Agent-based modelization for p2p networks

textbfJulien Siebert,
PhD advisors : Vincent Chevrier et Laurent Ciarletta
{julien.siebert, vincent.chevrier, laurent.ciarletta}@loria.fr

LORIA : MADYNES & MAIA teams - Campus Scientifique - BP 239 - 54506 Vandœuvre-lès-Nancy Cedex

March 4, 2008

A brief teams preview

MAIA

Intelligent and Autonomous MAchine.

MADYNES

MAnagement of DYnamic NEtworks and Security

- Distributed artificial intelligence.
- Multi-agent simulation.

PhD Thesis

Intersection of these domains.

PhD general context

- Dynamic, autonomous, ubiquitous networks and applications.
- Users' behaviour impacts the Quality of Services (QoS) and vice versa.

Problematic

Study mutual influences of *user's behaviour* and *QoS* in P2P systems.

P2P as a case study

- Expected advantages :
  - Robustness, fault tolerance.
  - Scalability (million users).
- Influence of user behaviour :
  - Resource availability.
    - Free-riding problem [AH00, HCW05].
    - Churn, Sessions durations.
  - Content.
    - Poisoning [Lia05].
    - Illegality.
- Influence of QoS :
  - Large bandwith consumption.
  - Downloading time.

Outline

Why simulations

- Realistic systems :
  - Large scale : from 10^3 to 10^6 users.
  - Real P2P protocols.
- Control experimental parameters from networks and from users.
- Repeatable experiments.
- Granularity : assess influences of :
  - Local users' behaviour on global network.
  - Local network parts on behaviours.

Identified issues

To make several levels of representation interact : users and QoS.

Modelling needs

- Realistic network parameters (QoS).
- Dynamic, heterogeneous behaviours, actions, interactions (Users).

Network modelling : state of the art

Issue # 1

Model (some) realistic network parameters.

- Test protocols, network dimensioning (objective measures/metrics).
- Users reduced to an input (as a packets/messages generator).

Our observation

User's parameters (dynamic behaviours, heterogeneousness) :

- Seldom represented in simulation tools ([NLB+07]).
- Difficult to integrate into network model.

Multi-agent modelling : brief presentation

- Complex distributed systems simulations (social sciences [PA07], animation [mas]).

http://www.massivesoftware.com

Principle

- Directly represents entities behaviours, actions, interactions.
- Impact of local behaviour on the global system.

User's modelling : state of the art

Issue # 2

Model dynamic, heterogeneous behaviours, actions, interactions.

- Game theory : free-riding model [FPCS04].
- Multi-Agent based systems : PeerSim [Hal04, Pee07].

Our observation

Networks parameters (protocols, delays, error rates...) :

- Seldom represented in simulation tools.
- Difficult to integrate into users' models.

State of the art synthesis

In short :

	network	application	user
Agent-based simulation	✗	✗	✓
Network simulation	✓	✓	✗

PhD starting point

- Study the mutual influence of users' behaviour and QoS.
- Models exist for each level of representation.
- Integrate them into a multi-model approach.

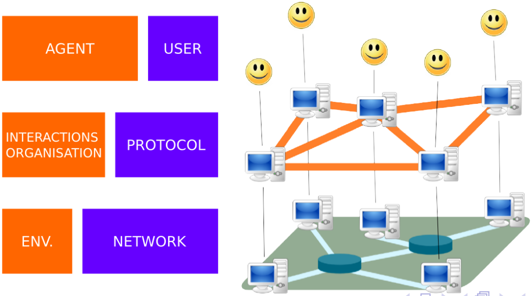
Outline

Linking together multi-agent and P2P

Idea

Use the multi-agent paradigm to integrate and make models interact.

- Agent (proactive, dynamic...) : user.
- Interactions : P2P protocols.
- Environment (interaction medium) : underlying physical networks.



Outline

Choosing a simulator

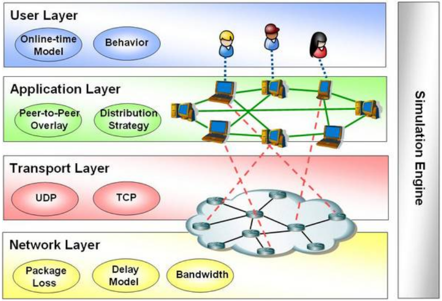
Case study problematic (Master thesis [SCC08])

Problems of sharing and pollution in a P2P file sharing network.

