



Trust and Reputation Mining in Professional Virtual Communities

9th ICWE, 24-26 June 2009, San Sebastian, Spain

Florian Skopik

Co-Authors: Hong-Linh Truong, Schahram Dustdar

Distributed Systems Group

Vienna University of Technology, Austria

skopik@infosys.tuwien.ac.at

http://www.infosys.tuwien.ac.at/prototyp/VieTE/VieTE_index.html

- Motivation
- Challenges
- Definition of Trust
- The Autonomic Cycle for Trust Inference
- Trust based on Discussions in PVCs
- Evaluation (Slashdot)
 - Global Reputation
 - Context-Aware Ranking
 - Individual Trust
- Conclusion

Motivation

- (Professional) Virtual Communities in SOA
 - humans **and** (Web) services
 - **joining/leaving** the environment **dynamically**
 - performing **activities** and **tasks**
- Massive Collaboration in SOA
 - large sets of **humans** and **services**
 - dynamic **compositions**
 - distributed communication and coordination
- Influences on collaboration partner selection?
 - → **TRUST**

Challenges

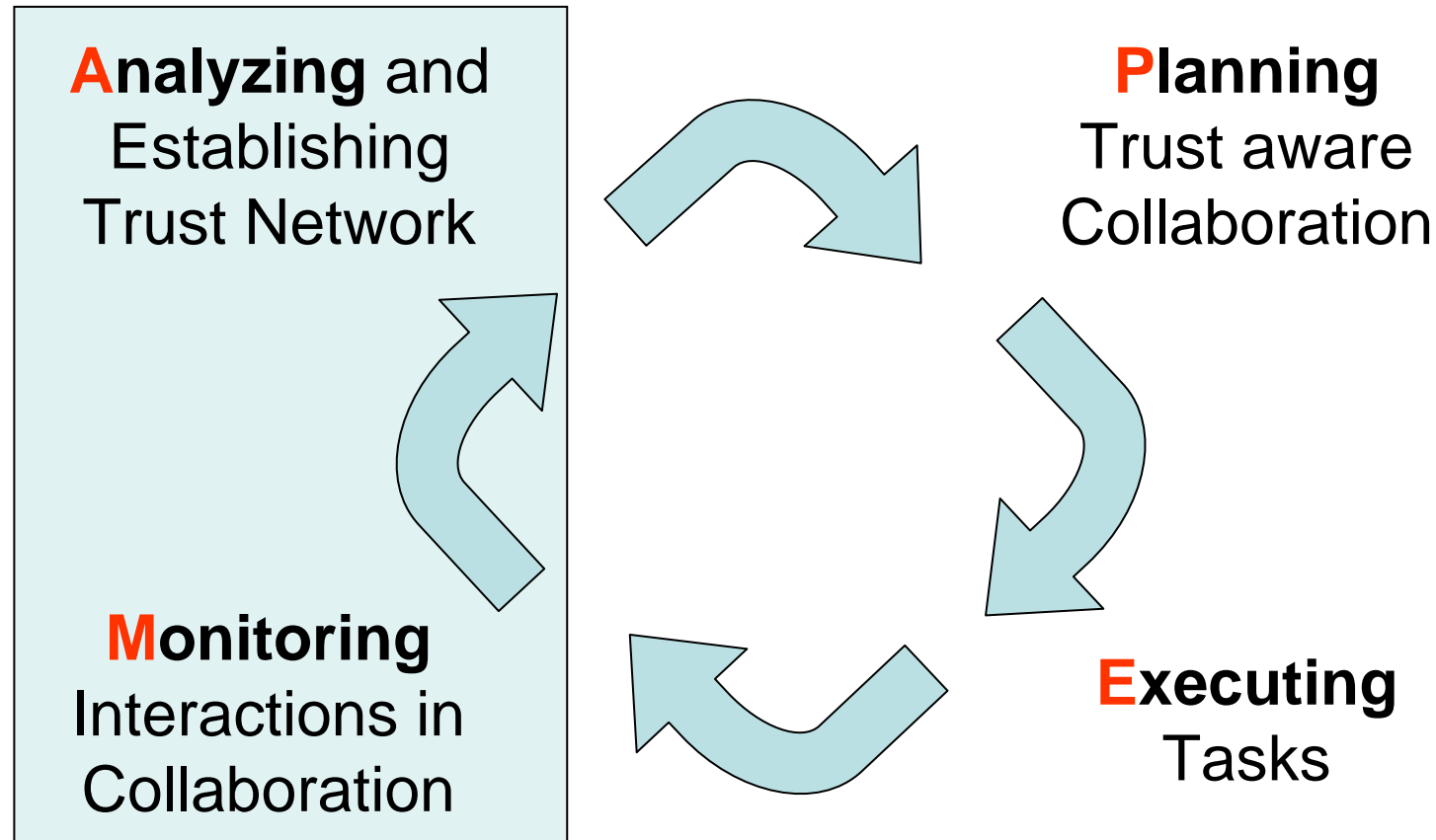
- Identify fundamental **concepts of trust in SOA**
- Enable **automatic determination of trust** due to the potential large sets of entities in SOA
- Identify **data influencing trust**
- **Approach** enabling collecting, managing and analyzing data for trust determination

