

Peer-Based Efficient Content Distribution in Ad Hoc Networks*

Seung-Seok Kang

Department of Computer Science
Seoul Women's University
Seoul 139-774, Korea (ROK)
msukang@swu.ac.kr

Abstract. Mobile devices pay the telecommunication cost of downloading Internet data proportional to the amount of data transferred. This paper introduces a special *ad hoc* network in which several mobile devices, called peers, to cooperate each other to reduce the overall cost to download Internet content. Each peer downloads a specific portion of the content over 3G connection and exchanges the portion with other peers over the *ad hoc* connection in order that all participating peers are able to reconstruct the whole content. This paper proposes a peer-based content distribution method, and compares its performance with a similar one named *per-packet* based distribution method. The simulation results indicate that *per-peer* based method outperforms the *per-packet* based method. In addition, approximately 90% of the telecommunication cost is saved with as few as 10 peers.

Keywords: ad hoc network, 3G, peers, content distribution.

1 Introduction

Current wireless telecommunication services provide high speed Internet access as well as voice service. Fig. 1 displays a situation in which many nearby mobile devices connect to the Internet via their wireless telecommunication links to their ISPs and then access their favorite content. Suppose the devices in the figure try to download the same file stored at a content provider (CP), such as mobile game programs, MP3 files, or movie clips. This situation may happen in many places. For example, a teacher may want to share some educational content with his/her students indoor and outdoor. Some friends may download an interesting mobile game program that they want to store on their mobile devices and play together interactively. In a sports stadium, spectators may want to retrieve the records of their home team and favorite players in a game. Since it is expected that the 3G connection cost to download data from the Internet is a function of the amount of data downloaded, the cost of the telecommunication connections to access the Internet may be reduced when each mobile device is assigned to download a given portion of the target file and shares its portion with other devices. This

* This work was supported by a research grant from Seoul Women's University (2006).

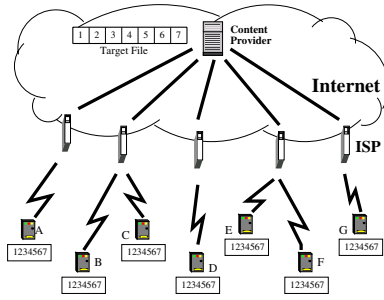


Fig. 1. Mobile devices download same target file in the Internet

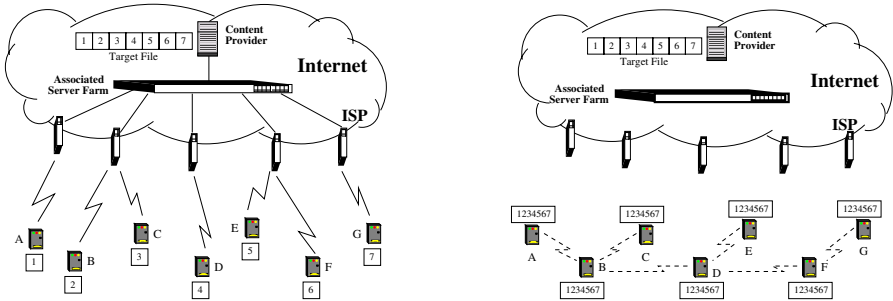


Fig. 2. Downloaded portions are exchanged over the *ad hoc* connection

paper proposes a low-cost sharing schemes in which mobile devices download in parallel their assigned portions of the target file via their 3G connections, and then build an *ad hoc* network (that has no fee for data transferred) to exchange the remaining portions of the file. This may be possible if each mobile device has both a 3G interface for wireless wide area network (WWAN) access such as UMTS or CDMA2000 and a wireless local area network (WLAN) interface, such as 802.11 or Bluetooth, to form an *ad hoc* network [1].

