A Scheme of Resource Allocation for Heterogeneous Services in Peer-to-peer Networks using Particle Swarm Optimization

Shiyong Li, Lin Jiao, Yue Zhang, Yan Wang, and Wei Sun

Abstract—**This paper considers reasonable resource allocation for heterogeneous services in peer-to-peer (P2P) networks, and presents a utility maximization model for bandwidth allocation of service providers. We firstly consider elastic services with concave utility functions and find that the optimal bandwidth allocation exists. Then we focus on heterogeneous services, where inelastic services with sigmoidal utility functions are coexisting with elastic services. The resource allocation model for heterogenous services is an intrinsically difficult problem of nonconvex optimization, resulting in difficulty of traditional gradient-based algorithm to obtain the optimum. In order to overcome it, we apply Particle Swarm Optimization (PSO) to revolve the model, propose a heuristic algorithm for resource allocation, and verify the results with simulation results.**

Index Terms—**P2P networks, heterogeneous services, utility function, PSO.**

I. Introduction

In peer-to-peer (P2P) networks, a service is offered and requested by several peers at the same time. A peer can choose freely the peers it provides service to and allocate resources for the service (e.g. upload bandwidth for P2P file-sharing applications)[1]. Thus, resources are used efficiently at a peer when at least one other peer requests a service from it. Meanwhile a requesting peer is normally serviced by several providers in parallel. Hence, its total service resource is accumulated over its providers.

Pricing mechanism for resource allocation in P2P networks was proposed in [2] to ensure a fair bandwidth allocation of service providers between requesters. The authors propose a rate-based bandwidth allocation algorithm where service providing peers adjust their service rates and service requesting peers adjust their price offers. The bandwidth allocation mechanism is further studied in [3] [4] and the weighted fairness is analyzed. Recently, a novel price-based resource allocation scheme is presented by applying the first order Lagrangian method and low-pass filtering scheme, so that a service provider can allocate its resources to its customers based on offered prices, achieving the efficient and fair resource allocation [5]. These research results mainly concentrate on how to allocate the resources of service providers between the service requesters fairly, however, they do not differentiate the types of services and assume there is only one service in the network.

Indeed there are a lot of network services [6][7][8][9]. These services can be divided into two types on the basis of the main shapes of utilities: elastic services and inelastic services. The former correspond to the traditional data services, such as file transfer. These services have the concave utility functions. The latter correspond to delay or rate sensitive real-time services, such as real-time streaming video services [10]. They have non-concave utility functions, e.g., the sigmoidal utility functions [7][8]. Resource allocation for elastic services in P2P has been investigated in [11], however, they do not consider resource allocation in scenario where inelastic services are coexisting with elastic ones.

In this paper we consider how to allocate the resource of service providers for heterogeneous services and propose the utility maximization model for resource allocation. It is different from that in [11] [5] where elastic services are only considered. Our model is to maximize the total utility of heterogeneous services, including not only elastic services, but also inelastic services. The resource allocation model for heterogeneous services is proven to be a difficult nonconvex optimization problem, which is hard to be resolved through traditional gradient-based resource allocation schemes, e.g., [11] [5]. We apply Particle Swarm Optimization (PSO) to resolve the problem and present a heuristic resource allocation algorithm, which is confirmed through simulation results.

The rest of this paper is summarized as follows: We introduce the resource allocation model for heterogeneous services in P2P networks in Section 2. In Section 3 we analyze the model for heterogeneous services through nonlinear programming theory. We propose a heuristic scheme by applying PSO to resolve the problem in Section 4 and give some simulation in Section 5. Finally we conclude this paper in Section 6.

II. Resource Allocation Model

Consider a P2P network which consists of a set of peers P and a set of services S. Each peer $p \in P$ is interested in one or several services, offers one or several services or is interested and offers different services at the same time. Thereby, a peer is not only a service customer, but also a service provider. For file-sharing applications in P2P, the upload capacity of one peer is used to transmit a file or a fragment of a file to remote peers which are interested in the file. Hence, the upload capacity of this peer is a scarce resource and other peers compete for it. Then an important

Manuscript received March 13, 2017; revised September 10, 2017. This work was supported in part by the National Natural Science Foundation of China (Nos. 71671159 and 71301139), the Natural Science Foundation of Hebei Province (No. G2016203236), the project Funded by Hebei Education Department (Nos. BJ2017029 and BJ2016063), the project Funded by Hebei Talents Program (No. A2017002108), and the China Postdoctoral Science Foundation (No. 2016M600196).

Shiyong Li, Lin Jiao, Yue Zhang, Yan Wang and Wei Sun are with School of Economics and Management, Yanshan University, Qinhuangdao 066004, China. Wei Sun is the corresponding author. e-mail: shiyongli@ysu.edu.cn, wsun@ysu.edu.cn.

problem is arising that how to allocate the resources of service providers between service customers.

Define the set of peers which offer at least one service s as the set of service providers $P(s)$. And also define the set of services that peer p provides as $S(p)$. Note that it is obvious that $p \in P(s)$ if and only if $s \in S(p)$. We model the resource allocation for heterogeneous services in P2P networks as the following optimization problem

$$\begin{array}{ll} \max & \sum_{s:s\in S} w_s U_s(y_s) \\ \text{subject to} & \sum_{p\in P(s)} x_{ps} = y_s, s \in S \\ & \sum_{s\in S(p)} x_{ps} \le C_p, p \in P \\ \text{over} & x_{ps} \ge 0, s \in S, p \in P \end{array} \quad (1)$$

where $U_s(y_s)$ is the utility function of service s when the total flow rate of this service is y_s, w_s is the weight of service s which can be used to achieve fair resource allocation between service customers, x_{ps} is the service rate that the service provider p provides for service s, and C_p is the upload capacity of service provider p.