One uniform platform to manage **trust** in **humans and in services** in distributed service-oriented environments.

- Trust reflects an **expectation**
 - based on **previous interactions**
 - one entity has about another's future behavior
 - to **perform activities** dependably, securely, and reliably
 - within a **specified context**.

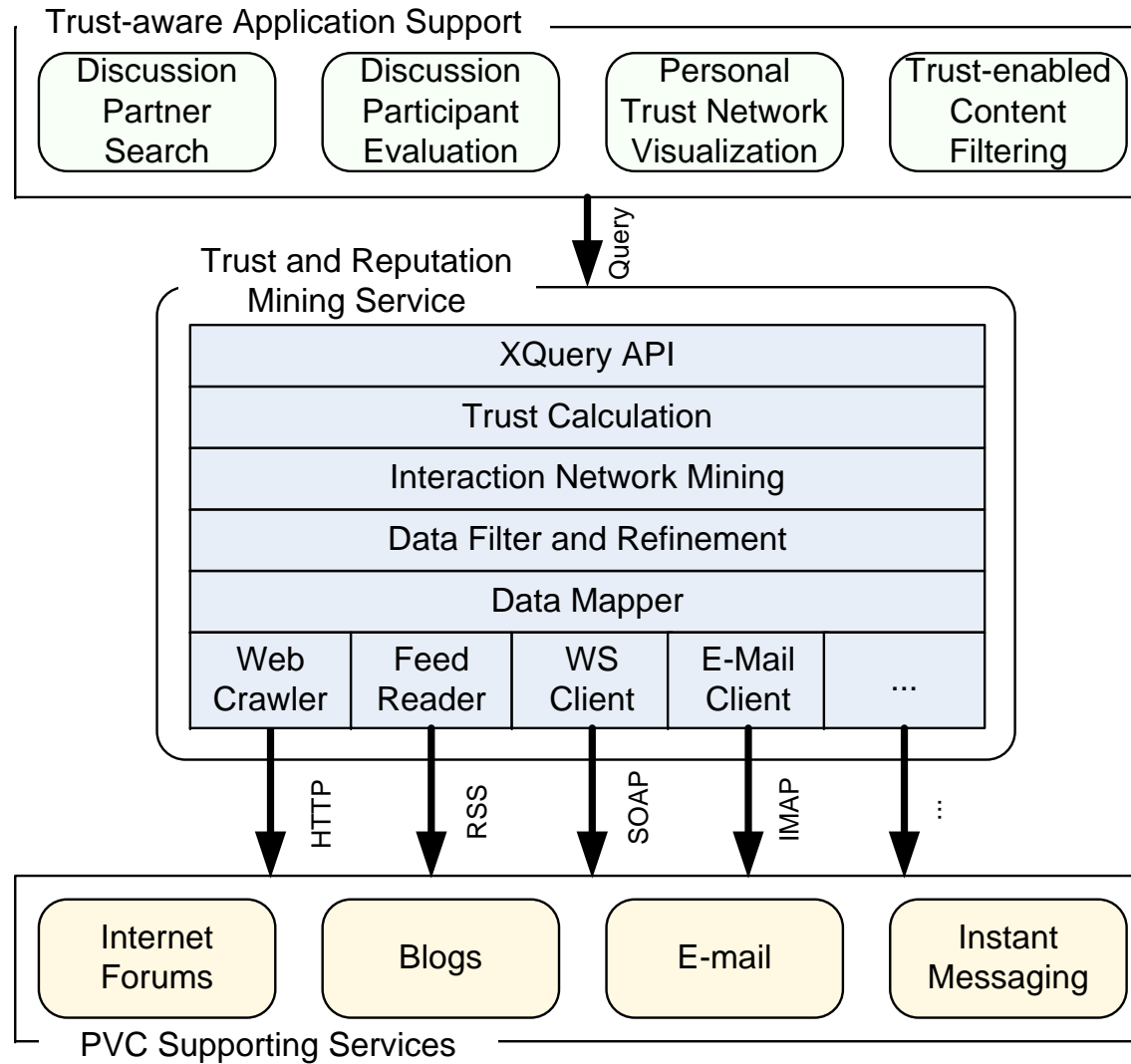
[WEBIST] F. Skopik, H.-L. Truong, S. Dustdar. *VieTE – Enabling Trust Emergence in Service-oriented Collaborative Environments*. 5th International Conference on Web Information Systems and Technologies (WEBIST). Lisbon, Portugal, 2009. Springer.

