The EigenTrust Algorithm for Reputation Management in P2P Networks

Sepandar Kamvar, Mario Schlosser, Hector Garcia-Molina

WWW 2003

Presented by –

Apurva Mohan

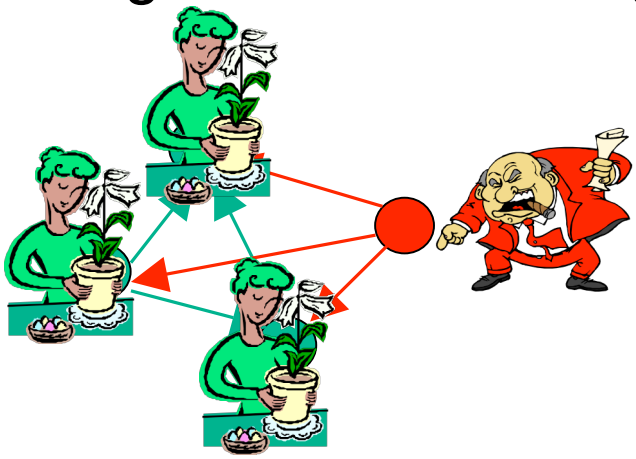
Aditi Dixit

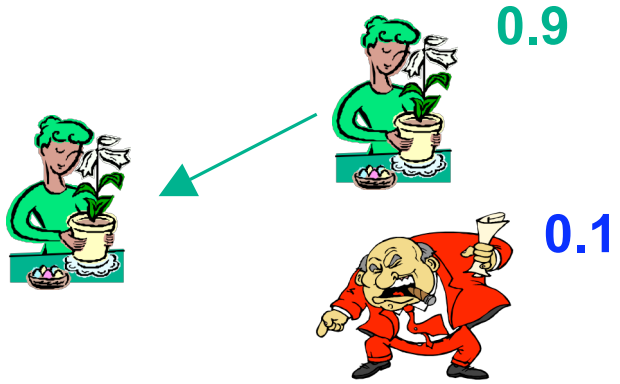
Contents

- Introduction
- Reputation Systems
- EigenTrust
- Secure EigenTrust
- Experimental Evaluation
- Threat Models
- Strengths and Weaknesses

Introduction – Problems in P2P File Sharing

- Anonymous nature -> Lack of accountability.
- Platform to spread malware, unauthentic files.
- Global trust value calculation for each peer is challenging.
- Malicious peers attempt to attack the trust model.
- Anonymization and encryption make it harder to ban illegal content sharing.



