Mehran GARMEHI, PhD Candidate
Computer Engineering School
Iran University of Science and Technology
Email: garmehi@iust.ac.ir
Associate Professor, Morteza ANALOUI , PhD
Computer Engineering School
Iran University of Science and Technology
Email: analoui@iust.ac.ir

A DISTRIBUTED MECHANISM FOR ECONOMIC MANAGEMENT OF TRANSMISSION INFRASTRUCTURE IN HYBRID CDN-P2P NETWORKS

Abstract. *Hybrid CDN-P2P networks blend CDN and P2P technology to benefit from the complementary advantages of these technologies. In these networks, a critical challenge is to construct and maintain multicasting trees to distribute the content from distribution servers to the edge servers, clients and peers. In this work considering tight relation between internal and P2P layers of these networks, a single form of multicasting tree construction problem in this domain is presented formally. NP-Hardness of the problems is proved and based on economic mechanism design theory, an economic direct privacy preserving mechanism is employed to solve the problem. Performance of the solution is studied and it is proved that the solution produces a two approximate outcome. Experiments confirm that the proposed solution produces near optimal results. These experiments also verify that the mechanism is able to maintain quality of results with presence of dynamic changes in multicasting groups.*

Keywords*: Hybrid CDN-P2P Network; Multicasting Trees; Economic Mechanism Design Theory; Content Distribution.*

JEL classification: L96; C71; L86; C63

Introduction

Content Distribution Networks overcome shortcomings of the Internet. These systems transfer content through the mostly congested intermediate paths of the internet and distributing it as close as possible to its destinations. In order to eliminate the need of users to pass their requests from long distance worldwide paths, these systems, most of the time, rely on a dedicated network infrastructure composed of powerful servers and wide bandwidth links. Relying on such architecture, these systems get the content from origin servers of content providers; relay it among their distribution and intermediate servers and across their reliable links to their edge servers close to clients. At the last step the content passes local edge paths (which have lower traffic intensity and price) with high performance to its audience

Despite the fact that CDNs benefit from many advantages like ease of management and high quality of service, they have some serious concerns including high running costs and low scalability. P2P networks benefit from high scalability and most of the times demand low startup and running cost but in these systems it is hard to guarantee quality of service and most of the times they are not reliable enough. In Hybrid CDN-P2P networks, Edge Servers of CDNs are equipped with P2P technology to benefit from advantages of both architectures. Dongyan and Kulkarni (2004) showed that for highly popular contents, Edge Servers of HCDN-P2Ps rely on P2P networks and as a result they are capable to handle much more requests in contrast to traditional CDNs. In other words, in some cases where the popularity of the content is higher than bootstrapping threshold of P2P network, it may be possible to distribute it using a few edge servers in some regional P2P network.

Combination of P2P networks and traditional CDN architecture in Hybrid CDN-P2P networks makes the management of content distribution among the servers and peers more challenging. The dynamic nature and churning rate of P2P networks makes it necessary to construct and maintain multicasting sets more dynamically. Application of these networks increases the variation of population of servers and peers in multicasting groups. As a result, the outcome of Replica Placement (RP) subsystem of the Hybrid CDN-P2P network, which is the list of edge servers and peers where the content must be replicated on, may change frequently. So in order to multicast the content through these networks, a tailor-made multicasting mechanism is required.

The high probability of changes in multicasting trees makes it necessary for multicasting mechanism to avoid recalculation of multicasting trees in case of variations in multicasting groups. In other words, not only it is essential to construct efficient multicasting trees but also it is important to keep these multicasting trees applicable after churning of peers or replication decisions. In Garmehi and Analoui (2011) we introduced a solution which only works for internal layer of these networks. It does not guarantee effectiveness of multicasting trees after changes in multicasting groups.

CDNs are conceived to be distributed systems. Nowadays CDNs like Akamai benefit from thousands of servers. A centralized solution like Garmehi and Analoui (2011) is only applicable in small size or academic CDNs and is not suitable and applicable in a scalable and reliable CDN.

In this work it is assumed that for the Hybrid CDN-P2P network there exists an efficient RP strategy like the solution introduced by Wang, et al. (2009). Outcome of RP algorithm is list of servers and peers which the content must be replicated on them. These outcomes are feed to the multicasting strategy for streaming the content from distribution servers to edge servers and peers. In order to increase the fault tolerance of the content distribution, this multicasting problem, assumes that in both internal and edge layers of the network, it is allowed to have more than one source for the content. Therefore, the content provider has the opportunity to deliver its content to more than one Distribution Server. Also since, Seyyedi and Akbari (2011) showed that employing multiple sources of flow increases the performance of content distribution significantly, this assumption makes it is possible to inject the flow of content into the P2P network using more than one Edge Server to increase the performance. These sets of multicasting trees obey latency constraint of the content and they also consider capacity constraint on the underlying links.

The paper is organized as follows. In section 2 a brief review of related work, is presented. The problem is introduced formally and its features are discussed in section 3. A short introduction to mechanism design theory comes next and based on this theory, an economic mechanism, as the solution to the problem, is introduced in section 4. Performance of the proposed solution is studied analytically in section 5. In section 6, the proposed mechanism is implemented and some experiments are performed, reported and analyzed. Finally the paper is concluded and future works are discussed.

RELATED WORK

Buyya et al. (2008) surveyed Content Distribution Networks, from fundamental concepts like Request Routing (RR) and Replica Placement (RP) techniques to pricing strategies. The architecture of Hybrid CDN-P2P networks and their benefits are discussed by Dongyan and Kulkarni (2004). Yin et al. (2009) presents the first commercial implementation of Hybrid CDN-P2P network. Despite the fact that there has been a significant body of research on Replica Placement, load balancing and multicasting in traditional CDNs, Oliveira and Pardalos (2005), the only published work on RP in HCDN-P2P is Jiang, et al. (2009) and except Garmehi and Analoui (2011) there exists no published work on load balancing and multicasting mechanisms in this field. Our previous work presents a centralized solution to the problem and suffers from scalability and reliability concerns.