The Autonomic Cycle for Trust Inference

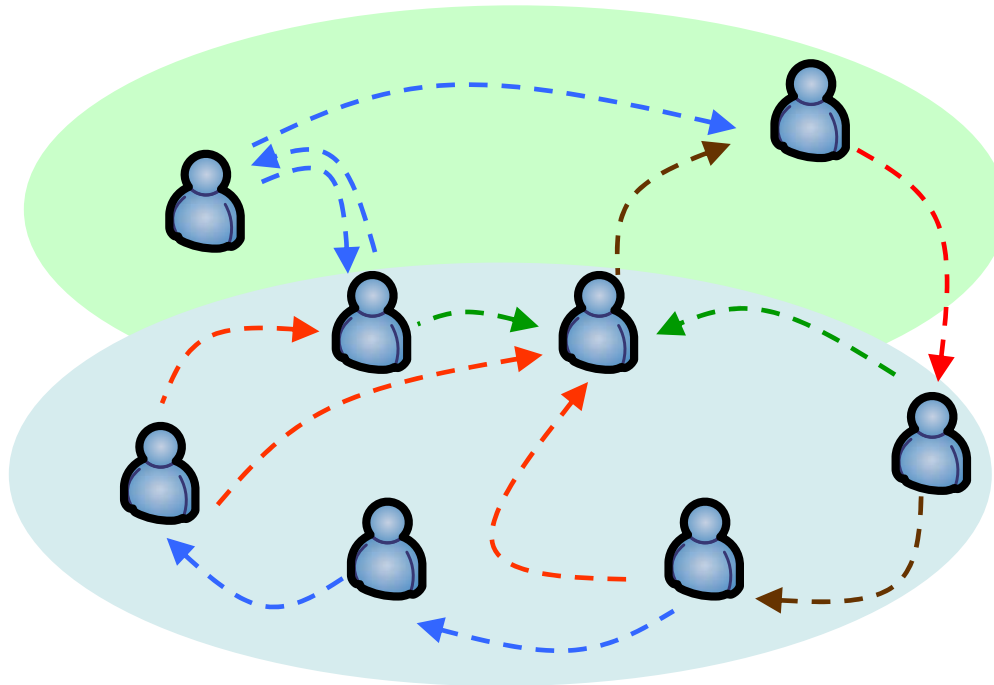


[SEAA] F. Skopik, D. Schall, S. Dustdar. *The Cycle of Trust in Mixed Service-oriented Systems*. 35th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). Patras, Greece, 2009. IEEE.