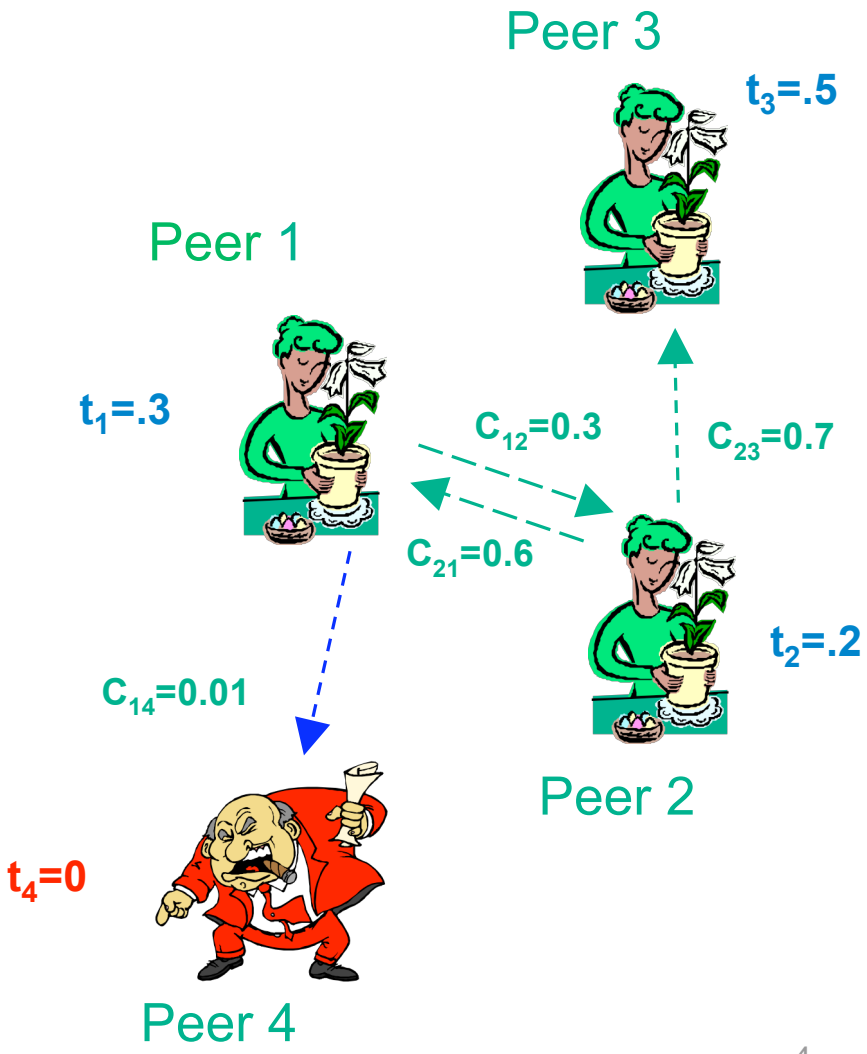


Reputation Systems

- Users collaborate to calculate reputation values.
- Example – eBay, PageRank, Advogato, Epinions.

$s_{ij} = sat(i, j) - unsat(i, j)$

- Transitive trust



EigenTrust

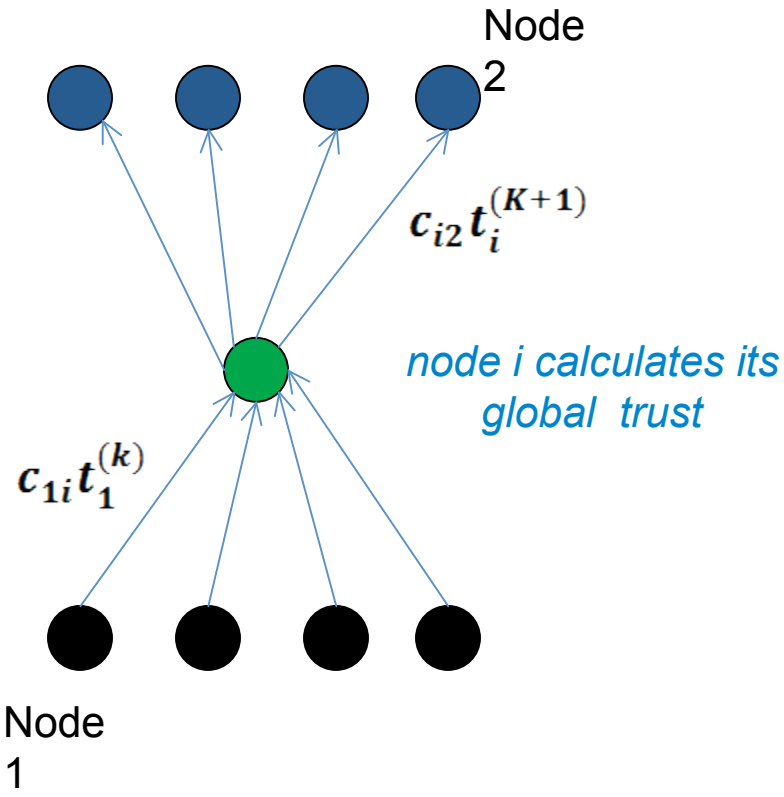
- Local vs. global trust values.
- Normalization $$c_{ij} = \frac{\max(s_{ij}, 0)}{\sum_j \max(s_{ij}, 0)}$$
- Aggregation $$t_{ik} = \sum_j c_{ij} c_{jk}$$
- Basic EigenTrust

