Mediation in the Context of a Peer-to-Peer Network

Tilo Luchs
LSI-Hannover

Dr. Susanne Bauere, Alexander Landen
Computing Institute Information Systems C9,
Technische Universität Berlin
19.02.03 Hannover

Examples
Scenario E-Learning

Observations
E-Learning: Paradigm shift

Instruction-based

Observations
E-Learning: Paradigm shift

Instruction-based

Self-directed, informal learning

Observations
Critical mass of information

- Existing information sources contain a critical mass of information
  - 60000+ based on 1250
  - new data bases
  - ...
- Lack of clarity and transparency
- High degree of distribution/heterogeneity

Observations
Unstructured distributed information sources
**Observations**

Introducing Structure

- Mediator-based Information Systems MBIS, see [Lesser 00, Neumann 01, Wiederhold 92, Garcia Molina 95, Busse 02]
- Super Peer Networks [Edutella - Neidl Wolpers Silbersk 02, Garcia Molina 02, P-Grid Alberer]

**Structure Elements**

Examples: Super Peer-Backbone

**Structure Elements**

Examples: Mediator based IS in E-Learning context

**Problem and Objective**

General Problem

- Problem
  - Software engineering of MBIS is complex
  - Development is costly

- Our assumption:
  - Increasing demand of MBIS
    - Low-cost but highly profitable
    - Easy to integrate
  - Demand for new evolutionary concepts for
    - Simply modeling and development phase
    - Shortening development time and investments for a stable mediator

**Problem and Objective**

General Problem

- Problem
  - Software engineering of MBIS is complex
  - Development is costly

- Our assumption:
  - Demand for low-cost but „mutable”, dynamic and adaptive mediators
  - Simply modeling and development phase
  - Shortening development time and investments for a stable mediator
Adaptive mediator

Focus in this talk

Mediation Correspondences
Super Peer Service Registry
Clustering Rules

Problem and Objectives

Specific Problem
- Selection of relevant, high quality information sources is an essential step in information management of MiBS
- Selection is a continuous process
- Problem: relevant sources may be "overseen"
- Demand for late binding of E. at MiBS
- Examples: new e-Learning provider, external mediator for used e-Learning materials

Objectives
- Investigation of concept technologies
  - Clustering rules for information sources
  - Supporting mediator's administrator

My Objective
- Investigation of concept technologies
  - Clustering rules for information sources
  - Supporting mediator's administrator
- Rules and regulations: support the administrator in identifying, evaluating, and classifying information sources
- Facilitate (semi-)automatic adaptation of the mediator in changing/creating existing information sources in the network
**Approach: subject-specific clustering rules**

**Example: E-Learning Infrastructure**

```
ON (Event) Enter
  IF
    Peer.Advertisement.Property query_scheme INCLUDE "DublinCore"
  }
  DO (Action) Approve(Peer)
```

---

**Approach: subject-specific clustering rules**

**Rule-Body**

```
ON (Event) Enter
  IF
    Peer.Advertisement.Property query_scheme INCLUDE "DublinCore"
  }
  DO (Action) Approve(Peer)
```

---

**Approach: subject-specific clustering rules**

**Rule-Body**

```
ON (Event) Enter
  IF
    Peer.Advertisement.Property query_scheme INCLUDE "DublinCore"
  }
  DO (Action) Approve(Peer)
```

---

**Approach: subject-specific clustering rules**

**Rule-Body**

```
ON (Event) Enter
  IF
    Peer.Advertisement.Property query_scheme INCLUDE "DublinCore"
  }
  DO (Action) Approve(Peer)
```

---

**Approach: subject-specific clustering rules**

**Rule-Body**

```
ON (Event) Enter
  IF
    Peer.Advertisement.Property query_scheme INCLUDE "DublinCore"
  }
  DO (Action) Approve(Peer)
```

---

**Technische Universität Berlin**
**Important Precondition**

Metadata describing information sources

- Important features of information sources
  - Query and result scheme
  - IP address
  - Number of average results
  - ...