In this optimization problem the objective is to maximize the aggregated utility of service rate y_s over all services in the network. Notice that, for service customer, the service rate y_s is the sum of the rates x_{ps} that service provider p provides, which is described by the equality in the resource allocation model. Meanwhile, the service rate of service provider p is constrained by its capacity, i.e. C_p, which is described by the inequality in the optimization problem above.

Recall that there are mainly two types of services in this network, elastic services with concave utilities and inelastic services with noncocave utilities (e.g., sigmoidal). We adopt the utilities described in [6][7][8]. As shown in Fig. 1, the concave utility function for elastic service s is given by $U_s(y_s) = c_s(\log(a_s y_s + b_s) + d_s)$ with flow rate y_s, and the sigmoidal utility function for inelastic service is given by $U_s(y_s) = c_s\left(\frac{1}{1+e^{-a_s(y_s-b_s)}} + d_s\right)$, where a_s, b_s, c_s and d_s are parameters of service s. The inelastic services correspond to delay and rate sensitive real-time services, such as real-time streaming video and audio services, and always require a high level of QoS[12].

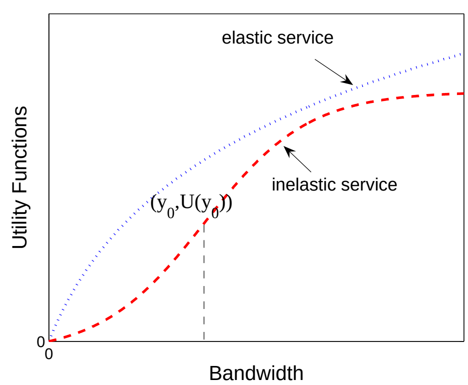

Fig. 1. Utility functions for elastic and inelastic services

Notice that the resource allocation model (1) is indeed a difficult nonconvex optimization problem, and it is hard to obtain the optimum through traditional methods. We will investigate this difficult problem, and find sufficient condition for existence of the global optimum to the model.

III. Model Analysis

A. Optimal Resource Allocation

The resource allocation model (1) equals to the following optimization problem, which is regarded as the primal problem

$$\begin{array}{ll} \max & \sum_{s:s\in S} w_s U_s\left(\sum_{p\in P(s)} x_{ps}\right) \\ \text{subject to} & \sum_{s\in S(p)} x_{ps} \le C_p, p \in P \\ \text{over} & x_{ps} \ge 0, s \in S, p \in P. \end{array} \quad (2)$$

We can obtain the Lagrangian of the primal problem as following

$$\begin{aligned} L(x,\lambda) = & \sum_{s:s\in S} w_s U_s\left(\sum_{p\in P(s)} x_{ps}\right) \\ & + \sum_{p:p\in P} \lambda_p \left(C_p - \sum_{s\in S(p)} x_{ps}\right), \end{aligned} \quad (3)$$

where $\lambda_p \ge 0$ is the Lagrange multiplier associated with the linear constrain on service provider p, and can be considered as the price charged by service provider p.

The Lagrange dual function $f(\lambda)$ is defined as the maximized $L(x,\lambda)$ over x for a given λ, that is

$$\begin{aligned} f(\lambda) = & \max_x \sum_{s:s\in S} w_s U_s\left(\sum_{p\in P(s)} x_{ps}\right) \\ & - \sum_{p:p\in P} \lambda_p \sum_{s\in S(p)} x_{ps} + \sum_{p\in P} \lambda_p C_p, \end{aligned}$$

and the optimal flow rate is denoted as

$$\begin{aligned} x^*(\lambda) = & \arg\max_x \sum_{s:s\in S} w_s U_s\left(\sum_{p\in P(s)} x_{ps}\right) \\ & - \sum_{p:p\in P} \lambda_p \sum_{s\in S(p)} x_{ps}, \end{aligned} \quad (4)$$

where $x^*(\lambda)$ is an optimal flow rate vector with the element $x^*_{ps}(\lambda_p)$, the optimal flow rate service provider p provides for service s, which is a function of the price λ_p charged by service provider p.

The Lagrange dual problem of the primal problem is

$$\begin{array}{ll} \min & f(\lambda) = L(x^*(\lambda),\lambda) \\ \text{subject to} & \lambda \ge 0, \end{array} \quad (5)$$

where the optimization variable is λ, a price vector with element λ_p.

The primal problem is how to allocate the scarce resource, i.e., upload capacity of service providers, to service customers, so as to obtain the requesting services. And the dual problem is how to charge the price the service customers should pay to service providers.

Let x^{opt} be the optimal solution to the primal problem, and $U^* = \sum_{s:s\in S} w_s U_s(\sum_{p\in P(s)} x^{\text{opt}}_{ps})$ be the optimal primal objective value. Also let λ^{opt} be the optimal solution to the dual problem, and $D^* = f(\lambda^{\text{opt}})$ be the optimal dual objective value. So we can obtain the following theorem.

