Temporal Visualization of Dynamic Collaboration Graphs of OSS Software Forks

Amir Azarbakht, Carlos Jensen

azarbaam@eecs.oregonstate.edu

cjensen@eecs.oregonstate.edu

Categories and Subject Descriptors

H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Keywords

Free/Libre Open Source Software Development, FLOSS, Social Dynamics, Temporal Analysis, FOSS, Forking, Socio-technical Interaction Networks.

Introduction

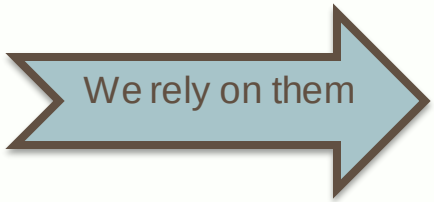
Oregon State University

Introduction

- Popular
  - End-users:

    Firefox, Wordpress, GNU/Linux OS
  - Infrastructure back-bone:

    Apache HTTP Server, FreeBSD

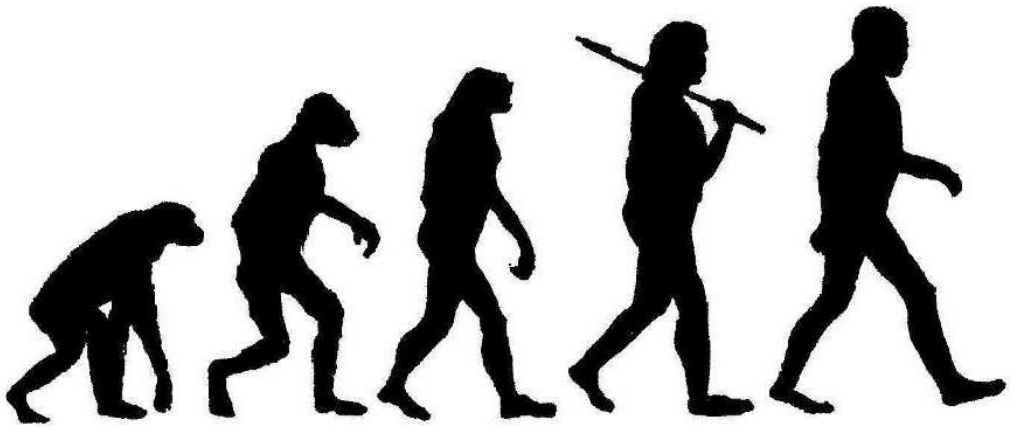


Should make sure these projects are sustained

Oregon State University

Related work

- Socio-Technical System for FLOSSD
  - Community and system co-evolve [Ye and Kishida 2003]
  - Evolution of the socio-technical network (out-degree centrality) [Scacchi 2007]
  - Social structure and dynamics of team communications [Crowston and Howison 2006]



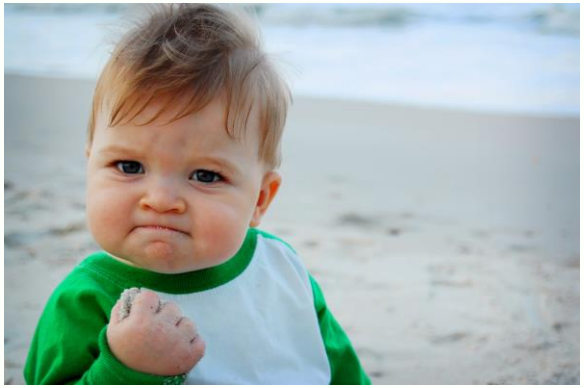
Related work

- FLOSSD informalisms [Scacchi 2007]
  - A forum to do collective activities
    - Establish social control & management [De Souza et al. 2005]
  - Why important? Sources of data to mine

Oregon State UNIVERSITY

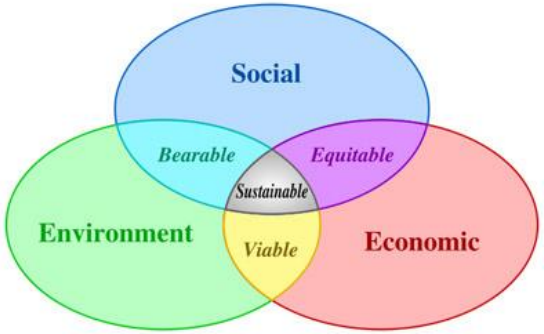
Proposed Future Work

[Scacchi 2007]