Focus: Monitoring Discussions



Interaction Graph Model



Interaction Types:

Internet Forums

Instant Messages

E-Mails

Dedicated services
(Web services)

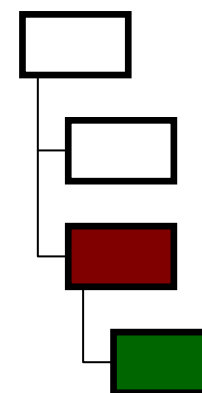
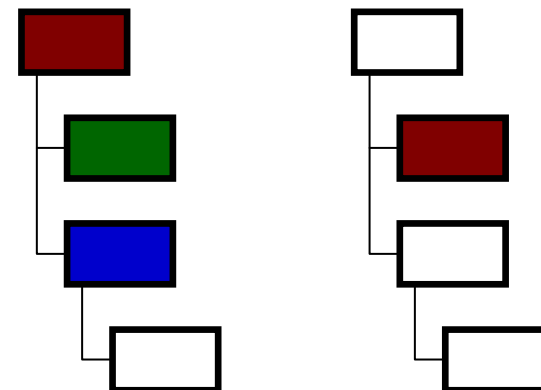
users V , relations E , $G_I=(V,E)$

$e \in E$, $e=(u,v,Ctx,metrics[name,value])$

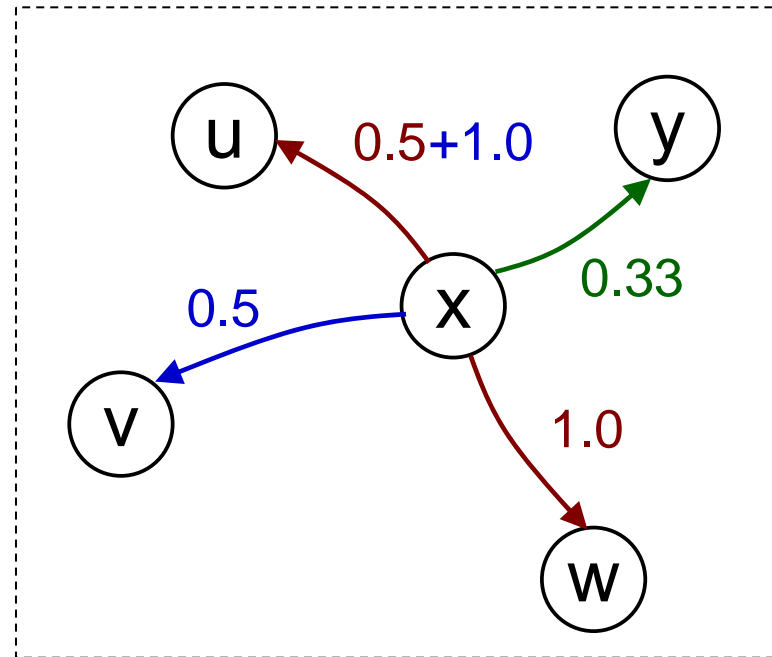
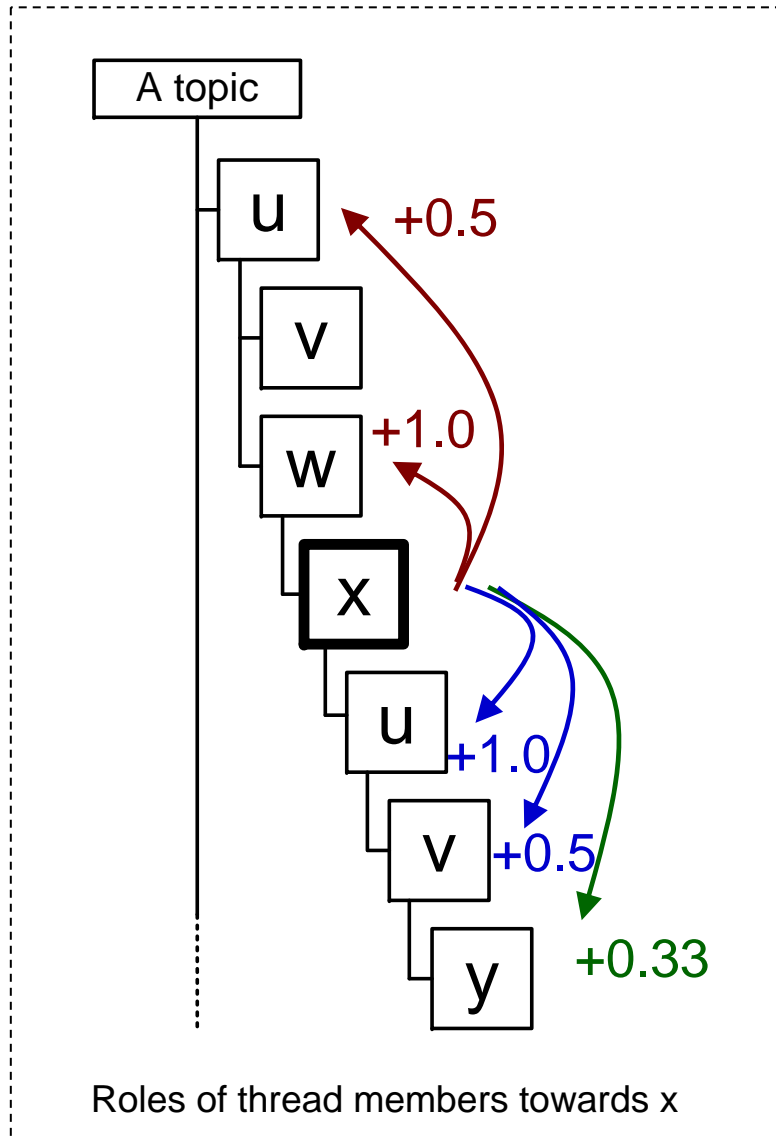
metrics: Rewards, Responsiveness, Reliability...

Roles in Threaded Discussions

- **Activator**
 - ‘Kicks off’ a discussion.
 - Attracts others to reply comments
 - Therefore: provides valuable contribution
- **Driver**
 - Replies to comments
 - Keep discussions alive
- **Affirmed Driver**
