Exercise
for Lecture “P2P Systems”

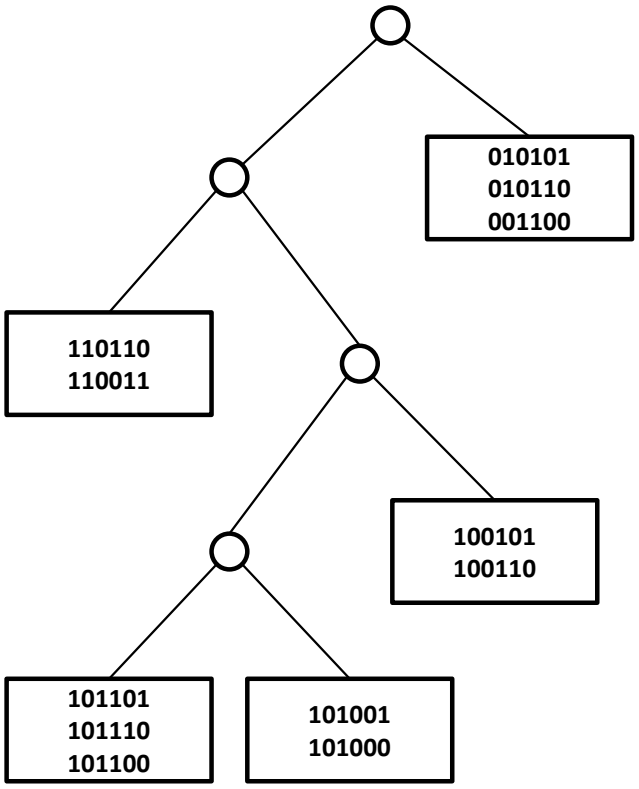
Prof. Dr. David Hausheer
Dipl.-Wirtsch.-Inform. Matthias Wichtlhuber, Leonhard Nobach, M. Sc., Dipl.-Ing. Fabian Kaup, Christian Koch, M. Sc., Dipl.-Wirtsch.-Inform. Jeremias Blendin

TECHNISCHE UNIVERSITÄT DARMSTADT

Summer Term 2015
Exercise No. 10
Published at: 23.06.2015, Submission date: 30.06.2015
Submission only via the Moodle platform in PDF, plain text, or JPG/PNG.
Contact: [mwichtlh|lnobach|fkaup|ckoch|jblendin]@ps.tu-darmstadt.de
Web: http://www.ps.tu-darmstadt.de/teaching/p2p/
– Example Solution –

Problem 10.1 - Kademlia

For the following questions, please consider the following Kademlia routing table for node 101010. Assume that the network has reached steady state. Also assume that all the nodes remain in the network for an unlimited amount of time, unless stated otherwise.



a) How much routing state information is stored in each Kademlia node at maximum in bytes? Assume 26 bytes for each entry consisting of IP address, UDP port, and node ID. The identifier space is 160-bit long. The variable k shall remain a parameter in the result. k is the number of entries per bucket.

Solution: Each node has 160 k-buckets.
Each bucket contains k entries.
Maximum routing state information: 160 * k * 26 => 4,160 bytes

b) What is the value of k (i.e. the maximum bucket size)?

Solution: $k = 3$

c) Which are the nodes present in the k-bucket range for the given node ID, 101010?

Solution: 101001 and 101000

d) Now, let us assume that three new nodes with IDs 001110, 111111, and 110101 join the network, in that order. How will each of these actions affect the routing table for node 101010?

Solution:

- After 001100: No change, node ID dropped; Reason: Node ID does not lie in the k-bucket range for node 101010.
- After 111111: Added to bucket with 110110 and 110011; Reason: $k = 3$ and bucket is not yet full.
- After 110101: No change, node ID dropped; Reason: The corresponding bucket is now already full, and not in k-bucket range.

e) Let us assume that the network has reached steady state after the above changes. We would now like to route a message from node 101010 to node 111100. If $\alpha = 2$, which node IDs are selected from the routing table in the first routing step?

Solution: Nodes 110110 and 111111 are selected.

Problem 10.2 - Distributed Hash Tables

3 Several DHT based approaches have been designed, each one with different limitations and characteristics. Explain in which scenarios you would select:

a) Chord over CAN
b) Kademlia over Chord
c) CAN over Kademlia

Solution:

a) Chord would be preferable over CAN in cases where the network size is unpredictable and the fixed dimensionality of CAN cannot be selected to provide a network with logarithmic diameter with respect to the network site.

b) Kademlia is in general more attractive than Chord because Kademlia's distance metric is symmetrical. This means that the same route between peers A and B may be used in both directions. Therefore, incoming queries may be used to build neighbour connections and in many cases, a number of alternatives exist to forward the incoming queries.

c) As CAN supports a multidimensional ID structure it is possible to map each dimension onto an application specific metrics, such that nodes fulfilling specific requirements can be found much faster.