- Modelling needs :
  - Behaviour : relatively simple (reactive agent).
  - Interactions : real life protocols.
  - Network : delays of messages/data transfers.
- Available tools :
  - Multi-agent simulator : network and protocol absent.
  - Packet level simulator : too detailed, scalability ?
  - **Overlay Simulator** : good compromise.

Peerfactsim.KOM (PFS) [Pee]

- Overlay Simulator, Darmstadt, Java.
- Protocols : Chord, **Kademlia**.
- Scalability.
    - Authors claimed 10^5 nodes [Dar05].
    - Our current experiments : 50000 nodes.
- Architecture : representation layers available



http://peerfact.org

Our adaptation of PFS

PFS is mainly an overlay simulator.

- Data model + Pollution rate.
- Concrete exchange of data.
- A model of user : present hooks but not instanciate.

A model of user based upon reactive agent

- Definition :
  - Few internal states.
  - Compact representation of the environment and the other agents.
- Instanciation :
  - Perceptions
    - Upload/download bandwidth : cost for sharing/downloading.
    - Number of available sources : scarcity of a resource.
    - Expected downloading time : patience attribute.
  - Actions
    - Connexion/disconnexion.
    - Searching for/downloading resources.
    - Sharing or not resources.
    - Pollution rate controlling.
  - Internal states.
    - Willingness to share [FC05].
    - Pollution awareness [LCC+06].

Outline

Experiments goals

Technical

- Bug free.
- Realistic behaviour.
- Scalability.

Conceptual

- First assessement of our approach.

Scenario of experiments

- Network of N users. Initially :
  - P initial publishers, f files each.
  - B_0 are polluted (bogus) G_0 are genuine. $B_0 + G_0 = P * f$.
- Step 1 : Searching for resources.
- Step 2 : Asking for download.
- Step 3 : Checking the download process. If download completed (4), otherwise (5).
- Step 4 : Controlling pollution. File ok : Sharing ? otherwise deletion and (5).
- Step 5 : Launch the whole process again ?

Results

Realistic behaviour + Scalability.

- Ideal case : no pollution $B_0 = 0$, no selfish user.
- Network size from $N = 25$ to $N = 50000$.
- Impact of the number of initial publishers P on the load per node.

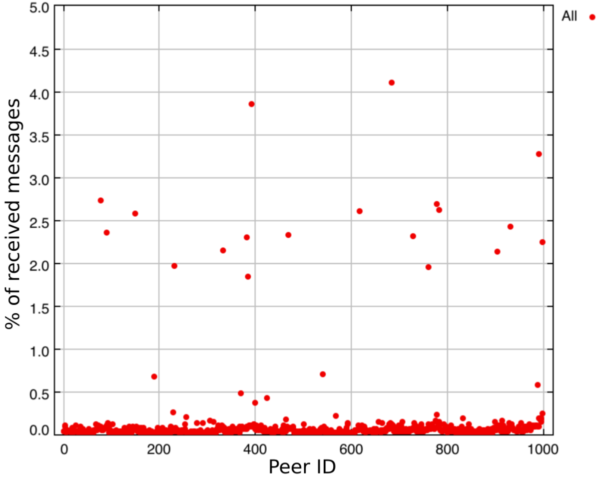
Expected results

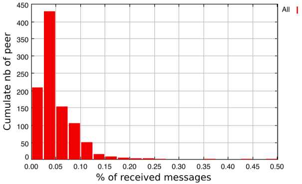
- when P increases, load equally spreads over all nodes.

Result example (1)

Realistic behaviour + Scalability.

$N = 1000$ $f = 1$ $P = 10$

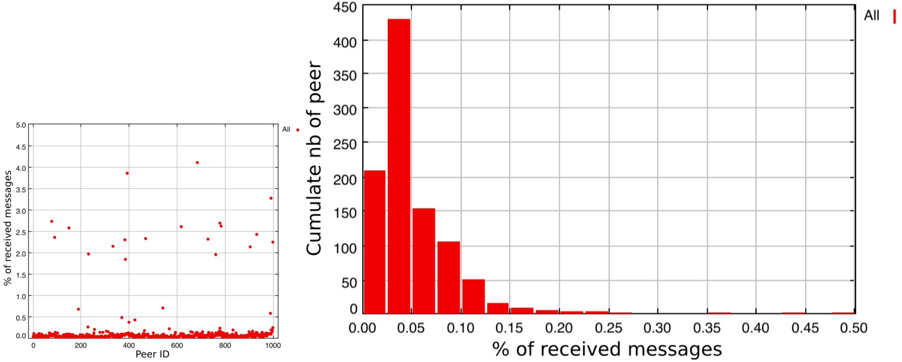




Result example (1)

Realistic behaviour + Scalability.