 - A driver’s comment is replied, therefore, its contribution is affirmed.



Role Mining Example: Reward user x



Rewarding contributions in discussion thread (left) and map to graph model (right).

Trust Network Model (1)

- **Confidence** of user u in user v to be a good discussion partner
 - Based on collected rewards (successful posts)
 - $c^{Ctx}(u, v) = f_c(G_p, u, v, Ctx)$
- **Reliability** of $c^{Ctx}(u, v)$ ($\in [0, 1]$)
 - Stable trends in discussion contributions
 - Based on number of interactions (or number of posts)
 - Based on variance of recently collected rewards
 - $r(c^{Ctx}(u, v)) = f_r(G_p, u, v, Ctx)$

Trust Network Model (2)

- Personal Trust

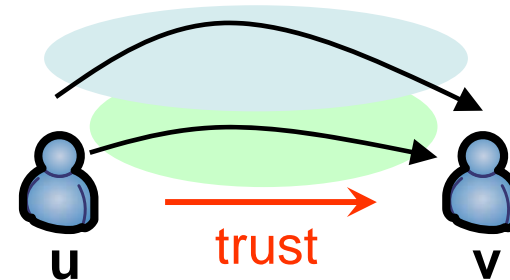
- $T^{Ctx}(u,v) = c^{Ctx}(u,v) \cdot r(c^{Ctx}(u,v))$