The left side of Fig. 2 illustrates the mechanism for partitioning the target file and downloading the assigned portions by the mobile devices. Each mobile device is assigned a portion of the target file to download. Each device connects to its favorite ISP with its 3G link and contacts the CP with the aid of its associated server in order to request its assigned portion of the file. The main difference between Fig. 1 and the left side of Fig. 2 is the thickness of the 3G connection, which represents the amount of data downloaded for each mobile device. In Fig. 1, each mobile device downloads the *entire* content from the CP. However, the thin line in the left side of Fig. 2 indicates that each device downloads only a *portion* of the target file. Upon completion of downloading the specified portion of the file, the mobile devices use their *ad hoc* connections to exchange

their content with other member devices. The right side of Fig. 2 shows that all mobile devices participate by exchanging their partial content with others in order to reconstruct the target file.

The idea of the cooperating *ad hoc* network concept is often described as *peer-to-peer* computing [2]. In [3], the peers download an assigned portion of a target file and distribute their portions to all other peers based on the *per-packet* based method. Fairness issues of the multiple sources to single receiver (multipoint-to-point) is studied in [4]. SplitStream [5] allows a source peer to split a file into k stripes and to multicast each stripe using k multicast trees. In BitTorrent [6], when many downloaders try to download a file(s) from a URL-specified location, the downloaders upload to each other concurrently to help reducing the load of the source. The authors in [7] proposed an approach for computing a schedule for coordinated data collection with avoiding congested links and maximizing the network resource utilization.

This paper focuses mainly on the data distribution method among participating peers. This method is based on a *per-peer* policy in which all peers are ordered to transmit their packets one peer at a time. They all know which peer is its predecessor. When a peer transmits some number of data packets, the content is delivered only to the sender's one-hop neighbors. At the end of the current peer's transmission, the peer sends a special packet that triggers the transmission of the next peer. The transmission process finishes when all peers receive the complete target file.

The rest of the paper is organized as follows. Section 2 deals with two distribution methods. Simulation results are explained in section 3. Section 4 draws conclusions.

2 Data Distribution Methods

2.1 Network Formation

The work in [3] describes the special *ad hoc* network formation. One of the device users, called a peer, initiates a connection to its ISP and contacts a special server on the Internet. The server becomes the associated server of the peer. Each peer needs to have its associated server. The associated server may be located at the ISPs, the CPs, or some other place within the Internet. A server may be associated with several peers. For some cases, all peers may connect to a single server that controls the operation of the *ad hoc* network and deals with any license issues and fees with CPs at a bulk rate on behalf of the peers. The associated server of the initiating peer plays an additional role of the master server. The master server deals with several management tasks for the special *ad hoc* network including the computation of the download scheduling that decides which peers to download which portions of the target file.

Each peer downloads only a portion of the target file, but needs to acquire complete content. In order to reconstruct the file, every peer becomes a sender as well as a receiver. However, if all peers, as senders, transmit their content in an uncontrolled manner, the broadcast storm problem [8] may arise. Each

peer needs a controlled way of broadcasting its received data to other member peers. A *per-packet* based distribution method [3] selects the rebroadcasting set of peers in order to propagate content to other peers because some peers are out of transmission range of others. When a peer transmits one unit-sized packet of downloaded data that is new to other peers, the data packet travels over the *ad hoc* network by rebroadcasting from some selected rebroadcasting peers, if needed. If all peers receive the packet, the scheduled next peer transmits another unit-sized packet of downloaded data. *Per-packet* transmission repeats until all peers receive the complete content.

2.2 Per-Peer Based Distribution

In the *per-packet* distribution method, one packet from one peer is forwarded to all other peers at a time. Then, the next scheduled peer takes a turn to transmit its next unit-sized packet. This delivery method, however, shows some inefficiency while transmitting packets. Peers located at the central area may experience heavy packet collisions. In addition, for some moment, non-data packets such as DONE packets may dominate the *ad hoc* network, which increases the completion time and degrades the transmission performance.

A *per-peer* based distribution method decreases the chance of packet collision and increases the time for more data packets transmitted, which results in shorter completion time, less 3G communication costs, and less power consumption than that of the *per-packet* based method. In the *per-peer* based distribution, only one peer has a chance to transmit several unit-sized data packets to its neighbors at a time. In addition, the data packets are not immediately forwarded, which reduces the possibility of packet collision. Each peer independently decides which unit-sized data to broadcast depending on the reception status of its neighbors. The neighbors do not rebroadcast any packet immediately, but wait for their transmission turn. When the current transmitting peer has finished broadcasting a given number of packets, it broadcasts a DONE packet. If a scheduled next peer is directly connected with the current peer, the next peer resumes its transmission after selecting which data to send. If the next peer is out of range from the current peer, the DONE packet contains a list of rebroadcasting peers that forward the DONE packet toward the next-scheduled transmitting peer.

