Interoperation, Mediation and Composition of Engineering Services

NSF Program: CMMI Information Technology and Infrastructure Systems

(Grant Number: CMS-0601167)

***Principal Investigators*: Prof. Kincho H. Law and Prof. Gio Wiederhold, Stanford University**

***Researchers*: Mr. Jack Cheng and Dr. Jiayi Pan, Dr. Gloria Lau and Dr. Chuck Han**

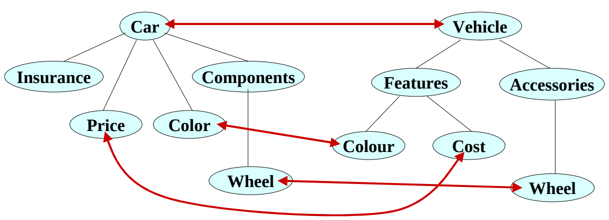
***Collaborators*: Prof. Hans Bjornsson (Chalmers University) and Dr. Albert Jones (NIST)**

Introduction

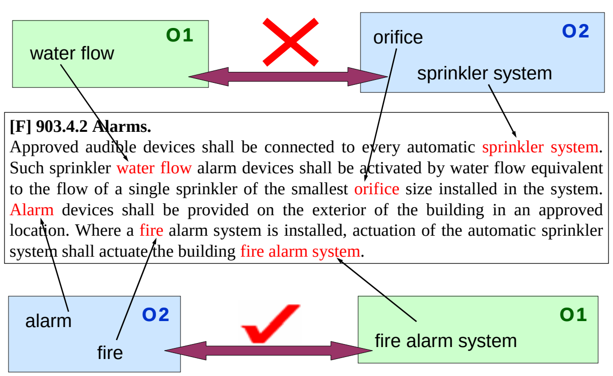
With the proliferation of computers and Internet technologies, the exchange of information is intrinsically free and instantaneous. However, in order to support meaningful exchange of information and to bring the applications together to solve complex engineering and business problems, software applications must be able to convey and understand the semantics among each other. Interoperability allows two or more distributed information sources to exchange information and re-use the information for further purposes, and, therefore, adds value to each individual source. As reported by the National Institute of Standards and Technology (NIST), the lack of information and service interoperability poses significant economic costs to both the manufacturing and construction industries [1]. This project addresses the issue of information interoperability and service integration in the facility engineering and management applications.

Ontology Mapping

Ontologies have been increasingly leveraged to describe the terminology and structure of the information sources. They act as a terminological basis for information exchange. Gartner Group forecasts that "By 2010, ontologies … will be the basis for 80 percent of application integration projects" [2]. Mapping among heterogeneous ontological standards is therefore essential for application and service integration. Ontology mapping is now mainly performed manually, which is a time-consuming and laborious process. Automated or semi-automated approaches based on linguistic similarity and ontology structures can significantly enhance the mapping process.



Knowledge-Driven Relatedness Analysis



Semantic Similarity Discovery

Advanced text mining and statistical analysis techniques are employed to discover the knowledge of semantic similarity among concepts. Three similarity analysis measures are used:

- **Cosine Similarity Measure**

$$Sim(i,j) = \frac{\vec{c}_i \cdot \vec{c}_j}{|\vec{c}_i| \times |\vec{c}_j|}$$

- **Jaccard Similarity Coefficient**

$$Sim(i,j) = \frac{|\vec{c}_i \cap \vec{c}_j|}{|\vec{c}_i \cup \vec{c}_j|} \quad \textit{or}$$

$$Sim(i,j) = \frac{N_{11}}{N_{10} + N_{01} + N_{11}}$$

- **Market Basket Model**

$$Sim(i,j) = \Pr(j\,|\,i) - \Pr(j) \quad \textit{or}$$

$$Sim(i,j) = \frac{N_{11}}{N_{11}+N_{10}} - \frac{N_{11}+N_{01}}{N_{11}+N_{10}+N_{01}+N_{00}}$$

where

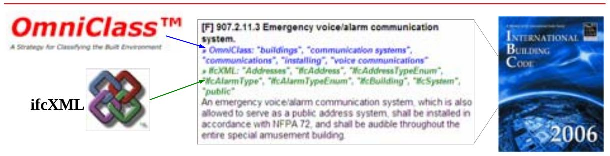
$\vec{c}_i$ is the frequency column vector of concept i

