Information Leaks in Structured Peer-to-Peer Anonymous Communication Systems

P. Mittal, N. Borisov
CCS 2008

Presented by B. Choi
cs6461 at MTU

Motivations

- Low latency anonymity systems
  - Tor!
  - Vulnerable for traffic analysis, now well known
- One more weakness?
  - Directory server
  - Any Tor node has to download the entire membership, out of which to create a path
  - The membership is readily available for anyone?
    - Excellent opportunity for attackers!

Ahh..., Membership!

- Yes, membership should not be easily and readily available to malicious nodes
- What the membership implies?
  - Big help for attackers to mount an attack
  - The anonymity set itself!
  - The anonymity set itself needs to be hidden
  - But how? Tor has only 400+ nodes, a vastly small subset of the Internet

Secure and scalable membership?

- One more concern with Tor is scalability
  - Can it grow to 1M nodes?
  - If so, the directory server needs to change
- Two obvious requirements
  - Scalability: distributed membership management
  - Security: lookup activities needs to be hidden
- One possible solution: peer-to-peer (P2P)

Two such P2P systems

- AP3: 2004
- Salsa: structured approach for large scale anonymity, 2006
- Commonality: set up and tear down an anonymous connection for each anonymous flow, similar to Crowds
- Structured?
  - Membership is distributed in a nicely (and hopefully securely) structured manner

Question

- Conventional wisdom: p2p based distributed membership management helps security and anonymity
- Doubts: lookup activities with the distributed membership may not help
- How prove or disprove the doubts?