$$t = C^T \vec{c_i} \qquad t = (C^T)^2 \vec{c_i} \qquad t = (C^T)^n \vec{c_i}$$

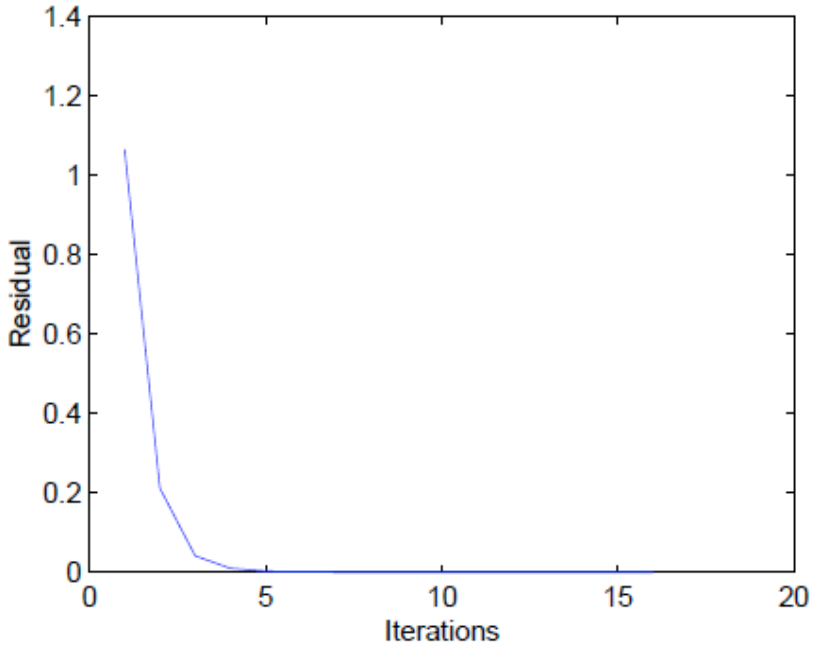
- Practical issues - trusted seed, inactive peers, Collectives.