Theorem 1 If the services are all elastic, that is, the utility functions are all concave, both the optimal flow rates for services and optimal primal objective value can be obtained. Furthermore, the optimal primal objective value is equal to the optimal dual objective value, i.e., $U^* = D^*$. However, the optimal service rate from each service provider, i.e., x_{ps}, may be not unique.

This theorem can be obtain from the convex optimization theory [13], since the primal objective function $\sum_s w_s U_s(y_s)$ is concave, and the constraint set of this problem is convex. The optimal rate y_s^{opt} for service s is unique, however, the rate that each service provider provides, i.e., x_{ps}^{opt}, may be not unique because of the equality $y_s^{\text{opt}} = \sum_{p\in P(s)} x_{ps}^{\text{opt}}$.

To realize the optimal resource allocation for elastic services in decentralized P2P architectures, distributed bandwidth allocation schemes should be designed. Recall that the resource allocation model (1) is a convex programming problem, thus gradient-based projection approach can be applied to the optimal bandwidth allocation scheme design, e.g., fair resource allocation mechanism proposed in [5] can be used to achieve the optimum of resource allocation model (1).

We analyze the prices charged by service providers and obtain the following result.

Theorem 2 At the optimal solution to the dual problem, i.e., λ^{opt}, for all service providers $p \in P(s)$ that provide resource for service s, prices charged by these service providers are equal, that is, for providers $i, j \in P(s)$ and $i \neq j$, $\lambda_i^{\text{opt}} = \lambda_j^{\text{opt}}$.

Proof The Lagrange dual function is

$$\begin{aligned} f(\lambda) = \quad & L(x^*(\lambda), \lambda) = \sum_{s:s\in S} w_s U_s(\sum_{p\in P(s)} x_{ps}^*) \\ & - \sum_{p:p\in P} \lambda_p \sum_{s\in S(p)} x_{ps}^* + \sum_{p:p\in P} \lambda_p C_p. \end{aligned}$$

At the optimal solution λ^{opt}, the Lagrange dual function is,

$$\begin{aligned} f(\lambda^{\text{opt}}) = \quad & \sum_{s:s\in S} w_s U_s(\sum_{p\in P(s)} x_{ps}^*) \\ & - \sum_{p:p\in P} \lambda_p^{\text{opt}} \sum_{s\in S(p)} x_{ps}^* + \sum_{p:p\in P} \lambda_p^{\text{opt}} C_p. \end{aligned}$$

The partial derivatives of $f(\lambda^{\text{opt}})$ on x_{is}^* and x_{js}^* are

$$\frac{\partial f(\lambda^{\text{opt}})}{\partial x_{is}^*} = \frac{\partial w_s U_s(\sum_{p\in P(s)} x_{ps}^*)}{\partial \sum_{p\in P(s)} x_{ps}^*} \frac{\partial \sum_{p\in P(s)} x_{ps}^*}{\partial x_{is}^*} - \lambda_i^{\text{opt}},$$

$$\frac{\partial f(\lambda^{\text{opt}})}{\partial x_{js}^*} = \frac{\partial w_s U_s(\sum_{p\in P(s)} x_{ps}^*)}{\partial \sum_{p\in P(s)} x_{ps}^*} \frac{\partial \sum_{p\in P(s)} x_{ps}^*}{\partial x_{js}^*} - \lambda_j^{\text{opt}},$$

respectively.

Since x^* is the optimal flow rate, $\partial f/\partial x_{is}^* = 0$, $\partial f/\partial x_{js}^* = 0$, then $\lambda_i^{\text{opt}} = \lambda_j^{\text{opt}}$. This theorem is completed.

B. Optimum for Elastic Services

We can construct a simple graph which is composed of the two sets S and P. Thereby, a node denotes a service provider or customer, and an edge denotes a service between a provider and a customer. Thus, we can separate the whole network into several regions according to the service relationship between providers and customers. In each one the prices charged by service providers that offer the same service are all equal at the optimum. In particular, if there is only one service offered by all providers, then the prices charged by these service providers are all equal. Suppose the bipartite graph is connected, then the optimal price charged by service providers is λ. Otherwise, if the bipartite graph is not connected, an optimization can be run for every disjoint connected subgraph separately.

Then, the Lagrangian (3) can be rewritten as

$$\begin{aligned} & L(x, \lambda) \\ = \quad & \sum_{s:s\in S} w_s U_s \left(\sum_{p:p\in P(s)} x_{ps} \right) + \sum_{p:p\in P} \lambda_p C_p \\ & - \sum_{p:p\in P} \lambda_p \sum_{s:s\in S(p)} x_{ps} \\ = \quad & \sum_{s:s\in S} \left(w_s U_s \left(\sum_{p:p\in P(s)} x_{ps} \right) - \sum_{p:p\in P(s)} \lambda_p x_{ps} \right) \\ & + \sum_{p:p\in P} \lambda_p C_p \\ = \quad & \sum_{s:s\in S} \left(w_s U_s \left(\sum_{p:p\in P(s)} x_{ps} \right) - \lambda \sum_{p:p\in P(s)} x_{ps} \right) \\ & + \lambda \sum_{p:p\in P} C_p \\ = \quad & \sum_{s:s\in S} (w_s U_s(y_s) - \lambda y_s) + \lambda \sum_{p:p\in P} C_p \\ = \quad & \widetilde{L}(y, \lambda). \end{aligned}$$