Traditional IP and overlay multicasting are well-known and developed fields in computer networks, Ramanathan (1996), Oliveira and Pardalos (2005). Most of the

works in this area model the underlying network using graph theory and try to solve problems mathematically or using various optimization approaches, e.g. Parsa, et al (1998). Mares (2008) showed that the algorithmic equivalent of the problem in this domain, most of the times, depending on constraints and goals, belongs to the set of NP-Hard problems. As a result of this fact it can be observed that most of the works in this area are concentrated around approximate and heuristic algorithms, e.g. Wu and Chao (2004). Farley, et al. (1999) and Connamacher and Proskurowski (2003) introduce solutions for multi source spanning trees.
Because of the fact that multicasting at IP layer, demands updating all or most of IP Routers and Switches which seems impossible, most of the works in multicasting are concentrated around overlay networks, Rezvani and Analoui (2010). In this work, a shared overlay infrastructure is considered as the underlying network and a multicasting mechanism inspired from Walrasian general equilibrium, is introduced. As these works just try to solve the problems in a shared overly network, application of a fully competitive mechanism is sufficient to overcome the selfishness of users. compared to the research performed by Rezvani and Analoui (2010), in this work, user selfishness problem is not an issue but there are several other constraints and goals in the network. So it can be said that, in this work the key reason for application of economic mechanism design theory is to overcome the high number of constraints and goals and also to provide a communication and processing efficient solution to the problem.

A comprehensive discussion on spanning trees and optimization problems is presented in Wu and Chao (2004). In this reference, some classical problems with direct relations with this problem are discussed. Table 1 reveals features of each problem and similarities and dissimilarities of each of them with this problem.

Table 1: Comparison of various related problems in spanning trees and optimization domain with our problem

Problem	Number of source nodes	Capacity Constraint	Latency constraint	Link cost constraint	Processing Complexity	Objective
SMT	All the nodes might be the source	No	No	Yes	NP-Hard	To construct a partially spanning tree spanning an exact set of nodes with minimum cost
MCCT	1	No	No	Yes	NP-Hard	To find a communication tree spanning some of the nodes with minimum communication cost
CST	1	Yes	No	Yes	NP-Hard	To find a Steiner minimum tree considering the link capacities
DCST	1	No	Yes	Yes	P	To find a spanning tree with minimumdiameter
Our Problem	N	Yes	Yes	Yes	NP-Hard ?	To find a set of trees spanning a subset of the nodes considering link capacities with minimum communication cost where the communication latency remain below a predefined threshold.

Considering table 1 it can be obviously seen that most of these classical problems are NP-Hard. Our problem seems to be a combination of these problems and it seems reasonable to make it a candidate to examine if it belongs to the set of NP-Hard problems or not. The best heuristic solutions to these problems most of the times, rely on Spanning Tree or General Routing approximation of the problem and as a result, these heuristic solutions at least demand an $O(n^3)$ of computation complexity in centralized implementation Wu and Chao (2004).

Because of the fact that the traditional multicasting problem is more complicated than P, there are many works dealing this problem with a game theoretic approach. In Nisan, et al. (2007) many game theoretic approaches with various objectives are discussed. These works lack study of communication and processing demands and as a result, most of them suffer from high network bandwidth consumption or unacceptable processing requirements.

Multicasting Problem in Hybrid CDN-P2P Networks

In traditional CDNs, like Akamai, users send their requests to a DNS based RR system. This system considering parameters like geographical location of the user, load of each Edge Server and probably traffic intensity on network paths, redirect the client to the most suitable server containing a replica of the requested content. Replica Placement system of CDNs work in concert with the RR system. This system decides to put replicas of the each content on which servers. The decision of replicating the content on servers is made considering several factors like popularity of the content, replication costs and available resources on each server.

The structure and a performance model of a Hybrid CDN-P2P architecture is introduced in Yin et al. (2009). Like traditional CDNs, in Hybrid CDN-P2P networks, two major cooperating subsystems (RR and RP) manage and control all the activities of the system but in these networks these activities must be done in two correlated layers, CDN and P2P. At CDN layer, RR mechanism may consider several factors like the geographical location of the users and servers, performance metrics of each server and traffic intensity on the links. At this layer, RP subsystem considering several parameters like popularity of content and load of servers and links, decides to put replicas of each instance of content on which Edge servers. The output of RP subsystem is a progressively updating mapping between servers and contents. The rate at which this list is updated depends on variation of the rate at which the replicas of the content on each server are referred by the clients. These variations may occur as a result of the changes in the popularity or changes in the working modes of edge servers as a result of switching between P2P and Client-Server modes Li, et al. (2009).

At the P2P layer, due to the limitations on storage capacity of peers, it is not possible to use full-site Replica Placement approach which is widely used in CDNs Lei, et al. (2011), instead, in this layer partial-site Replication must be considered. This fact makes it necessary to make an independent decision for replication of every object on peers and increases the load on both RP and RR subsystems. In addition to this fact, the dynamic nature and churning of peers in this layer makes it impossible and inefficient to use centralized mechanisms. Furthermore, in order to multicast the content among the peers, a distributed multicasting tree construction algorithm is needed.

