Credence:
Experience with an Object Reputation System for Peer-to-Peer Filesharing

Ennan Zhai

(ennan.zhai@yale.edu)

Roadmap

I   What's the problem ? (i.e., background)

II   What are the existing solutions ?

III   Credence ?

IV   My evaluations on Credence ?

Roadmap

I   What's the problem ? (i.e., background)

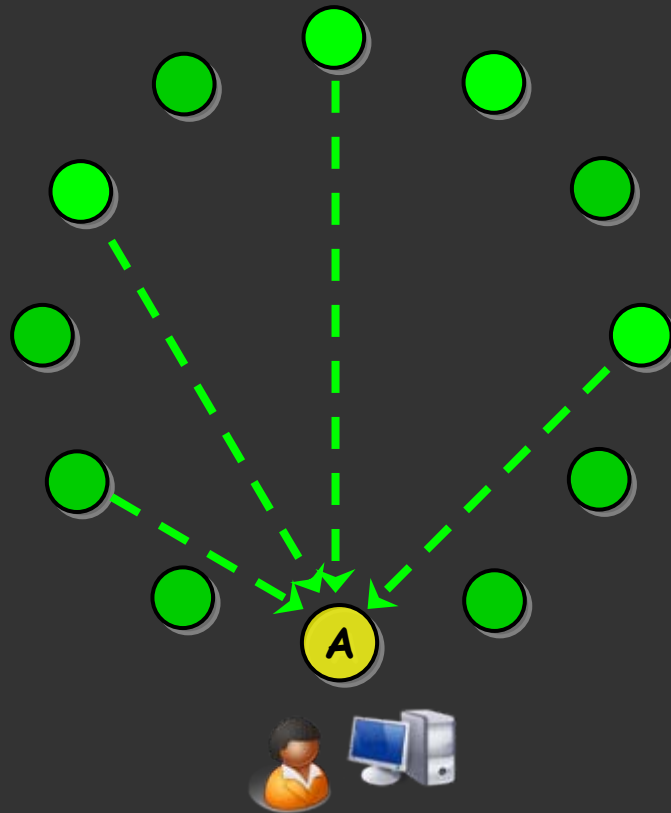**II**   What are the existing solutions ?
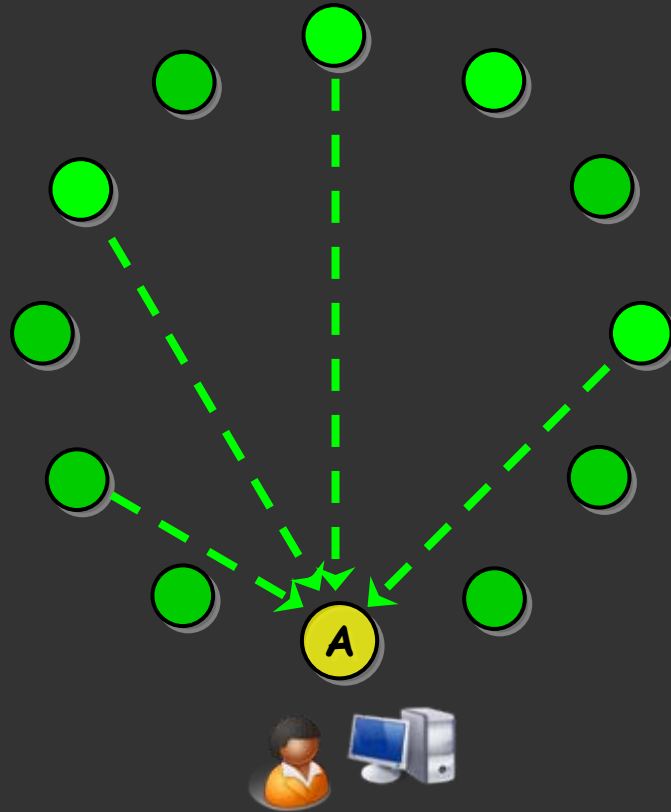
III   Credence ?

IV   My evaluations on Credence ?

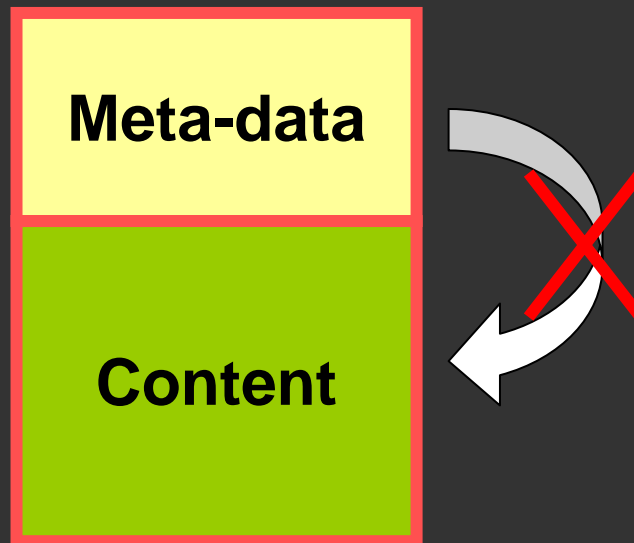# What's the problem?

