# A Sybil-proof DHT using a social network

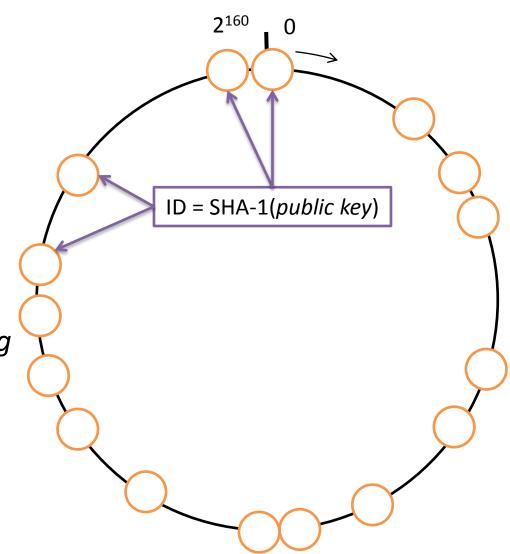
Socialnets workshop
April 1, 2008
Chris Lesniewski-Laas
MIT CSAIL

#### Overview

- Distributed Hash Tables
- The Sybil attack
- Model (network, adversary)
- Tool: random sampling from a social network
- Sybil-proof DHT protocols

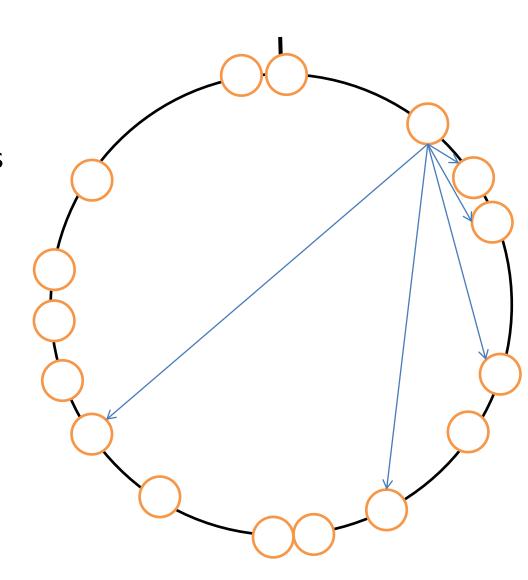
#### DHT routing in three slides

- Structured DHT: a layer in many P2P systems
- Used by requesting node to find another node by ID
  - IDs typically hash of public key: self-certifying
  - DHT maps ID to IP address



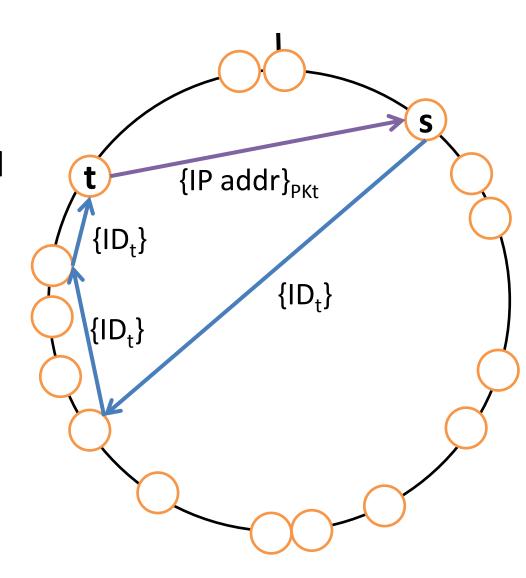
# DHT routing in three slides

- Sub-linear table size
  - Nodes need not keep track of all other nodes
  - Reduces bandwidth usage
  - Enables scaling



# DHT routing in three slides

- Routing via intermediate hops
- Result is authenticated
- Trade off table size versus routing hops