Like CDNs, Hybrid CDN-P2P networks have the responsibility of content distribution and this burden is put on their shoulder by content providers through a Service Level Agreement (SLA). Due to many technical limitations, content providers are restricted to choose just one CDN provider to contract a content distribution SLA Pathan and Buyya (2007). Therefore from client's point of view there is a monopoly in distribution of the content in Hybrid CDN-P2P networks and unlike traditional P2P networks, Analoui and Rezvani (2010), the selfishness of peers in content distribution is not an issue. At P2P layer of these networks the key challenge is to acquire the resources from the peers and allocate these resources in necessary multicasting trees. Multicasting of content in this layer must be performed considering many requirements including latency threshold for streaming media content, capacity of communication logical links between peers and traffic intensity on these links. In order to obey the SLA all the multicasting trees used to distribute the content among peers must respect the requirements listed before, but the Hybrid CDN-P2P as a selfish agent must try to maximize his utility. This utility maximization goal can be achieved by minimizing the expenses spent on delivery of the content at this layer. In other words at this layer there is a constraint optimization problem. Resources or expenses in this domain include network bandwidth on both the Edge server and peers. In this problem as the performance is defined as the constant throughput, required for distribution of streams, it does not play any role in the optimization problem and just puts a strict constraint on multicasting trees.

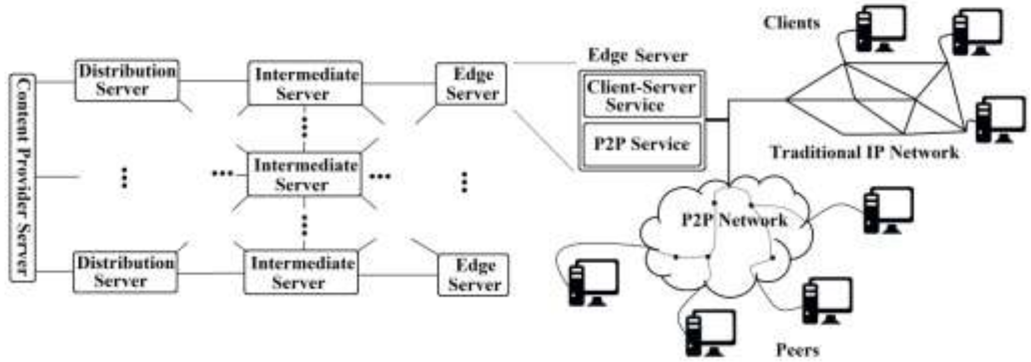

Figure 1. Architecture of a Hybrid CDN-P2P network

Considering figure 1, a content provider must provide some distribution servers of Hybrid CDN-P2P network with the flow of content from its origin server. As mentioned before, in order to preserve reliability concerns, most of the times the original content is delivered to more than one distribution server of Content Delivery System. Upon receiving the content at distribution servers, the responsibility of CDN to deliver the content to its audience begins. In order to carry the content to clients, first it should be transferred to edge servers. These edge servers are exactly those introduced by RP algorithm and to make the content available on these servers, it is unavoidable to use some multicasting trees within the internal layer of the Hybrid CDN-P2P network from distribution server with aid of intermediate servers and to edge servers. These multicasting trees should minimize the cost which the CDN provider pays for transmitting the content from the distribution servers to the edge servers.

At the edge servers, if it is possible to use P2P network for distribution of the content, the content must flow in a shared P2P network. At this layer, because of the fact that connected P2P networks in Hybrid CDN-P2P networks have higher performance Akbari, et al. (2008), it is assumed that it is possible for the content to be distributed using more than just one edge servers connected to the same P2P network. In this P2P network the edge server determines the amount of bandwidth each peer must pay to receive a predefined level of quality of service. The problem here is how to organize the P2P network to distribute the content from the P2P service at the edge servers to the peers.

In order to introduce the problem at both CDN and P2P layers, two problems (Problem 1 and 2) are defined and in order to give a uniform solution, they are mapped on a generalized and abstract form of the multi-source multicasting problem (Problem 3). Problem 1 defines the problem of content multicasting at internal layer of Hybrid CDN-P2P network:

Problem 1:
Given: Directed graph G=(V,E), the set of nodes in G, is partitioned to three isolated sets, D_G (Distribution Servers), I_G (Intermediate Servers) and E_G (Edge Servers). Flow of the content (C_i) with specified bandwidth (Bw_{ci}) and predefined threshold of latency (Max_Lat_{ci}) available on a predefined set of Distribution Servers ($D_{ci} \subset D_G$) which must become multicast with aid of Intermediate Servers (I_G) to Edge servers ($E_{ci} \subset E_G$) using links available in dedicated network ($L(v1,v2) \in E$) with known but not necessarily constant latency (Lat(v1,v2)), capacity (Max_Cap(v1,v2)), transmission price (Trans_Price(v1,v2)) and utilization (Util(v1,v2)).

Goal: To find and maintain, necessary multicasting trees rooted at D_{ci} set of sources, covering some of Intermediate Servers and all of Edge Servers in E_{ci}, considering latency constraint of content C_i such that no links in infrastructure witness traffic intensity or utilization more than threshold $\alpha < 1$. Among all possible answers the one with the least amount of total cost (Objective Function) is desired.

Objective Function: Total cost of content distribution at this layer is defined as the sum of transmission costs of content from links belonging to set of multicasting trees.

$$\text{Minimize:} \quad \text{Total\_cost}_{ci} = \sum_{l_{i,j} \in Multicasting_{trees}} Bw_{ci} . Trans\_price_{i,j} \qquad \text{(Eq.1)}$$

At the edge servers, considering the popularity of the content, one of P2P or Client-Server modes will be used for distribution of an instance of the content. If the popularity of the content be above the bootstrapping threshold, it will be possible and economic to deliver it using the P2P approach, Cheng, et al. (2008). As connected P2P meshes provide better performance, Akbari, et al. (2008), in P2P mode the content might be accessible for the clients from more than one source, Edge Server. At this layer also a same multicasting problem should be solved. Problem 2 defines the problem of multicasting the content at internal layer of Hybrid CDN_P2P network.

Problem 2:
Given: Flow of the content (or a chunk of that) (C_i) with specified bandwidth (Bw_{ci}) and predefined threshold of latency (Max_Lat_{ci}) from predefined set of Edge Servers ($E_i \subset E$) which must become multicast with aid of peers (P) to requesting peers ($P_i \subset P$) using links available in overlay network ($L(P_i,P_j) \in L$) with known but not

necessarily constant latency ($Lat(p_i,p_j)$), capacity ($Cap(p_i,p_j)$), transmission price ($Trans\_Price(p_i,p_j)$)and utilization ($U(p_i,p_j)$).