What's the problem?

However, there exists file pollution problem in current P2P file-sharing systems.

What's the problem?

What is the file pollution ?

For a given file, if its meta-data description (e.g., keyword) does not match its content, we say the file is polluted.
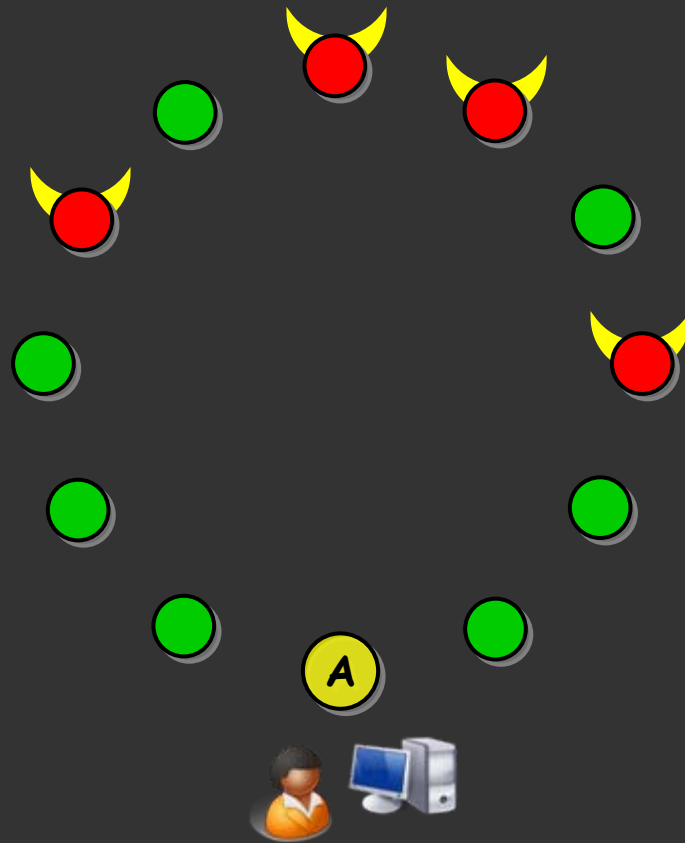
What's the problem?

Current P2P networks are full of polluted files [1]. Namely, pollution is a tough problem in current P2P networks.
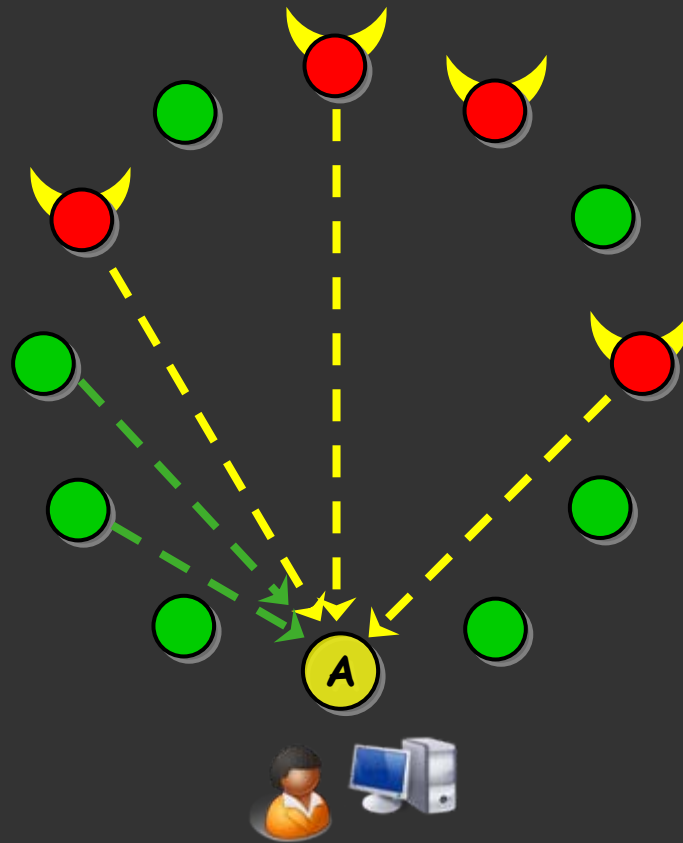
[1] J. Liang, Y. X. R. Kumar, and K. Ross, "Pollution in p2p file sharing systems," in Proceedings of IEEE Infocom'05.

What's the problem?