N_{10} is the number of sections that match concept i but not concept j

N_{11} is the number of sections that match both concepts i and j

N_{00} is the number of sections that do not match neither concept i nor concept j

Results and Discussions



- **By market basket model**

OmniClass concept *i*	IfcXML concept *j*	*Sim(i, j)*	*Sim(j, i)*
curtain walls	IfcCurtainWall	0.992849	0.992849
sound and signal devices	IfcSwitchingDeviceType	0.998808	0.998808
roof decking	IfcSlab	0.802344	0.370313
speakers	IfcAlarmType	0.883194	0.018024
gypsum board	IfcWallType	0.568832	0.029939
concrete	IfcSlab	0.119548	0.427615

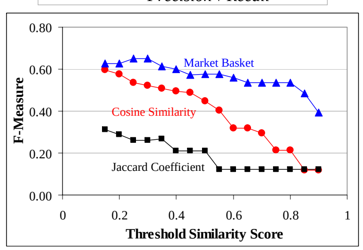
- **Precision, Recall and F-Measure**

	Actual Matched	Actual Unmatched
Predicted Matched	True positive (N_{TP})	False positive (N_{FP})
Predicted Unmatched	False negative (N_{FN})	True negative (N_{TN})

$$Precision = \frac{N_{TP}}{N_{TP}+N_{FP}}$$

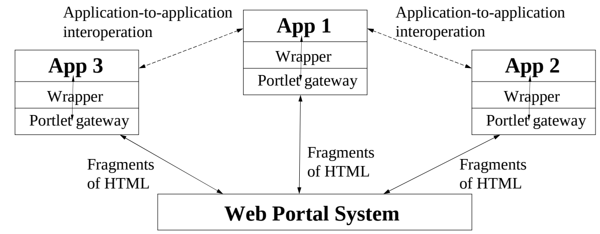
$$Recall = \frac{N_{TP}}{N_{TP}+N_{FN}}$$

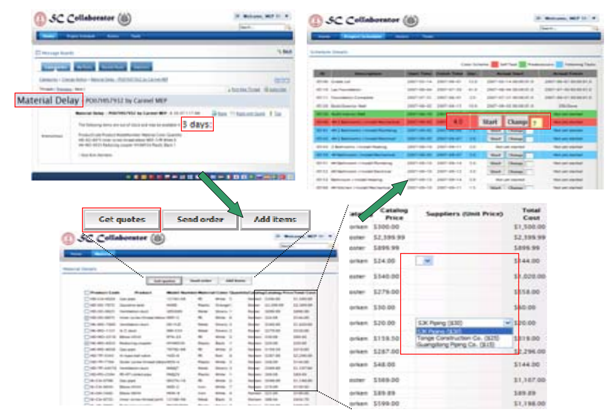
$$F\cdot Measure = \frac{2(Precision \times Recall)}{Precision + Recall} \in [0,1]$$



Service Integration

For any large scale, complex development project, information and applications are often loosely distributed among project participants and business partners. As Internet becomes ubiquitous and instantaneously accessible, the web services model allows the integration, sharing and re-use of services over the network.





Literature Cited

1. Gallaher, M., O'Connor, A., Bettbarn Jr., J., and Gilday, L. (2004). "Cost Analysis of Inadequate Interoperability in the US Capital Facilities Industry." Technical Report GCR 04-867, NIST.
2. Jacobs, J., and Linden, A. (2002). "Semantic Web technologies take middleware to next level." Research Note T-17-5338, Gartner, Inc.

Acknowledgements

The research team would like to acknowledge the supports by the National Science Foundation, the Center for Integrated Facility Engineering (CIFE) at Stanford University and the Enterprise Systems Group at the National Institute of Standards and Technology (NIST). They would also like to thank the International Code Council for providing the XML version of the International Building Code (2006).

Further Information at

http://eig.stanford.edu/supply_chain