Goal: To find and maintain, necessary multicasting trees rooted at E_i set of sources, covering some of intermediate peers and all of requesting peers, considering latency constraint of content C_i such that no links in infrastructure witness traffic intensity or utilization more than threshold $\beta < 1$.

Objective Function: Total cost of content distribution at this layer is defined as the sum of transmission cost of content from links belonging to set of multicasting trees.

Minimize:
Total cost of content distribution=

$$\sum_{l(p_i,p_j)\in Multicasting\_trees} Bw_{ci}.Trans\_Price(p_i,p_j) \qquad \text{(Eq. 2)}$$

As it can be obviously seen, these two multicasting problems are correlated and have a same nature and goal. So these problems are mapped to an abstract problem no. 3 and a same solution is provided for both of them.

Problem 3:
Consider a directed graph like G. In this graph nodes are denoted by $V=\{v_i\}$. Edges of this graph are denoted by the set $E=\{e(i,j)\}$. Each link of this graph like e(i,j) has four parameters, Lat_e, U_e, Cap_e, and $Price_e$. These parameters respectively indicate the latency, utilization, capacity of the link and price of a unit of the bandwidth on the link.

Flow of content C is available on set $S \subset V$, it demands bandwidth B_c and it should be delivered to the set $D_c \subset V$. It is allowed to use each of nodes in V-S-D as a relay of the flow. The content C can tolerate at most Max_Lat_c of latency along its path from edge servers to the requesting peers. As the links of the graph are abstraction of communication links, it is essential to avoid utilization near their capacity and a predefined constant like α, e.g. 0.95, is considered as the soft threshold of their utilization Analoui and Jamali (2008). The goal is to select a sub-graph of this graph in such a way that;

Constrains: for every member of D, a path from one of nodes in S be available such that sum of latencies of links on this unique path be below the latency threshold of C. No link on this path is allowed to have utilization above the threshold α.

$$\forall\, v_j \in D_c \quad \exists\, Path_{i,j} \,|\, v_i \in S_c \; And \; \sum_{e\in Path_{i,j}} Lat_e < Max_{Lat_c} \qquad \text{(Eq. 3)}$$

$$\forall\ e \in Multicasting\ Trees \quad U\_e \leq \ \alpha \qquad \text{(Eq. 4)}$$

Goal:

Minimize sum of costs on links present in this sub graph.

$$Minimize\colon\ Total\_cost_c = \sum\nolimits_{e \in Multicasting\ Trees\ for\ C} B_c . Price_e \qquad \text{(Eq. 5)}$$

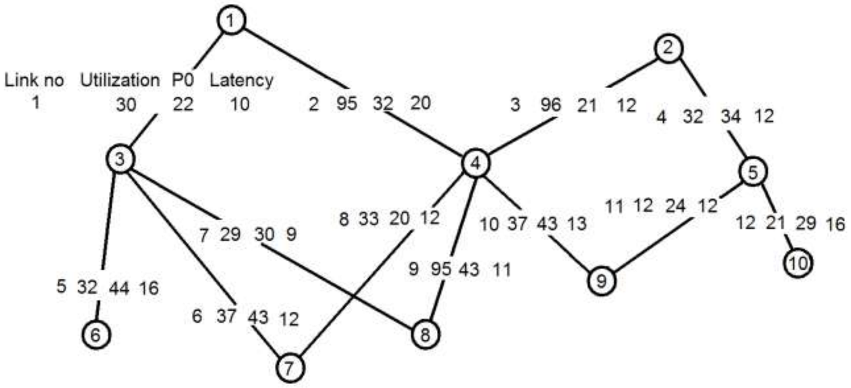

Figure 2. An example of the infrastructure

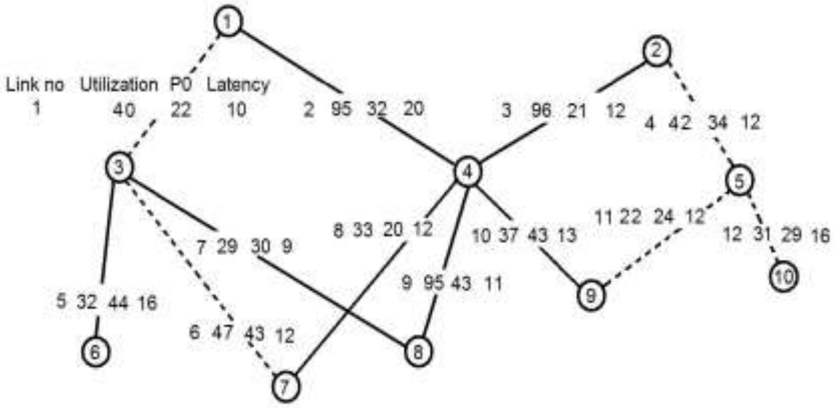

Figure 3. Multi-source multicasting trees; answer of example infrastructure; Dashed lines are included in multicasting trees

An instance of the problem is introduced in figure 2. This figure illustrates the graph of the underlying network. In this network, nodes 1 and 2 are assumed to be sources of the flow. Destinations of the flow introduce gradually to the mechanism, but at this moment they are assumed to be nodes number 7, 9 and 10. It is assumed that all of the links have equal capacities of 10 Mbps. The content is assumed to be a stream with CBR nature, 1 Mbps speed and maximum allowed latency of 30 msec. It can be observed that the answer must be a configuration illustrated in Fig. 3. It is notable that having a graph with more than a handful number of nodes it is impossible to find out the answer manually.

With this set of multicasting trees no link will be over utilized, at leafs and all the intermediate nodes of these trees the latency witnessed by the content remains below the threshold and minimum expense is spent to multicast the content.