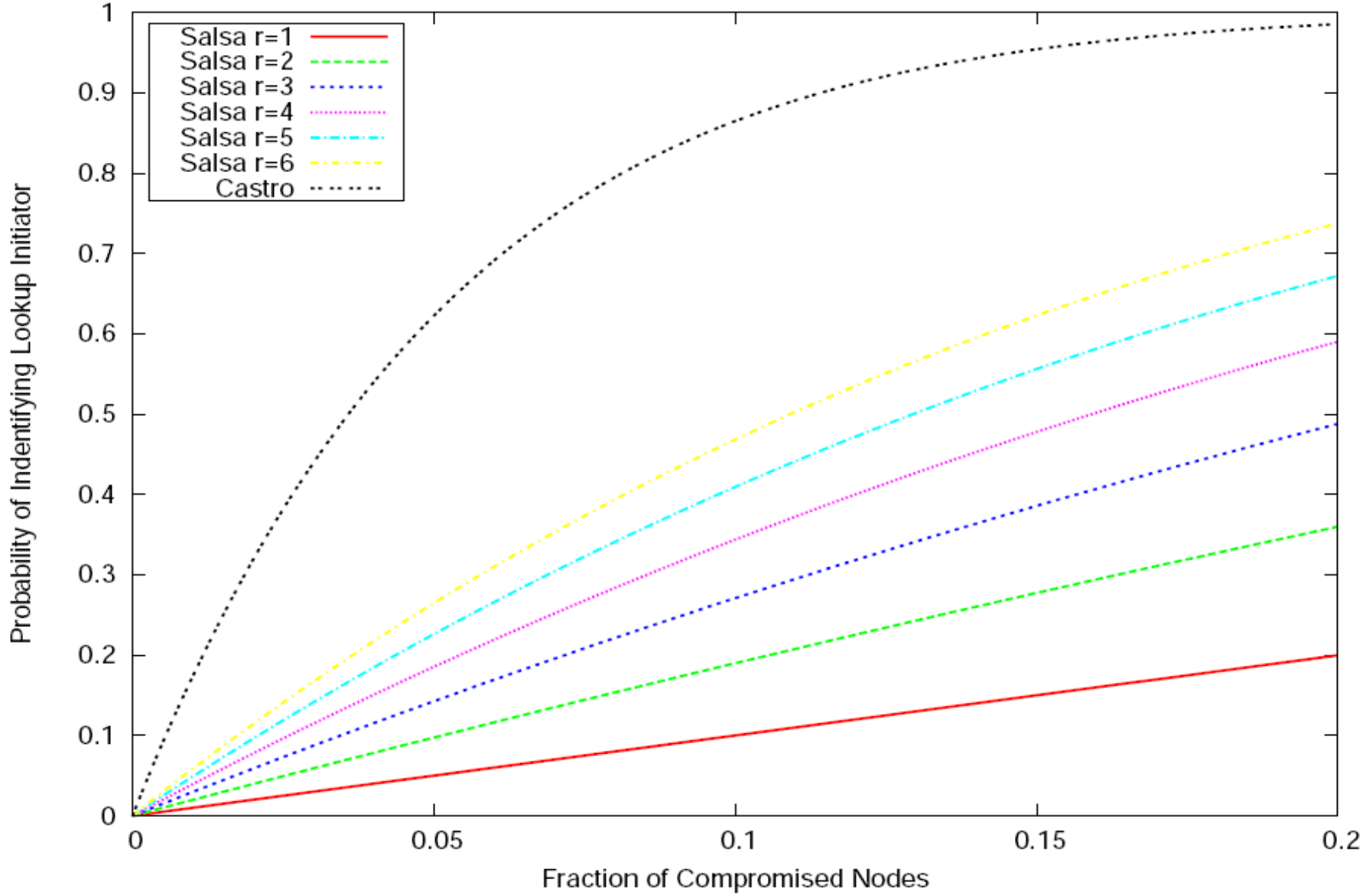
Threat model

- Partial, internal, active, static
  - Why partial? ..... since global may be difficult
- Collusion by the adversary
  - Why colluding? .... botnet

Secure lookup?

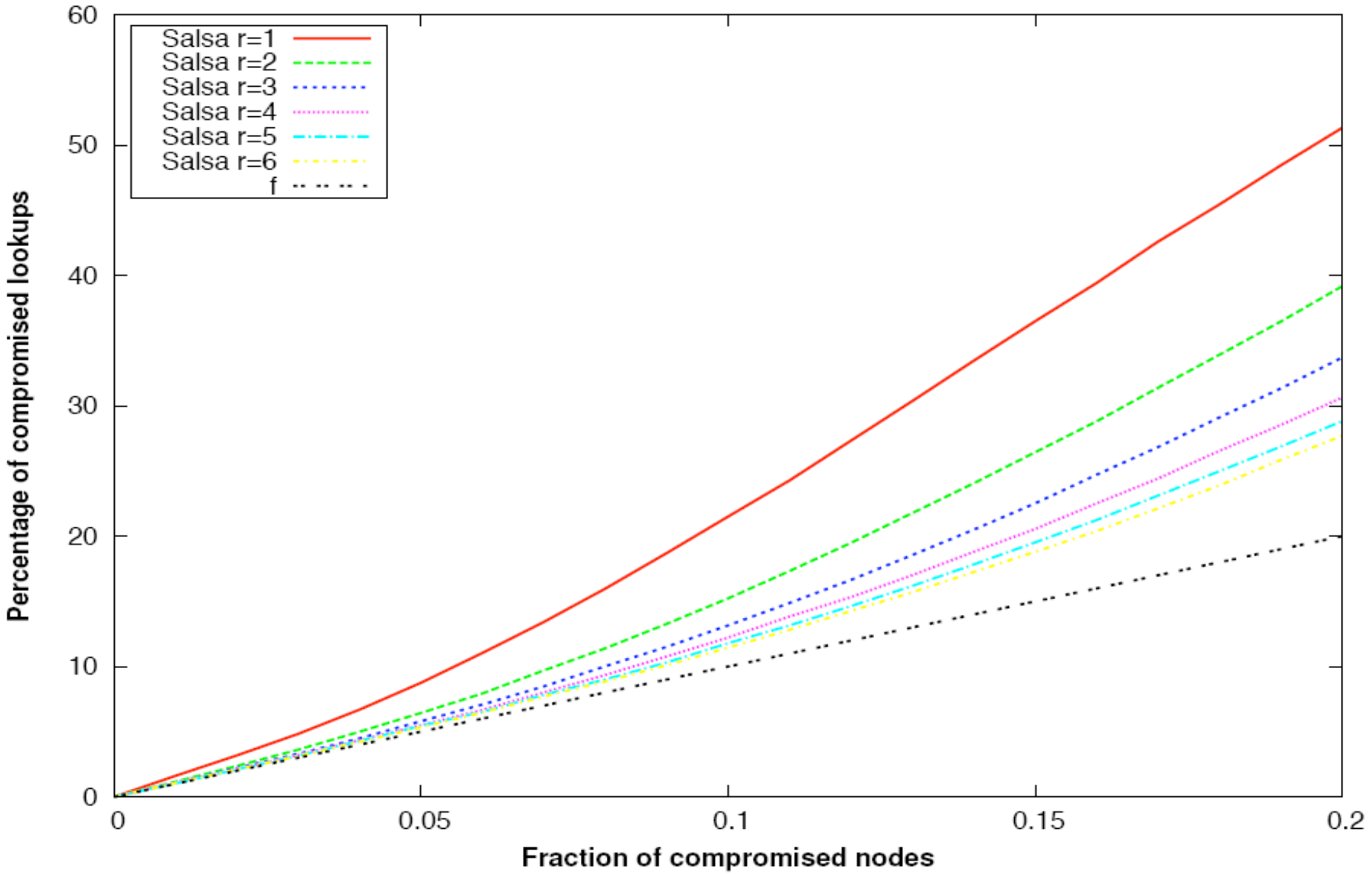
- AP3 and Salsa implement secure lookup by introducing redundant lookups
- Redundant lookups?
  - Assume a fraction of nodes is compromised
  - Distributed membership could help adversary narrow down the possible anonymous connections
  - Redundant lookups would confuse the compromised nodes, which are supposedly try to figure out the anonymous connection
  - The redundancy in return could compromise the identity of the lookup initiator?

