Peer-to-Peer Utility Maximization

Minghua Chen, Sudipta Sengupta, Miroslav Ponec, Philip A. Chou, and Jin Li

Abstract—In this paper, we study the problem of utility maximization in Peer-to-Peer (P2P) systems, in which aggregate utilities are maximized by running distributed algorithms on P2P nodes that are constrained by their uplink capacities. This may be understood as extending the seminal flow control framework in [1] and [2] from single-path unicast over general topology to multi-path multicast over P2P topology, with network coding allowed.

For single-rate multicast over certain popular P2P topologies, we show that routing along a linear number of trees per source can achieve the largest rate region that can be possibly obtained by (inter-session) network coding. This simplification result allows us to develop a new multi-tree routing formulation for the problem. Despite of the negative results in literature on convergence of Primal-dual algorithms under multi-path settings, we have been able to develop a delay-based Primal-dual algorithm to solve our multi-tree based utility maximization problem. We characterize the convergence behavior of the Primal-dual algorithm, and utilize our proposed sufficient condition to show its global convergence to the optimal solution under different P2P communication scenarios we study. We also discuss how to extend our solution for single-rate multicast to multi-rate multicast.

I. INTRODUCTION

The problem addressed in this paper is motivated by Peer-to-Peer (P2P) multi-party conferencing applications in which maximizing the quality of experience of all participants under network resource constraints is a crucial challenge. We measure the quality of experience of the conferencing participant by a utility function, which, for video conferencing, can be the Peak-Signal-to-Noise-Ratio (PSNR) of the decoded video.

In the past decade, network utility maximization have attracted significant attention ever since the seminal framework was introduced in [1] and [2]. In the framework, network protocols are understood as distributed algorithms that maximize aggregate user utility under wired or wireless network resource constraints. For the single-path unicast scenarios considered in [1] and [2], user's utility function is typically assumed to be strictly concave function of user rate, and the resource constraints set is linear. Various types of fairness across users can be warranted by choosing different utility functions [3]. This framework not only provides a powerful tool to reverse engineering existing protocols such as TCP [4], but also allows systematic design of new protocols, see [5] for a comprehensive review.

There have been work on extending the framework to multi-path unicast scenarios [6] [7] [8], as well as single-tree multicast scenarios [9] [10]. For utility maximization

in multi-path unicast scenarios, the utility function is non-strictly concave with respect to the individual path rate due to multi-path routing. The challenge is to design distributed algorithms to solve non-strictly concave optimization problems with provable fast convergence and easy implementation. Primal and Dual algorithms, and proximal approach are proposed to address such challenges [6] [7] [8].

For utility maximization in single-tree multicast scenarios where routers enable multicast functionality, the constraint set is non-linear, in particular, involving non-differentiable $\max(\cdot)$ terms. In [9] and [10], distributed Primal and Dual algorithms are proposed to maximize utility, under the assumptions that multicast trees are given and every session has a unique source. The challenge of dealing with non-differentiable max function in the constraints is approached by either using continuous and concave approximation of the max function [10], or introducing auxiliary variables and applying either Proximal or sub-gradient approaches [9].

There is also work focusing on multicast scenarios where routers can perform intra-session network coding [11] [12] [13]. The challenge is to deal with non-strictly concave optimization under non-linear constraints. By exploring the Proximal approach, or a slow timescale traffic engineering control approach, or expressing the constraints involving $\max(\cdot)$ terms with equivalent linear ones, distributed Primal, Dual subgradient and Primal-dual algorithms are proposed to maximize the sum of non-concave utility functions, or minimize the cost of using the network [11] [12] [13].

In this paper, we consider the utility maximization problem for multicast in a P2P setting, with multi-path delivery and inter-session network coding allowed. This setting differentiates our work from other existing work, and highlights the challenges we encounter. Our main contributions for single-rate multi-source multicast in P2P systems are as follows:

- **The Optimality of Routing over P2P Topology:** We focus on typical P2P topology where peer uplinks are the *only* bottleneck in the network. For multi-source multicast on P2P topology under certain assumption, we show that all feasible rates can be achieved by packing polynomial number of depth-1 and depth-2 Steiner trees. As such, routing is optimal even if the system contains Steiner nodes (helpers), and surprisingly there is no gain to perform (intra-session or inter-session) network coding. This result is a multi-source extension of the single-source result studied in [14].
- **Tree-rate-based Formulation and A Primal-Dual Algorithm for Multi-tree based Multicast:** We introduce a new multi-tree based formulation for utility maximization in P2P topology in which the variables are rates of individual *trees*. The largest achievable rate set is expressed by linear constraints in our tree based formu-

lation, thus avoiding the nonlinear $\max(\cdot)$ terms in path and link based formulations. Contrary to popular belief that Primal-dual algorithms in general fail to converge in multi-path delivery scenarios and supporting evidence for this in [8], we develop a queuing delay based Primal-dual algorithm that solves the utility maximization problem for multi-tree routing formulation under a general sufficient condition that holds in popular P2P settings.

