

# Virtual Community Dynamics

by Stefano Mazzocchi <[stefano@apache.org](mailto:stefano@apache.org)>

# Part I

## Virtual Communities

# Virtual Community (VC)

The concept that people form social aggregations based on common interest in an online environment.

# Examples of VC

- open software development groups
- corporate workgroups
- gaming networks

# Explicit Characteristics

- Interest
- Communication Media
- Accessibility

# Open Source VC

- Interest: creation, improvement, maintenance, or use of a software program.
- Communication: email, chat
- Accessibility: open to everyone

# Corporate VC

- Interest: related to the job
- Communication media: email, telephone calls, telephone conferences, videoconferencing
- Accessibility: restricted

# Gaming VC

- Interest: playing a game
- Communication media: mostly the game's own communication tools, but also forums, email, chat
- Accessibility: restricted or open, depending on the game

# Implicit Characteristics

- activity
- health
- collaboration

# Problems

- Measuring implicit characteristics objectively is impossible.
- Defining heuristics to estimate implicit characteristics creates social friction: nobody likes to be “rated”.

# Solutions

- Let humans judge, not machines
- Let machines do the tasks that do not require judgment

# Part 2

## Apache Agora

# What is Agora?

Agora is a graph visualizer bounded with processing scripts that extract topologic information from mail archives.

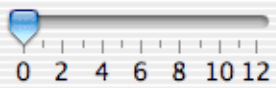
# Apache Agora 1.2

- Data
  - ▶ jakarta.apache.org
  - ▶ xml.apache.org
  - ▶ tcl.apache.org
  - ▶ httpd.apache.org
  - ▶ apache.org
  - ▶ db.apache.org
  - ▶ avalon.apache.org
  - ▶ ant.apache.org
  - ▶ ws.apache.org
  - ▶ incubator.apache.org
  - ▶ maven.apache.org
  - ▶ apachecon.com

calculation: 19 ms  
drawing: 202 ms  
nodes: 182  
edges: 888



Link Decay



Clear

Load

Controls

Drawing

Highlight

Parameters

- Antialias
- Nodes
- Edges
- Edge Values
- Groups
- Timing
- Background

# The Idea

- A message in reply to another message indicates some social connection between the sender and the receiver
- The more replies, the stronger the connection

# The Process

- the script processes mail archives and extracts information from the email headers about reply
- the application elaborates this information and visualizes the resulting graph

# How does it work?

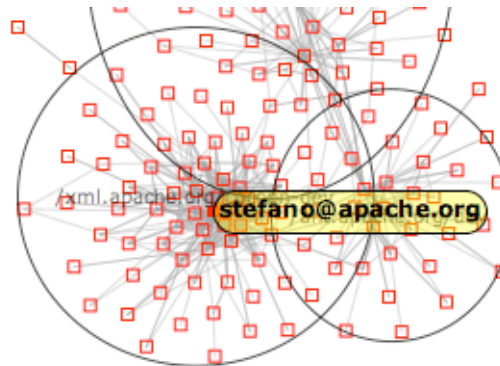
- each node has a repulsive force (this models the natural tendency of people to have a private space surrounding them)
- each node exhibits an attractive force with the linked nodes (the force is proportional with the strength of the link)

# Dynamic Visualization

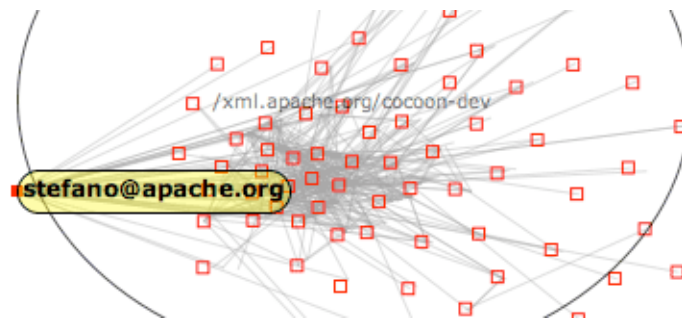
- Allows interactivity and what-if scenarios (empowers better human judgment)
- Explains global graph topology balance thru lack of motion
- ... but very computationally expensive! (grows with the square of the number of nodes)