What's the problem?
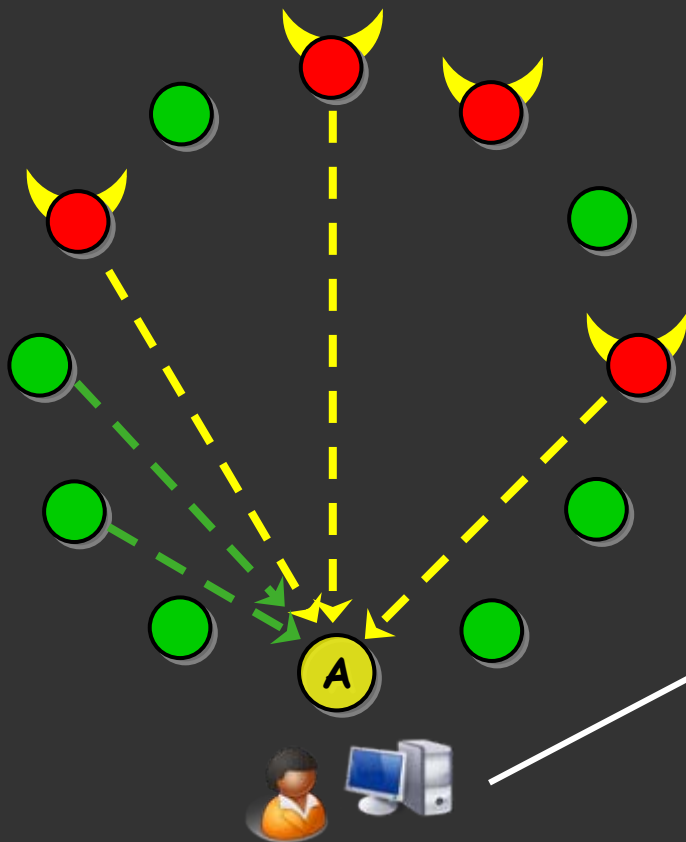
What's the problem?

It is high possible to select the polluted files !!!

Roadmap

I — What's the problem ? (i.e., background)

II — What are the existing solutions ?

III — Credence ?

IV — My evaluations on Credence ?

Roadmap

I What's the problem ? (i.e., background)

II What are the existing solutions ?
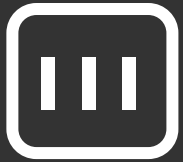
III Credence ?

IV My evaluations on Credence ?

Existing solutions?

Reputation-based Approaches:

* Peer-based: EigenTrust [WWW' 03], PeerTrust [P2PEc' 03], Scrubber [SAC' 06] … …

* Object-based: Credence [NSDI' 06] … …

* Hybrid: XRep [CCS' 02], X^2Rep [ACNS' 04], Hybrid Scrubber [P2P' 07] … …

Roadmap

I What's the problem ? (i.e., background)

II What are the existing solutions ?

III Credence ?

IV My evaluations on Credence ?

Roadmap

I   What's the problem ? (i.e., background)

II   What are the existing solutions ?

III   Credence ?

IV   My evaluations on Credence ?

Credence

Notice that it is a personalized reputation !!!

Request

Alice

P2P Networking

Credence

Name	Sources	Voters
F_{10}	P_2 P_6	P_2 P_4 P_6
F_{22}	P_2 P_6 P_8	P_2 P_7
F_4	P_2 P_4	P_2 P_4 P_7
F_6	P_{11} P_{13} P_{14}	P_{11}
... ...	... ...	... ...

Alice

Credence

Name	Sources	Voters
F_{10}	P_2 P_6	P_2 P_4 P_6
F_{22}	P_2 P_6 P_8	P_2 P_7
F_4	P_2 P_4	P_2 P_4 P_7
F_6	P_{11} P_{13} P_{14}	P_{11}
… …	… …	… …

Alice

Credence

$$\text{Rep}(F) = \frac{\sum\limits_{i=1}^{n} V_i \cdot \theta_{(\text{Alice,Voter\_i})}}{\sum\limits_{i=1}^{n} \left| \theta_{(\text{Alice,Voter\_i})} \right|} \qquad [\,\text{-1, 1}\,]$$

Credence uses weighted averaging to compute the reputation of an object.

Credence

The vote cast by voter$_i$ on F (+1 or -1)

$$Rep(F) = \frac{\sum_{i=1}^{n} V_i \cdot \theta(Alice, Voter\_i)}{\sum_{i=1}^{n} |\theta(Alice, Voter\_i)|}$$

The similarity between Alice and voter$_i$
The range is [-1, +1].

The # of voters who cast votes on file F

Credence

The vote cast by voter$_i$ on F (+1 or -1)

$$Rep(F) = \frac{\sum_{i=1}^{n} V_i \cdot \theta(Alice, Voter\_i)}{\sum_{i=1}^{n} |\theta(Alice, Voter\_i)|}$$

?

