TUM I 13
Chair of Application and Middleware Systems
Prof. Dr. H.- A. Jacobsen
Author: Stefano Turchetta, Martin Jergler

Tutorial
Distributed Systems (IN2259)
Pezhman Nasirifard
WS 2019/20

Issued on: 08.01.2020
Due on: 16/17.01.2020

Sample Solution: Exercises on Peer-To-Peer Systems (Part 2)

Exercise 1 - Pastry

Consider the following Pastry network with identifier bit-length $m = 8$ that contains the following nodes:

$$\{\mathbf{0023}, 0113, 0133, 0322, 1002, 1010, \mathbf{1132}, 1223, 2000, 2112, \mathbf{2210}, 2231\}$$

Every node ID $\in \{0,1,2,3\}^4$; hence, $B = 4$, which is the number of rows in the routing table. The leaf set L has size 4.

(a) Specify the routing table and the leaf set of nodes 0023, 1132, 2210 (in bold). The policy for choosing among candidate nodes is a random selection.

Solution: *Note that the tables below represent one possible solution out of many. For instance, the cell [0,1] for node 0023 could also contain 1002, 1132, or 1223 instead of 1010.*

- Node **0023**:

RT	0	1	2	3
0	0023	1010	2112	-
1	0023	0133	-	0322
2	-	-	0023	-
3	-	-	-	0023

Leaf-set: $L = \{0113, 0133, 2210, 2231\}$

- Node **1132**:

RT	0	1	2	3
0	0113	1132	2000	-
1	1002	1132	1223	-
2	-	-	-	1132
3	-	-	1132	-

Leaf-set: $L = \{1002, 1010, 1223, 2000\}$

- Node **2210**:

RT	0	1	2	3
0	0322	1132	2210	-
1	2000	2112	2210	-
2	-	2210	-	2231
3	2210	-	-	-

Leaf-set: $L = \{2000, 2112, 2231, 0023\}$

(b) Write the Pastry routing algorithm in pseudo-code.

Solution:
To handle a lookup message M addressed to a destination node D at a node A (where $R[p,i]$ is the element in row p and column i in the routing table at A):

// CASE 1: The destination is within the leaf set or is the current node.
if ($L_{-|L|/2} < D < L_{|L|/2}$) {
    forward M to the element L_i of the leaf set with ID closest to D, or stop at current node A.
}
// CASE 2: Use the routing table to dispatch M to a node with a closer ID.

else
{
find p, the length of the longest common prefix of D and A, and i is the $(p+1)$-th digit of D.
if ($R[p,i] \neq NULL$)
// route M to a node with a longer common prefix
forward M to $R[p,i]$
// CASE 3: There is no entry in the routing table.
else
{
forward M to any node in Leaf Set or Routing Table with a common prefix of length p but an ID that is numerically closer.
}
}

(c) Imagine node **1312** wants to join the network. For doing so, suppose it contacts node 2210. Illustrate the joining procedure. How many messages will peers exchange?

Solution:

(a) First, the join request reaches node 2210. (one message)

(b) Next, **1312** copies 2210's RT and adds its own identifier. The first row of 2210's RT becomes the first row in **1312**'s RT. (NB: This generates a legal row in **1312**'s RT as 2210 and **1312** share no common prefix.)

(c) Then, 2210 looks at its routing table and forwards the request to 1132. (one message)

(d) Next, **1312** copies 1132's RT and adds its own identifier. The second row of 1132's RT becomes the second row in **1312**'s RT. (NB: generates a legal row in **1312**'s RT as 1132 and **1312** share a common prefix of length one.)

(e) Then, 1132 looks at its RT (or leaf-set) and forwards the request to 1223. (one message)

(f) Node 1223 is the numerically closest existing node in the current Pastry network. Hence, **1312** copies 1223's leaf set (i.e., $L_{1223} = \{1010, 1132, 2000, 2112\}$) and adjusts it to $L_{1312} = \{1132, 1223, 2000, 2112\}$.

