



Neighbourhood Maps

Decentralised Ranking in Small-World P2P Networks

Matteo Dell'Amico

Dipartimento di Informatica e Scienze dell'Informazione
Università di Genova

UCL, September 28, 2006



Outline

- 1 Motivation
 - Free Riding Versus Cooperation
 - Issues with Reputation Evaluation
 - Related Work
- 2 Neighbourhood Maps
 - The Idea
 - Graph Distance
 - PageRank



Outline

- 1 Motivation
 - Free Riding Versus Cooperation
 - Issues with Reputation Evaluation
 - Related Work
- 2 Neighbourhood Maps
 - The Idea
 - Graph Distance
 - PageRank



Free Riding

- In P2P applications, peers are both **providers** (servers) and **consumers** (clients) of resources (storage, bandwidth, CPU, ...).
 - Usually, they are free to decide how much to allocate for others.
- Nodes like to be clients: they receive valuable service.
 - They don't want to be servers: resources are costly.
- Nodes are tempted to **free ride** by not providing resources to peers.



Free Riding

- In P2P applications, peers are both **providers** (servers) and **consumers** (clients) of resources (storage, bandwidth, CPU, ...).
- Usually, they are free to decide how much to allocate for others.
- Nodes like to be clients: they receive valuable service.
- They don't want to be servers: resources are costly.
- Nodes are tempted to **free ride** by not providing resources to peers.



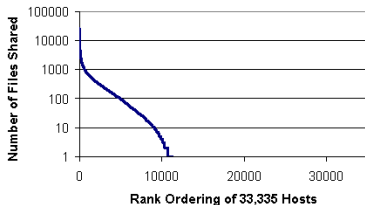
Free Riding

- In P2P applications, peers are both **providers** (servers) and **consumers** (clients) of resources (storage, bandwidth, CPU, ...).
 - Usually, they are free to decide how much to allocate for others.
-
- Nodes like to be clients: they receive valuable service.
 - They don't want to be servers: resources are costly.
-
- Nodes are tempted to **free ride** by not providing resources to peers.

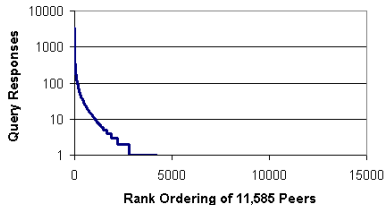
Free Riding in Gnutella

E. Hadar and B.A. Huberman, 2000

Rank Ordering of Peers by Number of Files Shared



Rank Ordering of Peers by Query Responses



- 70% of nodes does not share **any** file.
- Top 1% of nodes give **47%** of the service.



Reciprocity

- Goal: encouraging cooperation, so that selfish peers act in a socially acceptable way.
- Reciprocity: nodes that provide good service will receive good service.

Examples

- BitTorrent: “Tit-for-Tat” - a node uploads to peers that upload to it.
- eDonkey: credits - uploading peers obtain credits; those credits give priority in queues.



Reciprocity

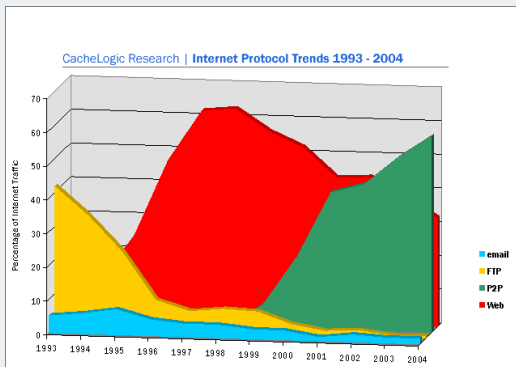
- Goal: **encouraging cooperation**, so that selfish peers act in a socially acceptable way.
- **Reciprocity**: nodes that provide good service will receive good service.

Examples

- BitTorrent: “**Tit-for-Tat**” - a node uploads to peers that upload to it.
- eDonkey: **credits** - uploading peers obtain credits; those credits give priority in queues.

Does Reciprocation Work?...

CacheLogic, 2005

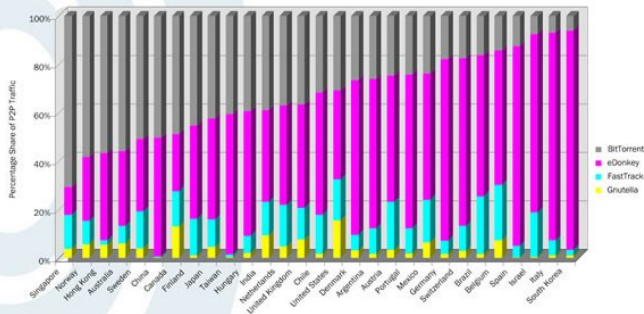


- Around 60% of Internet traffic comes from P2P applications...