The similarity between Alice and voter$_i$
The range is [-1, +1].

The # of voters who cast votes on file F
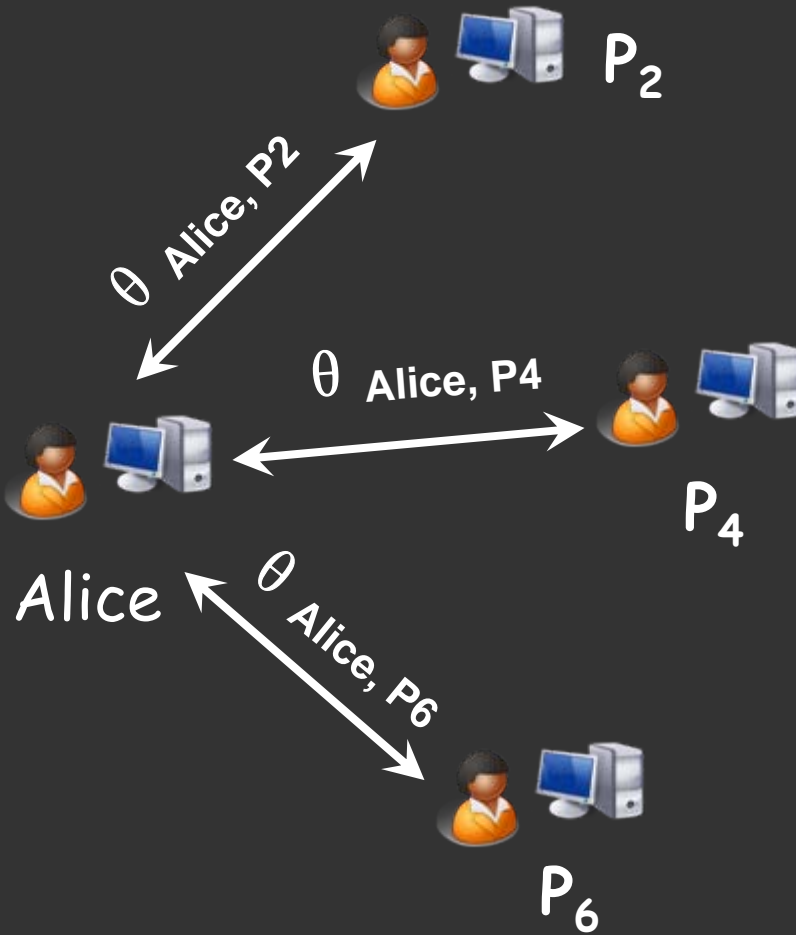
Credence

Credence

$$\theta = (p-ab) / \sqrt{a(1-a) \, b(1-b)}$$

For the overlapping voting set (e.g., S) between Alice and C_i :

$$a = \frac{\text{\# of positive votes cast by Alice on the files in S}}{\text{\# of all the votes cast by Alice on the files in S}}$$

$$b = \frac{\text{\# of positive votes cast by } C_i \text{ on the files in S}}{\text{\# of all the votes cast by } C_i \text{ on the files in S}}$$

$$p = \frac{\text{\# of positive votes cast by both Alice and } C_i \text{ on the files in S}}{\text{\# of all the votes cast by both Alice and } C_i \text{ on the files in S}}$$

Credence

Example ... ...

File 1	File 2	File 3	File 4
A: +1 B: +1	A: -1 B: -1	A: -1 B: +1	A: +1 B: +1

Credence

Current similarity mechanism cannot robustly evaluate the relationship between a client and peers having only few interests in common with the client (i.e., lack of overlapping voting sets).
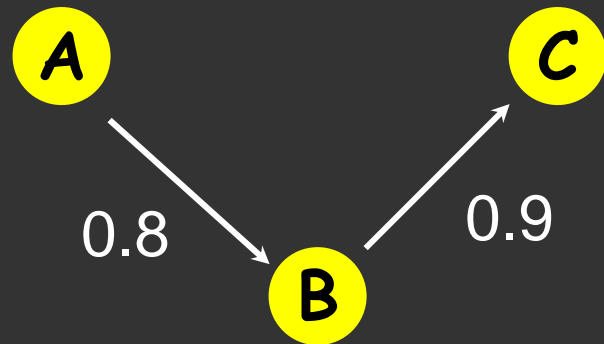
Credence

Current similarity mechanism cannot robustly evaluate the relationship between a client and peers having only few interests in common with the client (i.e., lack of overlapping voting sets).