Substituting the utility function $U_s(y_s) = c_s(\log(a_s y_s + b_s) + d_s)$ for elastic service s into $\widetilde{L}(y, \lambda)$, and setting $\partial \widetilde{L}(y, \lambda)/\partial y_s = 0$, we obtain the optimal resource allocation for elastic service s

$$y_s^* = \frac{w_s c_s}{\lambda} - \frac{b_s}{a_s}. \tag{6}$$

Substituting (6) into $\widetilde{L}(y, \lambda)$, then

$$\begin{aligned} \widetilde{L}(y, \lambda) = \quad & \sum_{s:s\in S} \left(w_s c_s (\log \frac{w_s a_s c_s}{\lambda} + d_s) - w_s c_s + \frac{b_s}{a_s} \lambda \right) \\ & + \lambda \sum_{p:p\in P} C_p. \end{aligned}$$

Setting $d\widetilde{L}(y, \lambda)/d\lambda = 0$, we can obtain the optimal price that the peers request service s should pay

$$\lambda^* = \frac{\sum_{s:s\in S} w_s c_s}{\sum_{p:p\in P} C_p + \sum_{s:s\in S} \frac{b_s}{a_s}}, \tag{7}$$

then with (6)

$$y_s^* = \frac{w_s c_s}{\sum_{s:s\in S} w_s c_s} \left(\sum_{p:p\in P} C_p + \sum_{s:s\in S} \frac{b_s}{a_s} \right) - \frac{b_s}{a_s}. \tag{8}$$

Particularly, if we choose the utility function $U_s(y_s) = \log(y_s + 1)$ for elastic service s, then the optimal price

$$\lambda^* = \frac{\sum_{s:s\in S} w_s c_s}{\sum_{p:p\in P} C_p + |S|},$$

and

$$y_s^* = \frac{w_s c_s}{\sum_{s:s\in S} w_s c_s}\left(\sum_{p:p\in P} C_p + |S|\right) - 1,$$

where $|S|$ is the number of services in the network. Therefore, the total flow rate of an elastic service depends on the number of services plus the total capacity of all service providers, and the total willingness-to-pay weighted by the willingness-to-pay of the service customer who requests this elastic service. We can also observe from (8) that the total optimal flow rate of each service is unique.

C. Capacity Guarantee for Inelastic Services

Now consider the optimization model with inelastic services which have the sigmoidal utility functions. For simplicity, we assume all the service providers can offer these inelastic services. For the inelastic service s, instruct a tangent from the origin to the sigmoidal function $U_s(y_s)$, and let $y^0_{s(\text{sig})}$ and λ^0_s be the flow rate and slope, respectively, at the point where the tangent from the origin interests the sigmoidal function.

Theorem 3 Consider the optimization model for resource allocation of heterogeneous services in P2Ps. The optimal flow rates for services can be obtained and the optimal primal objective value is also equal to the optimal dual objective value, i.e., $U^* = D^*$, if there exists a price vector $\lambda \geq 0$ with element $\lambda_p \geq 0$ satisfying the inequality $\lambda_p < \lambda^0_s, p \in P(s)$ and the capacity $C_p(\lambda_p)$ for each service provider $p \in P$, where

$$C_p(\lambda_p) = \sum_{s\in S(p)} x^*_{ps(\text{sig})}(\lambda_p) + \sum_{s\in S(p)} x^*_{ps(\text{con})}(\lambda_p). \quad (9)$$

Here, $x^*_{ps(\text{sig})}(\lambda_p)$ and $x^*_{ps(\text{con})}(\lambda_p)$ are the price-based flow rates (3) that service provider p offers for inelastic and elastic services, respectively.

Proof Note taht the optimal price for the dual problem is $\lambda^{\text{opt}} = \lambda$, i.e., $\lambda^{\text{opt}}_p = \lambda_p$ because the subgradient of $f(\lambda)$ is zero when $C_p = C_p(\lambda_p)$. Since $\lambda_p < \lambda^0_s$, $p \in P(s)$, the optimal price at service provider p is less than the critical price λ^0_s, i.e. $\lambda^{\text{opt}}_p < \lambda^0_s, p \in P(s)$. At the optimal point to the dual problem (4), the price seen by the service, i.e., λ_s, is equal to the price of any service provider that supports this service, i.e. $\lambda_p, p \in P(s)$. So for inelastic service, $y_{s(\text{sig})}(\lambda_s) = \sum_{p\in P(s)} x^*_{ps(\text{sig})}(\lambda_p) > y^0_{s(\text{sig})}(\lambda^0_s)$, which means that $y_{s(\text{sig})}$ lies in the concave part of the sigmoidal function. In this part, the optimal flow rate for the service supported by the service providers can be obtained, and the duality gap is zero. Since λ_p is the optimal price of the dual problem, $x^*_{ps(\text{sig})}(\lambda_p)$ and $x^*_{ps(\text{con})}(\lambda_p)$ are the flow rates that service provider p offers for inelastic and elastic services, respectively. $y_{s(\text{sig})} = \sum_{p\in P(s)} x^*_{ps(\text{sig})}(\lambda_p)$ and $y_{s(\text{con})} = \sum_{p\in P(s)} x^*_{ps(\text{con})}(\lambda_p)$ are the optimal flow rates for the inelastic and elastic services, respectively. This theorem is completed.