... Looks Like It Does!

Cachelogic, 2005

P2P Market Share by Country

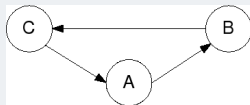


- ... and around 80% of P2P traffic (~50% of *all* Internet traffic!) uses reciprocative approaches.

... But ...

Asymmetry of Interest

- Direct reciprocation between A and B is only possible when A wants something from B *and* B wants something from A.
- There could be space for reciprocation, but nodes are not able to find it (Feldman et al., 2004).



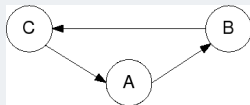
Idea

- Solution: indirect reciprocation. Nodes that have good behaviour build reputation; those that get a better reputation will be treated better.

... But ...

Asymmetry of Interest

- Direct reciprocation between A and B is only possible when A wants something from B *and* B wants something from A.
- There could be space for reciprocation, but nodes are not able to find it (Feldman et al., 2004).

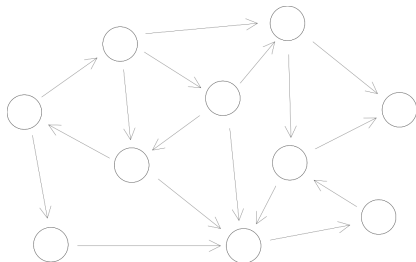


Idea

- Solution: **indirect** reciprocation. Nodes that have good behaviour build **reputation**; those that get a better reputation will be treated better.

Webs of Trust

- Peers start with local trust values (“recommendations”) about other peers.
- Nodes and trust values form a directed graph which we call a **web of trust**.



- Objective: assembling those values into **global reputation values**.



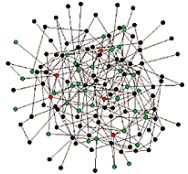
The OpenPGP web of trust

- Our test case: real data on trust relationships.
- An edge from A to B means “Owner of key A attests B’s key really belong to person B”.
- A *scale-free*, *small-world* network.

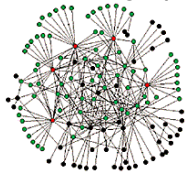
Graph Data

Nodes	27,398
Average Degree	8.99
Average Shortest Path	5.96
Clustering	0.371

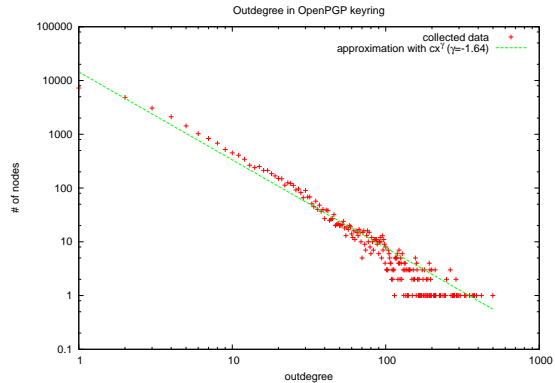
A Scale-Free Network



Random graph



Scale-free graph





Outline

- 1 Motivation
 - Free Riding Versus Cooperation
 - **Issues with Reputation Evaluation**
 - Related Work
- 2 Neighbourhood Maps
 - The Idea
 - Graph Distance
 - PageRank



Reputation in a P2P Network

The WWW as a Web of Trust

- To rank results, web search engines look at the graph structure of the WWW.
- Links can be seen as recommendations from a page to another.
- We can borrow some of those algorithms.

Main Challenges

- 1 Lack of global knowledge.
- 2 Cheap Identities.
- 3 Collusive Attacks.



Reputation in a P2P Network

The WWW as a Web of Trust

- To rank results, web search engines look at the graph structure of the WWW.
- Links can be seen as recommendations from a page to another.
- We can borrow some of those algorithms.

Main Challenges

- 1 Lack of global knowledge.
- 2 Cheap Identities.
- 3 Collusive Attacks.



No Global Knowledge

- Popular P2P networks have **large size** (millions of users).
- **High dynamicity**: nodes enter, exit, transactions are frequent.
- No node can have a global knowledge of the whole network.
- Solution: using either local knowledge or a distributed data structure (e.g., a DHT).



Cheap Identities

- New users can be created at low or no cost.
- *Whitewashing*: malicious users can create a new identity to discard history and erase their bad reputation.
- No point in punishing bad behaviour; it is nevertheless possible to reward cooperators (Friedman-Resnick, 2001).



Collusive Attacks

- Malicious users can introduce erroneous information.
- A group of malicious users can significantly alter the image of a node.
- *Sybil attack*: cheap identities are exploited in order to attack with many fake users (Douceur, 2002).