Evidence to the doubts (anonymity)



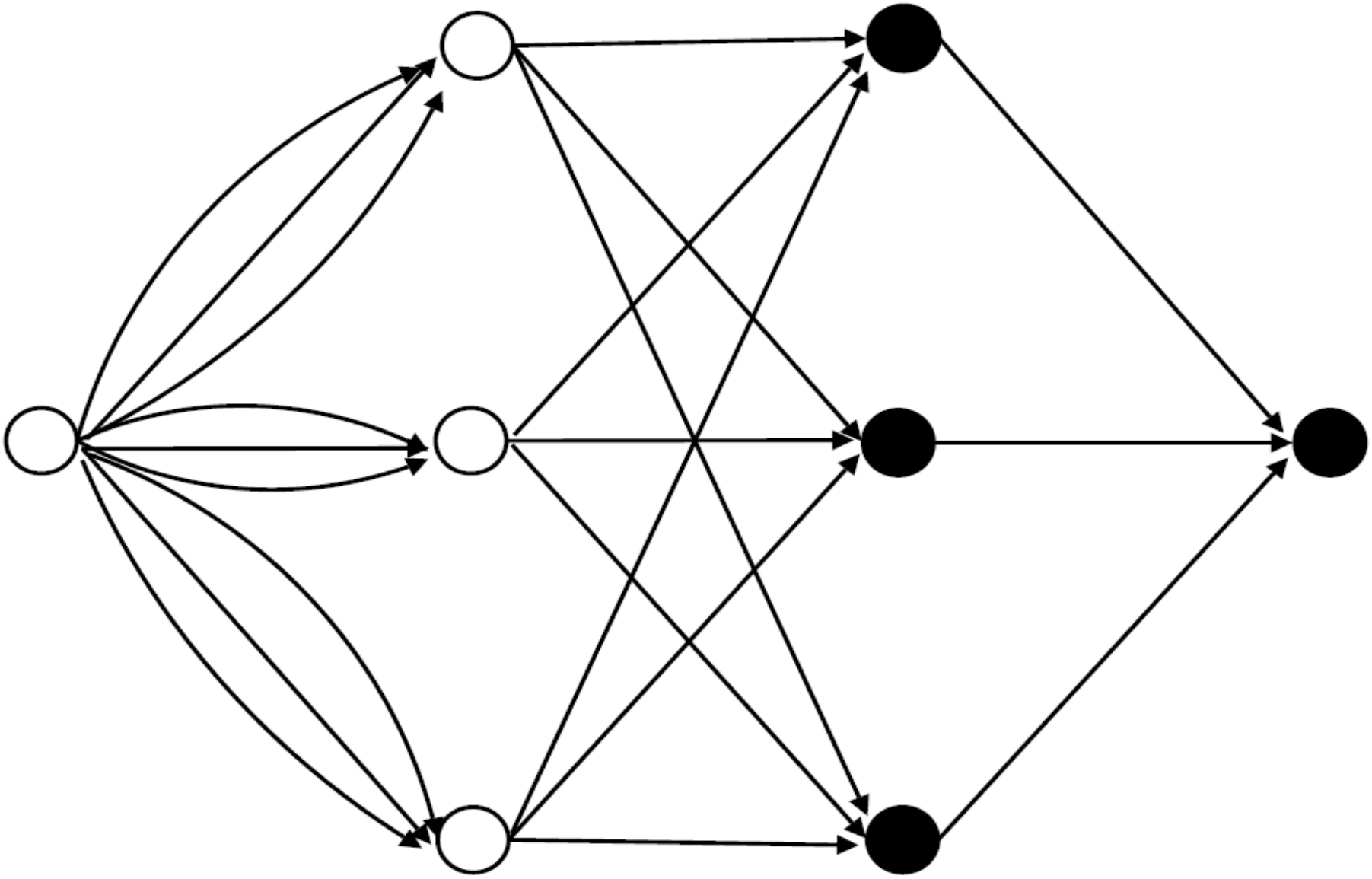
(a) Information leak from secure lookups

