Computer Networks 83 (2015) 100–117

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Content distribution by multiple multicast trees and intersession cooperation: Optimal algorithms and approximations

Xiaoying Zheng [a,*], Chunglae Cho [b], Ye Xia [c]

[a] Shanghai Advanced Research Institute, Chinese Academy of Sciences and Shanghai Research Center for Wireless Communications, China
[b] Electronics and Telecommunications Research Institute (ETRI), Daejeon, Republic of Korea
[c] Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL 32611, United States

ARTICLE INFO

Article history:
Received 27 August 2014
Received in revised form 1 March 2015
Accepted 4 March 2015
Available online 10 March 2015

Keywords:
Multi-tree multicast
Optimization
Tree packing
Universal swarming
Column generation
Subgradient algorithm

ABSTRACT

In traditional massive content distribution with multiple sessions, the sessions form separate overlay networks and operate independently, where some sessions may suffer from insufficient resources even though other sessions have excessive resources. To cope with this problem, we consider the universal swarming approach, which allows multiple sessions to cooperate with each other. We formulate the problem of finding the optimal resource allocation to maximize the sum of the session utilities and present a subgradient algorithm which converges to the optimal solution in the time-average sense. The solution involves an NP-hard subproblem of finding a minimum-cost Steiner tree. We cope with this difficulty by using a column generation method, which reduces the number of Steiner-tree computations. Furthermore, we allow the use of approximate solutions to the Steiner-tree subproblem. We show that the approximation ratio to the overall problem turns out to be no less than the reciprocal of the approximation ratio to the Steiner-tree subproblem. Simulation results demonstrate that universal swarming improves the performance of resource-poor sessions with negligible impact to resource-rich sessions. The proposed approach and algorithm are expected to be useful for infrastructure-based content distribution networks with long-lasting sessions and relatively stable network environment.

1. Introduction

The Internet is being applied to transfer content on a more and more massive scale. While many content distribution techniques have been introduced, most of the recently introductions are based on the *swarming* technique, such as FastReplica [1], Bullet [2,3], Chunkcast [4], BitTorrent [5], and CoBlitz [6]. In a swarming session, the file to be distributed is broken into many chunks at the source node, which are then spread out to the receivers; the receivers will then help each other with the retrieval of the missing chunks. By taking advantage of the resources of the receivers, swarming dramatically improves the distribution efficiency (e.g., average downloading rate, completion time) compared to the traditional client–server-based approach.

The swarming technique was originally created by the end-user communities for peer-to-peer (P2P) file sharing. The subject of this paper is how to apply swarming to infrastructure-based content distribution and make the

* Corresponding author.
E-mail addresses: zhengxy@sari.ac.cn (X. Zheng), clcho@etri.re.kr (C. Cho), yx1@cise.ufl.edu (Y. Xia).

http://dx.doi.org/10.1016/j.comnet.2015.03.004
1389-1286/

distribution more efficient. Compared with the dynamic end-user file-sharing situation, such infrastructure networks are often of medium size, consisting of hundreds of nodes, dedicated, and have much more stable nodes and links. In this setting, we will see that it is beneficial to view swarming as distribution over multiple multicast trees. There are more incentives to share resource among different sessions. This view allows us to pose the question of how to optimally distribute the content (see [7]).

The specific problem addressed in this paper is how to conduct content distribution more efficiently in a network where multiple distribution sessions coexist. A distribution *session* consists of a file to be distributed, one or more sources and all the nodes who wish to receive the file, i.e., the receivers. Different sessions may have heterogenous resource capacities, such as the source upload bandwidth, receiver download bandwidth, or aggregate upload bandwidth. For instance, there may exist some sessions with excessive *aggregate* upload bandwidth because their throughput bottleneck is at the source upload bandwidth, the receiver download bandwidth, or the internal network; at the same time, there may exist some other sessions whose throughput bottleneck is at their aggregate upload bandwidth. In the traditional swarming approach, the sessions operate independently by each forming a separate overlay network; this will be called *separate swarming*, which does not provide the opportunity for the resource-poor sessions to use the surplus resources of the resource-rich sessions. However, if we conduct *universal swarming*, that is, we combine multiple sessions together into a single "super session" on a shared overlay network and allow them to share each other's resources, the distribution efficiency of the resource-poor sessions can improve greatly with negligible impact on the resource-rich sessions, provided the resource-rich sessions have sufficient surplus resources. The paper examines algorithms and theoretical issues related to universal swarming.