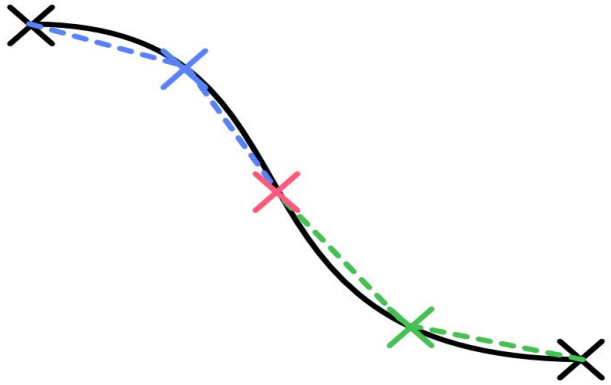


- Success
  - How to measure?
  - What attributes => success

- Sustainability
  - How long can be sustained

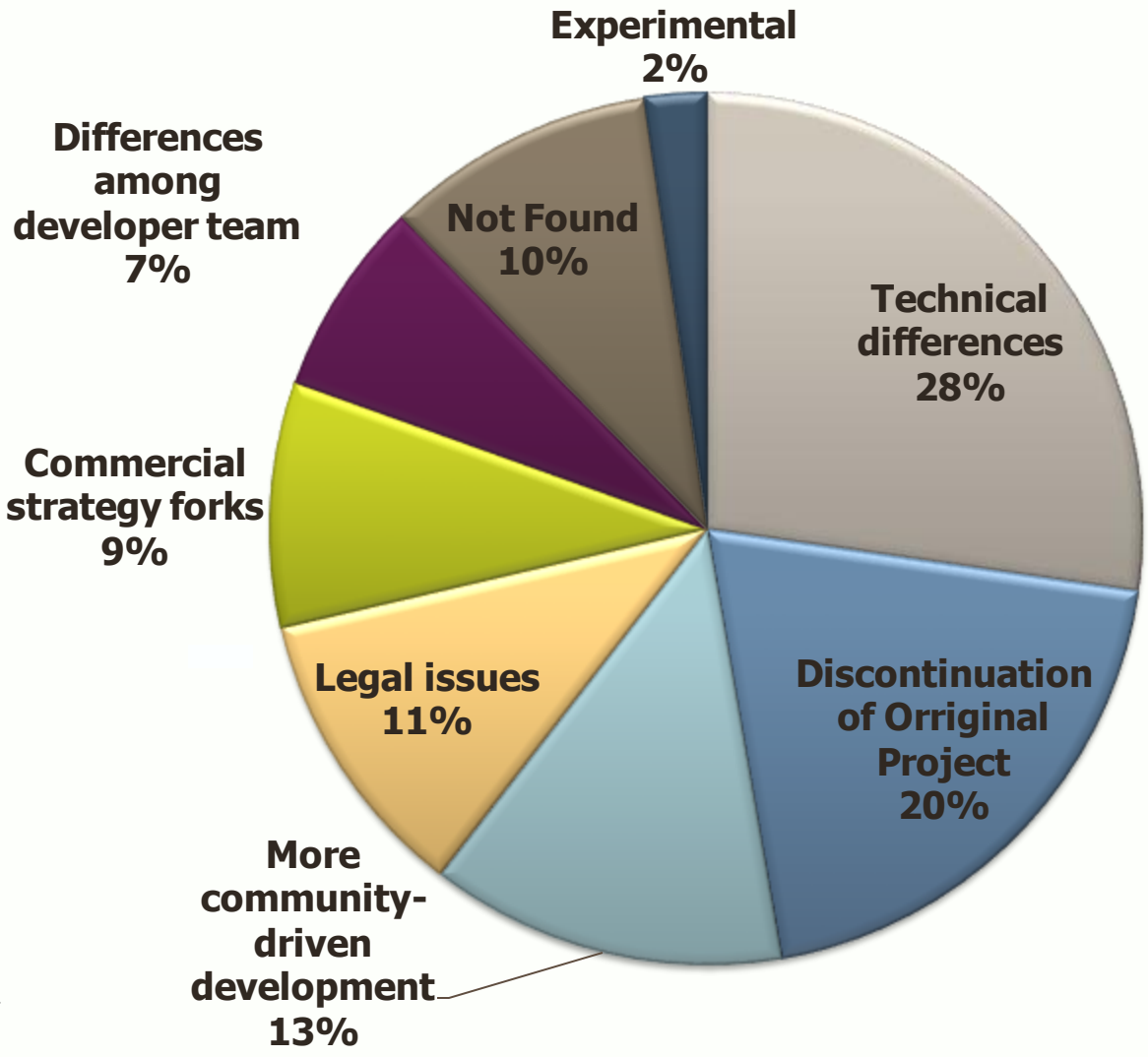


- Inflection points
  - What events/conditions => inflection



Forking and socio-technical interaction network

Main Reasons for forking



Oregon State UNIVERSITY

Methodology

Table 1: The main reasons for forking as classified by Robles and Gonzalez-Barahona [19]

Reason for forking	**Example forks**
Technical (Addition of functionality)	Amarok & Clementine Player