Distributed EigenTrust

Calculation



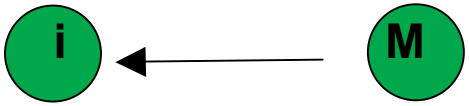
Convergence



EigenTrust convergence

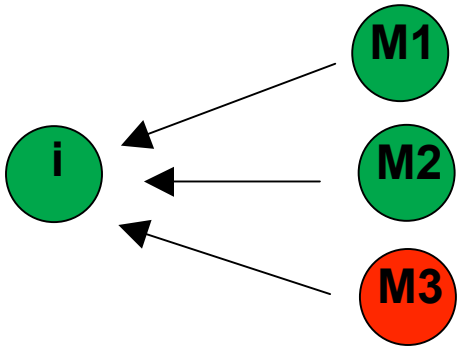
Secure EigenTrust

Score calculated by another peer



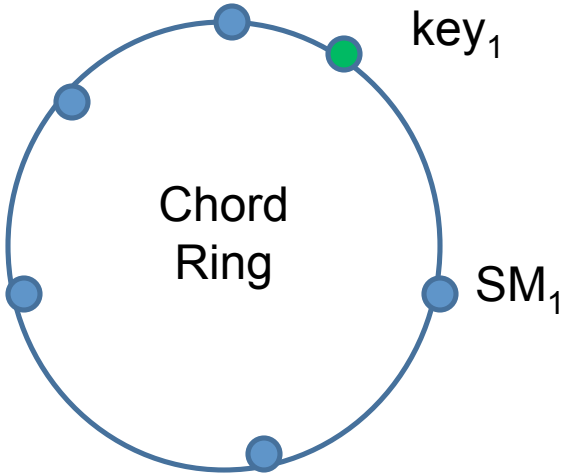
Score Manager

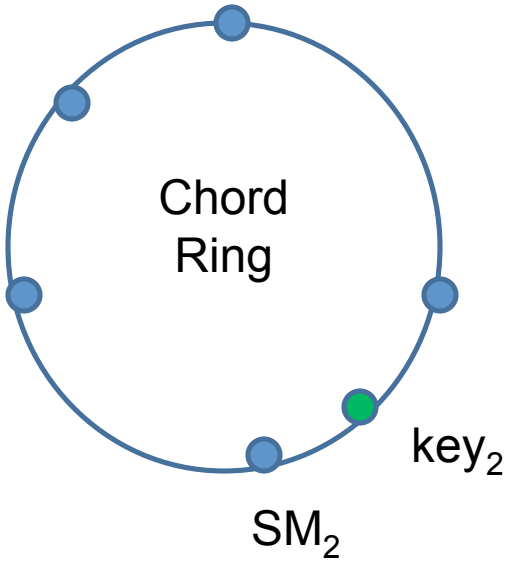
Score calculated by multiple other peers

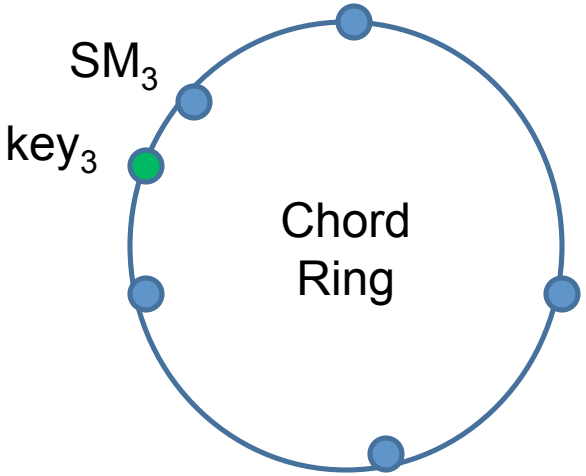


Multidimensional DHT's

IP Address + port = ID_1
$h_1(ID_1) = key_1$
$h_2(ID_1) = key_2$
$h_3(ID_1) = key_3$