The master server creates and maintains a transmission sequence as a circular list of peers using global topology information. Fig. 3 illustrates one example of a 3-hop *ad hoc* network. The master server constructs the list with the following rules. The server prefers the peer that has the largest number of neighbors. Then it selects the peer that is directly connected with the previously selected peer. If there is no unselected peers that have direct connection with the recently selected peer, the server chooses the peer with the minimum number of hops from the previously selected peer. Fig. 3 (b) is one possible example of the transmission sequence as a circular list. Peer B is selected first and subsequently peer E, A, and D are chosen. Because peer D and F have no direct connection, the link

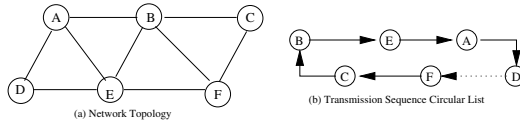


Fig. 3. A 3-hop network topology and its transmission sequence circular list

is shown as a dotted line. Peer D keeps a rebroadcasting set that includes peer E for rebroadcasting a DONE packet. In addition, each peer maintains two-hop neighborhood information. Much research have proposed distributed algorithms that utilize two-hop neighborhood information [9,10,11,12]. When a peer receives a packet, it immediately knows which peer also receives the same packet. This knowledge may obviate some unnecessary transmissions of data packets.

2.3 Benefit Value

One important value that each peer maintains for each unit-sized data segment is the number of neighbors that do not have the segment. Because the target file size is known and the data packet size can be defined, e.g., 500 bytes, each peer is able to compute how many unit-sized data are needed to construct the target file. Suppose the target file size is 1 Mbyte and the unit packet size is 500 bytes, 2048 data packets are necessary. Each peer maintains both a bitmap size of 2048 bits and the same-sized integer array called the *benefit* value. The bit in the bitmap indicates the existence of the corresponding data segment stored in the peer. The integer benefit value implies the number of neighbors that do not have the corresponding data segment. In case that a peer downloads some portion of the target file from the Internet, the benefit value(s) of the data segment(s) is(are) set to the number of its neighbors, because no neighbor has the data segment. The corresponding bit is also set, because the segment is in the peer. When each peer resumes transmission, it selects the unit-sized data that has the largest benefit value, because the corresponding number of neighbor peers do not have the data and they will get a benefit to fill the gap in the target file. This also leads to increases a chance to discourage unnecessary data transmission by other peers with the aid of the two-hop neighborhood information. In Fig. 3, assume peer B sends a data packet whose benefit value is four, and peer A, C, E, and F receive the packet. Peer B's benefit value for the data becomes zero after the transmission, because all its neighbor peers hold the data this time. Both peer A and E computes a value of 1 for the data's benefit value, because only peer D does not have the data. Further assume that peer E transmits the data in its turn. Now peer A's benefit value of the data becomes zero and peer A is unnecessary to transmit the data in case of its turn to transmit. If a peer has all zero benefit values, it immediately sends DONE packet when it is the peer's turn. This allows other peers with positive benefit values to transmit data.

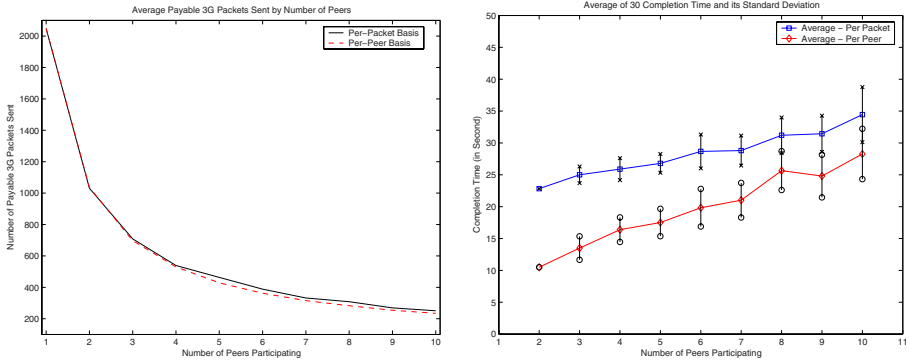


Fig. 4. Number of packets on the 3G link and completion time by varying number of peers