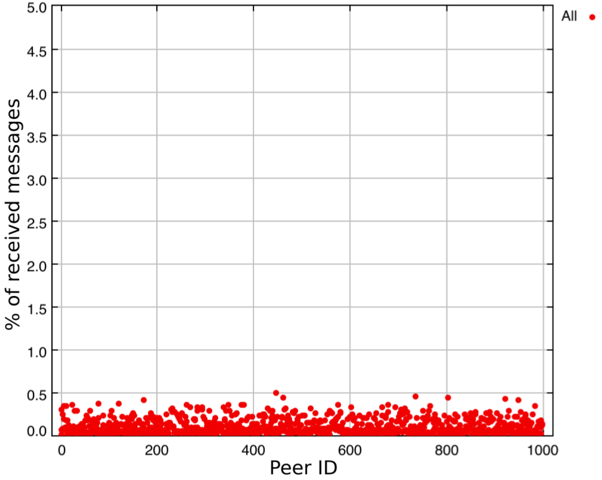
$N = 1000 \ f = 1 \ P = 10$

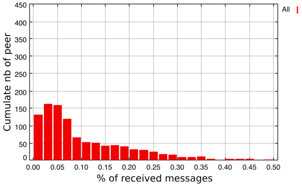


Result example (2)

Realistic behaviour + Scalability.

$N = 1000 \ f = 1 \ P = 500$

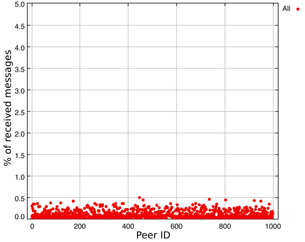


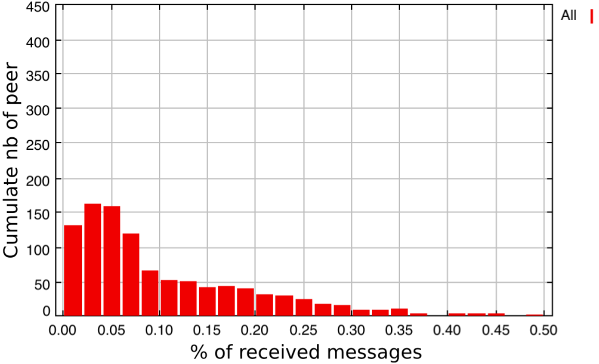


Result example (2)

Realistic behaviour + Scalability.

$N = 1000 \ f = 1 \ P = 500$

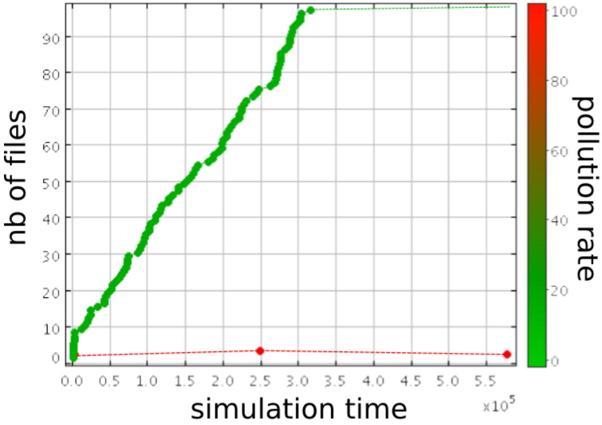




Result example with pollution

Same experiments with local polluters

$N = 100 \ f = 1 \ P = 20 \ B_0 = 2$

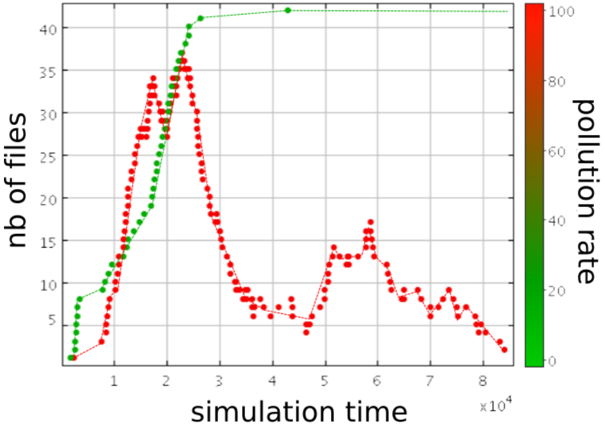


majority case[LCC+06]

Result example with pollution

Same experiments with local polluters

$N = 100 \; f = 1 \; P = 20 \; B_0 = 2$



unexpected results ?