More community-driven development	Asterisk & Callweaver
Differences among developer team	Kamailio & OpenSIPS
Discontinuation of the original project	Apache web server
Commercial strategy forks	LibreOffice & OpenOffice.org
Legal issues	X.Org & XFree

Methodology

Table 2: The project forks for which collaboration data was collected

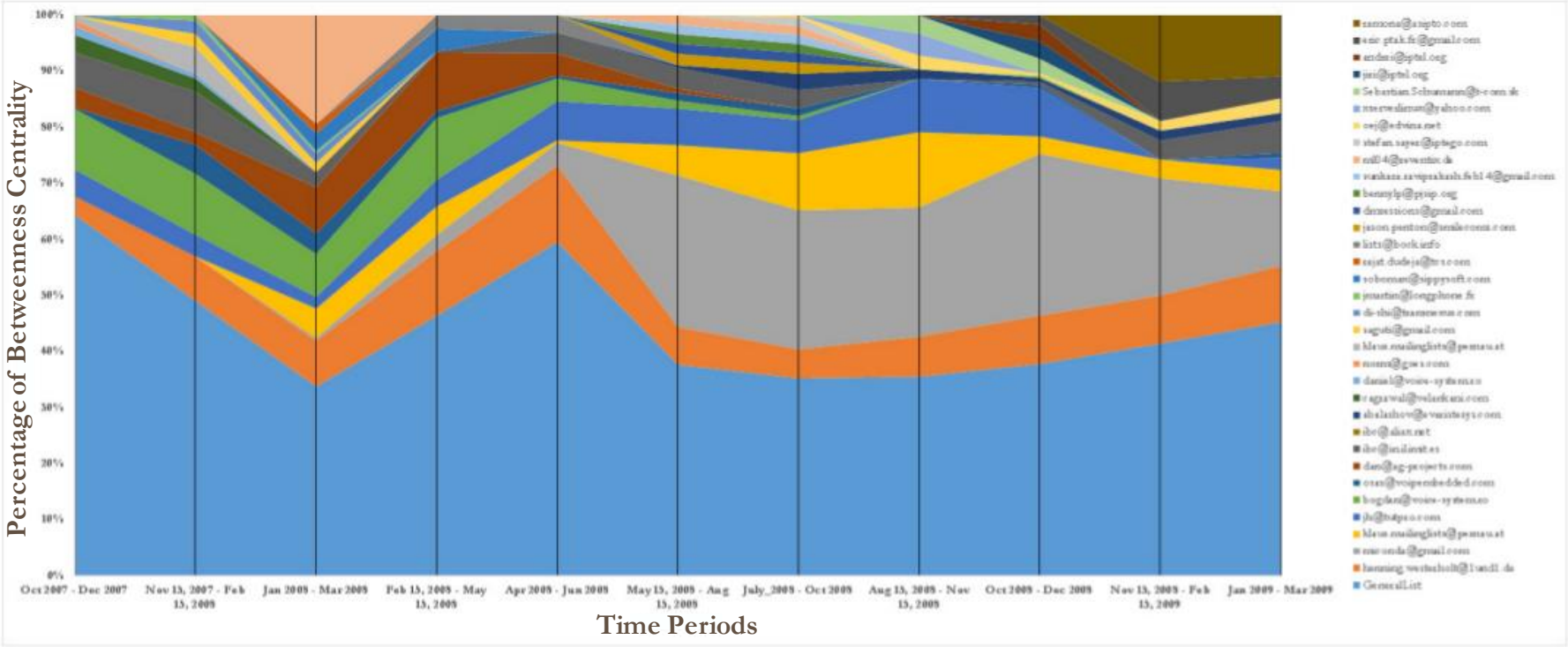
Projects	Reason for forking	Year
Amarok & Clementine Player	Technical (Addition of functionality)	2010
Asterisk & Callweaver	More community-driven development	2007
Kamailio & OpenSIPS	Differences among developer team	2008

