Mining and Composition of Emergent Collectives in Mixed Service-Oriented Systems

IEEE Conference on Commerce and Enterprise Computing (CEC)
10-12 November 2010, Shanghai, China

Daniel Schall and Florian Skopik

Distributed Systems Group
Vienna University of Technology, Austria
schall@infosys.tuwien.ac.at
skopik@infosys.tuwien.ac.at
Environment and Motivation

- Open and dynamic environment
  - humans and resources (e.g., services)
  - joining/leaving the environment dynamically
  - humans perform activities and tasks
- Massive collaboration in SOA/Web 2.0
  - large number of humans and resources
  - dynamic compositions
  - distributed communication and coordination
- Keep track of the dynamics to control
  - future interactions
  - resource selection
  - compositions of actors
  - activity and task assignments

[computational social network model]
Application Scenario: Expert Web

- Process model is based on tasks and flow structure
- Embedding of Web 2.0 collaboration tools
- Connected experts provide online help and support

Expert Discovery in Crowdsourcing

- How do actor **discovery and selection** mechanisms work?
- What is the technical grounding for the proposed system?
- How can actors **be flexibly involved** in a service-oriented manner?
- How do **interactions and behavior** influence future discovery?

![Diagram]

**Symbols:**
- expert
- expertise area
- network relation

**Expert Crowd**
Hubs and Authorities

- **On the Web**
  - Hubs: *pointing to* authoritative pages
  - Authorities: *are being referenced by* other important pages
  → Recursive definition

- **In social/collaborative networks**
  - Hubs: information “brokers” distributing work
  - Authorities: experts processing received work

Based on HITS algorithm (Kleinberg, *J. ACM* 1999)
Human-Provided Services

- Mixed System
  - Mix of human- and software services collaboration
  - Humans provide services using SOA concepts

- Human-Provided Services (HPS)
  - User contributions as services
    - Service description with WSDL
    - Communication via SOAP messages
  - Example: Document Review Service
    - Input: document, deadline
    - Output: review comments

Monitoring and Logging

Provisioning and Configuration IF

Interaction Metrics Calculation

Activity/Task Management

Distributed SOAP Interaction Monitoring

Metric Definitions

- Define metrics
  - emergency support: fast and reliable responses
  - neglect others, e.g., costs
- Calculate metrics in the scope of interactions (here: requests for support (RFSs))
  - average response time
    \[
    t_r^s = \frac{\sum_{rfs \in RFS} (t_{receive}(rfs) - t_{send}(rfs))}{|RFS|}
    \]
  - activity success rate
    \[
    sr^s = \frac{num(sRFS)}{num(sRFS) + num(fRFS)}
    \]
Delegation Behavior (1/2)

Context Tags:
- Applied to interaction links

Metrics:
- Responsiveness
- Success rate

... nodes
... social network relations (FOAF knows)
... delegation
... reply and rewarding
Delegation Behavior (2/2)

- ... nodes
- ... social network relations (FOAF *knows*)
- ... delegation
- ... reply and rewarding

Delegations and ratings relevant for a particular context
ExpertHITS Query $Q^A$: Required skills e.g., 
software engineering, compiler techniques

ExpertHITS Query $Q^B$: Required skills e.g.,
algorithms, social network mining
ExpertHITS (2/2)

Reputation

- Calculate reputation in scope of interaction contexts (expertise)
- Hub score
  - Rating through authorities based on delegation behavior:
    \[ H(u; Q) = \sum_{v: u \to v} w^Q_{vu} A(v; Q) \]
- Authority score:
  - Rating through hubs based on reliability in processing delegated tasks
    \[ A(v; Q) = \sum_{u: u \to v} w^Q_{uv} H(u; Q) \]
Generate artificial interaction data imitating real collaboration environment \textit{(preferential attachment)}

Small and medium scale networks (100 to 1000 nodes)

Concurrent processing time:

\begin{center}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
Requests & 50 & 100 & 200 & 300 & 400 & 500 \\
\hline
Small-scale & 9 & 16 & 20 & 21 & 24 & 25 \\
\hline
Medium-scale & 199 & 325 & 390 & 986 & 1432 & 1663 \\
\hline
\end{tabular}
\end{center}

For small-scale networks, on average 19 seconds can be expected under different load conditions (50-500 concurrent requests)
ExpertHITS Results (2/2)

- Measuring quality and impact of rankings *(see paper)*
  - Standard HITS algorithms versus ExpertHITS
- Ranking metrics
  - Relative ranking change
  - Quality

→ ExpertHITS promoting well-connected and rated hubs
→ Approach guarantees the discovery of reliable “entry points” to the Expert Web
Online Help and Support (1/2)
Formulate Query (Constraints)

One or more of specified skills:

Software/SE/Specifications/Languages OR Software/SE/Specifications/Analysis OR

Context parameters (optional):

- Expert is online (via Skype)
- HPS interaction (via Web services)
- Apply metric: Availability
  set value for minimum threshold (a number between 1..100)

ExpertHITS
Online Help and Support (2/2)

Discovery and Interactions

Trusted Online Help and Support

Trusted Online Help and Support offered by:

- Daniel Schall
  - Skills:
    - Software/SE/Specifications/Analysis
    - Software/SE/Specifications/Languages
  - [FOAF] Knows 20

- Florian Skopik
  - Skills:
    - Software/SE/Specifications/Languages
  - [FOAF] Knows 20

Click user to contact.
Thanks!

Daniel Schall and Florian Skopik

Distributed Systems Group
Vienna University of Technology, Austria

schall@infosys.tuwien.ac.at
skopik@infosys.tuwien.ac.at