Evidence to the doubts (security)



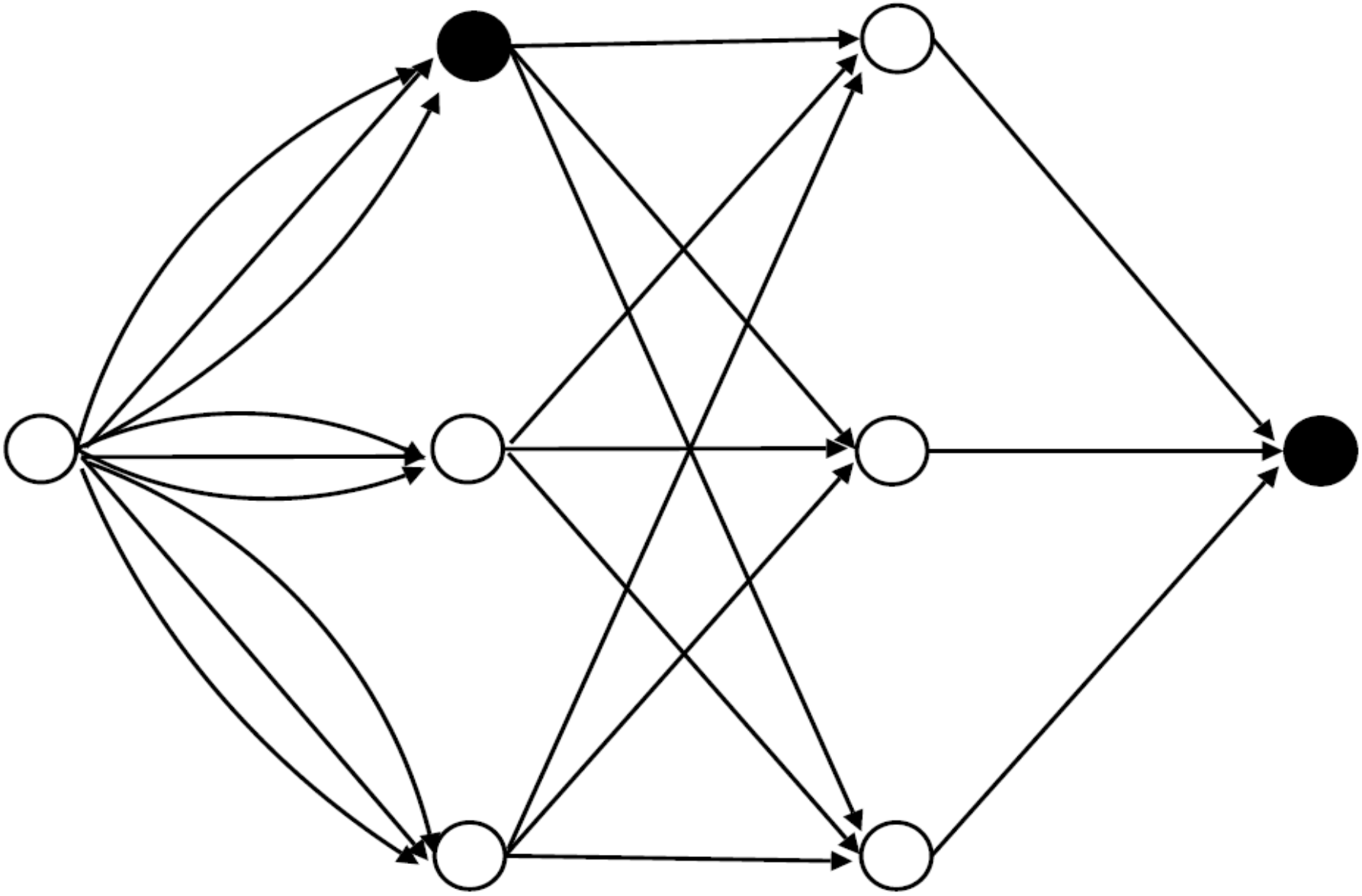
(b) Percentage of compromised lookups

Reasons



(a) Bridging an honest first stage

Reasons



(b) Bridging an honest stage