The inelastic services may require high QoS and the high QoS can be guaranteed by the service providers through the optimization model by adding the constraint $y_{s(\text{sig})} \geq y^0_{s(\text{sig})}$ for every inelastic service s. Here we assume the capacity of service provider p is no less than $\sum_{s\in S(p)} x^*_{ps(\text{sig})}(\lambda^0_s) + \sum_{s\in S(p)} x^*_{ps(\text{con})}(\lambda^0_s)$, where λ^0_s is the slope where the tangent from the origin interests the sigmoidal function. The model is modified to the following

$$\begin{aligned}
\max \quad & \sum_{s:s\in S(\text{sig})} w_s U_s(y_{s(\text{sig})}) + \sum_{s:s\in S(\text{con})} w_s U_s(y_{s(\text{con})}) \\
\text{subject to} \quad & \sum_{p\in P(s)} x_{ps(\text{sig})} = y_{s(\text{sig})}, \\
& \sum_{p\in P(s)} x_{ps(\text{con})} = y_{s(\text{con})}, \\
& \sum_{s\in S(p)} x_{ps(\text{sig})} + x_{ps(\text{con})} \leq C_p, \\
& y_{s(\text{sig})} \geq y^0_{s(\text{sig})}, \\
\text{over} \quad & x_{ps(\text{sig})} \geq 0, x_{ps(\text{con})} \geq 0.
\end{aligned} \quad (10)$$

where $x_{ps(\text{sig})}$ and $x_{ps(\text{con})}$ are the flow rates that service provider p offers for inelastic and elastic services, respectively.

Now for the resource allocation model (1) consisting of inelastic services which have the sigmoidal utility functions, it is a nonconvex optimization problem, and it is difficult to obtain the global optimum. Furthermore, the gradient-based resource allocation algorithm proposed for elastic services is not efficient to converge to the optimum. We will investigate the nonconcave optimization problem in next section and develop a heuristic algorithm by applying PSO approach.

IV. Resource Allocation Scheme Using PSO

A. Algorithm description

Since the resource allocation model is an intrinsically difficult problem of nonconvex optimization, thus the traditional gradient-based schemes are not efficient to converge to the optimum. In this short paper we apply PSO method to resolve the optimization problem and give a heuristic algorithm. PSO is an algorithm based on stochastic optimization algorithms, and owns some distinct advantages, for example, it does not need to calculate the gradients of the objective function, or even define the form of the objective function. PSO has been found useful to deal with very complicated optimization problems, such as power optimization [14][15] image segmentation [16][17], function optimization problems[18][19], and communication networks [8][9].

In the scheme each particle represents one available resource allocation solution for the model. Let x and v denote a particle coordinates (position) and its corresponding flight speed (velocity) in a search space, respectively. Therefore, the ath particle is represented as $X^a = (x^a_{11}, \ldots, x^a_{p1}; \ldots; x^a_{ps})$, and $V^a = (v^a_{11}, \ldots, v^a_{p1}; \ldots; v^a_{ps})$ in PSO-based resource allocation scheme. Next, let $X^a = (x^a_{pj})$ and $V^a = (v^a_{pj})$ for simplicity. And let $Pbest^a = (x^{aPbest}_{pj})$ and $Gbest = (x^{Gbest}_{pj})$ be the best position of individual a

and its neighbors' best position so far, respectively. Using the information, the velocity and position of individual a are updated by the following law

$$\begin{aligned} V^a(k+1) = &\ \omega V^a(k) + c_1 * \Re_1 * (Pbest^a(k) - X^a(k)) \\ &+c_2 * \Re_2 * (Gbest(k) - X^a(k)), \end{aligned}$$

$$X^a(k+1) = X^a(k) + V^a(k+1),$$

where ω is the inertia weight factor; c_1, c_2 are the acceleration constants; $\Re_1, \Re_2$ are uniform random values between 0 and 1; $X^a(k)$ is current position of individual a at the iteration step k; $V^a(k)$ is the velocity of individual a at the iteration step k, $V^a_{\min} < V^a(k) < V^a_{\max}$; $Pbest^a(k)$ is best position of the individual a at the iteration step k;$Gbest(k)$ is best position of the whole group.

In the iteration rule above, the parameters $V^a_{\min}$ and $V^a_{\max}$ determine the resolution, or fitness, with which regions between the present position and target position. The constants c_1 and c_2 represent the weighting of the stochastic acceleration terms which pull each particle toward *Pbest* and *Gbest* positions.

B. The Fitness Function

As for the fitness function in this scheme, notice that the optimization problem is subjected to inequality constraints, thus we adopt the PSO with penalty function in the scheme. Then, the fitness function in our scheme is

$$F_f = \begin{cases} f(X), & \text{if the solution is feasible} \\ f(X) + h(k)H(X), & \text{otherwise} \end{cases} \tag{11}$$

where $f(X)$ is the original objective function to be optimized, $h(k)$ is a penalty value, and $H(X)$ is a penalty factor.

Following the main idea for choosing fitness function above, we formulate the fitness function as the following form:

$$F_f = \begin{cases} \sum\limits_{s:s\in S} w_s U_s(\sum\limits_{p\in P(s)} x_{ps}), \text{if} \sum\limits_{s\in S(p)} x_{ps} \le C_p \\ \sum\limits_{s:s\in S} w_s U_s(\sum\limits_{p\in P(s)} x_{ps}) \\ +\sum\limits_{p:p\in P} \lambda_p(C_p - \sum\limits_{s\in S(p)} x_{ps}), \text{otherwise.} \end{cases} \tag{12}$$

Thus if one particle satisfies all constraints, the solution is feasible. Otherwise, an extra charge should be paid. It is proportional to the amount of violation with very large positive constant.

