Prof. Dr. Thorsten Strufe
Stefanie Roos
{strufe,roos}@cs.tu-darmstadt.de

Fachgebiet Peer-to-Peer Netzwerke
Fachbereich Informatik
Technische Universität Darmstadt

Exercise 5

for **Peer-to-Peer Networks** - winter term 2012/2013

(20.11.2012)

Deadline for submission: 27.11.2012 1:30 PM

Guidelines

- Exercises annoted by **G#** are intended to be discussed and solved in class without grading, whereas exercises annoted with **H#** are supposed to be solved in groups and be handed in for grading.
  This does not mean, that ungraded exercises are less important.
- Please submit your solutions until the beginning (1:30 PM) of the next exercise in the coming week. Solutions can either be dropped in the letterbox in front of A110 or handed in personally on the beginning of the exercise. **Electronic submission is no more allowed!**
- Note that points are only given if your solution is clearly legible. Unreadable submissions will not be rated! (Machine written submissions are allowed)
- Written assignments are to be solved in groups of 2-3 participants while programming assignments have to be done in groups of four to six participants.
- Always annotate your solutions on the handed in sheet with names and matriculation numbers. If you have privacy concerns, you are allowed to omit your name and tell it to us personally.
- Please subscribe to the mailing list:
  `https://mail.rbg.informatik.tu-darmstadt.de/mailman/listinfo.cgi/p2p-lecture-ws12`
- By submitting any processed exercises or program code you hereby commit to the "Grundregeln der wissenschaftlichen Ethik am Fachbereich Informatik" (see also `http://www.informatik.tu-darmstadt.de/de/sonstiges/plagiarismus/`).
  This especially means, that you should always write in your own words. However if you use external materials, you have to cite them correctly.
  We will not accept solutions that only rely on literal citations.

G#5.1 Chord

a) Finger Tables

Assume a Chord network with an 8-bit identifier space. Currently, 12 nodes are in the system. Their identifiers are: 0, 8, 13, 16, 17, 31, 36, 44, 49, 52, 55, 60. Compute and note the finger tables of nodes 16, 44, and 55. Use the format indicated by Table **??** for each finger table.

b) Node Join

A new Peer wants to join the Chord overlay described before. Its identifier is assumed to be 58. The peer knows the address of node 16 and contacts it with a join request. Note all exchanged messages until the new node knows its successor on the Chord ring using the format shown in Table **??**. Use links from finger tables as well as successor links to forward messages.

i	Target ID	Successor
1	...	...
2	...	...
...	...	...

Table 1: Finger table format

Sender	Receiver	Message
...	...	...
...	...	...
...	...	...

Table 2: Message format

c) Finger and Routing Table Size

Assume an n-bit Chord identifier space, i.e., id $\in [0, 2^n - 1]$. Further assume, that each IP address requires 32 Bit of storage while ports require 16 Bit. What is the size of a Chord finger table? What is the size of the complete routing table of a Chord node? Omit overhead for managing data structures as well as machine specific memory boundaries.

d) Duplicate Finger Table Entries

We already mentioned that there exist duplicate entries in the finger tables of most Chord nodes. In which scenario would there be no duplicate entries in the finger tables of all nodes in a Chord network? Justify your answer.

H#5.1 Chord

Consider a Chord network with a 4-bit identifier space. The network currently consists of 5 nodes with the following identifiers 0, 3, 8, 10, and 11.

a) Network structure (3 P.)

Sketch a graph which corresponds to the networks structure. Use normal arcs for the successor links (predecessor links should be omitted).
Insert finger links using dashed arcs. Draw duplicate links only once and use normal arcs if a link is both successor and finger link.

The resulting graph is planar, so draw it planary!

b) Characteristic Path Length (3 P.)

Calculate the characteristic path length in the drawn graph. Therefore you can use a table similiar to Table **??**, to denote the distances from node o to node p first.

Is a chord network always strongly connected?

Hint: The characteristic path length is equal to the *average length of shortest paths* between all pairs of nodes, but only in a connected graph. Consider $d(u,v)$ with $u, v \in V$ to be the shortest path from node u to node v and $d(u,v) = 0$ if $u = v$. Therefore the characteristic path length in a connected graph corresponds to $\frac{\sum_{u \in V} \sum_{v \in V} d(u,v)}{n^2 - n}$.

op	0	3	8	10	11
0	...	...	...	...	...
3	...	...	...	...	...
8	...	...	...	...	...
10	...	...	...	...	...
11	...	...	...	...	...

Table 3: Denoting path lengths

c) Average Routing Length (4 P.)

Now calcute the average routing length, meaning the mean average of hops that is needed to really *route* a package from o to node p.

Is the routing in chord done in an optimal way? If not, why can chord not simply be modified to always use the very shortest path?

H#5.2 Routing Complexity

a) Routing Complexity (2 P.)

Explain the difference of average case and worst case complexity on the example of routing in Chord.

b) Unstructured Routing (4 P.)

Consider an unstructured P2P system in which a node queries its neighbors sequentially. Let the node u have k neighbors, of which one has the desired file. In case u queries all its neighbors without repetitions, what is the average number of queries needed, what is the worst-case?
Which complexity class is that?

Assume now due to memory constraints, the node does not keep a list of queried nodes, but chooses a random neighbor each time. What is the number of queries in the worst case and in the average case? complexity class?

Hint: You can use that the expected value of geometrically distributed random variable with parameter p is $\frac{1}{p}$.

H#5.3 Comparing Distance Metrics (5 P.)

Assume we have an n-bit identifier space.
Compare the following distance metrics for nodes $u, v \in V$ with identifiers $id(u)$ and $id(v)$:

i. $d_{\oplus}(u,v) := id(u) \oplus id(v)$[1]

ii. $d_{\mathscr{P}}(u,v) := min\{|id(v) - id(u)|, 2^n - |id(v) - id(u)|\}$

$d_{\oplus}(u,v)$ is for example used by Kademlia and $d_{\mathscr{P}}$ is used by Pastry.

a) What is the maximal distance in both metrics?

b) Assuming a fixed node u. How many neighbours with distance 1 can u have in the different metrics?

c) Assume $n = 8$. Calculate the distances $d_{\oplus}(u,v)$ and $d_{\mathscr{P}}(u,v)$ for the following pairs of identifiers:

[1] $\oplus$ means bitwise XOR

i. $u = 11\ 111\ 110_2, v = 00\ 000\ 000_2$

ii. $u = 11\ 011\ 110_2, v = 01\ 011\ 010_2$

d) Can you find an example where $d_{\oplus}(u,v) < d_{\mathscr{P}}(u,v)$?