- Trust Aggregation

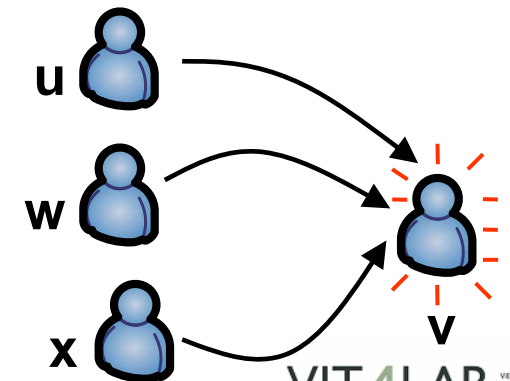
- General Trust

- Combine trust relations from u to v in different contexts to infer trust in a broader scope.



- Reputation

- Combine trust relations from $\{u,w,x,\dots\}$ in v in the same context.
 - Reflects somebody's standing within the community



- Data Characteristics*
 - Distinguishable users (anonymous posts skipped)
 - Performing threaded discussions
 - Posts rated by humans (unrated posts skipped)

Re: (Score:5, Funny) by [liam193 \(571414\)*](#) I believe your right. It appears that the first was, in fact, a shorten version of th

└ **Your attempt at humor** (Score:-1, Offtopic) by Anonymous Coward Lamell!Sincerely,

Re: (Score:4, Funny) by [chefmonkey \(140671\)](#) I believe your right.You believe his right to do what?

└ **Re:** (Score:1) by [DarthJohn \(1160097\)](#) He's not so sure about his left.

Re: (Score:2, Funny) by [elqatozorbas \(783538\)](#) Mom, there making fun of me... again!

Re: (Score:3, Funny) by [chefmonkey \(140671\)](#) I believe your right.You believe his right to do what?Mo

└ **Re:** (Score:1) by [elqatozorbas \(783538\)](#) See! You wouldn't believe me, but its through.

* Real data set: <http://slashdot.org>

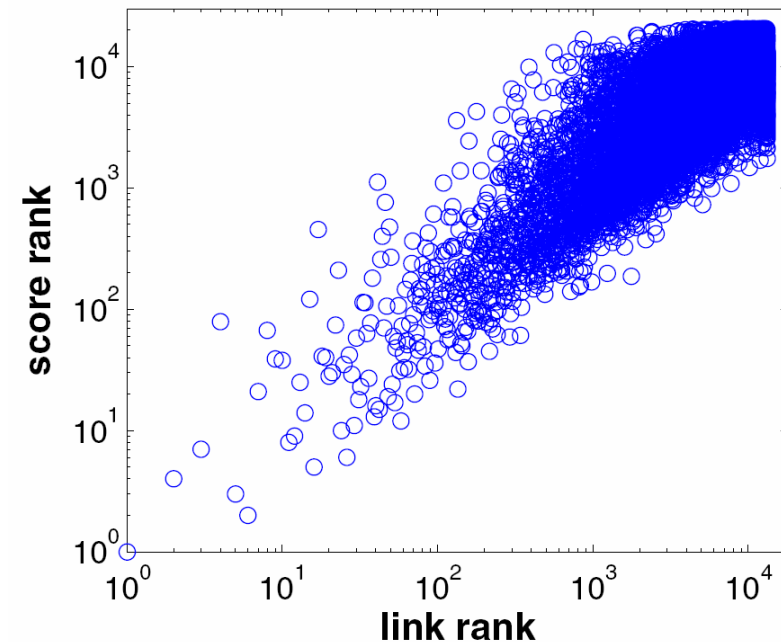
Evaluation Approach (2) - Ranking

- *Score rank (Slashdot):*
 - Average of a user's posting score (human feedback)
 - Interpretation
 - users tend to trust posters with high average posting scores more to deliver valuable contributions, than lower rated ones.