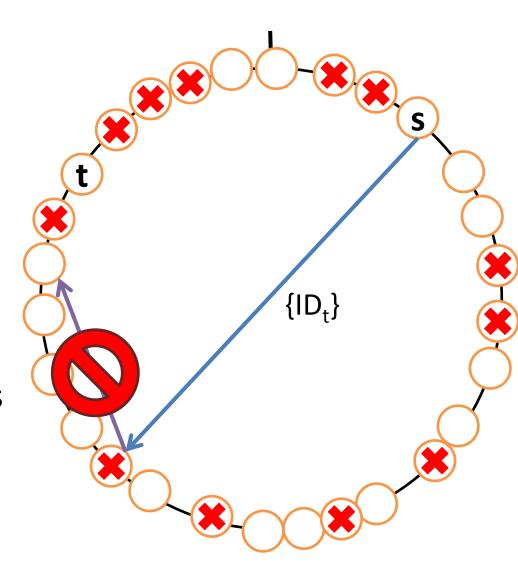
# The Sybil Attack

"One can have, some claim, as many electronic personas as one has time and energy to create."

Judith S. Donath

## DHTs are subject to the Sybil attack

- Attacker creates many pseudonyms
- Disrupts routing or stabilization
- Douceur, 2002:
   "without a logically
   centralized authority,
   Sybil attacks are always
   possible"

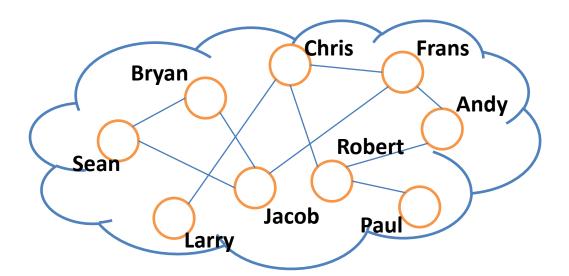


# Methods to limit the Sybil attack

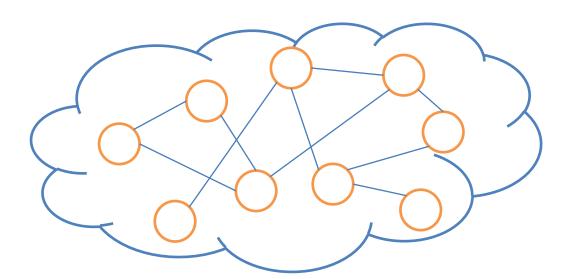
- Limit IDs per IP address
- Central CA issues IDs
  - Strong PKI
  - CAPTCHA
  - Cryptographic puzzles
- All methods have drawbacks
  - cost, compatibility, barriers to entry
- Adversary may have more resources

#### Social network can help

- Nodes have social links to other nodes
  - social links established outside of the DHT
  - provides additional information usable by DHT