Thus the proposed scheme is to obtain the optimal position (the optimal bandwidth allocation for peers) according to the fitness function (the objective function) through the velocity (the auxiliary variable).

V. Simulation Results

In this part we will give some numerical examples to verify the results above. In the simulation of PSO-based scheme, the PSO strategy parameters $c1$, $c2$ are both chosen to 2 and $\omega = 1$ in order to guarantee the convergence of the algorithm.

A. Elastic Services

Firstly, we consider a simple P2P network which only provides elastic services with utilities $U_s(y_s) = \log(y_s + 1)$. There are two service requesters and two service providers. The willingness-to-pay is $w = (w_1, w_2) = (2, 1)$, and the upload capacity of providers is $C = (C_1, C_2) = (4, 6)$Mbps. The swarm size is 20 and the maximum number of iterations is 100. The simulation result is shown in Fig. 2. We can observe that the optimum can be achieved within reasonable iterations (e.g., 50 iterations). Furthermore, the optimum obtained from the algorithm is equal to the value from (8). Thus, the algorithm is efficient in solving the resource allocation model and can achieve the optimum within reasonable convergence times.

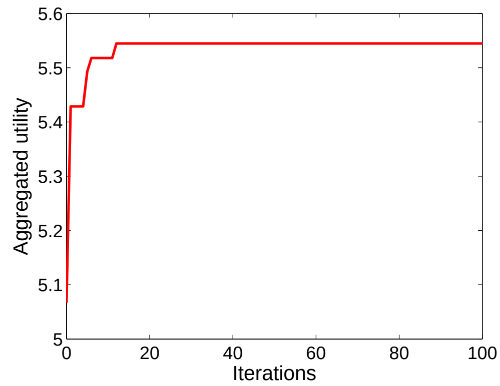

Fig. 2. Aggregated utility for elastic services

Now we consider a large scale network which consists of 100 service requesters and 100 service providers. The willingness-to-pay is 2 for the first 50 requesters and 1 for the rest. The upload capacities of service providers are all 5Mbps. We depict the performance of the proposed algorithm in this case in Fig. 3.

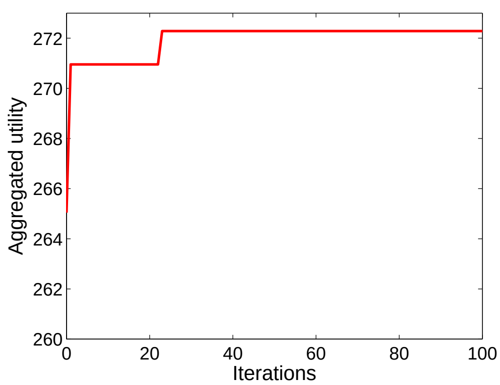

Fig. 3. Aggregated utility for elastic services in a large scale network

We observe that the size of the network (i.e., the number of peers) does not affect the convergence speed of the algorithm. The final objective increases with the number of peers but, this value is reached with almost the same number of iterations (e.g., 50 iterations). This is rather expected. The algorithm is synchronously operated, and different peers run the separate optimization steps in parallel. Thus the number of peers does not alter the convergence speed obviously. In fact, the convergence speed mainly depends on algorithm parameters other than the number of peers.

Now we consider the performance of the resource allocation algorithm with peer departures. The simulation setup is identical with the one above except that the peers are not static, that is, after a period of peer interaction 20 service requesters and 20 service providers are leaving the network. We depict the evolution of aggregated utility for the network in Fig. 4. We find that the algorithm behaves well after the transitional points of peer departures and is convergent to the optimum within reasonable iteration times.

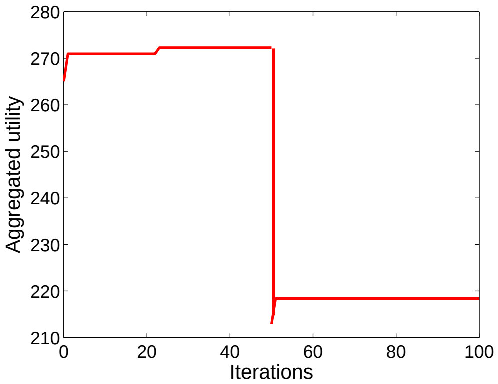

Fig. 4. Aggregated utility for elastic services with peers departure

B. Elastic and Inelastic Services

Now, we consider a simple network consisting of two different service requesters. The utility function of elastic service is $\log(y_1+1)$, and the sigmoidal utility function of inelastic service is $2/(1+e^{-(y_2-4)}) - 2/(1+e^4)$. For this inelastic service, $y_2^0 = 5.5398$ and $\lambda_2^0 = 0.2908$. These services are provided by two service providers.The upload capacities of these providers are 4Mb/s and 6Mb/s, respectively. The optimal bandwidth allocation is 3.9497Mb/s for the elastic service and 6.0503Mb/s for the inelastic service. And the optimal objective value is 3.3353, as shown in Fig. 5. The optimal price to the dual problem is $\lambda = (\lambda_1, \lambda_2) = (0.202, 0.202)$. The price at service provider 1 is equal to the price at service provider 2 because at the optimal point both these providers support the elastic service (or, the inelastic service). The optimal price seen by the inelastic service, i.e., 0.202, is less than the slope λ_2^0 where the tangent from the origin interests the sigmoidal function, i.e., 0.2908.