- *Link rank (Our Method):*
 - Mining of roles in discussions (automatically).
 - Calculating reputation (as defined before)
 - Interpretation
 - directly reflecting a user's willingness to share information and support others (driver role), and
 - attitude to highly recognized contributions (activator role).

Evaluation: Global Reputation

- Approach
 - Calculate score rank and link rank
 - Compare ranking results
- Findings
 - Strong correlation between both results
 - Similar sets of top ranked users
 - determined by humans who grade postings
 - determined by our algorithm based on structural analysis

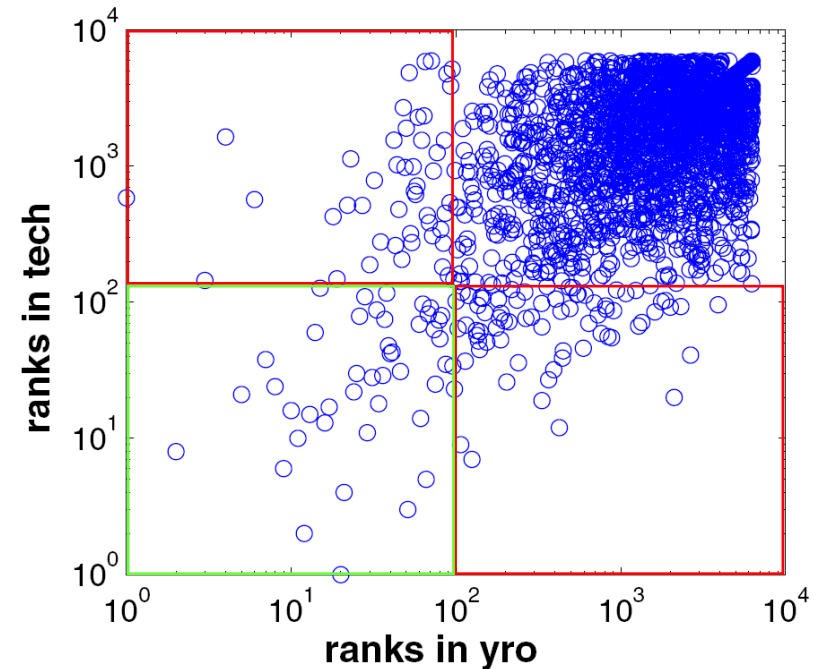


50k users, 670k posts

Pearson correlation coefficient = **0.77**

Evaluation: Context-Aware Ranking

- Approach
 - Rank users in the subdomains 'tech' a. 'yro' (your rights online)
 - Compare link ranks in both scopes (expressing expertise)
- Findings in Top100
 - 40% of users are top ranked in both fields (green box)
 - However, 60 percent are top ranked in one field (rank<100), but not in the other (rank>100) (red boxes)
 - Application: capturing the context of relations is beneficial for user profile mining (“trust in a user’s capabilities to be a good discussion partner”)



2 contexts: yro, tech

5939 users in yro

6288 users in tech

14793 distinct users

Conclusion

- Fundamental concepts of trust and reputation
 - **Automatic** inference of **trust** in large-scale SOA-based environments upon interactions
 - **Cyclic** approach
 - trust emerges from interactions
 - interactions are initiated based on trust
 - **Context** awareness
 - Personal **trust** v.s. global **reputation**

- Future Work
 - Extend the scope from discussions to more general interactions
 - Extend the role of time in the approach (long-term reputation)
 - for increasing attack resistance of the trust model
 - Evaluate the system in a real PVC use case (www.coin-ip.eu)

Thanks.

http://www.infosys.tuwien.ac.at/prototyp/VieTE/VieTE_index.html
skopik@infosys.tuwien.ac.at