Outline

- 1 Motivation
 - Free Riding Versus Cooperation
 - Issues with Reputation Evaluation
 - **Related Work**
- 2 Neighbourhood Maps
 - The Idea
 - Graph Distance
 - PageRank



Related Work

Cheng and Friedman, 2005

- No symmetric approach is resilient to sybil attacks
- Some nodes have to be better than others.

EigenTrust (Kamvar et al., 2003)

- EigenTrust: a P2P implementation of PageRank.
- Needs set of pre-trusted nodes.

Feldman et al., 2004

- Subjective evaluation: the evaluating node pre-trusts itself.
- Requires global knowledge.



Related Work

Cheng and Friedman, 2005

- No symmetric approach is resilient to sybil attacks
- Some nodes have to be better than others.

EigenTrust (Kamvar et al., 2003)

- EigenTrust: a P2P implementation of PageRank.
- Needs set of pre-trusted nodes.

Feldman et al., 2004

- Subjective evaluation: the evaluating node pre-trusts itself.
- Requires global knowledge.



Related Work

Cheng and Friedman, 2005

- No symmetric approach is resilient to sybil attacks
- Some nodes have to be better than others.

EigenTrust (Kamvar et al., 2003)

- EigenTrust: a P2P implementation of PageRank.
- Needs set of pre-trusted nodes.

Feldman et al., 2004

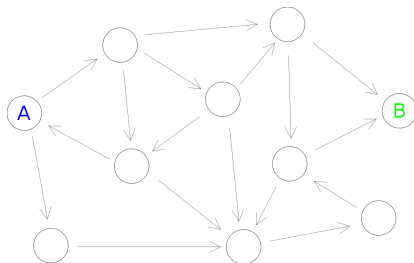
- **Subjective evaluation:** the evaluating node pre-trusts itself.
- Requires global knowledge.



Outline

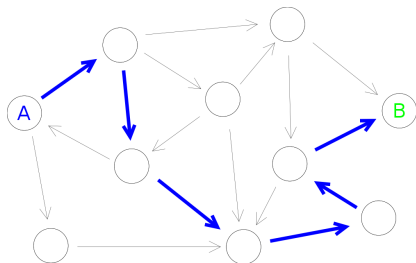
- 1 Motivation
 - Free Riding Versus Cooperation
 - Issues with Reputation Evaluation
 - Related Work
- 2 Neighbourhood Maps
 - **The Idea**
 - Graph Distance
 - PageRank

What Neighbourhood Maps Are



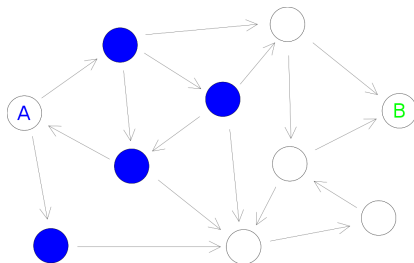
- Assumption: a web of trust, an evaluating node A and an evaluated node B.

What Neighbourhood Maps Are



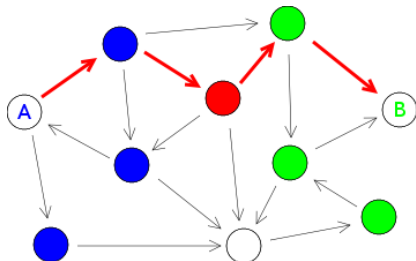
- We want to find **paths** from the evaluating to the evaluated node.

What Neighbourhood Maps Are



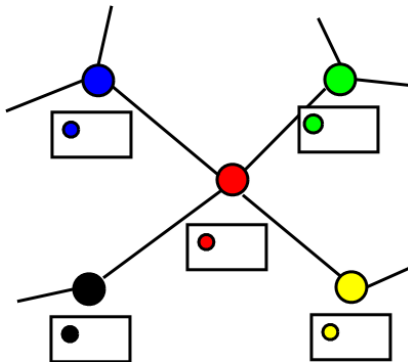
- A neighbourhood map for a node n is a view of the k nodes that are “closest” to n .

What Neighbourhood Maps Are



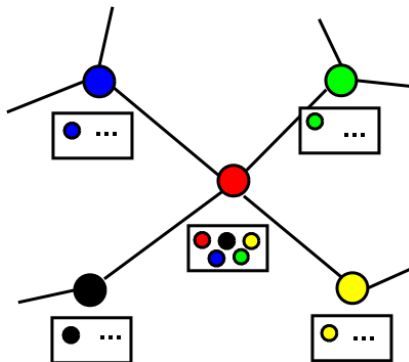
- Nodes appearing in the maps of evaluating and evaluated nodes provide us with data about paths from A to B.

How to Build Neighbourhood Maps



- Each node starts with knowledge about itself.