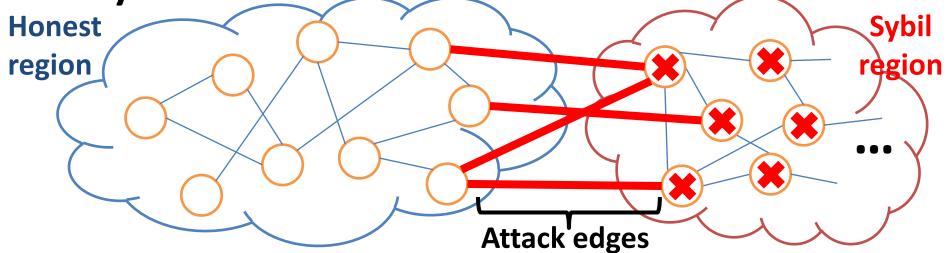


#### Social network model



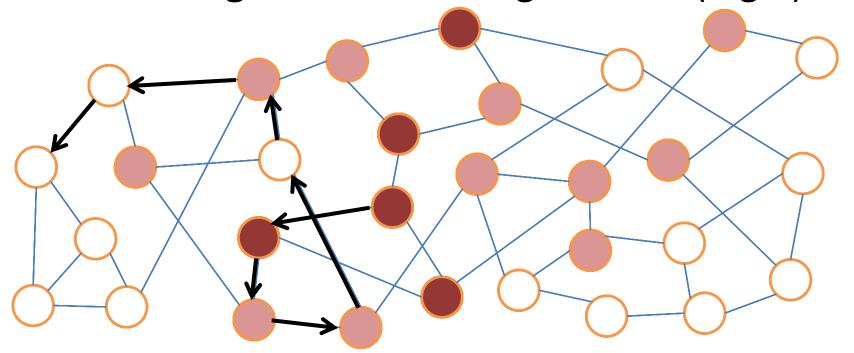
#### Social network model

- n = number of honest nodes
  - for this talk only, all nodes have ~same degree
- g = number of attack edges
  - $-g = o(n/\log n)$  tolerable by protocol
- Correctness is independent of number of Sybil nodes!



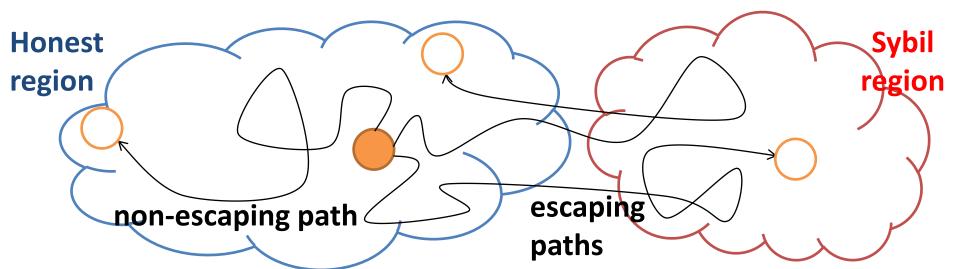
#### Mixing time

- Random walk: choose each hop randomly
- Mixing time: #hops until uniform probability
- Fast mixing network: mixing time = O(log n)



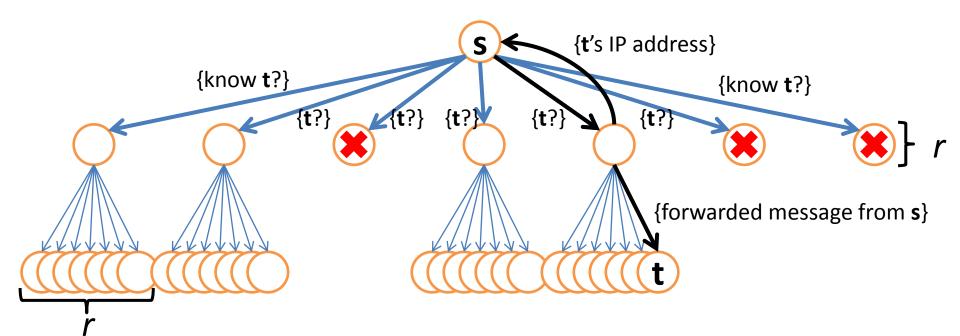
#### Sampling by random walks

- A random walk has o(1) chance of escaping\*
  - True when g bounded by  $o(n/\log n)$
  - Of r walks,  $(1-o(1))r = \Omega(r)$  end nodes are good!
  - Can't distinguish good from bad nodes in set



