GAP – Practical anonymous networking

Krista Bennett, Christian Grothoff
PET'03

Presented by B. Choi in cs6461
Computer Science
Michigan Tech

Introduction

- Applications of anonymous communication
  - Electronic payment, voting, auction, email, and web browsing
  - One still not mentioned: File Sharing!
- What's the differences?
  - Query and reply
  - Widely spread P2P application already
- Which platform?
  - GNUnet

P2P file sharing?

- File sharing, what is it?
  - You share I share
  - Equal rights and equal responsibilities
  - Napster in 2000
- P2P in general
  - Structured
  - Unstructured
  - Hybrid
- Security and trust are the primary concern

GNUnet (my impression)

- Unstructured system
  - How to join the system
    - Well known node (distributed registry)
    - Obtain partial membership from the registry
    - Leave? - when you want!
  - Query forwarding
    - Random selection of next nodes
    - Multiple forwarding at each forward
    - Time-to-live to remove loops
  - Reply
    - Encoded blocks
    - Content migration

Query and reply

- Fundamental difference from other applications
  - More query the higher chances to hit a copy
  - One-to-many (file sharing) vs. initiator-and-responder (other applications)
  - Potentially many replies with different blocks of the target file (movie or music titles)
- What needs to be anonymous?
  - Who (identity) is looking for which file?
  - Who (identity) is responding to which query?
  - Sender and receiver anonymity?

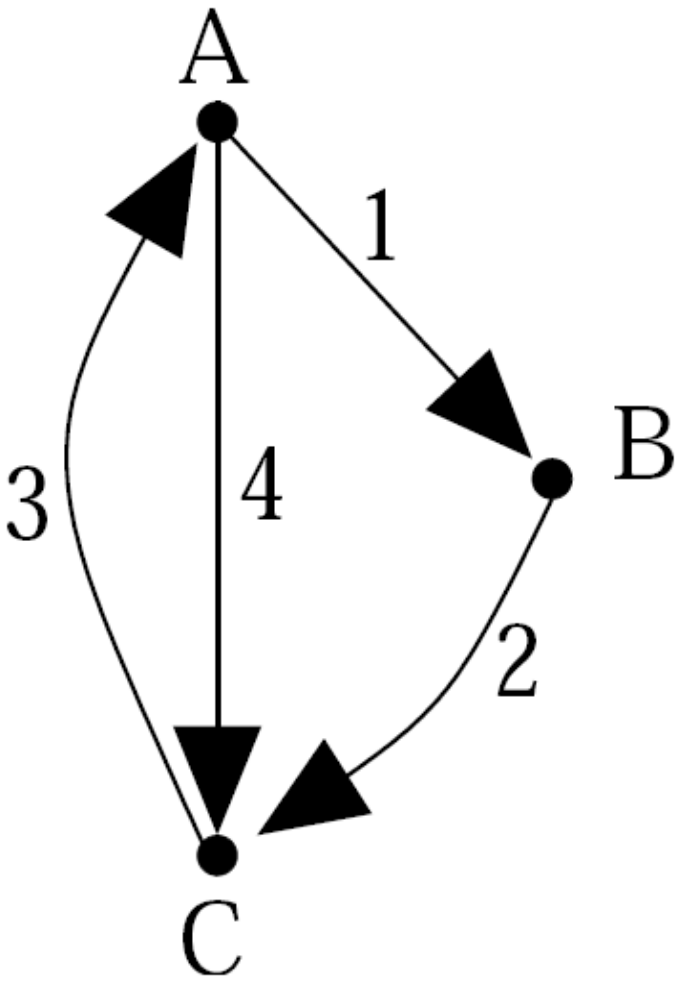
Basic decisions

- To have or not? Cover traffic
  - Not chosen as in many other P2P-based anonymity systems (Crowds, MorphMix, Tarzan?)
  - Why not?
    - Churn
    - Content migration
    - Probabilistic responding
    - Dynamism
- Adversary model
  - External passive, internal active (colluding nodes)

Main idea

- GAP: GNUnet anonymity protocols (my guess)
- For a given time window, a node
  - Creates n queries
  - Forwards m foreign queries
  - Indirect k foreign queries out of m
  - Anonymity of a node is n/(n + m – k)
  - Has to maintain a routing status for each indirect-ed foreign query (how long?)
  - Forwarding/indirecting to a random selection of nodes
  - Decision based on local situations (workload)