In universal swarming, a distribution tree not only includes all the receivers interested in downloading the file but may also contain nodes that are not interested in the file; the latter will be called *out-of-session nodes*, and the source and receivers are called *in-session nodes*. Thus, each distribution tree for a session is a *Steiner* tree rooted at the source covering all the receivers and the out-of-session nodes on the tree are Steiner nodes.[1]

To illustrate the main ideas, consider the toy example in Fig. 1(a). The numbers associated with the links are their capacities. Node 1, 2 and 3 form a multicast session, and a large file is distributed from the source node 1 to the receivers, node 2 and 3. Node 4 is an out-of-session node. Suppose the file is split into many chunks at node 1. To distribute a chunk to the receivers, the chunk must travel down some tree rooted at the source and covering both receivers. All possible distribution trees are shown in Fig. 1(b). Observe that, except the first tree, the other three trees include the out-of-session node. Fig. 1(b) shows an optimal rate allocation with respect to the objective of maximizing the total distribution rate, or equivalently, minimizing the required distribution time. The scenario is an example of universal swarming since the out-of-session node is used. For separate swarming, only the first tree can be used and the distribution rate is only one half as much.

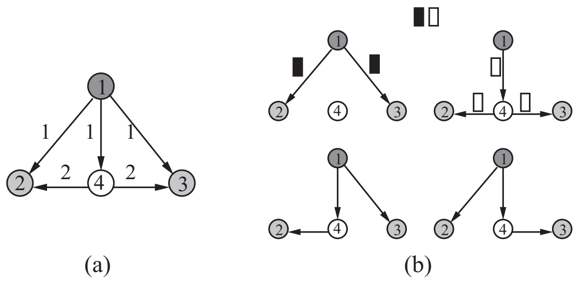

Fig. 1. An universal swarming example. (a) Node 1 distributes a file to node 2 and 3; node 4 is an out-of-session node. The numbers next to the links are the link capacities. (b) All possible distribution trees and the optimal solution for throughput maximization. The boxes represent file chunks. Three of the trees involve the out-of-session node 4.

With the tree-based model, the optimal distribution problem can be formulated as finding an optimal rate allocation on the multicast trees so that it achieves the optimal performance objective. A version of this problem was considered in [7] and its longer version [11] in the context of separate swarming. The rate-allocation problem in universal swarming, which this paper concerns, is substantially more difficult. The main reason is that, by the optimality condition, an optimal solution typically uses only the minimum-cost (min-cost) trees to distribute the file chunks. In [7], for each multicast session, an overlay network consisting of only in-session nodes is constructed above the underlay physical network, where the topology of each overlay network is pre-specified. The algorithm for separate swarming in [7] only needs min-cost *spanning* trees; finding a min-cost spanning tree is considered an easy problem since polynomial algorithms exist. In contrast, an optimal universal swarming algorithm usually involves repeatedly finding a min-cost *Steiner* tree, which is an NP-hard problem. How to cope with this difficult issue is one of the main themes in this paper. The approach proposed in [7] is unable to handle this difficulty. On the positive side, since universal swarming corresponds to a less restricted way of doing multicast than separate swarming, performance improvement is expected. The degree of improvement can sometime be large.

We present two solution approaches, which can be used in combination. First, we incorporate into our rate-allocation algorithm a column generation method, which can reduce the number of times the min-cost Steiner tree is computed. Second, we allow the use of approximate solutions to the Steiner-tree subproblem. Such approximate solutions on directed graphs can be found in [8–10]. When the above two methods are put together, the combined algorithm is rather difficult to analyze. For the most part, there are little standard results that can be used directly to prove algorithm convergence or to give an approximation ratio to the rate-allocation problem when approximate min-cost trees are used in each iteration step.

[1] Given a directed graph $G = (V, E)$, and a subset $V' \subseteq V$ of vertices, a Steiner tree is a connected and acyclic subgraph of G which spans all vertices of V'. More information can be found in [8–10].