Peer-to-peer networks – (due till May 20, 2009)

Exercise 10.1: Bubblestorm

a) A node with high bandwidth has to store

O more data
O less data
O on average as much as the other nodes

Explain how you come to this conclusion.

Solution:

The answer is “more data”. In the beginning, a node establishes a number of edges to itself. The number is proportional to the capacity of the network. A larger number of connections also means a larger number of neighbors. Thus, the likelihood that a neighbor will send copies of data resp. that a neighbor will include the node into its bubble is higher for nodes with a high network capacity.

b) What is a bubble?
c) How does the creation of the network differ from the creation of a bubble?

Solution:

In the beginning the network is created: This happens by Bubblestorm nodes including themselves into the network one or several times. Later, a node which wants to publish information will send copies to its neighbors and keep one copy for itself. Thus, the number of copies is decreased by one and fractions of the remaining copies are forwarded to the neighbors. The set of nodes owning a copy of the data is referred to as a bubble. The bubble is only a small subset of nodes of the entire Bubblestorm network.

Peer-to-peer networks – (due till May 20, 2009)

Exercise 10.1: Bubblestorm

d) What is the split factor in Bubblestorm?

The split factor refers to the number of neighbors copies are send to. In case of a split factor of 3, one third of the remaining copies are forwarded to each neighbor. (Of course no real duplicates are sent. The neighbors get only one and are requested to forward a certain number of copies themselves.)

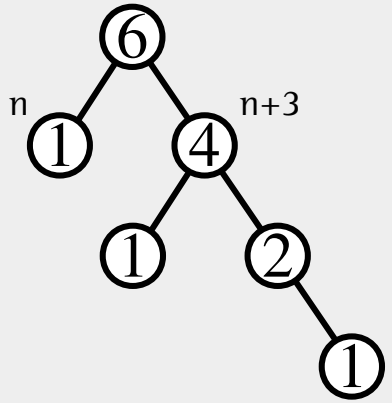
e) Is it possible that a nodes gets multiple copies and if yes, when does that happen resp. why is it impossible?

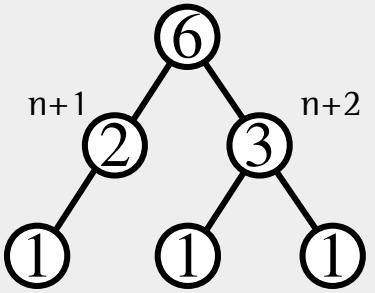
A node can and will include itself multiple times into a network, according to the number of connections it established to itself in the beginning. Since it appears at several places, it can also get the same copy of data several times from different neighbors. The authors assume that the connectivity of a node is large (it has many neighbors) and that the probability of obtaining multiple copies is small, respectively.

f) The authors of Bubblestorm claim that leave nodes of a bubble are only located on level n or (n-1). So the difference of the levels of leave nodes does never deviate by more than one. Explain why this holds true.

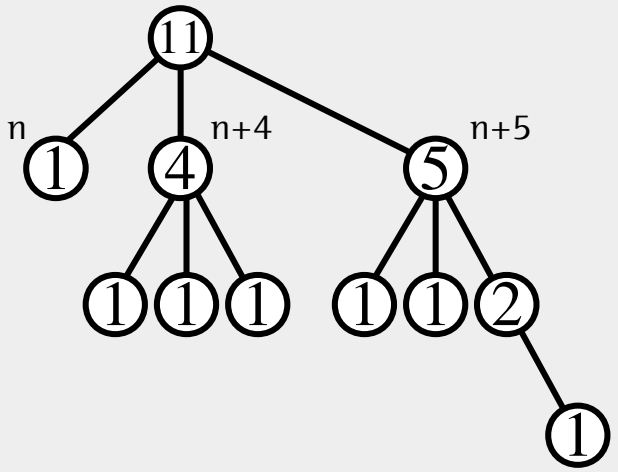
Peer-to-peer networks

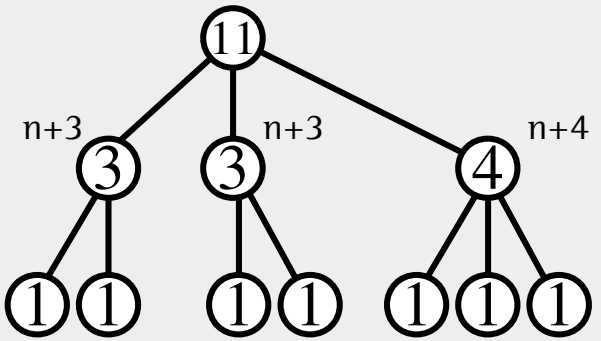
Exercise 10.1: Bubblestorm





split factor s = 2





split factor s = 3

Peer-to-peer networks

Solution:

The upper sketch shows on the left side the minimal configuration resulting in a difference of more than one level for the leave nodes in case of a split factor of 2. Without loss of generality, on the second level, two neighboring nodes spread 1 resp. 4 copies. According to the protocol of Bubblestorm, an equal number (deviating by one in case of remainders from the division) of copies are spread over the child nodes, so 2 and 3 would conform to the protocol as depicted on the right hand side. This obviously results in a more balanced tree with only one level difference. Note that balancing the tree would work likewise in even more unbalanced configurations.

The lower example shows an unbalanced tree, also with two levels of difference and a split factor s = 3. In order to yield a level-difference of two we can state that one node needs to distribute only one more packet (its own packet), another node must distribute s+2 packets (its own packet, another s for the next level and at least one for the lowest level). Only then, we will get a level-different of at least 2. However, the s+2 copies could be evenly distributed over the neighboring siblings of a node. The distribution over the siblings (nodes of the same level) obviously stops if the pairwise difference between neighbors is less than two. Otherwise, the node with the larger number of copies contributes to the node with a smaller number. As some nodes may remain unconsidered, a difference of one (copy) it possible and can not always be avoided.

If neighbors can not differ by more than one copy, the difference for the level of their children can also not differ by more than one.