As discussed before and like most of multicasting problems the Multi-source Multicasting Problem belongs to NP-Hard problems.

Theorem: The Multi-Source Multicasting Problem belongs to NP-Hard problems.

Proof: The proof of this theorem is inspired from the proof given for NP-Hardness of Steiner Minimal Tree problem by Wu and Chao (2004).

Suppose that the problem of multiple multicasting tree generation is not NP-Hard and there exist a P solution for it. In order to make it possible to use the proof given in Wu and Chao (2004), constraints on maximum allowed latency and capacities of links must be removed from the problem. The first step of this process is performed using a threshold of maximum allowed latency with a value more than sum of the latencies on maximum latency path in the network. In order to eliminate the capacity constraint it must be assumed that on all the links in the network the free capacity is more than the required bandwidth plus 1-α. In order to maintain the connectivity of the result, it is also assumed that the number of sources is limited to 1.

Having a look at features of the Steiner Minimum Tree Problem it can be concluded that this version of the solution is able to solve it. As the SMT problem belongs to the set of NP-Hard problems, and our hypothetic solution has solved it with a P complexity the contradiction appears. As a result the first assumption of existence of a P solution for the Multi Source Multicasting Problem is rejected.

The proposed solution

As mentioned before the problem of Multi-Source Multicasting Trees is composed of many NP-Hard problems and as proved above, belongs to the set of NP-Hard problems. As a result it is not possible to find a polynomial processing complexity solution to this problem. In fact even for small numbers of nodes, using exhaustive search it takes too long to find the answer of the problem. For example calculation of the solution of a problem of size 50 on a traditional PC, 3 GHz with 2 GB of RAM, takes around 10 hours.

Since the number of constraints and goals in this problem is considerably high, it does not seem possible to design effective heuristics for this problem. Economic mechanisms have shown to be effective means to solve complicated problems and in this work the foundations of the solution is based on them.

In Hurwicz and Reiter (2006), the problem of mechanism design is defined as: Given a class Θ of environments (here all the Multi source multicasting problems), an outcome space Z (Single dimension cost function), and a goal function F, find a privacy preserving mechanism π=(M, μ, h) that realizes the problem. M is the message space used to solve the problem and yield equilibrium, μ is the function used for generation of messages and h is a function which determines the outcome. In other words the problem is solved and the best mechanism is chosen if:

$$\forall\theta \in \Theta \quad F(\theta) = h(\mu(\theta)) \qquad \text{(Eq. 6)}$$

Mechanisms might be privacy preserving or not. Mechanisms with Privacy preserving feature demand minimum of communication among the agents. In other words when there is needed to minimize the communication among the agent implementing a mechanism, it is sufficient to rely on a privacy preserving version of the mechanism.

Economic mechanisms are classified to direct and indirect. As in direct mechanisms, all the decisions are made based on private information of each agent plus information which is transferred to it by other agents and no efforts are made to reveal private information of other agents, these mechanisms demand the least possible processing costs. In addition, it is possible to convert all the indirect mechanisms to their direct equivalent. In order to guaranty that in the implementation of the solution, minimum communication and processing costs are spent; the solution is based on a direct and privacy preserving mechanism.

In order to implement economic mechanisms, two approaches exist. Differential approach relies on differential equations and is more common but it is limited to problems with few variables and is suitable for most of the pure economic problems. Despite most of the existing applications of economic mechanisms, due to the high number of variables and constraints, design of the economic mechanism for this problem is made using the sets approach. The mechanism used to solve the problem is a competitive market based on the spanning tree approximation for Steiner Minimal Tree problem. In this mechanism relying on a pricing strategy, tuned by results reported by Analoui and Jamali (2008), the capacity constraint on the linked is embedded in the solution. The latency constraint of the problem affects the result by excluding unacceptable paths from the market as a macroeconomic decision.

In general, the pricing function depends only on the agents in the market and there is no direct function to predict the prices, Jehle and Reny (2011), but in this domain, like our previous work Garmehi and Analoui (2011) the price of a unit of bandwidth is changed based on its utilization. A link has a constant natural price for a unit of bandwidth, when its utilization is below the congestion control threshold. If the load of a link goes above the congestion control threshold its price will increase nonlinearly and in order to prevent the multicasting mechanism to include this link, its price will have an extremely high value near full link utilization. Considering the fact that the problem of multi source multicasting belongs to NP-Hard problems, having an applicable P solution for it is impossible. As discussed before this problem has tight relationship with the Steiner Minimum Tree problem and it seems reasonable to rely on existing heuristics for SMT to construct a solution for this novel problem.

```
//Process No 1: insert a node in $MSMT_c$
 if node $\in$ V-S                                  // If the node is not a source node
O(log|V($MSMT_c$|))
{
  Run shortest path algorithm from v$\in$ V($MSMT_c$) to the new node
                                   //Auction initiation: Shortest Path calculation     O(|V($MSMT_c$)|$^2$)
  Construct a sorted list of the paths based on their costs.
                                   // Competitive market Mechanism  O(|V($MSMT_c$)|log(|V($MSMT_c$)|))
  Remove paths with latency more than $Lat_c$        //Applying latency constrain on the market     O(|V($MSMT_c$)|)
  Determine the winner & Update Utilization and Bandwidth price on all the links present on the winner path
                                   // Decision on constrained market                O(|V($MSMT_c$)|)
                                   // Update Propagation on the path                O(|V($MSMT_c$)|)
  Add the winner paths nodes to  $MSMT_c$            // Update $MSMT_c$                            O(|V($MSMT_c$)|)
  }
else // node $\in$ S
  {
    Run shortest path algorithm and find new shortest paths from this node v$\in$ V($MSMT_C$)          //
O(|V($MSMT_c$)|$^3$)
    Cluster the nodes based on their distances to members of S                        //O(|V($MSMT_c$)|$^2$)
    Remove all the members of $MSMT_C$ which has classified in a group other than their original distribution server and
    put them in a set named Temp.                                                     //O(|V($MSMT_c$)|$^2$)
   Remove members of Temp one by one and insert them into $MSMT_c$
     //These nodes are not Sources and as a result the time complexity of this action will be     O(|V($MSMT_c$)|$^2$)
   Update the parameters of new $MSMT_c$ and its Links.                               //  O(|V($MSMT_c$)|)
   }


//Process No 2: Remove a node n from $MSMT_c$
{
 Construct a sorted list named Temp_list of  nodes in $MSMT_c$ obtaining the content from this node  // O(|V($MSMT_c$)|)
 Remove all the nodes obtaining the content c from this node                          // O(|V($MSMT_c$)|)
 Insert all the nodes of Temp_list into $MSMT_c$    //These nodes are not Sources and as a result, the time complexity of
                          // insertion of them into the $MSMT_c$ using Process No 1 will be     O(|V($MSMT_c$)|$^2$)
Update the parameters of new $MSMT_c$ and its Links.                                  // O(|V($MSMT_c$)|)
}
```