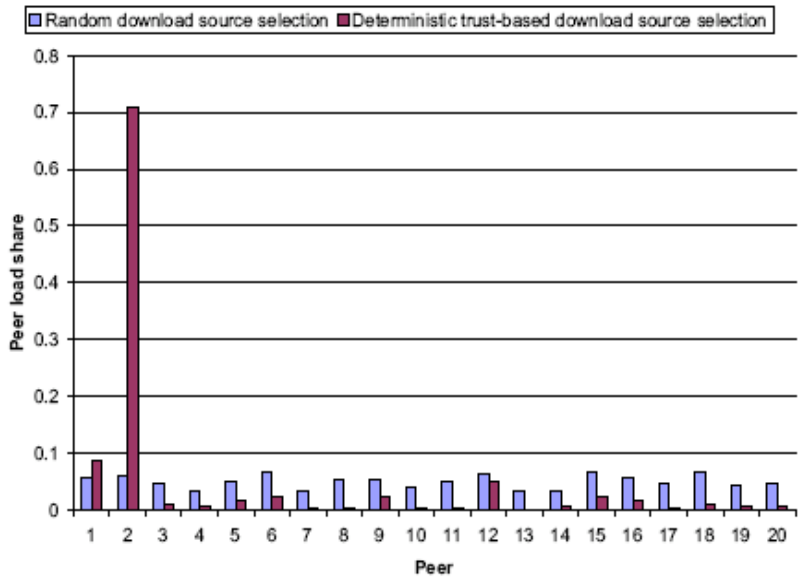
Simulation Settings

Network	#of good peers #of malicious peers #of pre-trusted peers #of initial neighbors of good peers #of initial neighbors of malicious peers #of initial pre-trusted peers #Time-to live for query message	60 42 3 2 10 10 7
Content Distribution	# of distinct files at good peer i set of content categories supported by good peer i	 Zipf distribution over 20 content categories
Peer Behavior	% of download requests in which good peer i returns inauthentic file % of download requests in which malicious peer i returns inauthentic file download source selection algorithm probability that peer with global trust value 0 is selected as download source	5% 0% probabilistic algorithm 10%
Simulation	# of simulation cycles in one experiment # of query cycles in one simulation cycle # of experiments over which results are averaged	30 50 5

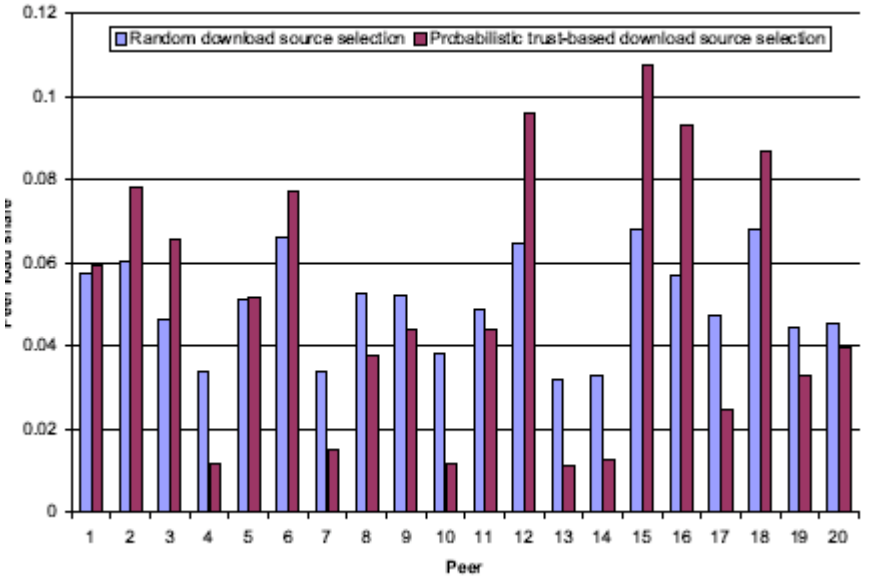
Load Distribution

- Deterministic Algorithm
  - Peers with highest trust value chosen among the peers responding to a query.
  - Leading to accumulation of reputation by a peer due to responding to a many queries.
- Probabilistic Algorithm
  - Peer i chosen as download source with probability p

Load Distribution(cont..)



Deterministic download source selection versus a non-trust based network.



Probabilistic download source selection versus a non-trust based network.