Indirecting (not new) vs. forwarding



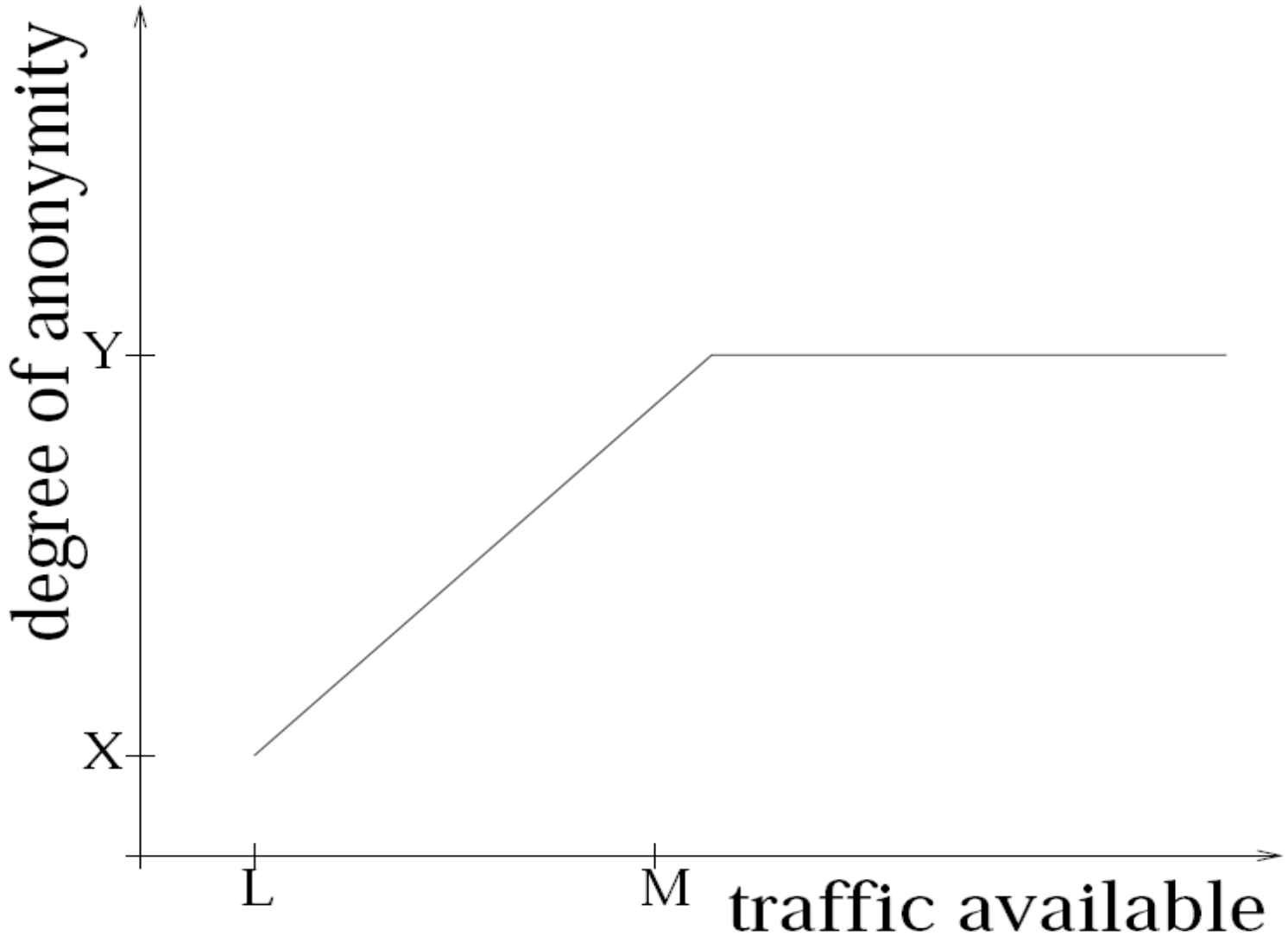
Thoughts on indirect/forward

- What anonymity is affected?
  - Originator?
  - Responder?
  - Forwarder?
- Why would one choose forward?
  - Better efficiency of the node
- What about the system in general?
  - Better efficiency?
  - Higher vulnerability (vs. easy content migration?)

Thoughts on hops-to-live

- Traditional hops-to-live would leak much information
- Solution: TTL --> time window to process a reply
  - TTL + the local time
  - Another hole for the adversary?
  - Delay to process exceptionally soon reply?

Measuring anonymity



Discussions

- GAP: individual node chooses whether to exchange portions of its own anonymity for its own efficiency without impacting the security of other nodes
- Statefulness!
- Any other application not studied in light of anonymity?
  - Social networks (Facebook, etc,..)
  - Skype!
  - Instant chatting?