Algorithm 1 pseudo code for two processes of insertion and deletion of nodes in MSMT mechanism

In order to minimize the processing and communication demands of the solution, the solution is embedded in a privacy preserving direct economic solution. As the solution is needed to be able to work online, the solution consists of two correlated insertion and deletion processes. Algorithm 1 describes these processes. These two processes cooperate to construct and maintain a multiple source multicasting tree for the flow of the content from its sources to the destinations.

Although the mechanism is designed based on SMT approximation of Steiner Minimum Tree but an essential variation makes it essentially different. At the insertion phase, (Process No. 1) the new node not only considers the entire leaves (Edge Servers or downloading peers) but also it considers all the intermediate nodes. As the number of these nodes in worst case is comparable with number of leaves, this decision does not change the time complexity of the process but it eliminates the need for performing a cycle finding and removal process after each insertion. Considering the fact that breaking the cycles of the graph demands visiting all the nodes, this decision not only simplifies the insertion process but also it decreases the processing demand. This feature of the solution also has a positive effect on the depth of the generated trees.

Decentralized version of the solution relies on a simple message passing process among nodes of the graph, as the economic agents. In order to manage and coordinate this process for a MSMT, the first source is selected as the controller. This node receives the request of insertion and deletion and distributes it among the other nodes. All the nodes of the graph have ability to implement distributed version of shortest path algorithm. They also maintain statistics of all output transmission links, Price, Utilization, Latency and Capacity of them and a list of all the available source nodes for each instance of content. These agents also are capable to issue insertion and deletion requests to the controller node. In order to make it possible to maintain the tree like structure for upward and downward traversal, these agents are equipped with ability to maintain list of the child nodes in $MSMT_c$ which it should feed them with flow of the content and also they know their parent node. The ability to receive the list of available paths from nodes in MSMT and capability to report the cost and intermediate nodes of a winner path makes it possible for an agent to connect itself to the $MSMT_c$ when it has requested to join the multicasting mechanism.

This multi agent mechanism is implemented using object oriented C++ programming language and is compiled using Borland C++ version 5.02 compiler. All the mechanism details are implemented as member functions and variables of the classes. In order to verify the correctness the implementation of the solution a simple exhaustive search solution for the problem is implemented also. As the problem is NP-Hard and it might be expected, the time takes to obtain the results using exhaustive search version increases dramatically with increasing the number of nodes to more than a handful number. So verification is made only against small size inputs (below 50 nodes).

Analytical Performance Evaluation

Resource consumption is the first concern with distributed algorithms. Hurwicz and Reiter (2006) show that many game theoretic solutions lack from study of their implementation resource consumption and many of them are not possible to implement. Also Hurwicz and Reiter (2006) introduce ways to calculate the minimum dimension of message space as a measure of communication demands of economic mechanisms.

The solution provided in this paper for the Multi Source Multicasting Trees problem is a privacy preserving solution because an agent in this mechanism transmits only necessary information to the market. As it is also visible in the description of the introduced mechanism for an insertion request, every node only transmits the information of its connected links. The only transmitted parameter between agents is transmission cost of the flow from a neighboring link. Because of the fact that a market mechanism is employed to solve the problem and in this market the quantity and quality of the product (required bandwidth and maximum latency) is constant, the dimension of the message space is one. In other words a privacy preserving solution will only transmit messages with just one parameter. This fact is visible at the solution and proves that the solution is privacy preserving. In the terminology of algorithms it can be said that the algorithm (solution) demands a complexity of O(n) of communication resources and in other words from this point of view it is scalable.

The solution is a direct mechanism. In other words, in the provided mechanism all the agents rely just on the information provided directly by the other agents and they do not try to interpret this information or forecast other parameters of other agents. These agents just sum the costs of data transfer from the path. Considering the fact that through the solution, no processing but a decision making is made on the information provided by the agents. Therefore it seems reasonable to classify the solution as a direct mechanism and as a result it consumes as less as possible processing resources as possible. Considering time complexities provided in front of each line of the pseudo code of the algorithm 1, the time complexity of the solution on each agent belongs to $O(n^2)$ and it is of $O(n^3)$ processing complexity in total. Considering other multicasting problems which are less complicated than this problem, it can be said that the decentralized Multi Source Multicasting solution provided here is of the less possible time complexity.

It is reasonable to expect to have a minimum cost MSMT when the capacity and latency constraint are not preventing multicasting trees from using desired paths. However, if congested or high latency paths where present in the network infrastructure, the next applicable and low cost path must be selected.

Theorem: Results obtained from MSMT when the links are not congested and latency constraint does not prevent, are at most 2 times more than optimum.