Threat Models

Threat Model A

- There are individual malicious peers

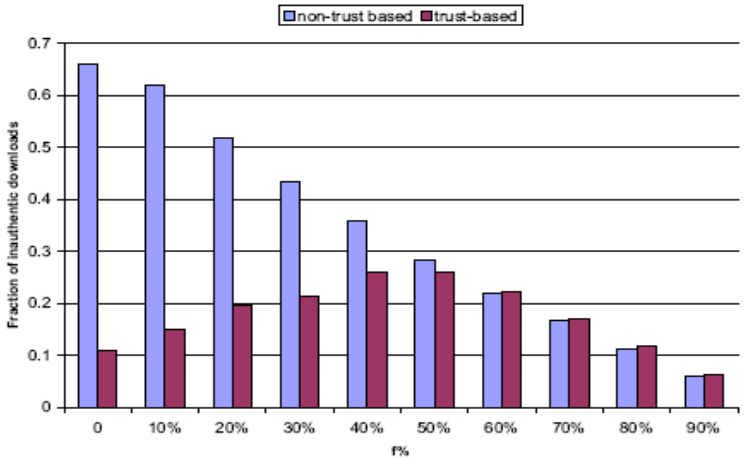
Threat Model B

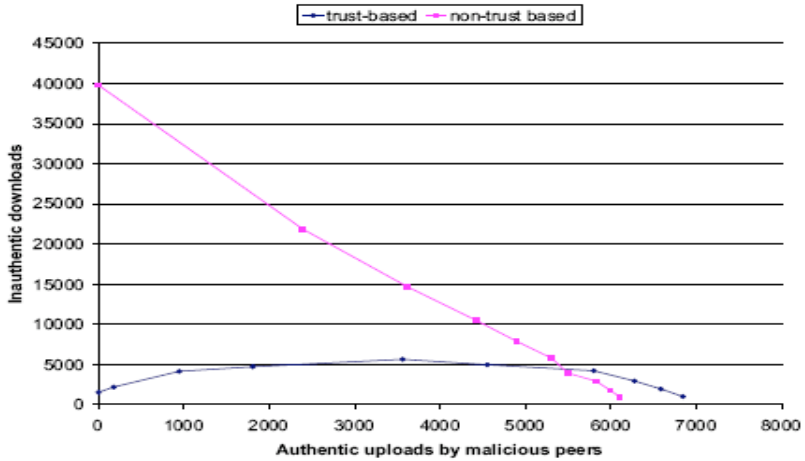
- A malicious collective i.e malicious peers aware of each other

Threat Models(cont..)

Threat Model C

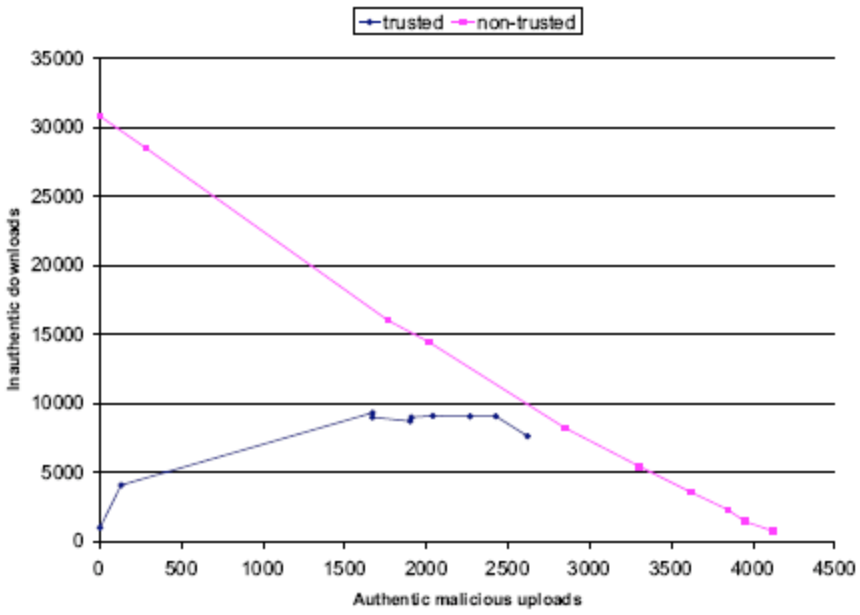
- Malicious collectives with camoflouge
- Malicious peers try to get high local trust values from good peers by providing authentic files in some cases





Threat Models(cont..)

- Threat Model D
  - Malicious Spies
  - One malicious group provides only authentic files.
  - Boosts the trust value of malicious group providing inauthentic files.



Strengths

- Self Policing
- Minimizes the impact of malicious peers on the performance of P2P systems
- Randomization
- No profit to newcomers.
- Probabilistic Algorithm
- Redundancy

Weakness

- Implications of secure score management like reputation management , incentive systems etc.
- Using Deterministic algorithm isolates malicious peers but causes overloading of highly trusted peers.
- Resistance to sybil attack
- Virus disseminators