We also discuss extending the above results to multi-rate multicast scenarios. Compared to single-rate multicast, multi-rate multicast allows different receivers of same source to receive content such as video at different rates using, for example, scalable layered coding. This introduces additional challenge as it is not clear how to obtain the largest achievable rate region when inter-layer network coding is allowed. Interestingly, we found that if all receivers for the same source have the same utility function, then an optimal solution allocates equal rates to all receivers for the same source, and no layer coding is needed. That is, performing single-rate multicast is optimal in such case, and all of the above results directly apply.

II. Problem Formulation

We consider a network presented by a directed graph $G = (V, E)$, where V is the set of vertices, i.e., nodes in the network, and E is the set of edges, i.e., links in the *physical* network. Note that by using a node splitting transformation, we can accommodate node uplink capacities in this model. Let $n = |V|$.

In the P2P systems we consider, some source $s \in S$ sends its content to a set of receivers, denoted by R_s. A set of helper nodes, denoted by H, are willing to help in distributing the content. In this paper, we assume a deterministic fluid model for sending rates of nodes and ignore packet dynamics. This assumption is reasonable when the timescale of rate control is sufficiently larger than that of packet dynamics.

Let z_s be the multicast rate of source s, and $z = [z_s, s \in S]$. We first consider single-rate multicast in P2P systems; that is, all members in R_s receive s's stream at this rate. We discuss multi-rate multicast scenarios in Section IV. Let $U_s(z_s)$ be the utility upon receiving the content from s at rate z_s. To prevent abusing resources of helpers, sources and receivers should use helpers' resources only after they have used up their own. Putting this into consideration, we associate a cost, denoted by $G_h(z)$, with using a helper $h \in H$ to distribute a content.

The utility maximization problem in P2P systems is given by

$$\max_z \sum_{s \in S} |R_s| U(z_s) - \sum_{h \in H} G_h(z), \text{ s. t. the constraints of } \{z_s\}.$$

Before formulating the problem further, we need to understand the constraint set for $\{z_s\}$ and how to achieve it.

A. Network Coding vs. Routing

The achievable rate region of single-source multicast scenario is characterized as the minimum of the min-cuts between the source node s and all nodes in its receiver set R_s [15], i.e., $\min_{t \in R}$ min-cut (s, t). If network coding is allowed, then the single-source multicast rate region can be achieved for arbitrary topology with polynomial complexity [16].

The achievable rate region for inter-session multicast scenarios was recently implicitly characterized in [17], but currently no scheme is known to achieve it. It is believed that information from different multicast groups should be coded in a nonlinear fashion in order to achieve the rate region (inter-session coding). However, doing such mixing and coding is complex and largely open.

Another way to explore the achievable rate region is by routing. Each source s packs directed Steiner trees rooted at s and reaching all receivers in R_s. For the *general case of arbitrary topologies*, this approach of routing brings up the following difficulties:

1) For a given source, the maximum rate achieved by routing can be a factor of up to $\log |V|$ lower than that achieved by network coding [16].
2) To achieve the maximum rate for routing, the problem of packing directed Steiner trees is $\mathcal{NP}$-hard [18]. Moreover, the number of Steiner trees used in an optimal solution may be exponential.

As such, routing can not achieve the largest rate region in a general topology and its cost is prohibited. However, the fact that our problem involves practical P2P topology where peer uplinks are the only bottlenecks allows us to tackle all of the above difficulties in a surprisingly elegant manner.

B. Impact of P2P Topology

In P2P topology, we assume peer uplinks are the only bottlenecks in the whole network, and every peer can connect to every other peer through routing in the overlay. In the overwhelming majority of residential broadband connections, bottlenecks typically are at the edge of the access networks rather than in the middle of the Internet. Furthermore, it is common to have the uplink capacity of a peer to be several times smaller than the downlink capacity, thus justifying the practicality of our assumption on P2P topology.

