-5286
16. Zhang XF, Yin BQ (2017) Performance Analysis of CDN-P2P Network based on Processer-sharing Queues. International Conference on Software Engineering and Service Science, pp 24-27
17. Ma ZY, Zhang CZ, Zhang LY, Wang SZ (2021) Energy Saving Strategy and Nash Equilibrium of Hybrid P2P Network. Journal of Parallel and Distributed Computing 157(10):145-156
18. Trunfio P (2015) A two-layer model for improving the energy efficiency of file sharing peer-to-peer networks. Concurrency Computation: Practice and Experience 27(13):3166-3183
19. Jin SF, Wang XS, Yue WY (2018) A Task Scheduling Strategy with a Sleep-delay Timer and a Waking-up Threshold in Cloud Computing. Queueing Theory and Network Applications, Springer, Cham, pp 115-123
20. Stoica I, Morris R, Karger DR (2001) Chord: A Scalable Peer-to-Peer Lookup Protocol for Internet Applications. Acm Sigcomm Computer Communication Review 31(4):149-160
21. Tian NS, Yue DQ (2002) The quasi birth and death process and matrix-geometric solution. Beijing, Science Press (in Chinese)