



HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Overlay and P2P Networks

Worked Example on Bloom Filters

Prof. Sasu Tarkoma

22.1.2013



Worked Example

- Gnutella uses Bloom filters to store and disseminate keyword indexes
- 1-hop replication in the flat ultranode layer, much improved design over flooding
- An ultrapeer maintains about 30 leafs and thus 30 Bloom filters, one for each leaf
- One leaf has about 1000 keywords in our example
- Assuming false positive rate of 0.1, for 1000 keywords we need 4793 bits. For 30 000 keywords we need 143 776 bits.
- There is overlap (some keywords are popular)!
- Gnutella uses 2^{16} (65536) bits that is sufficient even when aggregating leaf BFs
- Experiments report ultrapeer BF 65% full and leaf BF 3% full
- Today having hundreds of KBs in the BF is not a problem, Gnutella design is old and today's networks are much faster

Linear growth if p is fixed!

$$m = -\frac{n \ln p}{(\ln 2)^2}.$$