Flow-based Reputation:

$$\theta_{ac} = \theta_{ab} * \theta_{bc}$$
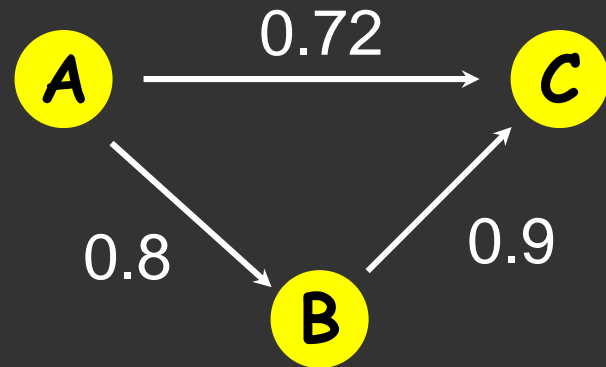
A   C

0.8   0.9

B

Credence

Current similarity mechanism cannot robustly evaluate the relationship between a client and peers having only few interests in common with the client (i.e., lack of overlapping voting sets).

Flow-based Reputation:

$$\theta_{ac} = \theta_{ab} * \theta_{bc}$$

Credence

$$Rep(F) = \frac{\sum_{i=1}^{n} V_i \cdot \theta \, (\text{Alice,Voter\_i})}{\sum_{i=1}^{n} \left| \theta \, (\text{Alice,Voter\_i}) \right|}$$

Back to previous reputation equation … …

Credence

Why it works ?

Credence

Name	Sources	Voters
F_{10}	P_2 P_6	P_2 P_4 P_6
F_{22}	P_2 P_6 P_8	P_2 P_7
F_4	P_2 P_4	P_2 P_4 P_7
F_6	P_{11} P_{13} P_{14}	P_{11}
... ...	... ...	... ...

Alice

Credence

Name	Sources	Voters
$F_{10} = 0.8$	P_2 P_6	P_2 P_4 P_6
F_{22}	P_2 P_6 P_8	P_2 P_7
F_4	P_2 P_4	P_2 P_4 P_7
F_6	P_{11} P_{13} P_{14}	P_{11}
… …	… …	… …

Alice

Credence

Name	Sources	Voters
$F_{10} = 0.8$	P_2 P_6	P_2 P_4 P_6
$F_{22} = 0.5$	P_2 P_6 P_8	P_2 P_7
F_4	P_2 P_4	P_2 P_4 P_7
F_6	P_{11} P_{13} P_{14}	P_{11}
… …	… …	… …

Alice

Credence

Name	Sources	Voters
$F_4 = 0.9$	P_2 P_4	P_2 P_4 P_7
$F_{10} = 0.8$	P_2 P_6	P_2 P_4 P_6
$F_6 = 0.6$	P_{11} P_{13} P_{14}	P_{11}
$F_{22} = 0.5$	P_2 P_6 P_8	P_2 P_7
… …	… …	… …

Alice

Roadmap

I   What's the problem ? (i.e., background)

II   What are the existing solutions ?

III   Credence ?

IV   My evaluations on Credence ?

Roadmap

I — What's the problem ? (i.e., background)