In the context of P2P topology with the above uplink constraint assumptions, a powerful theorem established in the Mutualcast paper [14] states the following. Consider a network with P2P topology consisting of a source s, a set of receivers R_s, and a set of helpers H. Then, the min-cut capacity for source s and receivers R_s can be achieved by packing at most $1 + |R_s| + |H|$ Mutualcast trees as follows:

- One depth-1 tree rooted at s and reaching all nodes in R_s.
- $|R_s|$ depth-2 trees, each rooted at s and reaching all other receivers in R_s via different $r \in R_s$.
- $|H|$ depth-2 trees, each rooted at s and reaching all receivers in R_s via different $h \in H$.

This result extends and simplifies Edmonds' theorem [19] for P2P topology. It allows Steiner nodes (helper) and uses only depth-1 and depth-2 Steiner trees. Fig. 1 shows all 12 Mutualcast trees for a three peers and one helper scenario where each peer multicast its content to the other two.

The Mutualcast Theorem is for single-source multicast scenario, it states routing is optimal and intra-session network
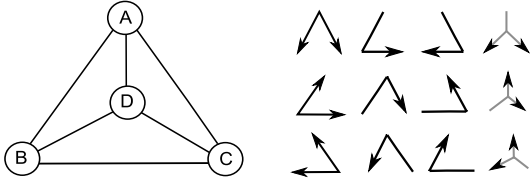
Fig. 1.   Multicast trees for three nodes (A, B, C) and one helper (D).

coding is not necessary. We extend Mutualcast Theorem to multi-source scenario in the following theorem, stating that routing is optimal and *inter-session network coding* is not needed, if we require multicast groups to be pairwise identical or disjoint . Such scenario is common in P2P systems.

Theorem 1: Consider a P2P topology in which peer uplinks are the only bottleneck. Consider multiple multicast sessions given by source nodes $s \in S$, receiver set R_s, and helper nodes $H = V - \sum_{s \in S} (\{s\} \cup R_s)$. Further, assume that multicast groups to be pairwise identical or disjoint, i.e., $R_s \cup \{s\} = R_{s'} \cup \{s'\}$ or $(R_s \cup \{s\}) \cap (R_{s'} \cup \{s'\}) = \emptyset$ for all s and s' in S, and a node in a multicast group connects only to every other node in the same group and every helper node in H. Then, the largest achievable rate region $z = \{z_s, s \in S\}$, can be obtained by routing along $1 + |R_s| + |H|$ Mutualcast trees for each source s independently.

In contrast to the known results that inter-session coding is needed to achieve the maximum rate region in a general topology, the unique structure of the P2P topology we consider in this paper allows us to achieve the maximum rate region by packing only linear number of Steiner trees per source, if each receiver is part of every multicast session. The result is not only surprising but also elegant.

We summarize the advantages and disadvantages of using network coding and packing (directed) Steiner trees to achieve multicast rate region in Table I.

C. Tree-rate-based Formulation

With the packing Mutualcast trees approach, each source $s \in S$ builds a set of depth-1 and depth-2 Mutualcast trees to send data to all receivers in R_s. We denote this set of trees also as s, and a source is identified by the set of trees of which it is the root. A tree $m \in s$ is a set of links and peer or helper nodes that the tree passes through; all receiver nodes on a tree receive the same content at the same rate. We denote the rate of tree m as x_m. Rates of the trees rooted at source s sum up to the source rate z_s, i.e., $\sum_{m \in s} x_m = z_s$, $\forall s \in S$.

Let J be the set of uplinks associated with each node, one link per node, and assume each link $j \in J$ has a finite capacity C_j. The injecting rate of link j is the aggregate rate of the trees that pass through link j, denoted by y_j, and is given by,

$$y_j \triangleq \sum_{s \in S} \sum_{m \in s:j \in m} b_j^m x_m, \quad \forall j \in J, \tag{1}$$

where b_j^m is the number of tree m's branches that pass through physical link j. Since different branches of a tree in the overlay can pass through the same physical link in the underlay, the tree rates might be counted multiple times when computing

the injecting rate of a link, hence the multiplication by b_j^m. Similarly, define the forwarding rate of a helper node h as

$$y_h \triangleq \sum_{s \in S} \sum_{m \in s:h \in m} b_h^m x_m, \quad \forall h \in H, \tag{2}$$

where b_h^m is the out-degree of helper node h in multicast tree m. Denote $y^H = [y_h, h \in H]$.