## Basic one-hop DHT design

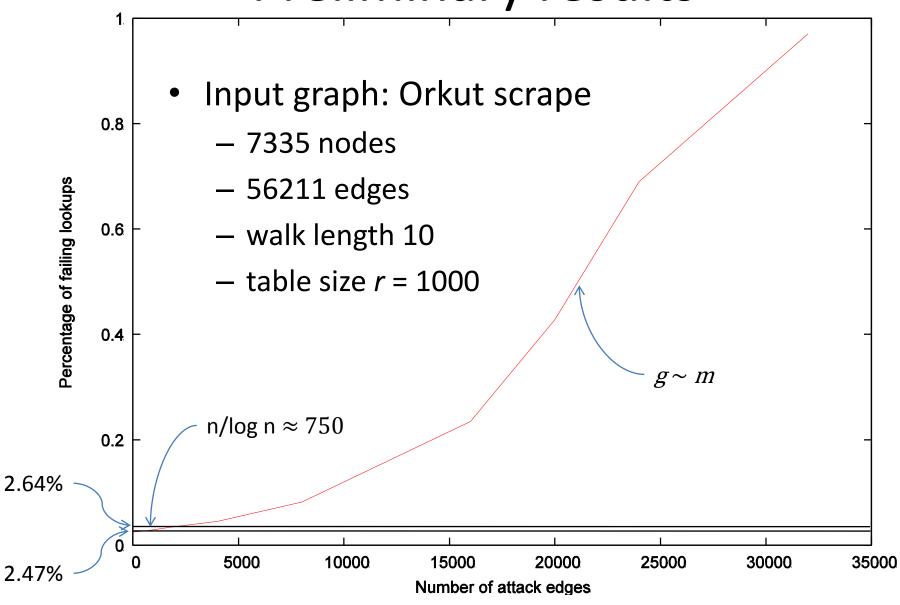
- Construct finger table by r random walks
- Route to t by asking all fingers about t— If  $r = \Omega(\sqrt{n \log n})$ , some finger knows t WHP
- Adversary cannot interfere with routing



#### Properties of this solution

- Finger table size:  $r = O(\sqrt{n \log n})$
- Bandwidth to construct: O(r log n) bits
- Bandwidth to query: O(r) messages
- Probability of failure: 1/poly(n)

Preliminary results

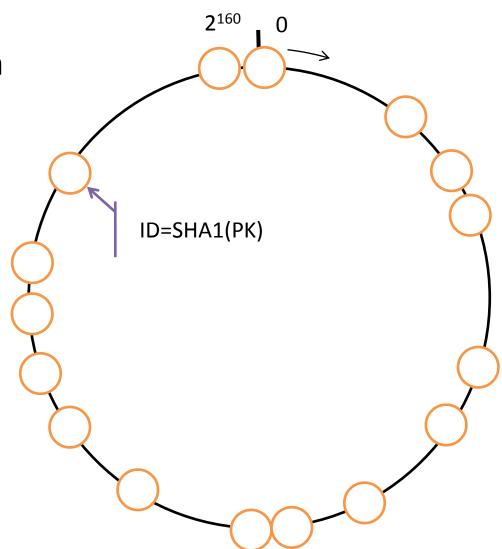


#### Properties of this solution

- Finger table size:  $r = O(\sqrt{n \log n})$
- Bandwidth to construct: O(r log n) bits
- Bandwidth to query: O(r) messages
- Probability of failure: 1/poly(n)

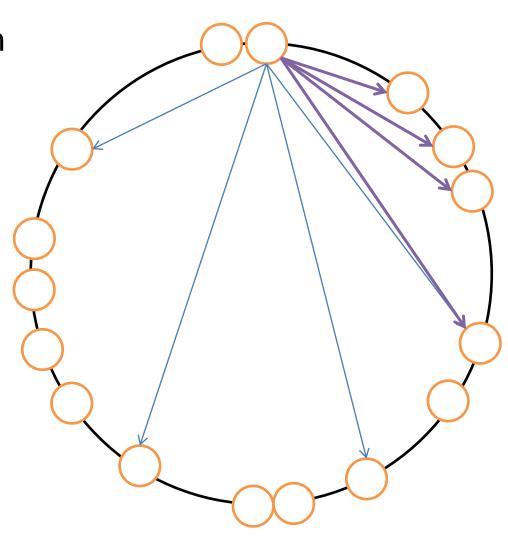
#### Structured one-hop DHT

- Goal: reduce bandwidth used by routing lookup
- Method: add Chord-like structure to DHT
- Assign hash IDs on ring



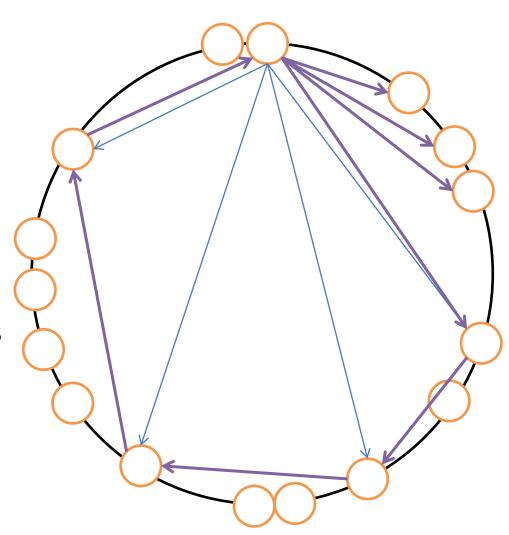
#### Structured one-hop DHT

- Goal: reduce bandwidth used by routing lookup
- Method: add Chord-like structure to DHT
- Assign hash IDs on ring
- Already have finger tables
- Need successor tables



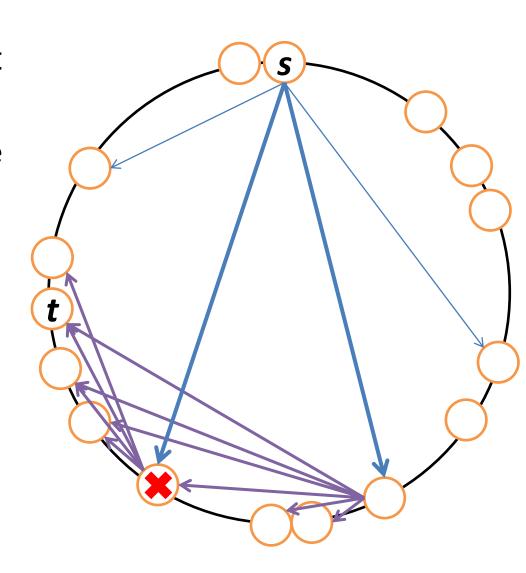
#### Constructing successor tables

- Construct finger tables
- Sort finger table by ID
- Tell each finger about its successors in your finger table
  - costs extra messages
  - send O(log n) successors
- Each node learns its r successors WHP



## Using successor tables

- To route, query closest finger to target
  - finger's successor table should contain target
- If failed, finger may be evil or simply unlucky
- Try next closest finger
  - Expect O(1) tries



## Hard to extend to O(log n) hops

- Would like to have smaller routing tables
  - this requires more hops per lookup
- First finger table (from random walks) has o(1) fraction of bad fingers
- Successive refinements (closer successors in ID space) using Chord stabilization: fraction of bad nodes grows at each step
- Tricky? Yes. Impossible? Unclear.

#### Summary

- DHTs are subject to the Sybil attack
- Social networks provide useful information
- Created a Sybil-resistant one-hop DHT
  - Resistant to  $g = o(n/\log n)$  attack edges
  - Table sizes and routing BW  $O(\sqrt{n \log n})$
  - Uses O(1) messages to route
- This is important: enables fully decentralized and secure peer-to-peer systems

#### The "Tom" attack



Tom has 230357403 friends.

#### The "Tom" attack

From: Flickr Mail <mail@flickr.com>

Subject: [Flickr] You are aameesh's newest contact!

Date: 29 Mar 2008 08:00:19 +0000

To: ctl-flickr@mit.edu

Hi Chris Lesniewski,

You are aameesh's newest contact! If you don't know aameesh, aameesh is probably a fan of your photos or wants a bookmark so they can find you again. There is no obligation for you to reciprocate, unless you want to. :)