We depict the performance of the resource allocation scheme using PSO in Fig. 6. Here we choose the swarm size 20 and the maximum number of iterations 100. We can observe that the algorithm can converge to the optimum within reasonable iteration times.

Now we increase the number of both elastic and inelastic services to 50 and consider the performance of the scheme. There are 100 service providers. Each one has the upload capacity 10Mb/s. As shown in Fig. 7, the resource allocation scheme still converges to the optimum at about 50 iterations in this large scale network.

Now we consider the performance of the resource allocation algorithm with peers departure and arrival. The simulation setup is identical with the one above except that the peers are not static, that is, after a period of peer interaction some service requesters and providers are leaving the network while other new service requesters and providers join. That is, at iteration 50, 20 service requesters (10 peers requesting elastic services and the others requesting inelastic services) and 20 providers leave the network and after a period time at iteration 100, 40 new service requesters (20 peers requesting elastic services and the others requesting inelastic services) and 40 new service providers join the network. The evolution of aggregated utility for the network is shown in Fig. 8. Obviously, the algorithm behaves well after the transitional points of peer departures and arrivals, and it is still convergent to the optimum within reasonable

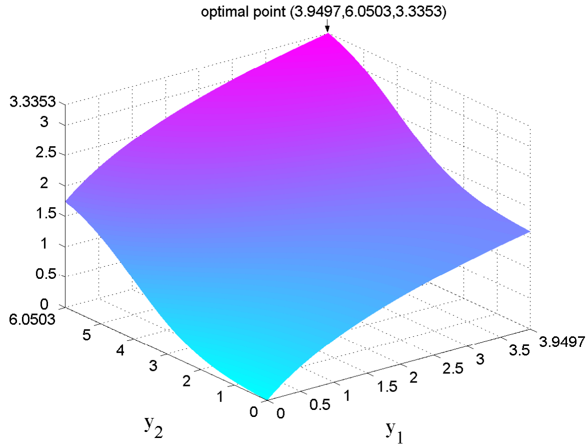

Fig. 5. Optimum for resource allocation model

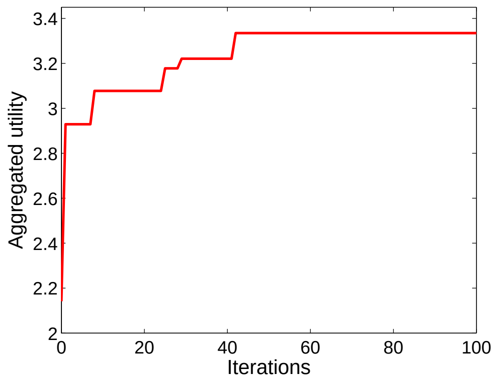

Fig. 6. Aggregated utility of heterogeneous services

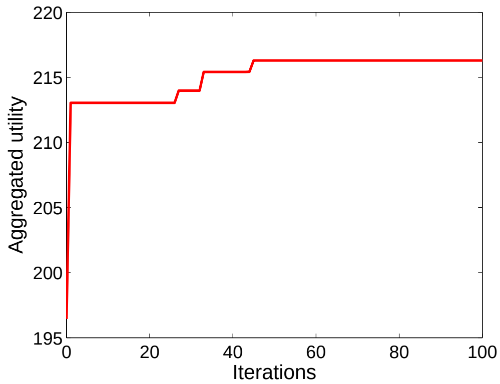

Fig. 7. Aggregated utility of heterogeneous services in a large scale network

iteration times.

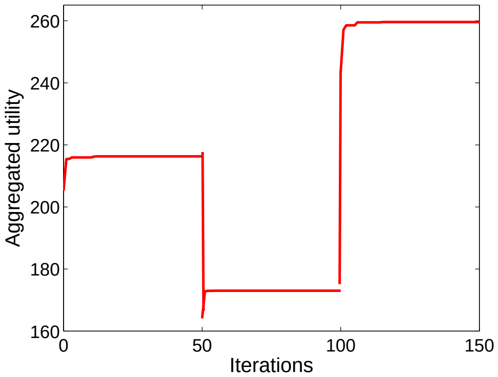

Fig. 8. Aggregated utility for heterogeneous services with peers departure and arrival

Finally we also analyze the affect of the swarm sizes on the behavior of the resource allocation scheme using PSO and depict the result in Fig. 9. We observe that the convergence speed of the proposed algorithm is slightly improved by increasing the swarm size from 20 to 100. Thus, as we have observed in Figs. 3 and 7, the convergence speed mainly depends on algorithm parameters other than the number of peers.

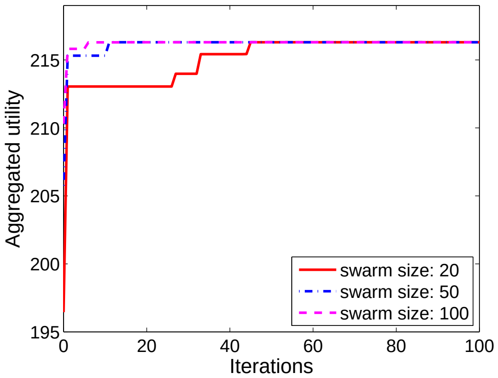

Fig. 9. Aggregated utility of heterogeneous services with different swarm sizes

VI. Conclusions

In this short paper we present the utility maximization model for resource allocation of heterogeneous services in P2P networks. Firstly, we only consider the model for elastic services. The optimal resource allocation for each service can be obtained since the utilities of these services are all concave. For the model with heterogeneous services, it is a difficult nonconvex optimization problem, and is hard to handle through traditional methods such as gradient-based algorithms. We apply PSO approach to revolve the optimization problem and develop a heuristic algorithm. Finally, we verify the performance of the scheme with some numerical examples.