The utility maximization problem in P2P systems can be formulated based on tree rates $\{x_m\}$ as follows:

$$\max_{\{x_m\}} \quad \sum_{s \in S} |R_s| U_s \left(\sum_{m \in s} x_m \right) - \sum_{h \in H} G_h(y_h) \tag{3}$$
$$s.t. \quad y_j \leq C_j, \quad \forall j \in J,$$

where $|R_s| U_s \left(\sum_{m \in s} x_m \right)$ is the aggregate utility of a group R_s upon receiving content at rate $\sum_{m \in s} x_m = z_s$, and $G_h(y_h)$ is the cost of using helper node h to deliver content at rate y_h. As discussed earlier, this cost is to prevent peers from abusing resources from helpers – sources and receivers should use helpers' uplink capacities only after they use up their own. Formally, if the optimum objective function value can be achieved without using (or using lower) helper uplink capacities, then this should be preferred. We assume that the utility functions $U_s(\cdot), s \in S$, are strictly concave, and the cost functions $G_h(\cdot), h \in H$ are strictly convex.

The optimization problem in (3) is a concave optimization problem with linear constraints. It might have more than one optimal $\{x_m\}$, since the objective function is not strictly concave with respect to $\{x_m\}$. However, the optimal aggregate rate associated with each source $\{z\}$ is unique. This is because the objective function is strictly concave with respect to $\{z_s\}$, and the rate constraint region of $\{z_s\}$ can be shown to be a polyhedron by eliminating the tree-rate variables x_m (for example, by Fourier-Motzkin elimination [20]).

For the optimization problem shown in (3), interior-point and simplex based algorithms can be applied to solve the problem in a centralized manner [21]. However, centralized solutions may put a huge burden on the central solver and it requires the central solver to know the up-to-date topology, peer uplink rates, cross traffic, and the utility function of each peer. Tracking this information may not be feasible in practice and it is therefore desirable to have a distributed algorithm that can be deployed in practice.

III. A DISTRIBUTED PRIMAL-DUAL ALGORITHM FOR MULTI-TREE BASED MULTICAST

The optimization problem we consider in (3) is a non-strictly concave optimization due to multi-tree routing between sources and their receivers. There are in general three ways to approach such a problem in a distributed manner, namely Primal algorithms, Dual sub-gradient algorithms, and Primal-dual algorithms. The advantages of Primal algorithms are their wide applicability and fast convergence in multi-path/multi-tree routing scenarios. The down side of the Primal algorithms is that they typically only generate approximate solutions.

Due to the non-strictly concave objective function, standard Dual gradient algorithms fail to work since the gradient is

TABLE I
COMPARISONS OF APPROACHES TO ACHIEVE MULTICAST RATE REGION

	single-source multicast (P2P topology)	single-source multicast (general topology)	multi-source multicast (P2P topology)	multi-source multicast (general topology)	complexity
network coding	optimal	optimal	? (open)	? (open)	polynomial in the known optimal cases
packing Steiner trees	optimal	suboptimal	optimal in certain cases	suboptimal	$\mathcal{NP}$-hard in general, polynomial in the known optimal P2P cases

not everywhere defined. Alternatively, dual subgradient algorithms [9] [12] and dual proximal algorithms [9] [7] are proposed to solve the problem. However, convergence of dual variables in these approaches can be slow, and recovering optimal primal variables from optimal dual variables requires solving another optimization problem.

In this paper, we focus on Primal-dual algorithms. The advantage of Primal-dual algorithms is that they can be implemented by utilizing the delay measurements between peers, which makes it particularly attractive in practice. The challenge of using such algorithms lies in its unclear convergence capability in multi-path routing scenarios. In this paper, we provide a sufficient condition for Primal-dual algorithms to converge in multi-path routing scenarios. We show that the condition can be easily satisfied in two popular P2P settings. Furthermore, we show that this delay-based Primal-dual algorithm converges exponentially fast.

Lagrangian of the optimization problem in (3) is given by

$$L(x,p) = \sum_{s \in S} |R_s| U_s (z_s) - \sum_{h \in H} G_h(y_h) - \sum_{j \in J} p_j (y_j - C_j),$$

where p_j is the Lagrangian multiplier, and can be interpreted as the price of using link j. Since the original problem in (3) is a concave optimization problem with linear constraints, strong duality holds and there is no duality gap. Any optimal solution of the problem in (3) and one of its corresponding Lagrangian multiplier is a saddle point of L over the set $\{x \geq 0, p \geq 0\}$, and vice versa. Further, (x,p) is one such saddle point of L if and only if it satisfies the Karush-Kuhn-Tucker conditions [21]: $\forall s \in S, \forall m \in s, \forall j \in J,$

$$p_j \geq 0, \quad y_j \leq C_j, \quad p_j (y_j - C_j) = 0, \quad (4)$$
$$|R_s|U_s'(z_s) - \sum_{h \in m} b_h^m G_h'(y_h) - \sum_{j \in m} b_j^m p_j = 0. \quad (5)$$

The first equation is the complementary slackness condition. The optimal Lagrangian multiplier can be nonzero only if the capacity constraint of link j is activated, i.e., $y_j = C_j$. We denote the set containing all (x,p) that satisfy the above conditions by E. As the original problem has at least one solution, E contains at least one point and is not empty.