Proof: Assume a MSMT for content c, like MSMT_c in a network without high traffic intensity and with a maximum latency threshold more than the maximum latency on the longest path in the network (normal situation). In this graph select one of the source nodes. Connect all the source nodes (roots of the trees) with bidirectional, zero cost, unlimited capacity and zero latency links to the selected node. This action does not change the total cost, structure and goal of the problem. This simple change in the input date, transforms the results to a connected Steiner tree. This Steiner three includes all the source nodes, covers all the destination nodes and some of other nodes. When the traffic intensity of the links and latency constraints are not preventing the process to choose lowest price links, the result will be equal to a spanning tree approximation of the Steiner tree problem. The proof introduced by Wu and Chao (2004), duplicates the links and relies on the minimum Hamiltonian cycle of the graph to prove the theorem. As in this problem the links of the graph are abstractions of communication links and the graph is assumed to be directed, this assumption makes no problem and as a result it can be said that this proof is valid in this domain and the outcome of the mechanism in normal situations in worst case costs two times of the optimum (but unknown) result.

Table 3: Results for the example 1

Distribution Servers are servers with ID 1 and 2	Link ID▼	Inserted or deleted Edge server ID in each step (+ means insert and – means remove) ▼					
		+7	**+9**	**+10**	**-10**	**-7**	**-9**
		Utilization of links in each step ▼					
The content is a 1 Mbps flow with Maximum allowed latency of 30 msec All the links have bandwidth of 10 Mbps For more details Look at Fig. 2 and Fig. 3	**1**	40	40	40	40	40	30
	2	95	95	95	95	95	95
	3	96	96	96	96	96	96
	4	32	42	42	42	42	32
	5	32	32	32	32	32	32
	6	47	47	47	47	37	37
	7	29	29	29	29	29	29
	8	33	33	33	33	33	33
	9	95	95	95	95	95	95
	10	37	37	37	37	37	37
	11	12	22	22	22	22	12
	12	21	21	31	21	21	21
Variance of link utilization		911.5	834.3	791.5	834.3	869.7	940.9
STD Dev of link utilization		30.2	28.9	28.1	28.9	29.5	30.7
Average of link utilization		47.4	49.0	49.9	49.0	47.4	45.7
Maximum Latency (msec)		22	24	28	24	24	0
Total cost of Multicasting		650	1230	1520	1230	580	0

Experimental Evaluation

The outcome of MSMT for the problem introduced previously in Fig 2 is presented in table 3. As it can be obviously seen, the result produced for this problem equals with the manual result of the problem in Fig 3. This example not only verifies that the solution works well but also considering the standard deviation of the utilization of links it can be concluded that application of the solution exploits links in a balanced manner.

In order to study the performance of the solution 100 realistic experiments is made. For each experiment a realistic topology is generated using Igen, Quoitin, et al. (2009). Igen is a topology generator which is able to produce realistic internet like topologies with traditional links like STM1, 100 base T etc and spanning various geographical areas. Every topology has 500 nodes. The degree of core nodes is assumed 4 and border nodes are assumed to have degree equal to 2. The geographical area is assumed to be North America. It is assumed that the links already are carrying a self similar traffic with average intensity of 50%. The link latencies are approximated by their length. Also the initial bandwidth prices for the links are assumed to be their length divided by their bandwidth.

As the content, a 1Mbps flow with 200 msec latency threshold is considered. Among the 500 nodes of each network, 100 nodes are selected randomly. The first node is assumed to be a source node. Remaining nodes are assumed to be source with probability of 10% and with 90% probability destination. Remaining nodes may or may not be included in the MSMT.

These 100 nodes are inserted in the MSMT one by one and averaged total cost of 100 MSMTs is reported in figure 4. In order to study the dynamic behavior of the solution, after insertion of 100 nodes for 100 iterations, one of 100 nodes is selected randomly, removed from the MSMT and another node from outside of the MSMT is selected randomly and inserted into the MSMT. Averaged results of 100 iterations of this experiment are reported in figure 5.

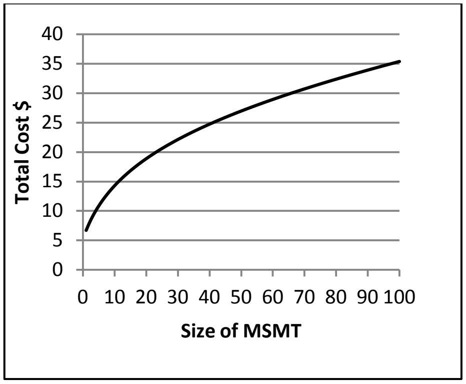

Fig. 4. Total cost of multicasting increasing the size of MSMT to 100.

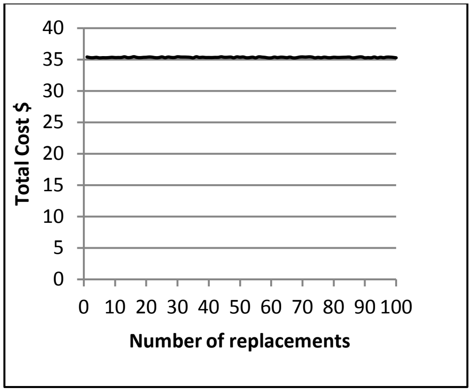

Fig. 5. Averaged Total cost of multicasting during 100 replacements in MSMT with size of 100.

Considering the curve illustrated in Fig 4 it can be said that as it is expected the cost spent for content distribution increases with the number of participants. When the size of multicasting group increases the total cost curve becomes smoother. In figure 5, after the initial phase when 100 nodes inserted into the multicasting group for 100 iterations the nodes are replaced. As it can be seen, the average of cost spent for content distribution in this phase remains approximately constant. Based on this observation it can be concluded that the solution works well in dynamic situation and it does not make it necessary to compute the MSMT again after churning of peers or breakdown of servers.