Analysis

Technical feasibility

- Platform assessment.
- Scalability : 50000 nodes in 5min, 4Gb RAM, java 1.6.

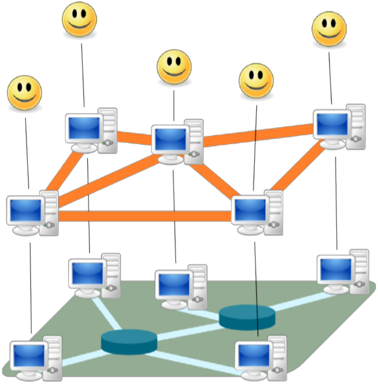
Conceptual approach

- Model + Tool + strengthen convictions.
- But results not mature, for the moment.
- Reach the limits of the tool (programming).

Outline

Take home messages

- Problem : study mutual influences of users' beahviour and QoS.
- Proposition : multi-model approach.
  - Different levels of representations.
  - agent + application + network (model).
- First implementation : adaptation an existing tool.
- First experiments : proof of concept.



Currents and future directions

- Study mutual influences of users' beahviour and QoS (P2P).
- Continue experiments.
- Multi-model issues : different time and space scales.
- Build a generic framework (dynamic, autonomous networks).

Thank you for your attention

Questions ?

E. Adar and B. Huberman.
Free riding on gnutella, 2000.

Vasilios Darlagiannis.
Overlay Network Mechanisms for Peer-to-Peer Systems.
PhD thesis, Vom Fachbereich Informatik der Technischen Universität Darmstadt, 2005.

Michal Feldman and John Chuang.
Overcoming free-riding behavior in peer-to-peer systems.
SIGecom Exch., 5(4):41–50, 2005.

Michal Feldman, Christos Papadimitriou, John Chuang, and Ion Stoica.
Free-riding and whitewashing in peer-to-peer systems.
In *PINS '04: Proceedings of the ACM SIGCOMM workshop on Practice and theory of incentives in networked systems*, pages 228–236, New York, NY, USA, 2004. ACM Press.

David Hales.

From selfish nodes to cooperative networks " emergent link-based incentives in peer-to-peer networks.
In *P2P '04: Proceedings of the Fourth International Conference on Peer-to-Peer Computing (P2P'04)*, pages 151–158, Washington, DC, USA, 2004. IEEE Computer Society.

Daniel Hughes, Geoff Coulson, and James Walkerdine.
Free riding on gnutella revisited: The bell tolls?
IEEE Distributed Systems Online, 6(6):1, 2005.

Uichin Lee, Min Choiz, Junghoo Choy, M. Y. Sanadidiy, and Mario Gerla.
Understanding pollution dynamics in p2p file sharing.
In *5th International Workshop on Peer-to-Peer Systems (IPTPS'06)*, Santa Babara, CA, USA, February 2006.

Jian Liang.
Pollution in p2p file sharing systems, 2005.

http://www.massivesoftware.com/.

S. Naicken, B. Livingston, A. Basu, S. Rodhetbhai, I. Wakeman, and D. Chalmers.
The state of peer-to-peer simulators and simulations.
Accepted for editorial publication in ACM SIGCOMM journal for Computer Communication Review, 2007.

D. Phan and F. Amblard.
Agent-Based Modelling and Simulation in the Social and Human Sciences.
Bardwell Press, Oxford, 2007.

PeerfactSim.KOM: A Simulator for Large-Scale Peer-to-Peer Networks.
http://www.peerfactsim.com.

http://peersim.sourceforge.net/, 2007.

Julien Siebert, Vincent Chevrier, and Laurent Ciarletta.
Modélisation multimodèle des réseaux dynamiques : cas des réseaux pair-à-pair (in french).

In *JDIR'08 - 9èmes Journées Doctorales en Informatique et Réseaux*, Villeneuve d'Ascq, France, 2008.