# Features

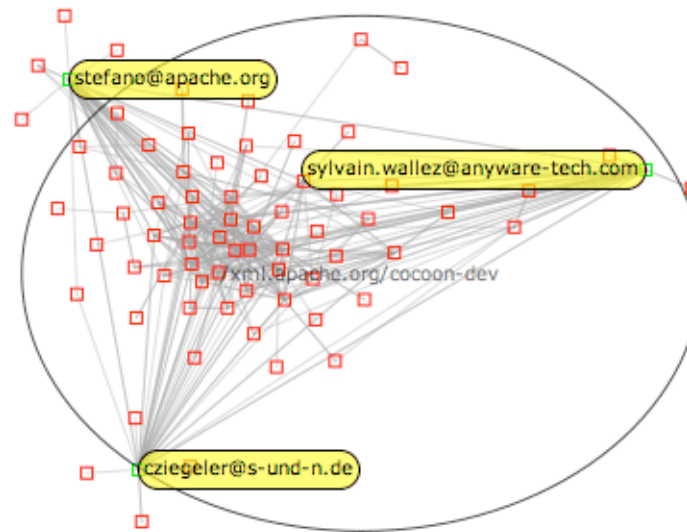
# Selecting a node



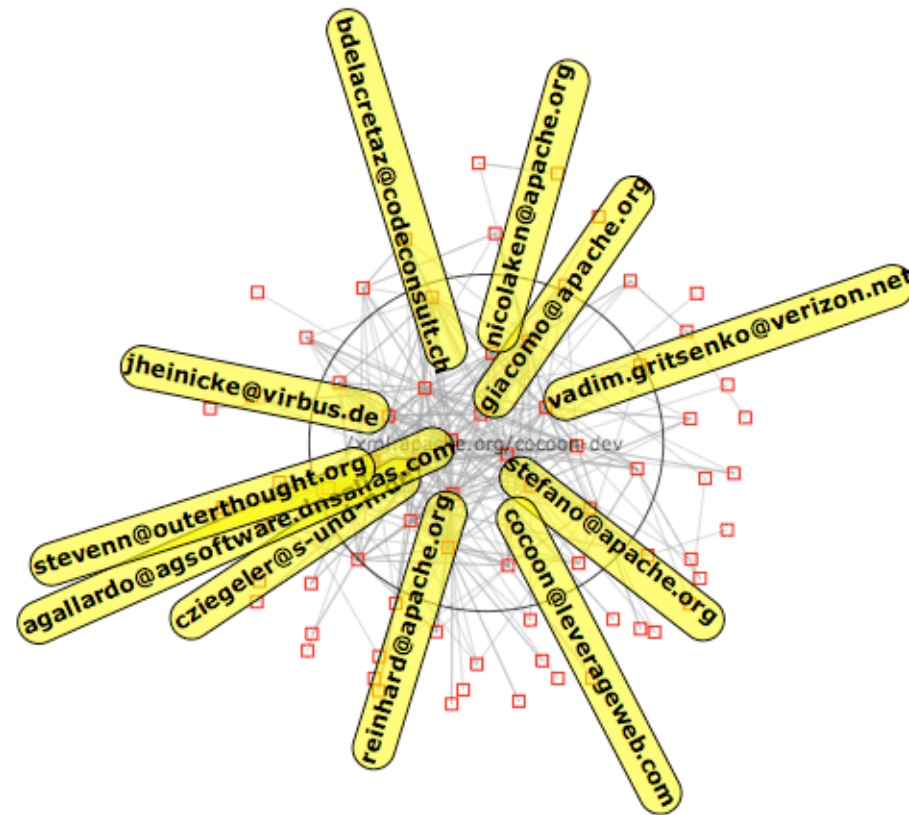
# Dragging a node



# Fixing nodes



# Zooming clusters



**Demo**

# Technical Issues

- Algorithmical complexity grows with the square of the number of nodes (with more than 1000 nodes interactivity is impractical)
- Java is not the faster language around (should be rewritten in a faster language, but loses portability)

# Social Issues

- So intuitive that people could think that Agora paints the absolute truth about a VC
- Could be abused for analyzing private information (privacy)

# About privacy

- the mail analysis script that comes with Agora does **NOT** use the content of the email messages, only the headers!

# Potential other uses

- Visualization of a web of trust by analyzing public key signing
- Visualization of a social network thru the analysis of FOAF data
- ??? (agora can visualize any graph you can come up with!)

# Where is it?

<http://www.apache.org/~stefano/>

and follow the links  
(since location might change in the future)

**Thanks!**

Q&A