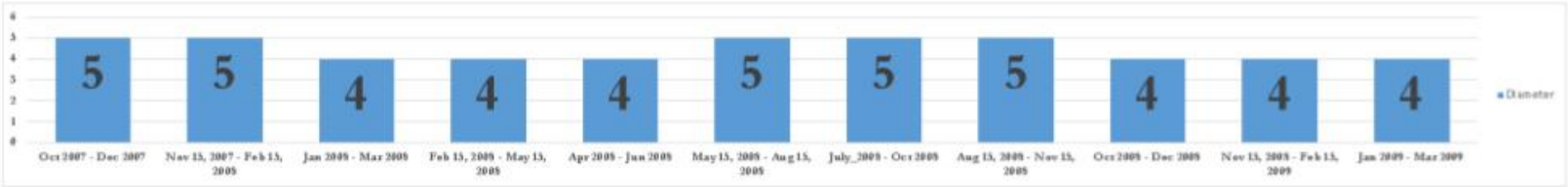
- What we expect to see?
- Can we predict? vs. post-mortem

Methodology

- Data collection
- Communication graph
- Measuring network properties
- Graph and metric visualization

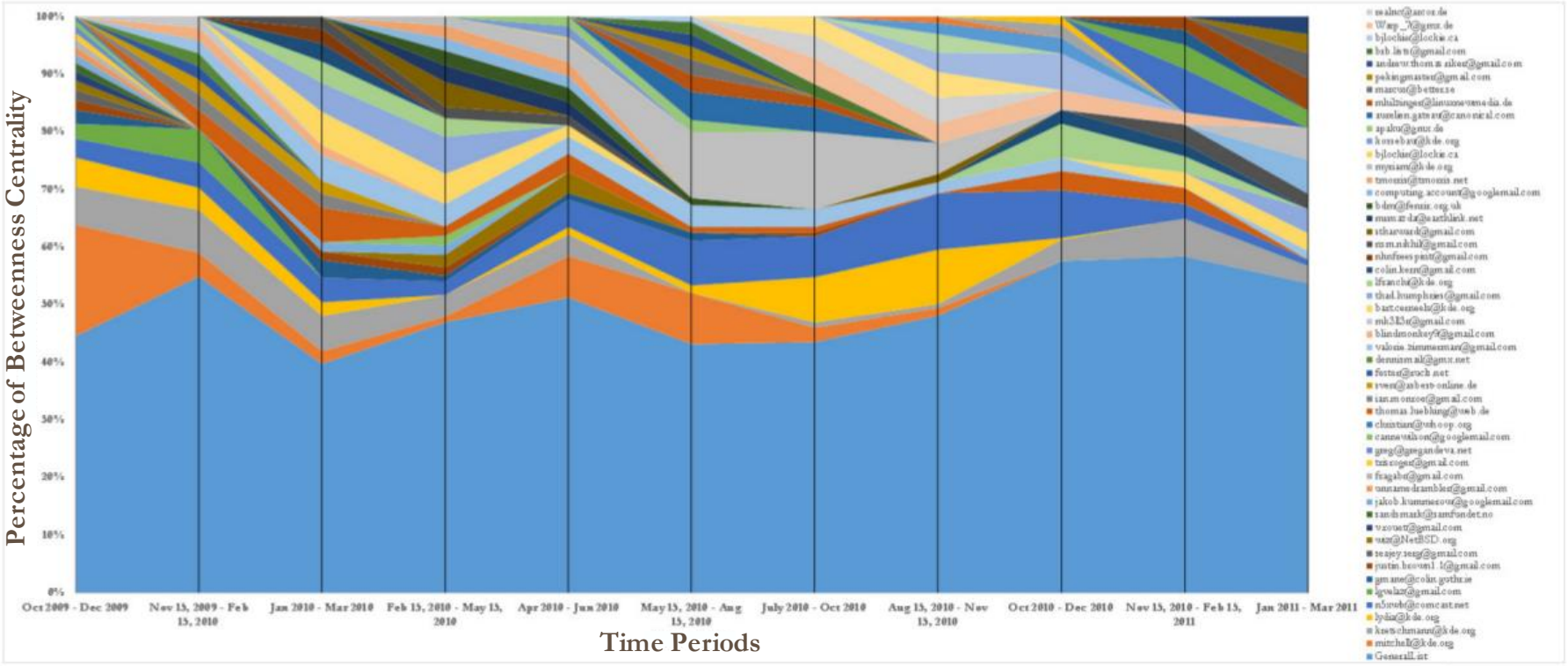
Temporal visualization: Kamailio

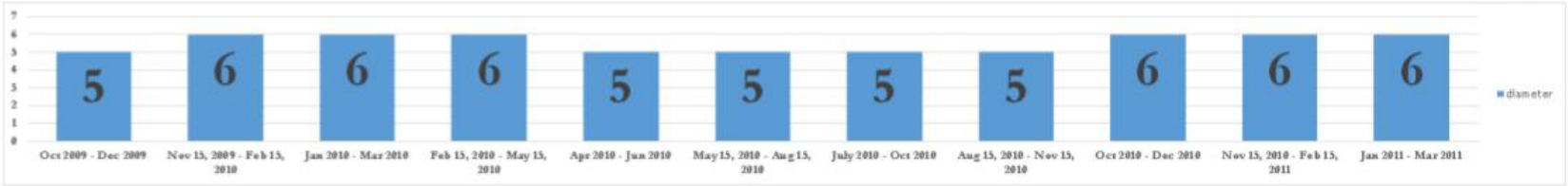




Oregon State UNIVERSITY