To pursue one of the saddle points of L, we consider the following Primal-dual algorithm, over the set $\{x \geq 0, p \geq 0\}$: $s \in S, \forall m \in s,$ and $j \in J,$

$$\dot{x}_m = k_m \left(|R_s| U_s'(z_s) - \sum_{h \in m} b_h^m G_h'(y_h) - \sum_{j \in m} b_j^m p_j \right) \quad (6)$$
$$\dot{p}_j = \frac{1}{C_j} (y_j - C_j)_{p_j}^+, \quad (7)$$

where k_m is a positive constant controlling the adaptation rate of tree m and $(y_j - C_j)_{p_j}^+ = y_j(t) - C_j$ if $p_j > 0$, and

is $\max(0, y_j - C_j)$ otherwise. It is known that p_j adapted according to (7) can be interpreted as queuing delay [22]. Every saddle point of L is an equilibrium of the above system.

Whether the Primal-dual algorithm can be applied to multi-path/multi-tree routing scenarios is an open problem. Served as a negative result, it is shown that (x,p) following (6)-(7) can oscillate indefinitely in common multi-path unicast scenarios [8, Section 2.5]. Consequently, to our best knowledge, almost no solution for multi-path routing utilizes the Primal-dual algorithm.

In this paper, we give a general sufficient condition for the Primal-dual algorithm in (6)-(7) to converge to the optimal solution, regardless of unicast or multicast, single path or multipath routing. To our best knowledge, this is the first attempt to characterize the applicability of the Primal-dual algorithm. We believe its applicability is beyond the P2P systems we study in this paper.

We give the definitions and notations to be used in later analysis. Let A be the connectivity matrix, where the (i,j) entry is the number of branches of tree j passing through link i. This is different from traditional connectivity matrix (for unicast) as its entries can take values other than 1 or 0. Similarly, let A_H be the helper connectivity matrix whose entries being the number of branches of a tree passing through a helper. Let $K = \text{diag}\{k_m, m \in s, s \in S\}$, $C = \text{diag}\{C_j, j \in J\}$, where J is assumed to contain only the bottlenecks, without loss of generality. Let B be the matrix representing the relation of source rate, rate passing through helpers and the tree rate, with the (i,j) entry being 1 if tree j belongs to source i, being b_i^j if tree j passes through helper i, and 0 in any other cases.

The following Lemma shows that the nonlinear system in (6)-(7) converges to an invariant set, over which the nonlinear system turns into a *linear* one.

Lemma 1: All (x,p) trajectories of the system in (6)-(7) converge to an invariant set, denoted by $V_0 = \{(\bar{x}, \bar{p}) : [\bar{z}, \bar{y}^H]^T = B\bar{x} = \text{const}\}$, over which the following is true:
- $\bar{z}$ and $\bar{y}^H$ are the unique solution to the problem in (3);
- the nonlinear system reduces to a linear one:

$$\begin{cases} \dot{\bar{x}} = KU' - KA_H G' - KA^T \bar{p} \\ \dot{\bar{p}} = C^{-1} A\bar{x} - \mathbf{1} \end{cases} \quad (8)$$

where U' and G' are constant matrices;
- the above linear system is marginally stable, and all its trajectories do not converge and form limit cycles.

Shown by the above theorem, (x,p), trajectories of the system in (6)-(7) converge to a set V_0 where the source rates $\bar{z}$ are optimal. Clearly, all saddle points of L belong to V_0, and $E \subseteq V_0$. If we also have $V_0 \subseteq E$, then the Primal-dual algorithm solves the problem in (3).

However, it is possible that V_0 contains some $(\bar{x}, \bar{p})$ that are not in E; $\dot{x}$ and $\dot{p}$ are not zero. If (x,p) moves onto these

points, then they will keep oscillating and never converge. This is exactly the challenge of using the Primal-dual algorithm in multi-path/multi-tree routing scenarios, and explains the oscillations in rates and delays discussed in [8].