How to Build Neighbourhood Maps



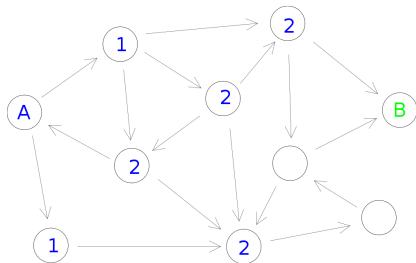
- Iteratively, they collect neighbour's maps until the desired size is reached.



Outline

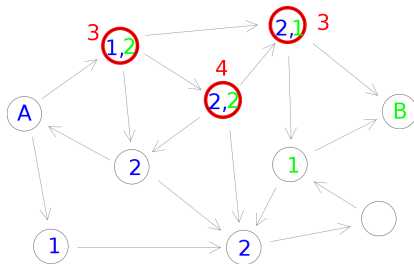
- 1 Motivation
 - Free Riding Versus Cooperation
 - Issues with Reputation Evaluation
 - Related Work
- 2 Neighbourhood Maps
 - The Idea
 - **Graph Distance**
 - PageRank

Calculating Graph Distance



- Straightforward implementation: each node keeps track of the k closest neighbours.

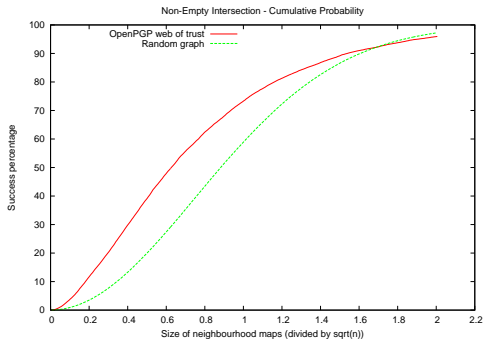
Calculating Graph Distance



- Distance for paths is the smallest sum of distances to intersections.

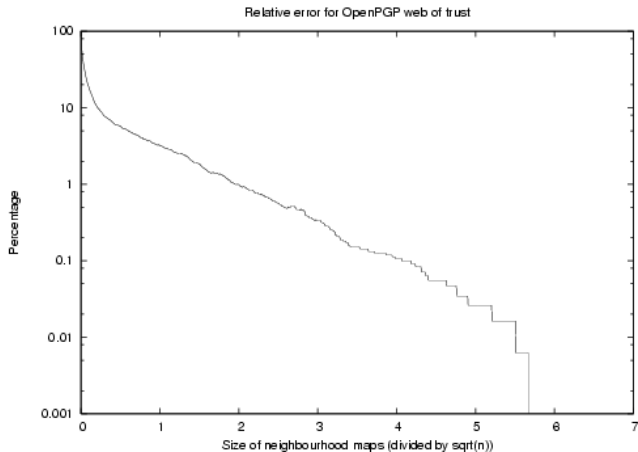


Non-Empty Intersections



- In order to have non-empty intersections, we need $O(\sqrt{n})$ nodes in maps.

Experimental Results





Outline

- 1 Motivation
 - Free Riding Versus Cooperation
 - Issues with Reputation Evaluation
 - Related Work
- 2 Neighbourhood Maps
 - The Idea
 - Graph Distance
 - PageRank



What PageRank Is

- Models a random walk on a network.
- With probability α , the walk stops. Otherwise, it chooses a random outgoing link.
- Nodes are ranked according to the probability of such walk to end in them.
- Used for ranking web search results in Google.
- In our case, the random walk starts at A ; we want to evaluate the probability of ending in B .

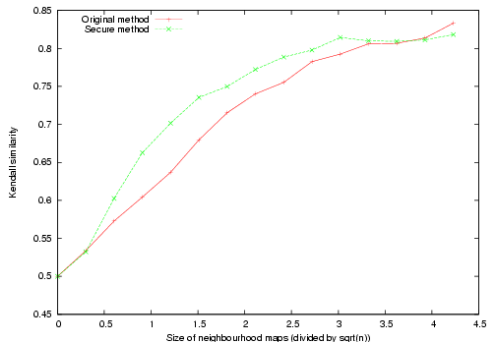


Neighbourhood Maps for PageRank

- A 's “outgoing” map: the k nodes with higher probability of being the endpoint of a random walk starting at A .
- B 's “incoming” map:
 - 1 the k nodes with a higher probability of a random walk starting from them and arriving at B ;
 - 2 the k nodes with a higher probability of a random walk starting at B , arriving at them and returning back to B .
- Ranking value: the sum of products for values in intersection.



Experimental Results



- Kendall similarity: probability that two rankings agree on which node is higher given a random pair.



Summary

- Reputation systems **encourage cooperation** and help solve the free rider problem.
- Neighbourhood maps are a feasible **decentralised** way to evaluate reputation metrics.
- Outlook
 - Performance enhancement and other metrics.
 - Evaluation of impact on real tasks.