3 Simulation Results

The *ns2* network simulator [13] is used in this simulation. The simulation model assumes each peer downloads a portion of the target file using its 3G connection that the master server schedules to download, while it exchanges the portions with other peers using its *ad hoc* network. The peers do not experience buffer underrun when they transmit their portions to other peers. The unit packet size of the 3G connection is set to 500 bytes. The *ad hoc* network uses the same size of data packets. The peers download the target file size of 1 Mbytes that consists of 2048 unit packets. Each peer has the 802.11 MAC with the transmission range of 250 meters. The number of peers varies from 2 to 10. Because of the short amount of completion time described in [3], the mobility of peers is not considered. All peers are located in a 400 meter by 400 meter grid unless specified otherwise. In case of *per-peer* based distribution, each peer transmits the maximum of 90 packets at a time because the packet buffer size is set to 100 in the simulation. Each peer uploads its new neighborhood information in a second. In all cases, an *ad hoc* network is connected. That is, each peer is connected to at least one other peer.

The left side of Fig. 4 displays the average number of 3G packets used by the peers that may have a fee charged by a telecommunication provider. Each value in the figure is the average of 30 runs. As the number of peers increases, the number of fee-based packets decreases substantially. However, an additional peer only reduces a marginal cost when there are already enough peers, whereas, the new peer receives the same cost reduction. Approximately 90% of the telecommunication cost is saved with as few as 10 peers. The *per-packet* based method uses slightly more packets on the 3G telecommunications link than that of *per-peer* based method. This is due mainly to the recovery process in which each peer uploads its bitmap reporting any data gaps to its associated server. *Per-peer* based distribution resolves the gap by exchanging bitmaps with its neighbors over the cost-free *ad hoc* connection.

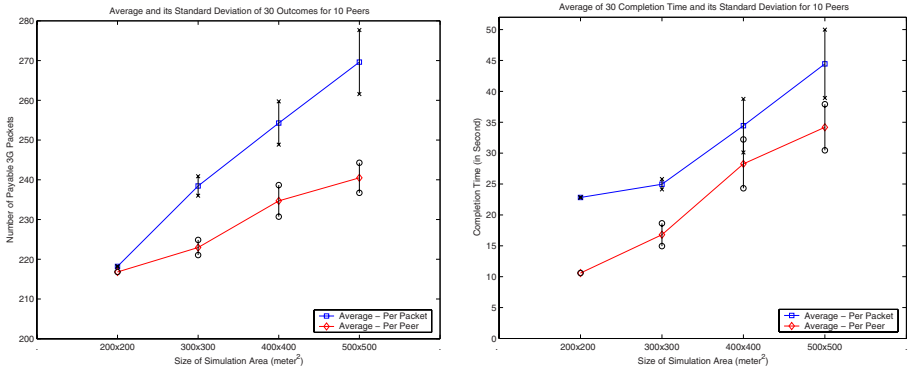


Fig. 5. Number of packets on the 3G link and completion time by varying size of simulation area

The completion time of each simulation is shown in the right side of Fig. 4. Both horizontal lines are the average of 30 runs and the short vertical lines represent the standard deviation of the 30 results. The average completion time increases slowly as more peers participate. This is due to the increasing number of hops in the network. As the number of peers increases, they may disperse over the grid, which may increase the number of hops. More hops cause more time to complete the distribution. The *per-peer* based distribution completes the content distribution faster than that of the *per-packet* based one. One important point is that as the number of peers increases, the difference of the two completion times decreases. This result indicates that the number of hops in a network is a more dominant factor than the density because *per-peer* based distribution hardly experiences any collision.