One way to guarantee $V_0 = E$ is to utilize the fact that $B\bar{x}$ is constant to explore the conditions for V_0 to not include those singular points, as explored in the following theorem.

Theorem 2: All trajectories (x, p) of the system in (6)-(7) converge globally asymptotically to one of its equilibria and $V_0 = E$, if $\bar{p}$ is completely observable from $(\bar{z}, \bar{y}^H)$ through the linear system in (8). Equivalently, $V_0 = E$ if for any eigenvalue of $C^{-1}AKA^T$, denoted by λ,

$$\text{rank} \begin{pmatrix} C^{-1}AKA^T - \lambda I \\ BKA^T \end{pmatrix} = |J|. \tag{9}$$

Furthermore, we can access a stronger convergence result for the Primal-dual algorithm in (6)-(7), if the above condition is satisfied.

Theorem 3: If the Primal-dual algorithm in (6)-(7) converges globally asymptotically, then the following is also true: there exists a compact set Ω such that any compact set containing Ω is a positive invariant set of the system in (6)-(7). Further, if (x, p) are bounded within one such compact set, then the system trajectories (x, p) converge to the equilibria globally exponentially.

The Primal-dual algorithm described in (6)-(7) can be implemented by each link generating its queuing delay and each source adjusting the rates of its trees by observing sum of the queuing delays introduced by using the trees. As such, the algorithm can be implemented in a distributed manner.

Regardless of the nice convergence properties and the easy implementation of the Primal-dual algorithm in (6)-(7), it is possible that with some network settings, the condition in (9) is not satisfied and the Primal-dual algorithm may not converge. One example is shown in [8, Section 2.5].

Interestingly, the unique structure of Mutualcast and the P2P topology allows us to prove that the sufficient condition can be easily satisfied for two typical P2P systems - P2P Content dissemination systems and multi-party conferencing systems, as explored in the following subsections.

1) P2P Content Dissemination Scenarios: Consider a P2P data dissemination system with n peers, among which there are sources, receivers, and helpers. Every source distributes its content to its receivers, with or without helpers' assistance. A receiver can receive contents from multiple sources simultaneously, while sources are servers that only distribute data but do not receive contents. These scenarios correspond to popular P2P file distribution and P2P streaming scenarios in practice.

Assume the first n_s number of the nodes are sources. Define $R_i, 1 \le i \le n_s$, to be the set of peers that want to receive contents from source i. Define $H_i, 1 \le i \le n_s$ be the set of helpers that help distributing the content of source i. For the ease of explanation, we assume there is no helper in the following analysis, i.e., $H_i = \emptyset$. The analysis can be straightforwardly extended to the case where $H_i \ne \emptyset$.

Let $n_i = |R_i| + 1$, $1 \le i \le n_s$. Each source uses total $n_i + 1$ Mutualcast trees to deliver its content. For the sake of simplicity, we use the following notation when stating the

result. We denote x_{ij} as the rate of tree j of source i passing through node j in the level one, with $1 \le j \le n$ and $1 \le i \le n_s$. Let k_{ij} represent how fast the tree rate x_{ij} adapts, and $k_{ij} = 0$ if $j \notin R_i$ and $j \ne i$. This is equivalent to having exactly $n_i + 1$ trees for source i. Since a source is not receiver for other sources in our P2P data dissemination scenarios, we also have $k_{ij} = 0$ for all $1 \le j \le n_s$ and $j \ne i$.

The following theorem gives a sufficient condition for the Primal-dual algorithm in (6)-(7) to converge to the saddle points of L in P2P data dissemination systems.

Theorem 4: For P2P data dissemination systems in P2P topology, all (x, p) trajectories of the system in (6)-(7) converge to one of its equilibria globally asymptotically, if the following conditions are satisfied:

- For all $1 \le i \ne j \le n$, $\xi_i \ne \xi_j$, where

$$\xi_l = \begin{cases} \frac{(n_l - 1)n_l}{C_l} k_{ll}, & 1 \le l \le n_s; \\ \frac{1}{C_l} \sum_{j:l \in R_j} (n_j - 1)^2 k_{jl}, & \text{else.} \end{cases}$$

- $k_{ii} < \frac{C_i}{2C_j} k_{ij}$, for all $1 \le i \le n_s$ and $n_s < j \le n$.

In practice, these conditions are in fact easy to satisfy with high probability. For example, source i can generate $k_{ij}, j \ne i$ in a suitable range *randomly* such that the first condition is satisfied with a good chance.