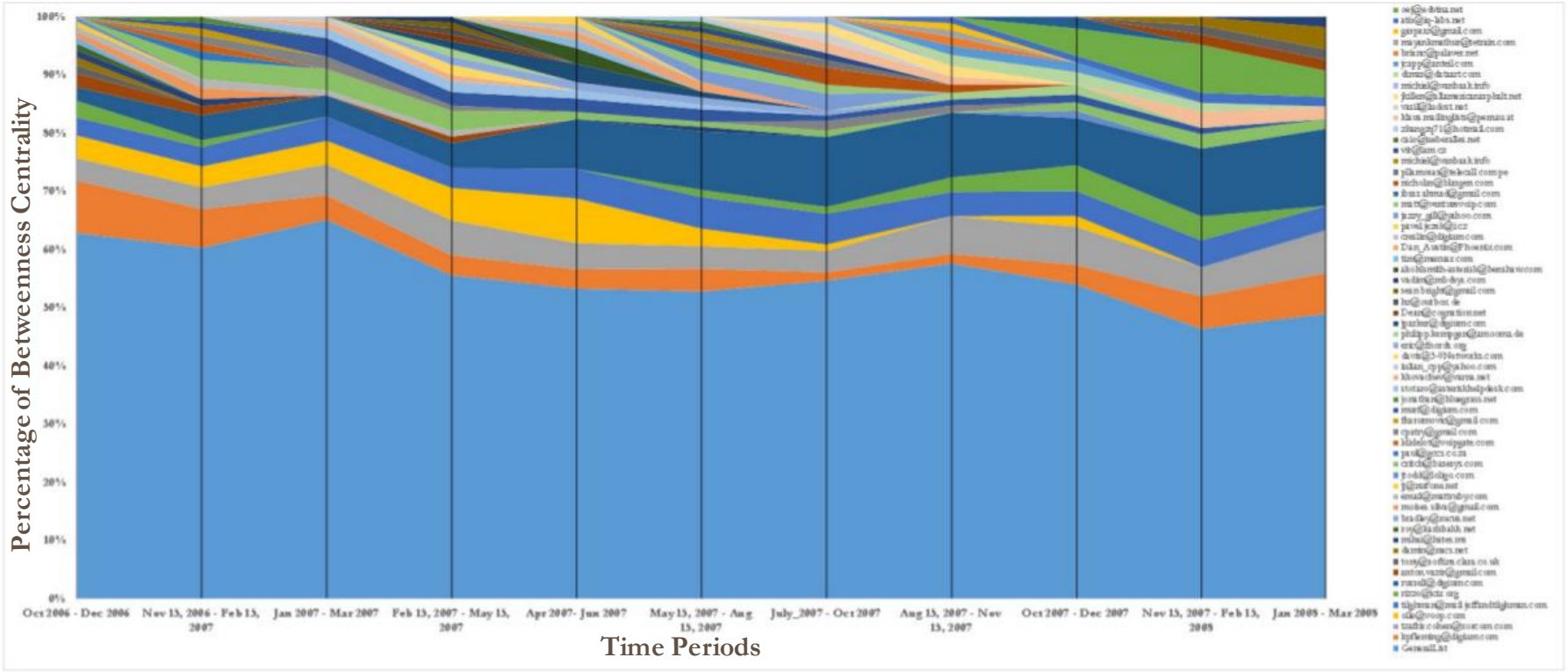
Temporal visualization: Amarok

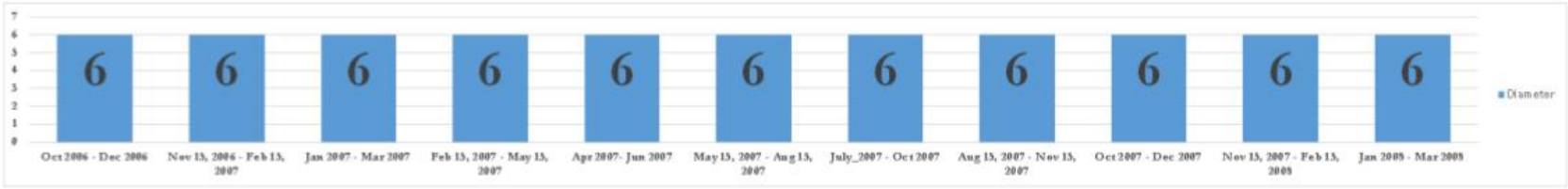




Oregon State University

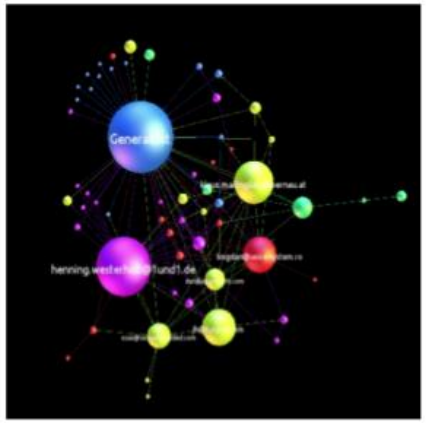
Temporal visualization: Asterisk



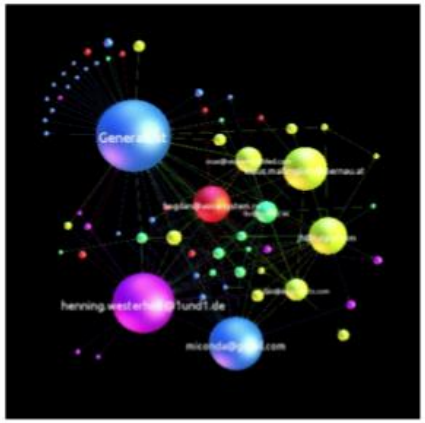


Oregon State UNIVERSITY

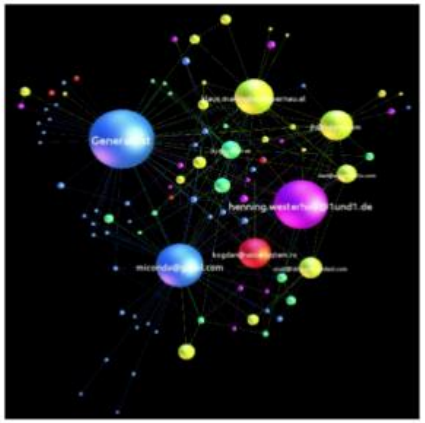
Temporal visualization