When all peers are located within the transmission range of each other, no collisions are expected because only one peer transmits at a time. Fig. 5 illustrates how the performance changes due to the change of the grid size. The simulation runs with 10 peers and repeats 30 times with 30 different network topologies. The left side of Fig. 5 shows the number of packets on the 3G telecommunications link as the size of simulation area increases. A larger area may result in a larger number of hops in a network, and more time to take to complete the content distribution. Because peers upload their neighborhood information periodically, longer completion time leads to increase the number of packets on the 3G link as well. In addition, peers upload their bitmaps and download the recovery instructions for *per-packet* based distribution, which consumes packets on the 3G link. Due to more collisions in a larger area, the *per-packet* based method uses more 3G packets than that of the *per-peer* based method. The right side of Fig. 5 shows the completion time by varying the simulation area size. The largest number of hops was 3 among 120 different topologies used in the figure. As the simulation size increases, so does the number of hops, which causes longer completion time. Overall, the *per-peer* based distribution method

outperforms the *per-packet* based method in terms of the completion time, and uses less packets on the 3G link.

4 Conclusion

Telecommunication cost may be one of the crucial factors for mobile device users to access Internet content. This paper describes a special *ad hoc* network in which mobile peers save telecommunication cost by sharing their partially downloaded data with other peers. Each peer agrees to download a specified portion of the target file located in the Internet using its fee-based WWAN connection. Each participating peer distributes its downloaded portion to all other member peers over the cost-free WLAN *ad hoc* connection so that all participating peers can generate the complete target file. The *per-peer* based distribution method utilizes the reception status of its neighbors using 2-hop neighbor information and the benefit value, and outperforms the *per-packet* based method. In addition, both distribution methods save approximately 90% of the telecommunication cost with as few as 10 peers.

References

1. Xiao, Y., Leung, K., Pan, Y., Du, X.: Architecture, mobility management, and quality of service for integrated 3G and WLAN networks. *Wireless Communications and Mobile Computing* **5** (2005) 805–823
2. Oram, A.: *Peer-to-Peer : Harnessing the Power of Disruptive Technologies*. 1st edition edn. O'Reilly & Associates (2001)
3. Kang, S., Mutka, M.: Efficient Mobile Access to Internet Data via a Wireless Peer-to-Peer Network. *IEEE Int'l Conference on Pervasive Computing and Communications* (2004) 197–205
4. Karbhari, P., Zegura, E., Ammar, M.: Multipoint-to-Point Session Fairness in the Internet. In: *Proceedings of INFOCOM 2003*. (2003) 207–217
5. Castro, M., Druschel, P., Kermarrec, A., Nandi, A., Rowstron, A., Singh, A.: Split-Stream: High-Bandwidth Multicast in Cooperative Environments. (2003) 298–313
6. BitTorrent: The BitTorrent file distribution system, <http://www.bittorrent.org>. (2006)
7. Cheng, W., Chou, C., Golubchik, L., Khuller, S., Wan, Y.: Large-scale Data Collection: a Coordinated Approach. In: *Proceedings of INFOCOM 2003*. (2003) 218–228
8. Tseng, Y., Ni, S., Chen, Y., Sheu, J.: The Broadcast Storm Problem in a Mobile Ad Hoc Network. *Wireless Networks* **8** (2002) 153–167
9. Lim, H., Kim, C.: Multicast Tree Construction and Flooding in Wireless Ad Hoc Networks. (2000) 61–68
10. Qayyum, A., Viennot, L., Laouiti, A.: Multipoint Relaying for Flooding Broadcast Messages in Mobile Wireless Networks. (2002) 3866–3875
11. Peng, W., Lu, X.: AHBP: An Efficient Broadcast Protocol for Mobile ad hoc Networks. *Journal of Science and Technology (JCST) - Beijing, China* **16** (2001) 114–125
12. Calinescu, G., Mandoiu, I., Wan, P., Zelikovsky, A.: Selecting Forwarding Neighbors in Wireless Ad Hoc Networks. (2001) 34–43
13. ns2: The network simulator, <http://www.isi.edu/nsnam/ns>. (2006)