References

[1] Y. Zheng, F. Lin, Y. Yang, and T. Gan, "Adaptive resource scheduling mechanism in P2P file sharing system," *Peer-to-Peer Networking and Applications*, vol. 9, no. 6, pp. 1089-1100, 2016.

[2] K. Eger, and U. Killat, "Resource pricing in peer-to-peer networks," *IEEE Communications Letters*, vol. 11, no. 1, pp. 82-84, 2007.

[3] K. Eger, and U. Killat, "Fair resource allocation in peer-to-peer networks (extended version)," *Computer Communications*, vol. 30, no. 16, pp. 3046-3054, 2007.

[4] S. Li, W. Sun, C.-G. E, and L. Shi, "A scheme of resource allocation and stability for peer-to-peer file-sharing networks," *International Journal of Applied Mathematics and Computer Science*, vol. 26, no. 3, pp. 707-719, 2016.

[5] S. Li, and W. Sun, "A mechanism for resource pricing and fairness in peer-to-peer networks," *Electronic Commerce Research*, vol. 16, no. 4, pp. 425-451, 2016.

[6] J.W. Lee, R.R. Mazumdar, and N.B. Shroff, "Non-convex optimization and rate control for multi-class services in the Internet," *IEEE/ACM Transactions on Networking*, vol. 13, no. 4, pp. 827-840, 2005.

[7] P. Hande, S. Zhang, and M. Chiang, "Distributed rate allocation for inelastic flows, *IEEE/ACM Transactions on Networking*, vol. 15, no. 6, pp. 1240-1253, 2007.

[8] S. Li, W. Sun, and N. Tian, "Resource allocation for multi-class services in multipath networks," *Performance Evaluation*, vol. 92, pp. 1-23, 2015.

[9] S. Li, W. Sun, and C. Hua, "Optimal resource allocation for heterogeneous traffic in multipath networks," *International Journal of Communication Systems*, vol. 29, no. 1, pp. 84-98, 2016.

[10] M. Chen, M. Ponec, S. Sengupta, J. Li, and P.A. Chou, "Utility maximization in peer-to-peer systems with applications to video conferencing," *IEEE/ACM Transactions on Networking*, vol. 20, no. 6, pp. 1681-1694, 2012.

[11] S. Li, L. Shi, and C. E, "Utility-based resource allocation for elastic services in peer-to-peer networks," *ICIC Express Letters*, vol. 8, no. 10, pp. 2711-2717, 2014.

[12] A.E. Mae, W.K. Kwee, P.Y. Han, and L.S. Hoe, "Resource allocation for real-time multimedia applications in LTE's two-level scheduling framework," *IAENG International Journal of Computer Science*, vol. 43, no. 4, pp. 427-437, 2016.

[13] D. Bertsekas, Nonlinear Programming, 2nd ed. Athena Scientific, 1999.

[14] P.K. Sharma, and P. Garg, "An optimal power distribution algorithm for two time slot cooperative diversity protocols," *Engineering Letters*, vol. 21, no. 3, pp. 119-122, August 19, 2013.

[15] H.A. Gabbar, Y. Labbi, L. Bower, and D. Pandya, "Performance optimization of integrated gas and power within microgrids using hybrid PSO-PS algorithm," *International Journal of Energy Research*, vol. 40, no. 7, pp. 971-982, 2016.

[16] R.M. Swarna Priya, S. Prabu, and V.S. Dharun, "F-SIFT and FUZZY-RVM based efficient multi-temporal image segmentation approach for remote sensing applications," *Automatic Control and Computer Sciences*, vol. 50, no. 3, pp. 151-164, 2016.

[17] Y. Liu, C. Mu, W. Kou, and J. Liu, "Modified particle swarm optimization-based multilevel thresholding for image segmentation," *Soft Computing*, vol. 19, no. 5, pp. 1311-1327, 2015.

[18] J.-S. Wang, and J.-D. Song, "A hybrid algorithm based on gravitational search and particle swarm optimization algorithm to solve function optimization problems," *Engineering Letters*, vol. 25, no. 1, pp. 22-29, February 22, 2017.

[19] J.-S. Wang, and J.-D. Song, "Application and performance comparison of biogeography-based optimization algorithm on unconstrained function optimization problem," *IAENG International Journal of Applied Mathematics*, vol. 47, no. 1, pp. 82-88, 2017.

Shiyong Li received his Ph.D. degree from Beijing Jiaotong University, Beijing, China, in 2011. Currently he is an associate professor in the School of Economics and Management at Yanshan University. He is the (co)author of more than 50 papers in mathematics, technique, and management journals. He has been a principal investigator/co-investigator on several research projects supported by the National Natural Science Foundation of China, the National Education Committee Foundation of China, the China Postdoctoral Science Foundation, and other foundations. His research interests include resource allocation in networks, electronic commerce, and economics of queues.