Results

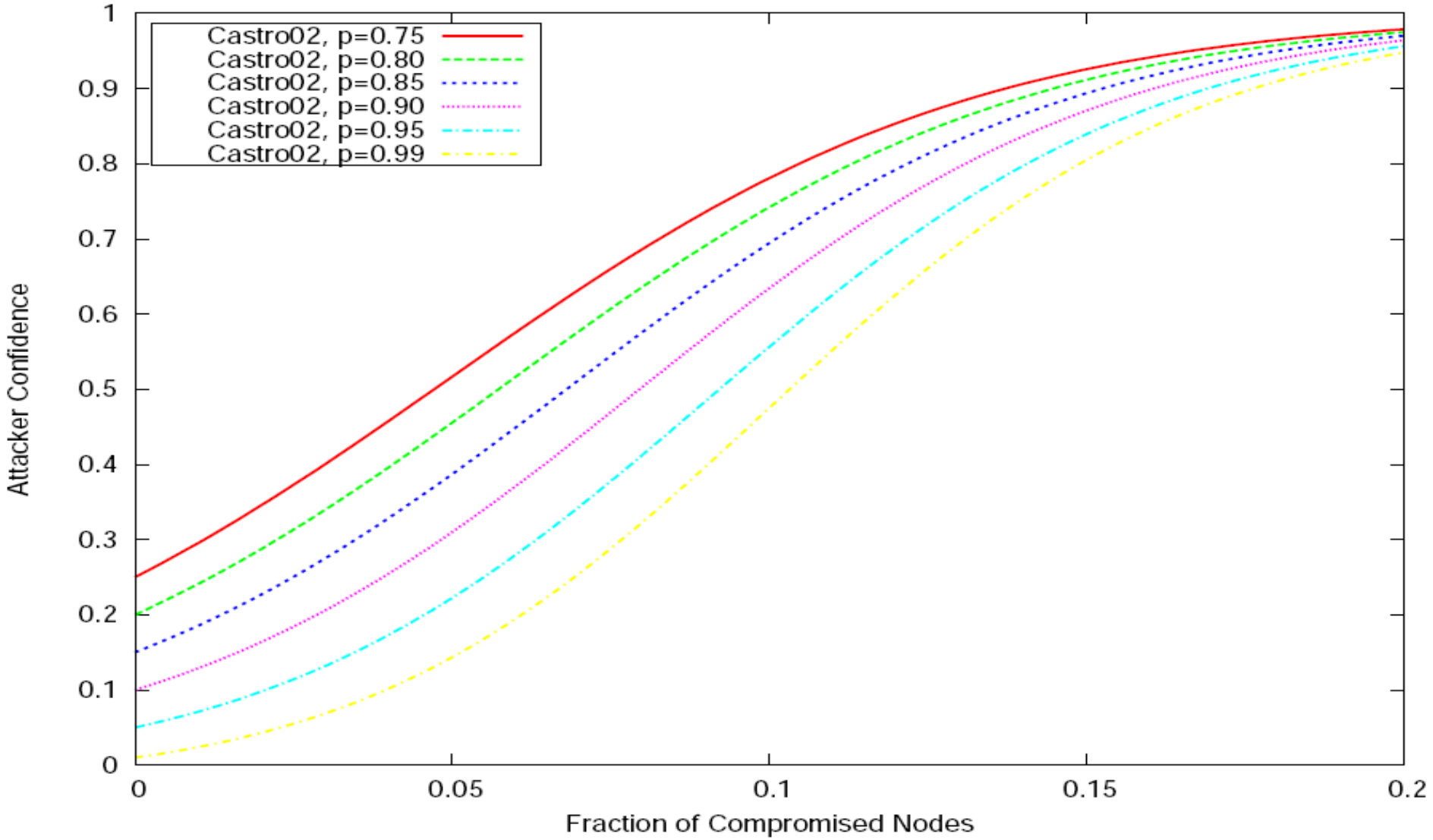


Figure 3: $P(I|E_1)$

Results

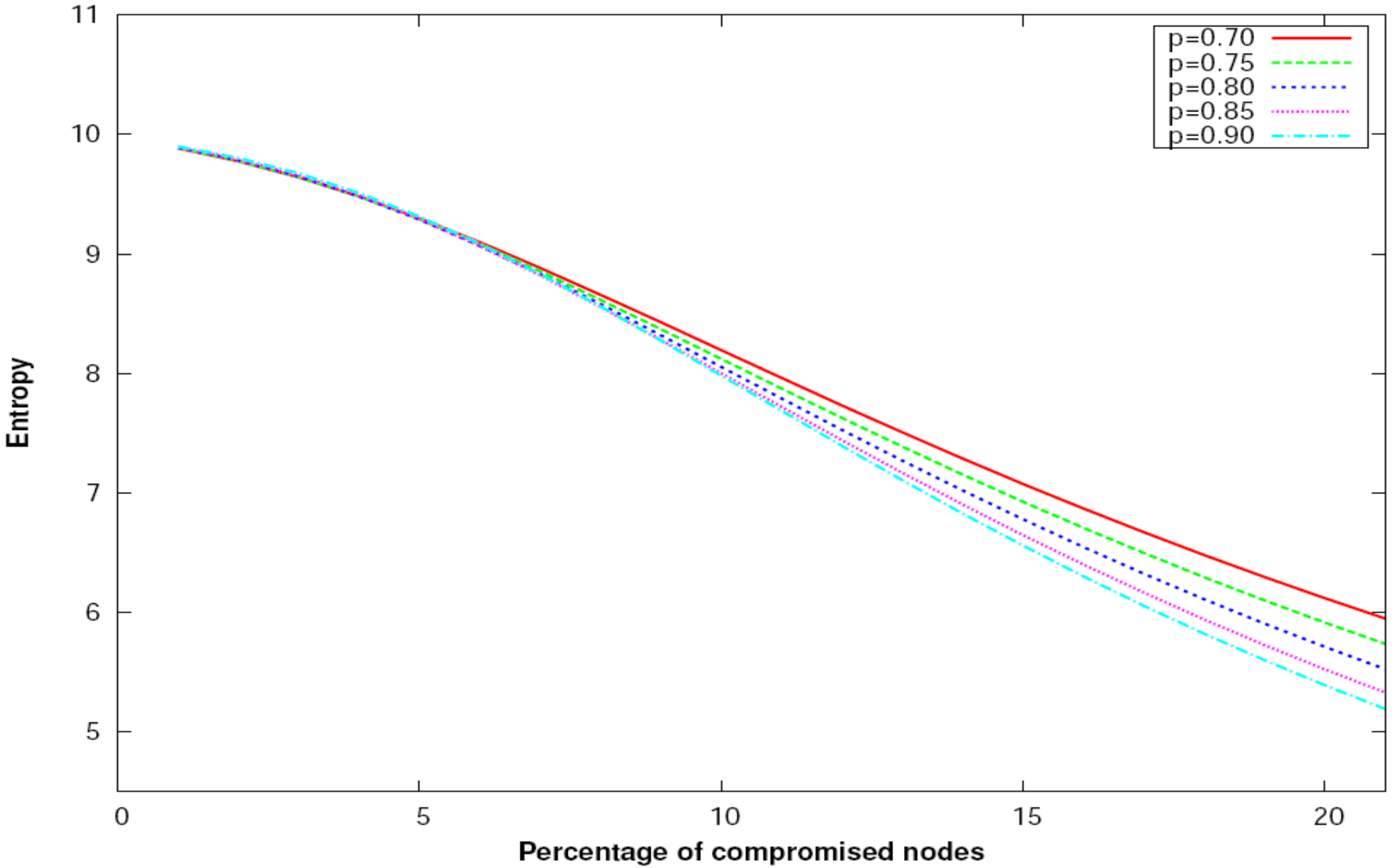
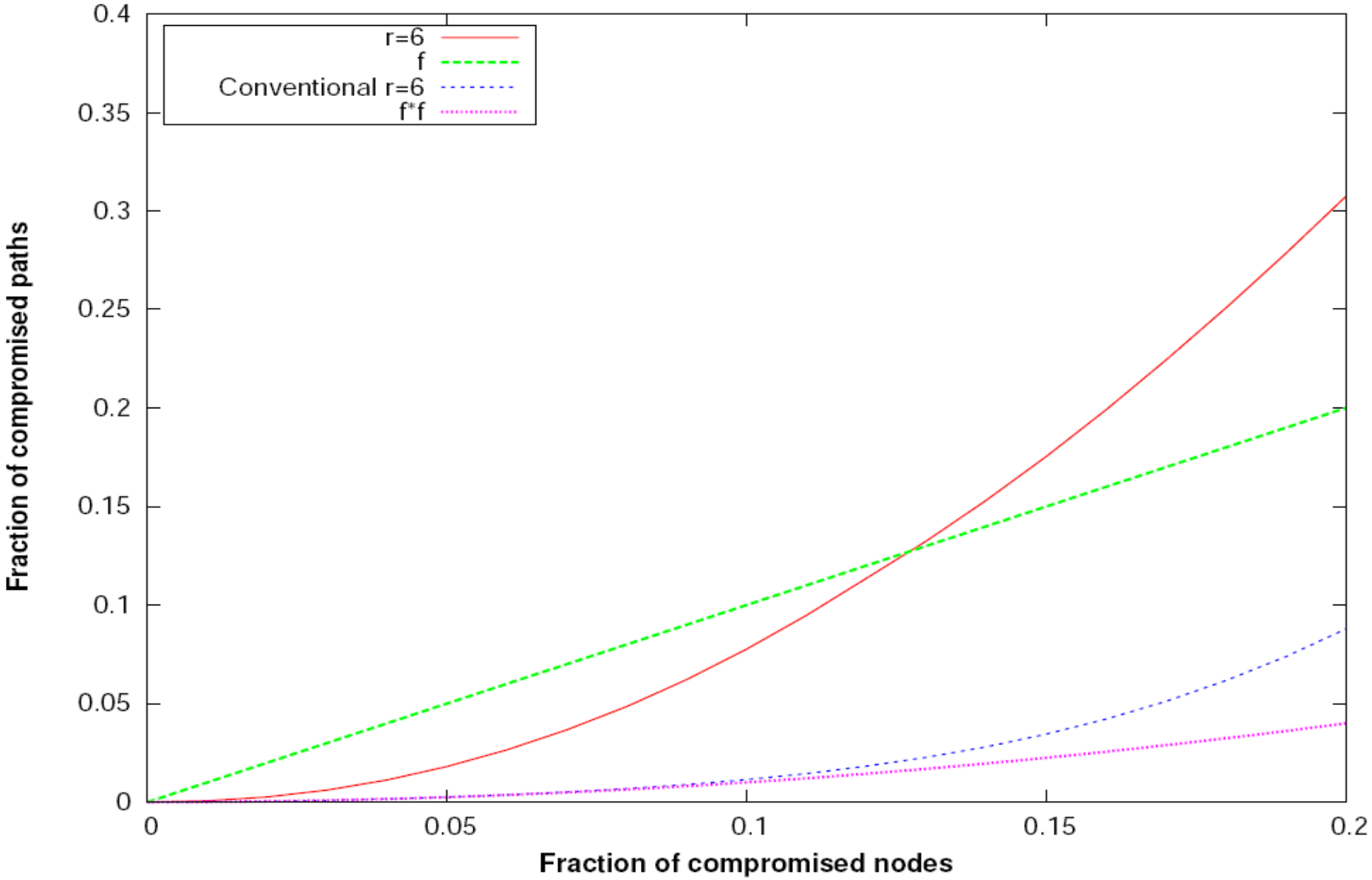


Figure 5: Entropy as a function of f.

Results


r=6
f
Conventional r=6
f*f
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
0.05
0.1
0.15
0.2
Fraction of compromised paths
Fraction of compromised nodes