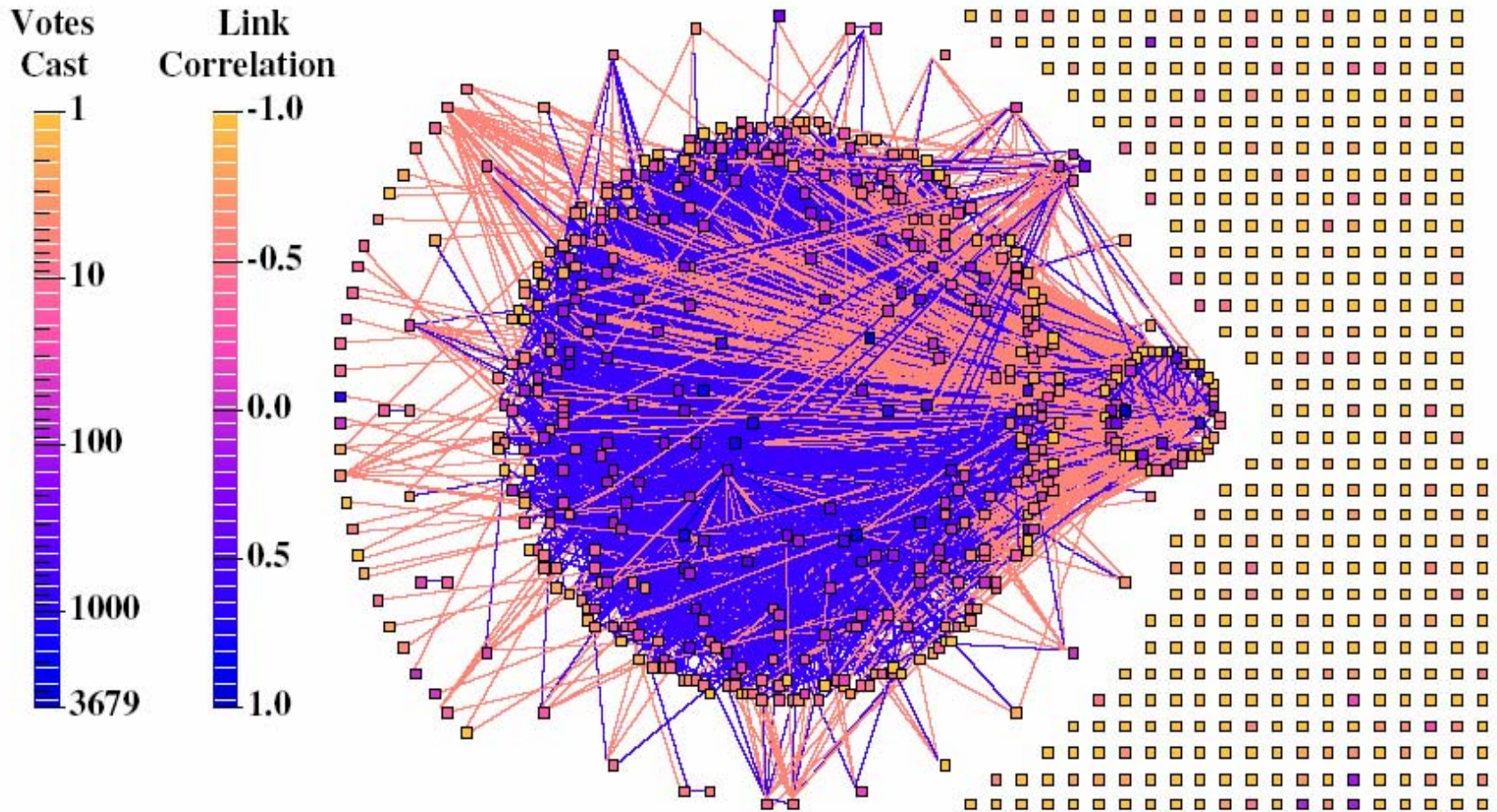**II** — What are the existing solutions ?

III — Credence ?

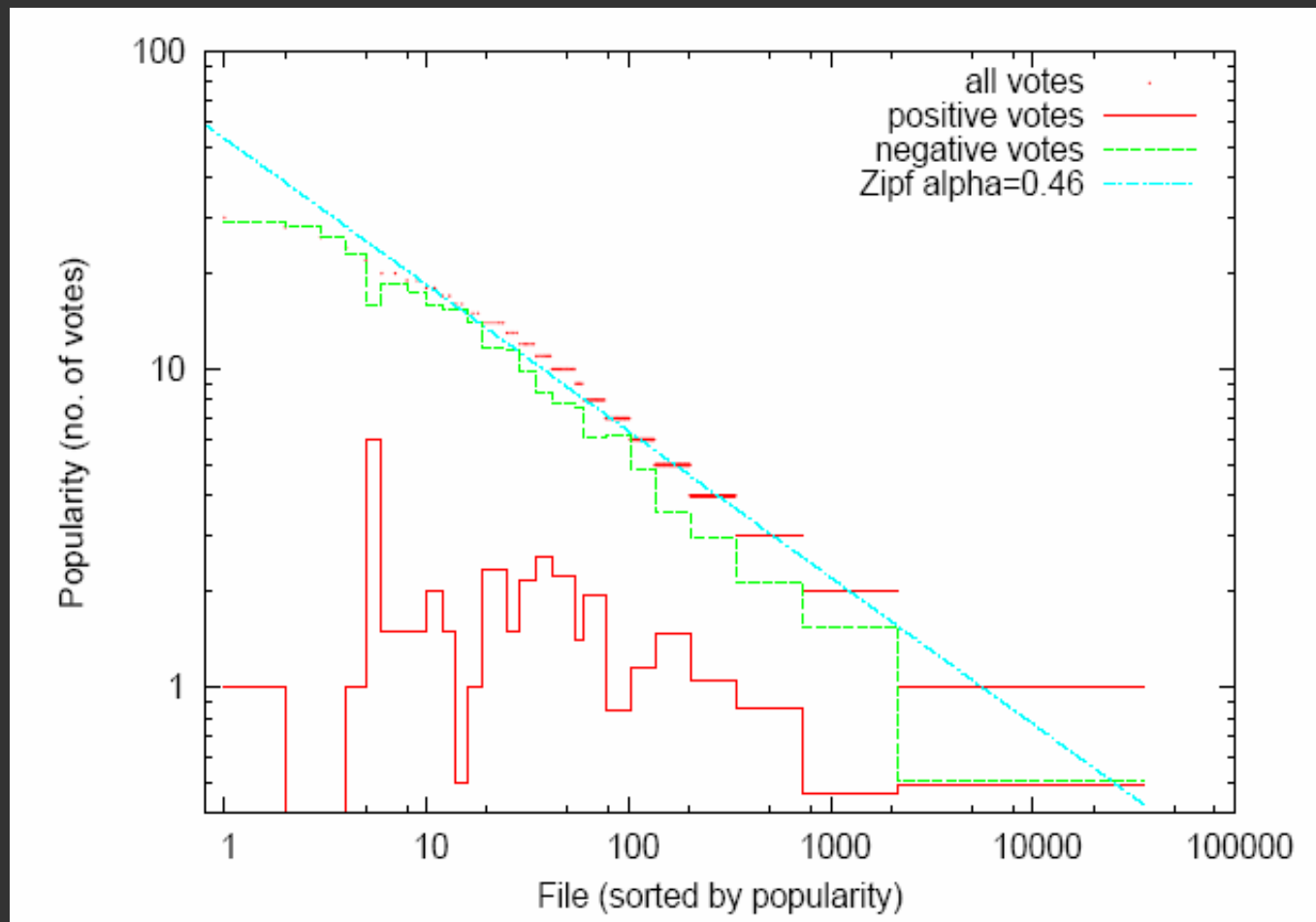**IV** — My evaluations on Credence ?