Source i can then select k_{ii} such that the second condition is satisfied under practical relationship between C_i, normally the server bandwidth, with C_j, normally peers' (home users') uplink bandwidth. For instance, we can assume practically $\min(\frac{C_i}{C_j}) = 1$ and set k_{ii} to be less than $\frac{1}{2} \min_j(k_{ij})$ in a random fashion.

In practice, satisfying this condition also forces the source to adapt the depth-2 Mutualcast trees with high priority, indicating source nodes, normally the server, will adapt quickly to the network condition changes in receivers, as compared to the response to its own uplink condition change.

2) Multi-party Conferencing Scenarios: Consider a P2P multi-party conferencing system with the first n_s of them being participants and the rest n_h peers being helpers. Every participant wants to receive contents from all other participants. The following theorem gives a sufficient condition for the Primal-dual algorithm in (6)-(7) to converge to the saddle points of L in P2P multi-party conferencing systems.

Theorem 5: For multi-party conferencing systems in P2P topology, all (x, p) trajectories of the system in (6)-(7) converge to one of its equilibria globally asymptotically, if for source s, all its $k_m, m \in s$ are the same.

The requirement of having all k_m to be the same for all $m \in s$ implies every source should adjust its trees at the same pace, which is convenient and easy to satisfy in practice.

IV. MULTI-RATE MULTICAST IN P2P SYSTEMS: THE CASE OF RECEIVER-INDEPENDENT UTILITIES PER SOURCE

In multi-rate multicast, receivers for the same source can receive at different rates. Let $U_r^s(\xi_r^s)$ be the utility associated with receiver r for receiving content from source s at a rate ξ_r^s. Let $\xi = [\xi_r^s, s \in S, r \in R_s]$. The utility functions U_r^s are assumed to be strictly concave and increasing. Let the set of

sources be S and let the set of receivers for each source $s \in S$ be R_s. Then, the utility maximization problem is stated as:

$$\max_{\{\xi_r^s\}} \sum_{s \in S} \sum_{r \in R_s} U_r^s(\xi_r^s) - \sum_{h \in H} G_h(\xi),$$

subject to $\{\xi_r^s\} \in$ feasible region.

Compared to single-rate multicast, multi-rate multicast allows different receivers of same source to receive content such as video at different rates using, for example, scalable layered coding. Suppose for a given source s, the receiver rates are ordered as $\xi_{i_1}^s \le \xi_{i_2}^s \le \ldots \le \xi_{i_{|R_s|}}^s$.

Then, the rate $\xi_{i_1}^s$ can be interpreted as a base layer (say, layer 1) received by all nodes in R_s. The next higher layer, layer 2, has rate $(\xi_{i_2}^s - \xi_{i_1}^s)$ and is received by all nodes in $R_s - \{i_1\}$. In general, layer ℓ, $1 < \ell \le |R_s|$ has rate $(\xi_{i_\ell}^s - \xi_{i_{\ell-1}}^s)$ and is received by all nodes in $R_s - \{i_1, i_2, \ldots, i_{\ell-1}\}$.

In such case, inter-layer network coding may be performed for content of different layers corresponding to the same source. Note the extended Mutualcast results in Theorem 1 fails to apply to inter-layer coding scenario because the multicast group for higher layers are subsets of that of lower layers, violating the assumptions required in Theorem 1. As such, with inter-layer coding, how to obtain the largest achievable set of $\{\xi_r^s\}$ is still a challenging open problem.

Interestingly, we found that the structure of receiver utility functions may have positive effect in resolving such challenging problem. In particular, we consider the special case when all receivers for the same source have the same utility function. Then, it can be expected that an optimal solution allocates equal rates to all receivers for the same source. This is, in fact, true and is established in the following theorem.

Theorem 6: If $U_r^s = U^s \forall r \in R_s, s \in S$ (receiver utility functions are identical for the same source) and the utility functions U^s are strictly concave, then in every optimal solution, we have $\xi_r^s = \xi^s \forall r \in R_s, s \in S$. (receiver rates are identical for the same source).

The above theorem says the following: even if we don't know the actual achievable region of $\{\xi_r^s\}$, it is sufficient to consider only part of the region in order to solve the utility maximization problem. This is because structure of utility functions restricts the structure of optimal solution to the maximization problem, allowing us to solve the problem over certain structural achievable rate region that contains the optimal solutions we pursue.

Noticing the resulting structural rate region is exactly the one of single-rate multicast, we can directly apply our solution for single-rate multicast to solve multi-rate multicast with receiver-independent utility functions per source.