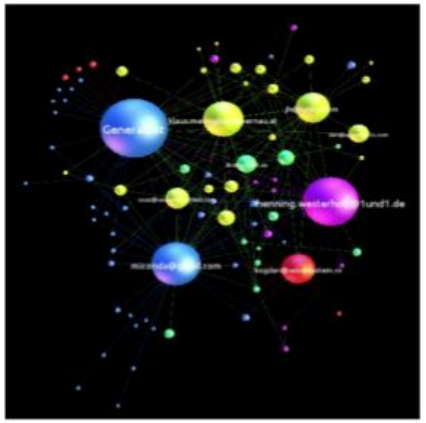
(a)



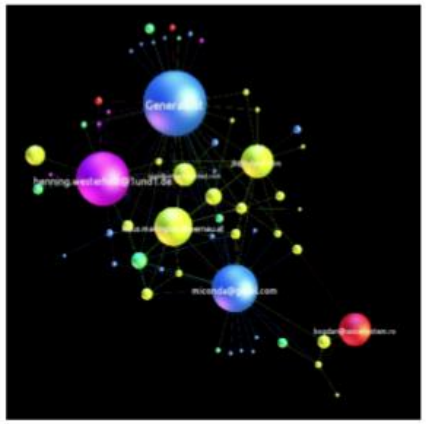
(b)



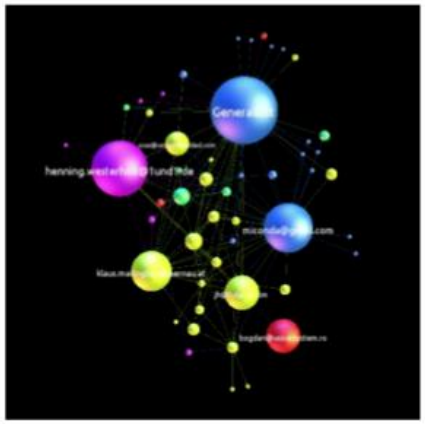
(c)



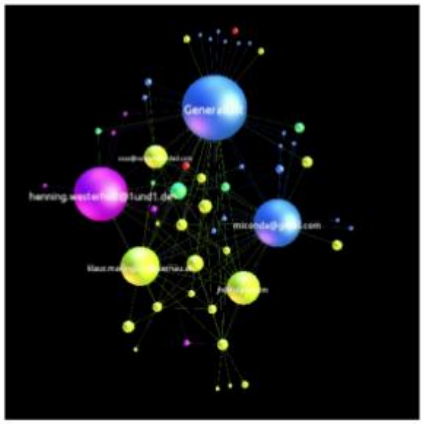
(d)



(e)



(f)



(g)

Video: Kamailio

Conclusion

- What we did:
  - Visualized the temporal change in the open source project community
- What we can do:
  - See the structure of the project member relationships
  - Predict a fork pattern
  - Identify significant members who leave
  - Identify when change happened
- Work in progress:
  - Early prediction
  - Which network metrics' are early indicators
  - Identify best practices to amend/deal with change

THANK YOU

Questions? Comments? Concerns? Ideas?

Oregon State UNIVERSITY