Results

Positive and negative correlations

Results

More negative votes than positive votes

Evaluations

Credence has several problems:

- Cold start

- Lack of overlapping vote sets

- Collusive attacks

Evaluations

Credence has several problems:

- Cold start

- Lack of overlapping vote sets

- Collusive attacks

Evaluations

We built Green, a social network-based P2P reputation system [SIGCOMM'09 poster].

Green client is able to overcome the challenges from cold start and lack of overlapping vote sets by leveraging the information of its social networking.

Evaluations

Credence has several problems:

- Cold start

- Lack of overlapping vote sets

- Collusive attacks

Evaluations
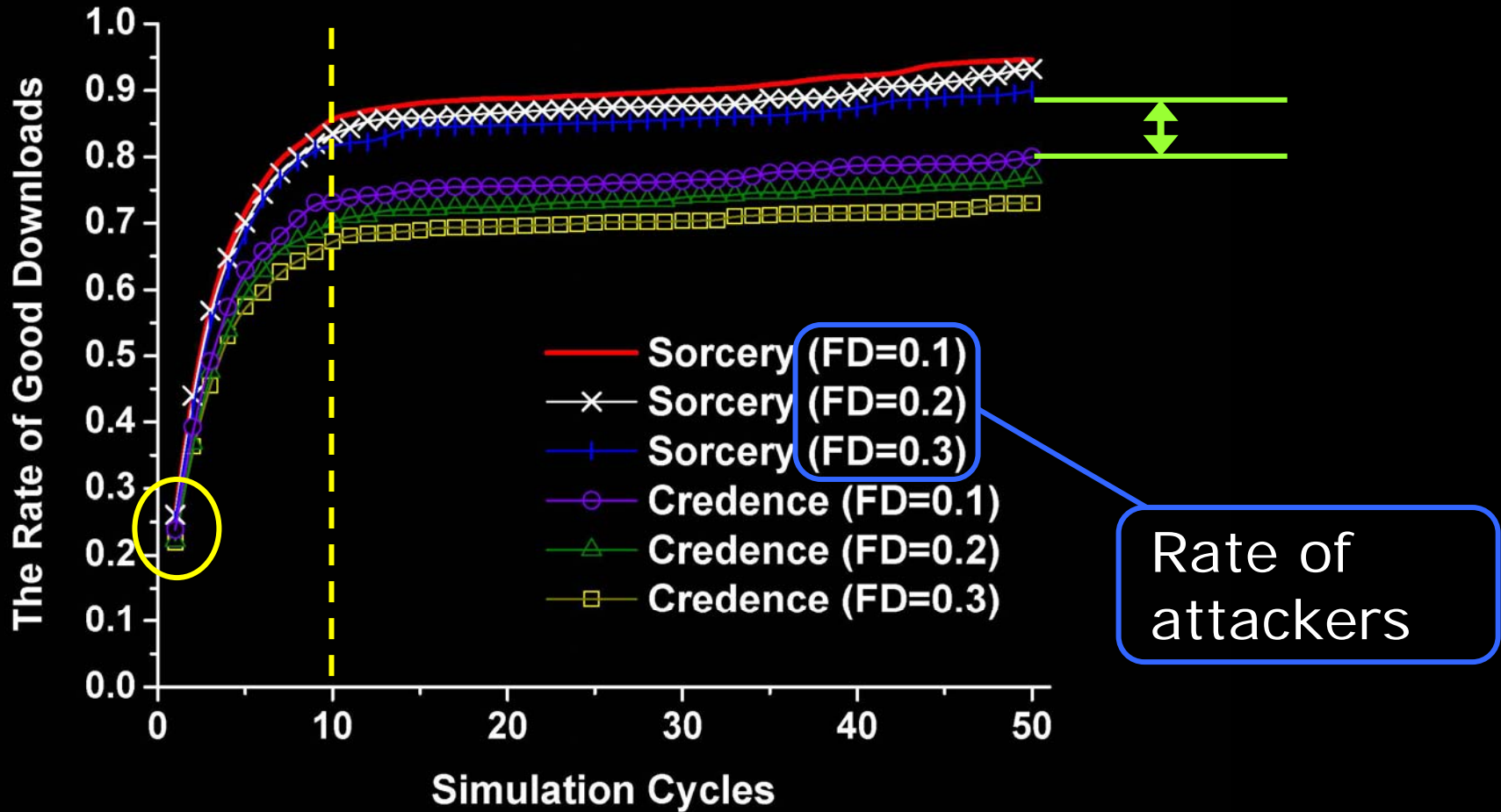
Credence has several problems:

- Cold start

- Lack of overlapping vote sets

- Collusive attacks

Evaluations

We proposed Sorcery, a challenge-response mechanism based on dominant information strategy in Game Theory [P2P'09 & PPNA].
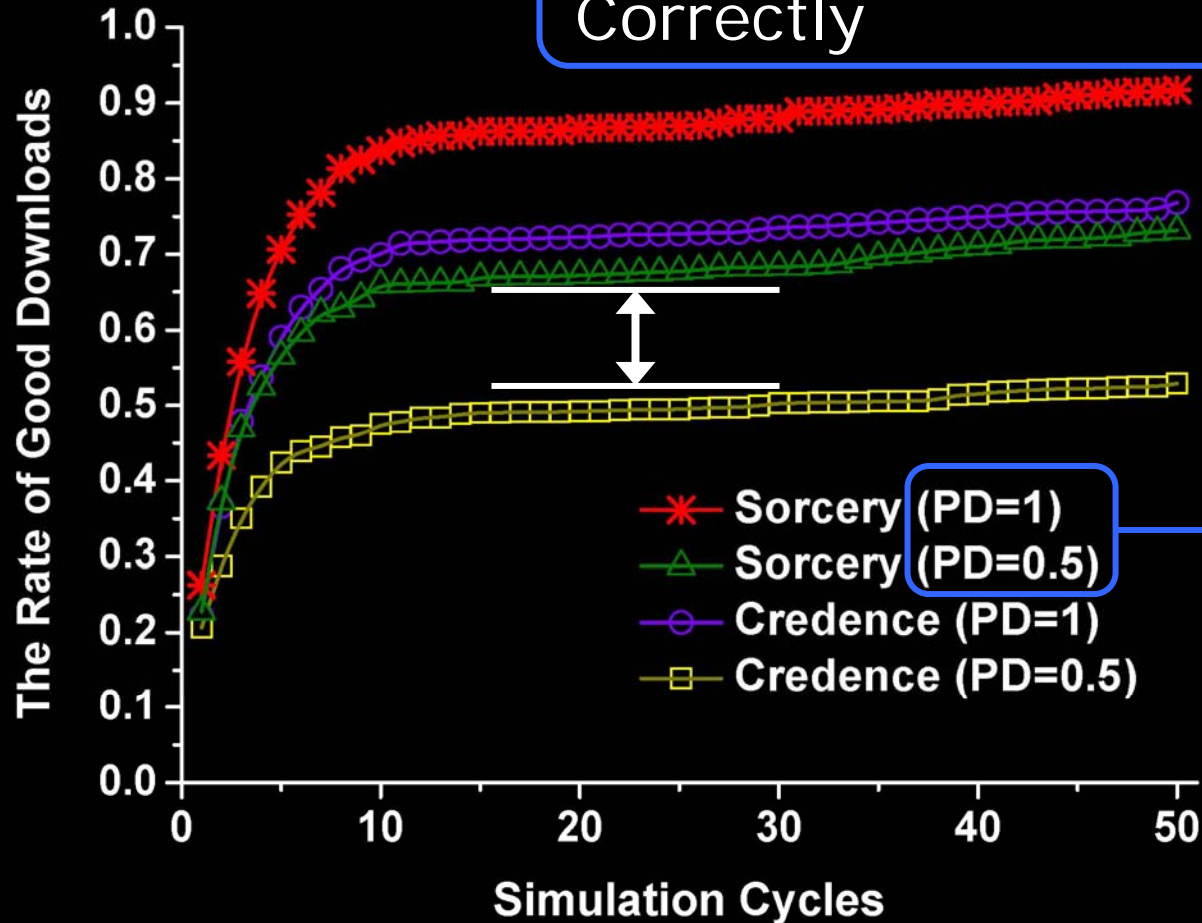
Sorcery can address the problem of strong collusion attacks in P2P networks.

Normal Attackers

Tricky Attackers

Thank you !!!

Appendix