V. CONCLUSION AND FUTURE WORK

We investigate the utility maximization problem in P2P systems. The nature of P2P topologies allows us to tackle difficulties arising in the general network case in a surprisingly elegant manner. We show that the maximum utility region can be achieved by considering only a linear number of multicast trees (in the overlay network) per source, and give a delay-based Primal-dual distributed algorithm that can be used by each source to determine the transmission video coding rate and to split that rate across multiple trees. Our approach naturally accommodates helper nodes within the optimization framework. The developed algorithms are practical and easy to implement in a P2P overlay over the current Internet.

We are investigating scalability of the proposed approach in large P2P networks. Another area of future work would be to consider general multi-rate multicast in P2P systems, with receiver dependent utility functions per source.

REFERENCES

[1] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness, and stability," *Journal of the Operationl Research Society*, pp. 237–252, 1998.

[2] S. H. Low and D. E. Lapsley, "Optimization flow control, i: Basic algorithm and convergence," *IEEE/ACM Trans. Networking*, no. 6, pp. 861–875, Dec. 1999.

[3] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Networking*, no. 5, pp. 556 – 567, Oct. 2001.

[4] F. P. Kelly, "Fairness and stability of end-to-end congestion control," *European Journal of Control*, pp. 159–176, 2003.

[5] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition:a mathematical theory of network architectures," in *Proceedings of the IEEE*, vol. 95, no. 1, Jan. 2007, pp. 255–312.

[6] H.Han, S. Shakkottai, C. Hollot, R. Srikant, and D. Towsley, "Multipath TCP: A joint congestion control and routing scheme to exploit path diversity in the internet," *IEEE/ACM Trans. Networking*, 2006.

[7] X. Lin and N. B. Shroff, "Utility maximization for communication networks with multi-path routing," *IEEE Trans. Automat. Contr.*, vol. 51, pp. 766–781, May 2006.

[8] T. Voice, "Stability of congestion control algorithms with multi-path routing and linear stochastic modelling of congestion control," Ph.D. dissertation, University of Cambridge, Cambridge, UK, May 2006.

[9] K. Kar, S. Sarkar, and L. Tassiulas, "Optimization based rate control for multirate multicast sessions," in *Proceedings of IEEE INFOCOM*, Anchorage, Alaska, Apr. 2001.

[10] S. Deb and R. Srikant, "Congestion control for fair resource allocation in networks with multicast flows," *IEEE Trans. Automat. Contr.*, no. 2, pp. 274–285, Apr. 2004.

[11] L. Chen, T. Ho, S. H. Low, M. Chiang, and J. C. Doyle, "Optimization based rate control for multi-cast with network coding," in *Proceedings of IEEE INFOCOM*, Anchorage, Alaska, May 2007.

[12] Y. Wu and S.-Y. Kung, "Distributed utility maximization for network coding based multicasting: a shortest path approach," *IEEE J. Select. Areas Commun.*, no. 8, pp. 1475–1488, Aug. 2006.

[13] D. S. Lun, N. Ratnakar, M. Medard, R. Koetter, D. R. Karger, T. Ho, E. Ahmed, and F. Zhao, "Minimum-cost multicast over coded packet networks," *IEEE Trans. Inform. Theory*, no. 6, pp. 2608–2623, June 2006.

[14] J. Li, P. A. Chou, and C. Zhang, "Mutualcast: an efficient mechanism for content distribution in a p2p network," in *Proceedings of Acm Sigcomm Asia Workshop*, Beijing, China, Apr. 2005.

[15] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, no. 4, pp. 1204–1216, July 2000.

[16] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhiuzen, "Polynomial time algorithms for multicast network code construction," *IEEE Trans. Inform. Theory*, vol. 51(6), pp. 1973–1982, 2005.

[17] X. Yan, R. W. Yeung, and Z. Zhang, "The capacity region for multisource multi-sink network coding," in *ISIT2007*, Nice, France, June 2007.

[18] K. Jain, M. Mahdian, and M. R. Salavatipour, "Packing steiner trees," in *14th ACM-SIAM Symp. on Discrete Algorithms (SODA)*, Jan. 2003.

[19] J. Edmonds, "Edge-disjoint branchings," *Combinatorial Algorithms, R. Rustin, ed.*, pp. 91–96, 1973.

[20] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*. Athena Scientific, 1997.

[21] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.

[22] S. H. Low, L. Peterson, and L. Wang, "Understanding vegas: A duality model," *Journal of ACM*, vol. 49, no. 2, pp. 207–235, Mar. 2002.