(g) Last, **1312** tells the leafs in its newly generated leaf-set that it joined ($|L| messages$). 1132 updates its leaf set to $L = \{1002, 1010, 1223, \mathbf{1312}\}$ (i.e., one message) and, all other neighbors/leafs $\{1223, 2000, 2112\}$ update their leaf-sets accordingly. (i.e., 3 more messages)

Number of messages for adding 1312:
- one message to node 2210
- one message to node 1132
- one message to node 1223
- $|L|$ messages to the leaf set

The routing state of node **1312**:

RT	0	1	2	3
0	0322	1312	2210	-
1	1002	1132	1223	1312
2	-	1312	-	-
3	-	-	1312	-

Leaf-set: $L = \{1132, 1223, 2000, 2112\}$

(d) Execute a lookup for key $k = 1000$ from node 2210.

Solution:
2210's routing table redirects us to 1132.
Node 1132 checks the leaf set to look up the key 1000, since 1000 is not within the range of leaf set, 1132 looks 1000 up in the table and discovers that 1002 is the closest, and 1132 forwards 1000 to 1002, which is the node responsible for 1000.

EXERCISE 2 - BITTORRENT

(a) Consider the piece selection strategy used by BitTorrent. What is the advantage of following a rarest first policy? Why is the first piece not rarest first?

Solution: Increase the availability of file pieces. As soon as you download a piece, you will contribute to the file-sharing community by uploading a requested piece to other fellow peers. BitTorrent would like to have file availability as much as possible. It is easy to see that downloading a rare piece first is a good way of spreading content and increasing the availability.
The first piece is random because a peer should have something to upload as soon as possible and usually the download of a rare piece is slower than a common one.

(b) Consider the endgame mode used by BitTorrent when approaching the end of a download. Why is this strategy used within this context? Why not for the entire download?

Solution: BitTorrent wants to have as many seeds (peers that have 100% of some data) as possible, so when a peer is about to complete a download it forwards requests to all its neighbors in order to finish sooner. Of course, you cannot do this for the entire duration of the download because it is redundant and expensive.

EXERCISE 3 - GNUTELLA

To answer the following questions, assume the topology of a Gnutella network shown in Figure 3.1 where all peers maintain state information about messages they have already sent or received. Supposed that all messages are sent and received in a synchronized, round-based manner.

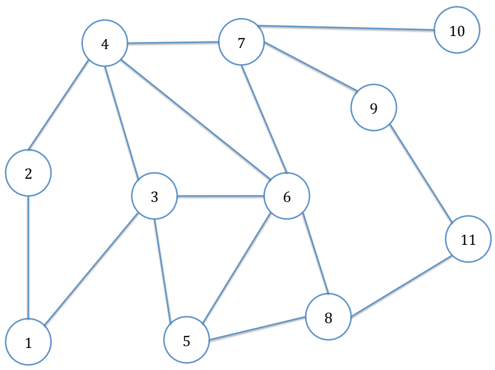


Figure 3.1: Gnutella network.

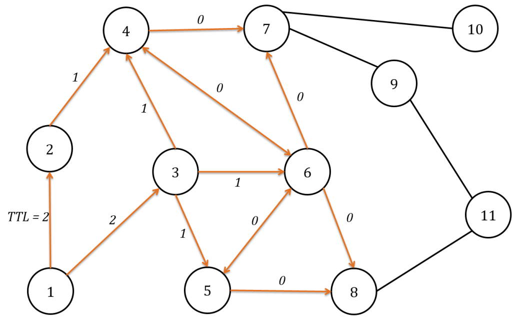
(a) Peer 1 attempts to search for a key k using a $TTL = 3$. Note down all QUERY messages that are sent throughout the network until the TTL expires. Specify the respective round, the sender, the receiver and the TTL attached to each message.

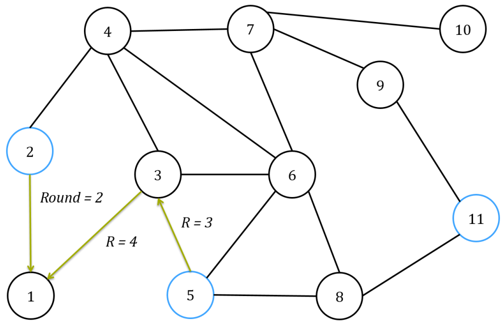
Solution:

(b) Assume that peers 2, 5, and 11 have files matching the keyword k. List all QUERYHIT messages that are sent after the search, described above, has been started. Specify round, sender and receiver.

Solution:

(c) What would happen during the query routing if the peers would not store information about messages that they already received? Describe the problem using an example in the topology given in Figure 3.1.





Solution: