Towards measuring anonymity

C. Diaz, S. Stefaan, J. Claessens, B. Preneel
PET'02

Presented by B. Choi, cs6461
Computer Science
Michigan Tech

Introduction

- Applications
  - Electronic payment, electronic voting, electronic auctions, email and web browsing
- No well established framework to assess the degree of anonymity by 2001?
  - Some proposals: anonymity set size
  - Shannon's theory of entrophy in 1948 could be a vehicle to approach the problem

System model

- Anonymity set fixed ans static
- Same number of sent messaged by all senders
- Senders behavior modeled as a Poisson process
- Mix-net anonymity system
- Attack model
  - The degree of anonymity depends on the probabilities that the users have sent a particular message: these probabilities are assigned by the attacker

Proposed measurement model

- Anonymity is the state of being not identifiable within a set of subjects, the anonymity set.
- Consider only sender anonymity
- Degree of anonymity provided by the system
  - The quality of the system
  - Depends on the distribution of probabilities and not on the size of the anonymity set

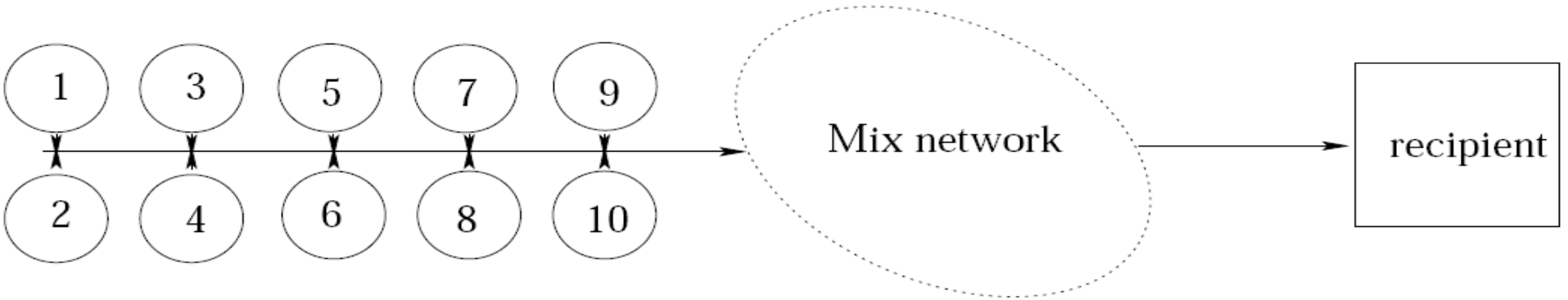
Proposed measurement model

$$H(X) = -\sum_{i=1}^{N} p_i \log_2(p_i)$$

$$H_M = \log_2(N)$$

$$d = 1 - \frac{H_M - H(X)}{H_M} = \frac{H(X)}{H_M}$$

Example

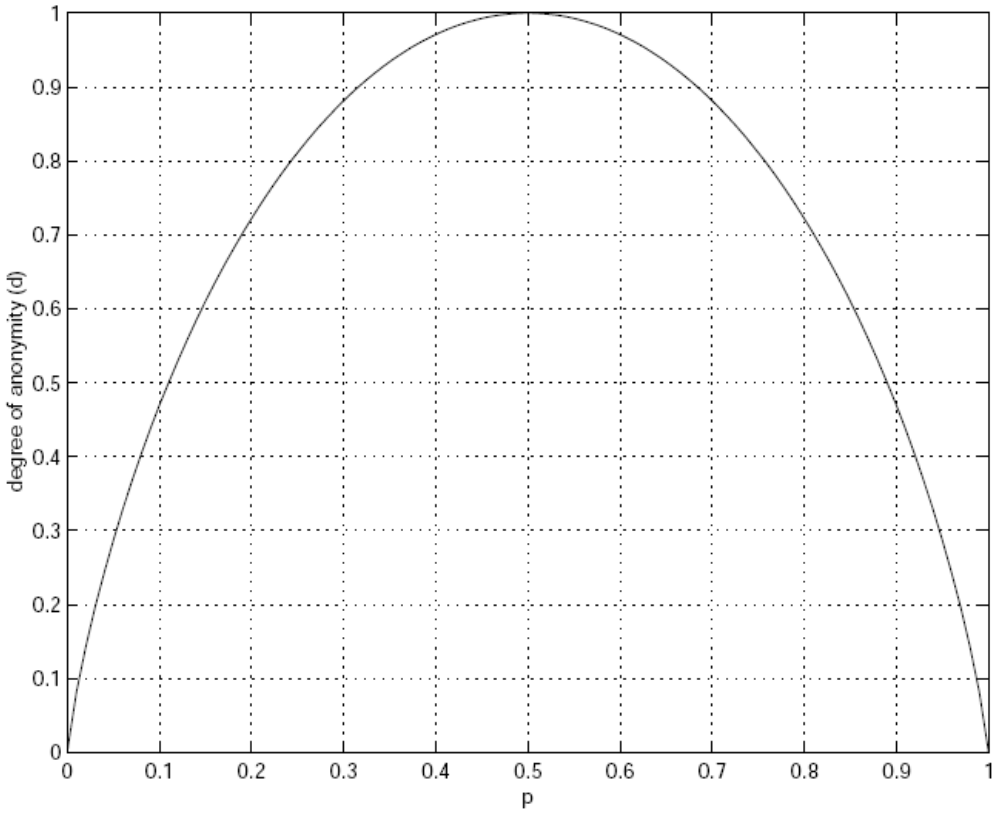


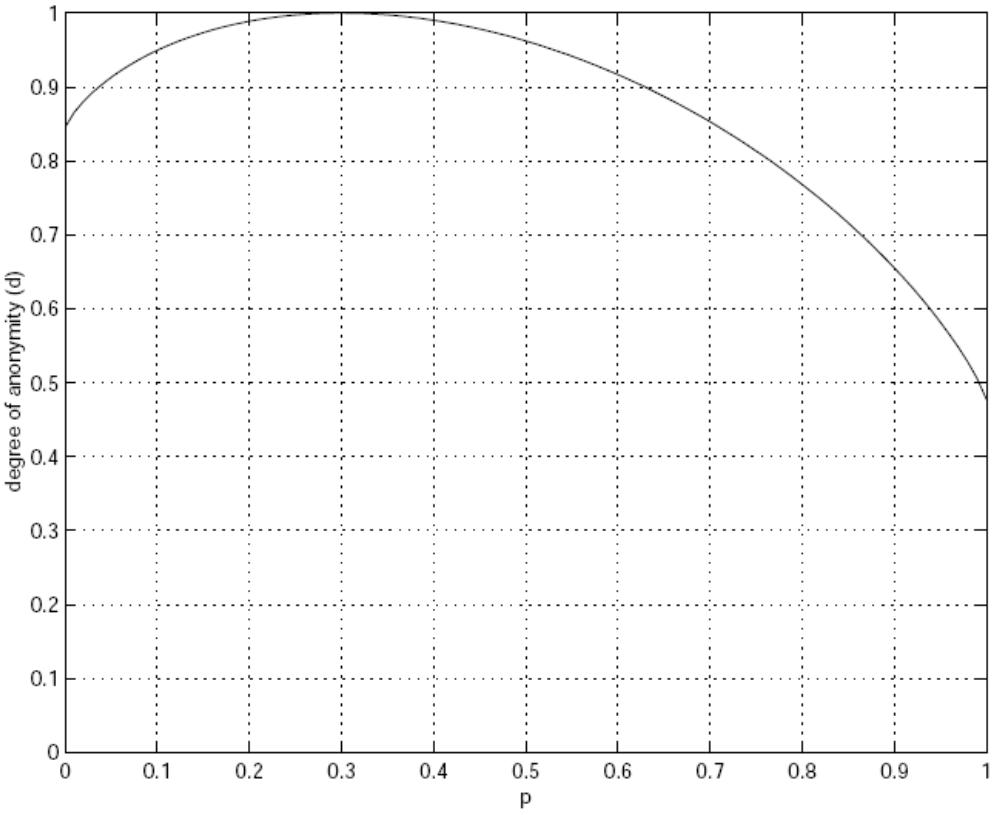
$$p_1 = p \ ; \quad p_2 = 1 - p \quad H_M = log_2(2) = 1$$

$$H_M = log_2(10)$$

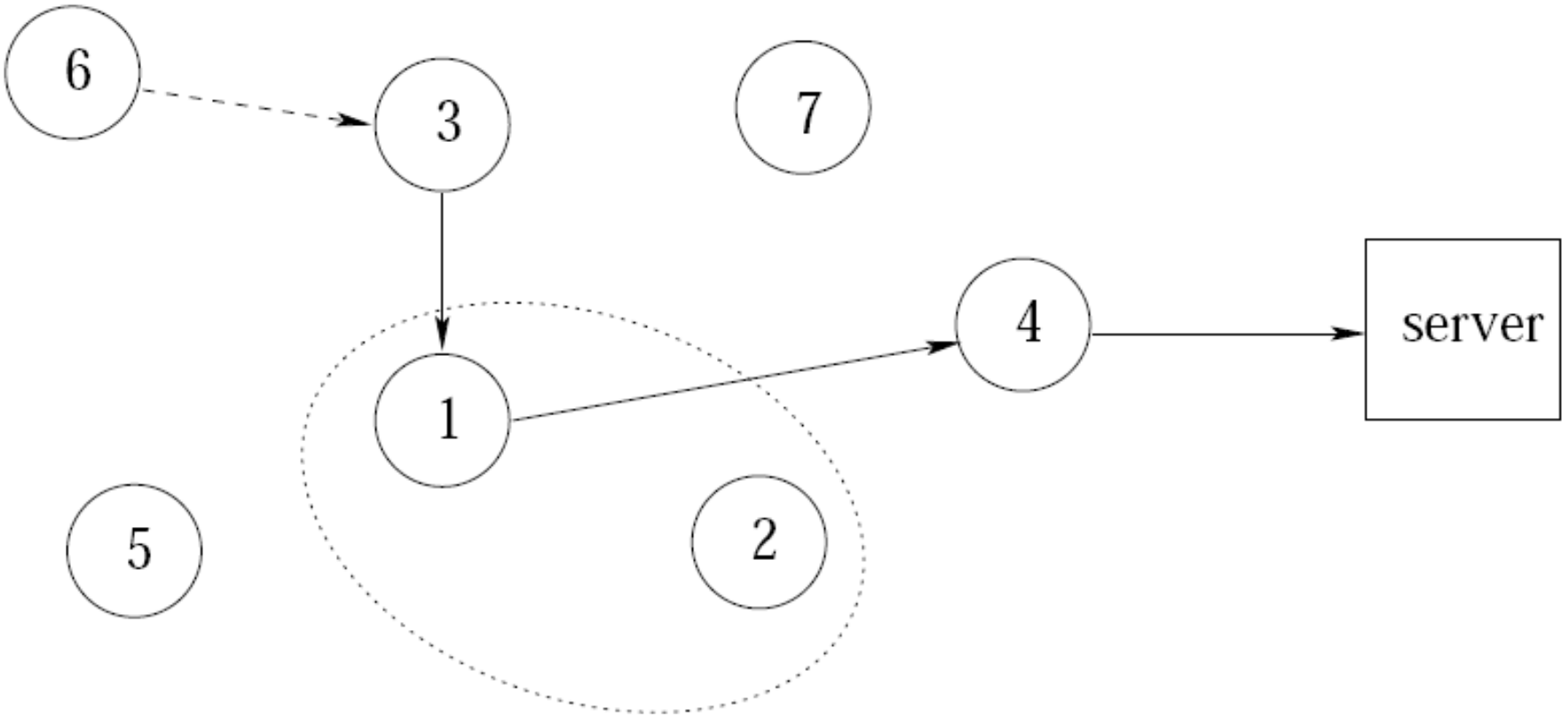
$$p_i = \frac{p}{3} \ , \quad 1 \leq i \leq 3 \ ; \quad p_i = \frac{1-p}{7} \ , \quad 4 \leq i \leq 10$$

Example


degree of anonymity (d)
p
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


degree of anonymity (d)
p
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Case study: Crowds



Crowds

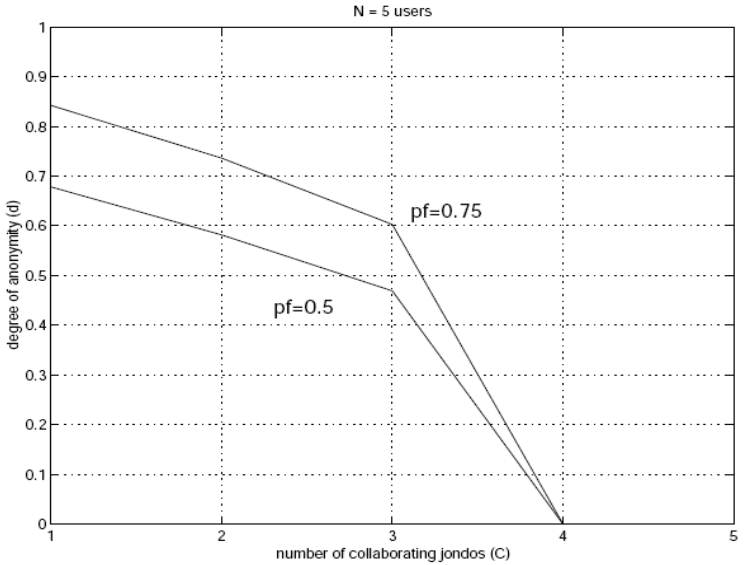
$$H_M = \log_2 (N - C)$$

$$p_{C+1} = \frac{N - p_f(N - C - 1)}{N} = 1 - p_f \frac{N - C - 1}{N}$$

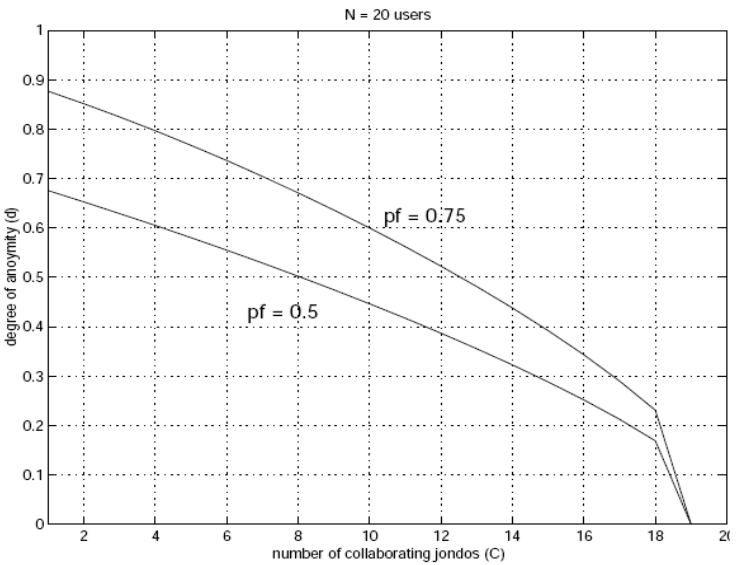
$$p_i = \frac{1 - p_{C+1}}{N - C - 1} = \frac{p_f}{N} \quad , \quad C + 2 \leq i \leq N$$

$$H(X) = \frac{N - p_f(N - C - 1)}{N} \log_2 \left[\frac{N}{N - p_f(N - C - 1)} \right] + p_f \frac{N - C - 1}{N} \log_2 \left[\frac{N}{p_f} \right]$$

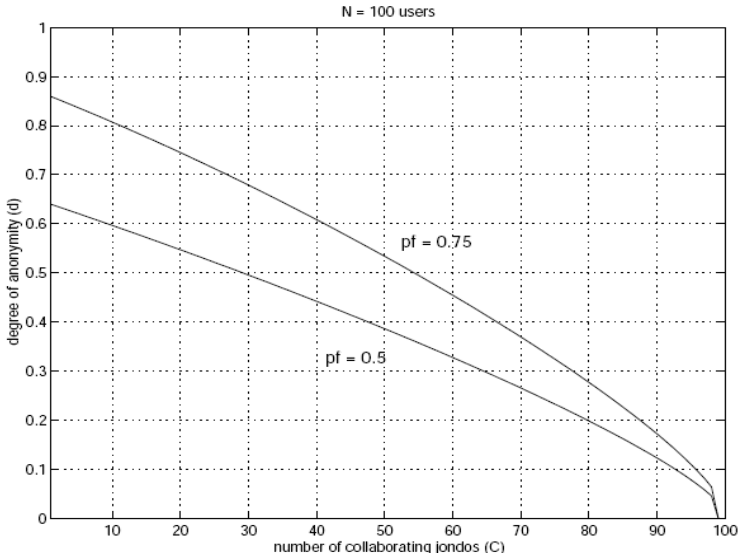
Crowds


N = 5 users
degree of anonymity (d)
number of collaborating jondos (C)
pf=0.75
pf=0.5

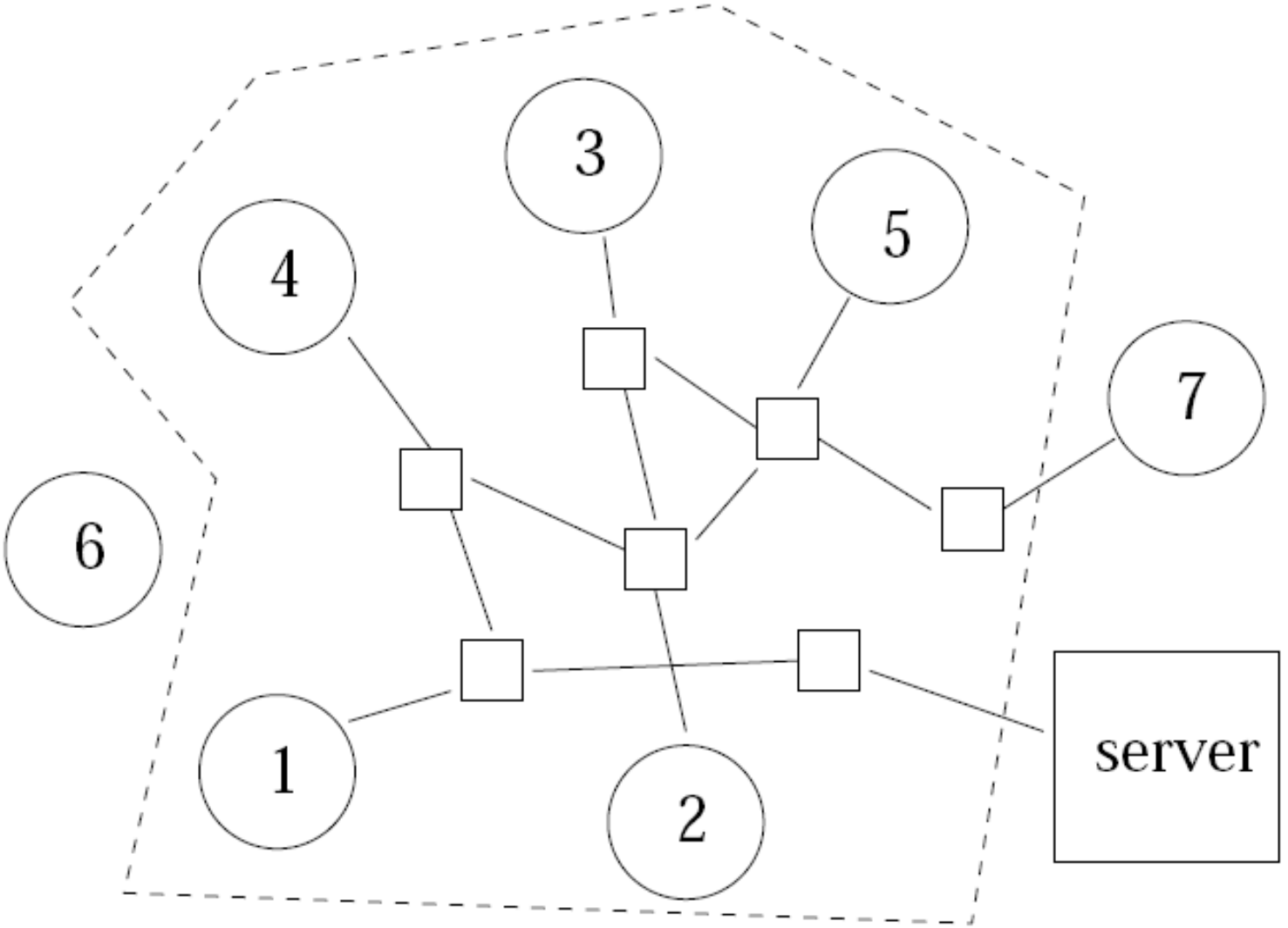
(a)


N = 20 users
degree of anonymity (d)
number of collaborating jondos (C)
pf = 0.75
pf = 0.5

(b)


N = 100 users
degree of anonymity (d)
number of collaborating jondos (C)
pf = 0.75
pf = 0.5

Onion routing



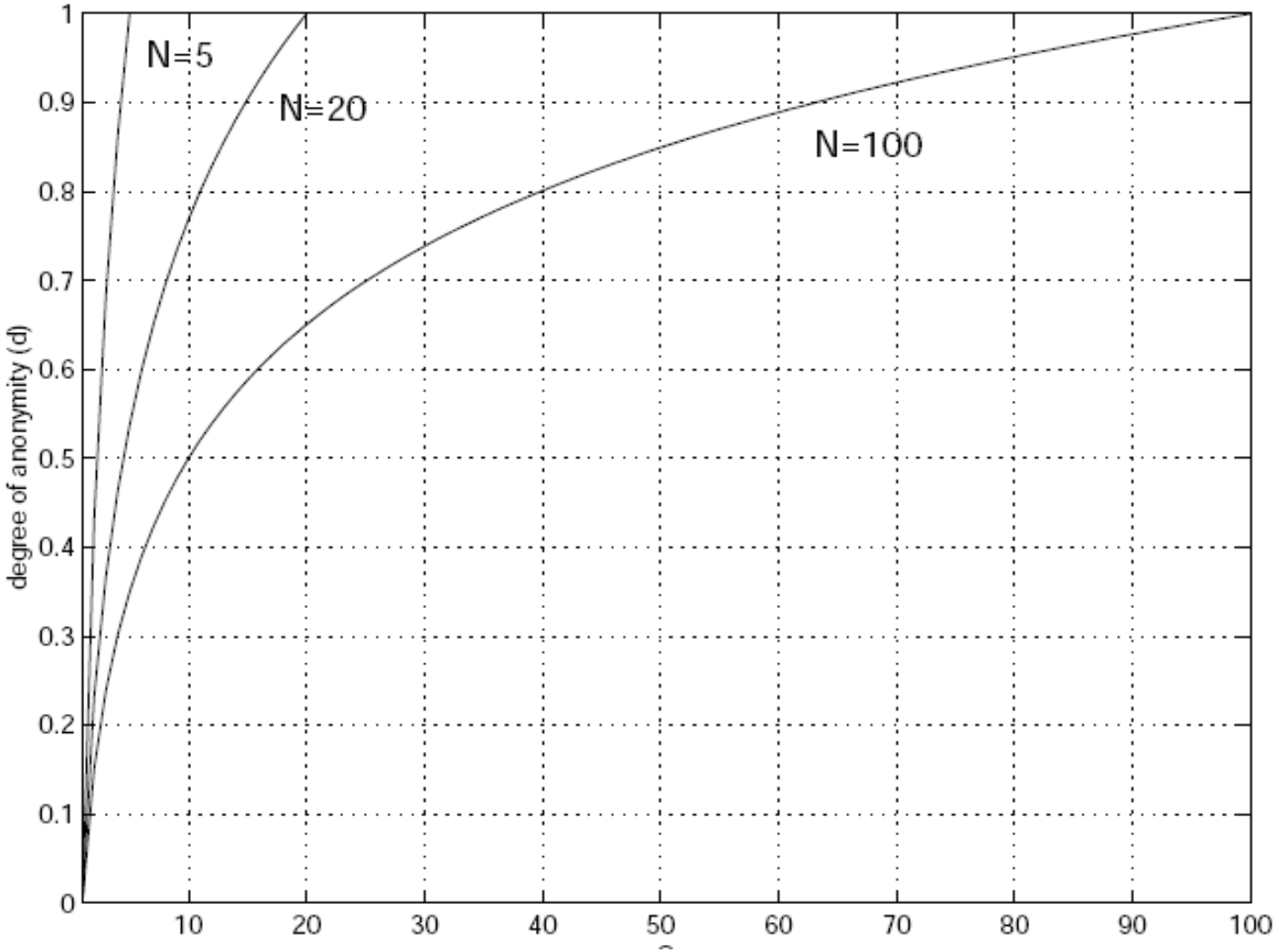
Onion routing

$$H_M = \log_2(N)$$

$$p_i = \frac{1}{S} \ , \quad 1 \leq i \leq S \ ; \quad p_i = 0 \ , \qquad S+1 \leq i \leq N$$

$$H(X) = \log_2(S) \ , \quad d = \frac{H(X)}{H_M} = \frac{\log_2(S)}{\log_2(N)}$$

Onion routing



Open problems

- How to find the probability distribution in real situations?
- Understanding of real attacks?
  - We still don't know how hard or easy to monitor part of or entire of an anonymity system
    - Core router-based
    - P2P-based
- Degree of anonymity is then relative to attackers
  - Any standardized absolute degree possible?