Conclusion

In this work, considering Hybrid CDN-P2P networks, a novel form of multicasting problem introduced. Based on neighboring problems in graph theory domain, proved that the problem belongs to the NP-Hard problems. Relying on high potentials of economic mechanism design theory and heuristic solutions of Steiner tree problem, a communication and processing efficient solution to the problem introduced. It proved that in normal situations the distributed solution provides a 2 approximate result for any given problem. Experimental results vividly show that the solution provides acceptable and balanced results without the need to recalculate the Multi Source Multicasting Trees after continuous changes in multicasting groups.

In this paper we solved the multicasting problem in Hybrid CDN-P2P Networks. In these networks Replica Placement and Request Routing are still open problems and remain as future works.

Acknowledgements
This research was supported by Iran Telecommunication Research Center (ITRC), Contract Number, 19251/500.

REFERENCES

[1] **Akbari, B., H. R. Rabiee, and M. Ghanbari (2008), *An Optimal Discrete Rate Allocation for Overlay Video Multicasting.*** *The Journal of Computer Communications, Elsevier* 12:4, 551–562;

[2] **Analoui, M. and S. Jamali (2008), *Congestion Control in the Internet: Inspiration from Balanced Food Chains in the Nature.*** *The Journal of Network and Systems Management Springer*, 16:1, 1-10;

[3] **Buyya, R., M. Pathan, and A. Vakali (2008), *Content Delivery Networks*.** *Published by Springer*;

[4] **Connamacher, H. S. and A. Proskurowski (2003), *The Complexity of Minimizing Certain Cost Metrics for k-source Spanning Trees.*** *The Journal of Elsevier Discrete Applied Mathematics*, 131:1, 113 – 127;

[5] **Dongyan, X. and S. Sunil Kulkarni (2004), *A CDN-P2P Hybrid Architecture for Cost-Effective Streaming Media Distribution*.** *The Journal of Elsevier Computer Networks*, 44:3, 353-382;

[6] **Farley, M., P. Fragopoulou, P. Sud, D. Krumme, A. Proskurowski, and D. Richards (1999), *Multi-Source Spanning Tree Problems.*** *The Journal of Interconnection Networks***,** 1:1, 61-71;

[7] **Garmehi, M. and M. Analoui (2011)**, ***An Economical Mechanism for Multicasting of Content among Servers of Hybrid CDN-P2P Networks.*** *6th International Conference on Internet Technology and Secured transactions. Abu Dhabi, United Arab Emirates*, 566-571;

[8] **Yin, H., Liu, X., Zhan, T., Sekar, V., Qiu, F., Lin, C. & Li, B. (2009), *Design and Deployment of a Hybrid CDN-P2P System for Live Video Streaming: Experiences with LiveSky.*** *In Proceedings of the 17th ACM international conference on Multimedia*, 25-34;

[9] **Jehle, G. A. and P. J. Reny (2011), *Advanced Microeconomic Theory.*** *Prentice Hall;*

[10] **Jiang, H., H. Wang, K. Wong, J. Li, and Z. Li (2009), *A Replica Placement Algorithm for Hybrid CDN-P2P Architecture.*** *IEEE Computer Society 15th International Conference on Parallel and Distributed Systems. Shenzhen, Guangdong, China*, 123-131;

[11] **Hurwicz, L. and S. Reiter (2006), *Designing Economic Mechanisms*.** *New York: Cambridge University Press*;

[12] **Lei, J., G. Camarillo, G. Yingjie, D. Montuno, and L. Xie (2011)**, ***Survey of P2P streaming applications.*** *Draft, IETF, Huawei*;

[13] **Mares, M. (2008), *The Saga of Minimum Spanning Trees.*** *The Journal of Elsevier Computer Science Review*, 2:3, 165-221;

[14] **Nisan, N., T. Roughgarden, E. Tardos, and V. V. Vazirani (2007)**, ***Algorithmic Game Theory.*** *Cambridge University Press*;

[15] **Oliveira, Carlos A.S., and Panos M. Pardalos (2005), *A Survey of Combinatorial Optimization Problems in Multicast Routing.*** *The Journal of Elsevier Computers & Operations Research,* 32:8 1953–1981;

[16] **Parsa, M., Q. Zhu, and J. J. Garcia (1998), *An Iterative Algorithm for Delay-Constrained Minimum-Cost Multicasting.*** *The Journal of IEEE/ACM TRANSACTIONS ON NETWORKING,* 6:4, 461-475;

[17] **Pathan, A. M. and R. Buyya (2007), *Economy-based Content Replication for Peering Content Delivery Networks.*** *IEEE 7th International Conference on Cluster Computing and. Rio de Janeiro, Brazil, May 14-17,* 887-892;

[18] **Quoitin, B., V. Van den Schrieck, P. Francois, and O. Bonaventure (2009), *IGen: Generation of Router-level Internet Topologies through Network Design Heuristics.*** *In 21st International Teletraffic Congress, ITC 21 , Paris*, 1–8.

[19] **Ramanathan, S. (1996), *Multicast Tree generation in network with Assymetric Links.*** *The Journal of IEEE/ACM Transactions on Networking,* 4:4, 558–568;

[20] **Rezvani, M. H., M. Analoui (2010), *An Economic Model for Multi-service Overlay Multicast Networks Based on Walrasian General Equilibrium.*** *Economic Computation and Economic Cybernetics Studies and Research,* 44:2, 139-161;

[21] **Seyyedi, S., B. Akbari (2011), *Hybrid CDN-P2P Architectures for Live Video Streaming: Comparative study of connected and unconnected meshes.*** *International Symposium on Computer Networks and Distributed Systems (CNDS). Tehran, Iran,* 175–180;

[22] **Wang, Z., H. Jiang, Y. Sin, J. Li, and J. Liu (2009), *A k-coordinated Decentralized Replica Placement Algorithm for the Ring-based CDN-P2P Architecture.*** *IEEE Symposium on Computers and Communications (ISCC). Beijing, China,* 811-816;

[23] **Wu, B. Y. and K. M. Chao (2004), *Spanning Trees and Optimization Problems.*** *New York: Cahpman & Hall/CRC Press*.