- Determined implicitly
  - See [Haarmann 02, Strengler & Wang 97, Gade & Dijkhuizen 11]

- Technologies: DAML+OIL, WSDL, RDF, XML, Java ToolKit

**Intended Proof of Concept**

Existing Peer-to-Peer based infrastructure

- EduMedia extends SUNS-DORA with query discovery, replication, mediation and clustering services
- Introduce at the WWW02
- Multi-staged effort to scope, specify, architect and implement an IDL-based metadata P2P infrastructure
- Main goal: achieving interoperability between heterogeneous metadata-driven e-learning systems

**Intended Proof of Concept**

My contribution

- Concepts for a rule-based clustering of information sources with identical or similar features
- Approaches to a technological integration of a rule-based language into distributed information systems (MDB, P2P)

**(Expected) Benefits**

- Simplified modeling of the information need requirements of the mediator
- Reuse of correspondences in MIBs (Jander 96, MOCAS [Busse 02]) and local schema transformations [Allen 97]
- Late binding at mediator
- Higher scalability of distributed ad hoc networks by „mainstream“ structures
- ...

**Adaptive Mediator**

Mediation correspondences

**Problem and Objective**

Specific Problem

- Sources use different standards describing their information
- Sources have specific interfaces (query scheme, query capabilities)
- Sources have overlapping contents
- Queries need combinations of sources
Problem and Objectives

Objectives

- Concept/Technology for semantic integration
- Schema transparency
- Query language transparency
- Data integration
- Source combination
- Facilitate (semi) automatic adaptation of the mediator in changing/implementing existing information sources in the network.

Approach: Correspondence-based mediators

Query Processing

Meta data Integration

Data Description

Mediator Schema

Export Schema

Autonomous Data Sources

Approach: Correspondence-based mediators

Query Processing

Meta data Integration

Data Description

Mediator Schema

Export Schema

Autonomous Data Sources

Approach: Correspondence-based mediators

Query Processing

Meta data Integration

Data Description

Mediator Schema

Export Schema

Autonomous Data Sources

Approach: Correspondence-based mediators

Query Processing

Meta data Integration

Data Description

Mediator Schema

Export Schema

Autonomous Data Sources

Approach: Correspondence-based mediators

Query Processing

Meta data Integration

Data Description

Mediator Schema

Export Schema

Autonomous Data Sources

Approach: Correspondence-based mediators

Correspondence Specification

<table>
<thead>
<tr>
<th>Meas</th>
<th>Wert</th>
<th>Datum</th>
<th>Staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lenten</td>
<td>12345</td>
<td>2023-03-01</td>
<td>9876</td>
</tr>
</tbody>
</table>

Approach: Correspondence-based mediators

Correspondence Spec.: GaV, LaV, VaV

- Global-view (GaV)
  - each mediator schema element is defined as view on integrated data sources
- Local-view (LaV)
  - each source element is defined as a view on the mediator schema
- View-view (VaV, GLaV)
  - view on the mediator schema is related to a view on data sources

<table>
<thead>
<tr>
<th>Understandability</th>
<th>GaV</th>
<th>LaV</th>
<th>VaV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguity</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Approach: Correspondence-based mediators**

**Correspondence Spec.: Query Capabilities**

- **Origin**
  - 
  - 

- **Schema**
  - 
  - 

- **Interface**
  - 
  - 

**Approach: Correspondence-based mediators**

**Correspondence Spec.: Intension vs. Extension**

- **Mediator**
  - 
  - 

- **Source**
  - 
  - 

**Approach: Correspondence-based mediators**

**Mediator Development**

1. Define infrastructure
   - Technical integration
   - Common data model, query language
2. Define mediator schema
   - Unit common correspondences
   - For each source:
     1. Define export schema and query capabilities
     2. Define correspondences or reference metadata standard
     3. Register at the mediator peer

**Intended Proof of Concept**

**Contribution and Benefit**

- Concepts for semantic integration of overlapping information sources
- Correspondence-based mediator for meta data - based integration
  - Understandable modeling
  - Correspondence-based implementation
- Integration of mediator concepts to P2P infrastructures

**Adaptive Solution for Semantic Integration**

**Conclusion**

**Selected publications**

- Information Integration in Information-Based Peer-to-Peer Networks, 7th IEEE/ACM Conf. on Coop. 2002, Herrenberg
- Collaboration and Coercing Strategies for RDF-Based Peer-to-Peer Networks, 2nd Conf. on WSDM, Longbeach
- "Biological" Evaluation of Semantic SharePoint
- Learning Objects and Metadata for Online Courses, Online Ed. 2000, pp. 1-4
- Adler, Strey, Hofmann
- Metadata for Online Courses, FTZ 2002, Munich
- Maiwald, Adler, Schuberth, Hofmann
- Metadata correspondence for the knowledge-driven development of network-based information systems, 7th IEEE, 2002, pp. 1-4