Ontology Matching in Context-driven Collaborative Recommending Systems

Alexander Smirnov, Alexey Kashevnik and Nikolay Shilov
St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 39, 14 Line, 199178, St. Petersburg, Russia

Keywords: Context-driven Collaborative Recommending System, Ontology Matching, Ontology Alignment Pattern.

Abstract: The paper proposes an approach to building a context-driven collaborative recommending system, and concentrates on the ontology matching algorithm and ontology alignment patterns. The designed collaborative recommendation system is based on application of such technologies as user and group profiling, context management, decision mining. It enables for self-organisation of user groups in accordance with changing user profiles and the current situation context. Utilizing of the developed ontology alignment patterns considerably accelerates the ontology fusion and matching processes due to typification of fusion and alignment schemes.

1 INTRODUCTION

Integration of different Information Technologies (IT) systems in flexible supply networks (FSN) requires semantic integration of their data and workflow models. This problem is even more evident if IT systems of different enterprises (FSN members) are considered. Developing frameworks, with appurtenant models, needs to be based on solid foundations. The alignment problem requires a common ontology capturing business as well as IT (Lind and Seigerroth, 2010).

Ontologies have shown their usability for this type of tasks (e.g., (Bradfield et al., 2007); (Chan and Yu, 2007); (Patil et al., 2005)). Integration of different enterprise aspects into an ontology has been also researched in a number of works. For example, socio-instrumental pragmatism (Goldkuhl and Röstlinger, 2002) incorporates human, organizational, and IT enabled actions in a coherent ontology. The concern of theorizing actions has also been acknowledged by actor-network theory (Latour, 1991), where technology and people are both seen as social actants.

The service-oriented architecture (SOA) is a step towards information-driven collaboration. This term today is closely related to other terms such as ubiquitous computing, pervasive computing, smart space and similar, which significantly overlap each other (Balandin et al., 2009).

The main idea of the approach is to represent FSN elements by sets of services provided by them. This makes it possible to replace the configuration of FSN with that of distributed services. For the purpose of interoperability the services are represented by Web-services using the common notation described by a common application ontology (AO). Depending on the problem considered the relevant part of AO is selected forming the abstract context that, in turn, is filled with values from the sources resulting in the operational context. The operational context represents the constraint satisfaction problem that is used during self-configuration of services for problem solving.

An intensive collaboration requires strong IT-based support of decision making so that the preferences from multiple simultaneous users could be taken into account satisfying both the individual and the group (McCarthy et al., 2006). Collaborative recommending systems are aimed to solve this problem. Recommending / recommendation /recommender systems are widely used in the Internet for suggesting products, activities, etc. for a single user considering his/her interests and tastes (Garcia et al., 2009), in various business applications (e.g., (Hornung et al., 2009); (Zhena et al., 2009)) as well as in product development (e.g., (Moon et al., 2009); (Chen et al., 2010)). Collaborative recommending is complicated by the necessity to

Smirnov A., Kashevnik A. and Shilov N..
Ontology Matching in Context-driven Collaborative Recommending Systems.
DOI: 10.5220/0004126801390144
In *Proceedings of the International Conference on Knowledge Management and Information Sharing* (KMIS-2012), pages 139-144
ISBN: 978-989-8565-31-0

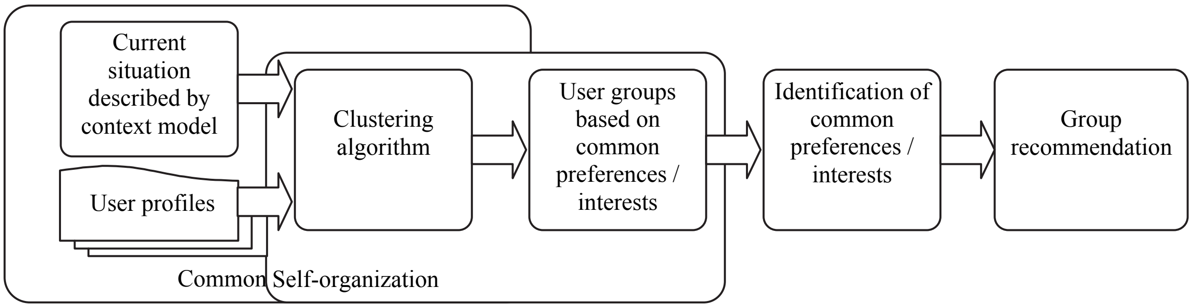

Figure 1: Collaborative recommendation system architecture.

take into account not only personal interests but to compromise between the group interests and interests of the individuals of this group. In literature (e.g., (Baatarjav et al., 2009); (Middleton et al., 2003)) the architecture of the collaborative recommending system is proposed based on three components: (i) profile feature extraction from individual profiles, (ii) classification engine for user clustering based on their preferences (e.g., Romesburg, 2004)), and (iii) final recommendation based on the generated groups. Development of clustering algorithms capable to continuously improve group structure based on incoming information enables for self-organisation of user groups (Flake et al., 2002).

2 CONTEXT-DRIVEN COLLABORATIVE RECOMMENDING SYSTEM

The developed context-driven collaborative recommendation system architecture is presented in Figure 1. It is centralized around the user clustering algorithm (Smirnov et al., 2005) originating from the decision mining area area (Smirnov et al., 2008); (Rozinat and van der Aalst, 2006); (Petrusel and Mican, 2010). The proposed clustering algorithm is based on the information from user profiles. The user profiles contain information about users including their preferences, interests and activity history. A detailed description of the profile can be found in (Smirnov et al., 2009). Besides, in order for the clustering algorithm to be more precise, this information is supplied in the context of the current situation (including current user task, product(s) she/he works with, time pressure and other parameters. The semantic interoperability (consistent notation and terminology) between the profile and the context is supported by the common ontology.

The user profiles are considered to be dynamic and, hence, the updated information is supplied to the algorithm from time to time. As a result the algorithm can run as updated information is received and update user groups. Thus, the development of the algorithm has made it possible to be used for building self-organisation mechanism for user group formation.

Since, in the company considered as a case study (Smirnov et al., 2011), the major difference between users is the group of products they work with, the generated groups are expected to be product related. However, in other environments this is not necessary to be the case and groups can be process-oriented, resource-oriented or other.

When groups are generated the common preferences/interests of the groups can be identified based on the results of the clustering algorithm. These preferences can be then generalized and analyzed in order to produce collaborative recommendations.

3 ONTOLOGY MATCHING

In order to analyze the existing ontology matching techniques an extensive state-of-the-art review has been done, which covered about 20 systems/approaches/projects related to ontology matching (Smirnov et al., 2010).

All the similarity metrics in the performed state-of-the-art review are based on the two information retrieval metrics of precision and recall. The balance between these is achieved via choosing the right threshold value. The possibility of choosing the right threshold value has to be taken into account in the development of the matching models.

Since in enterprise information systems most of services are problem-oriented, the usage of reusable ontology patterns for the common ontology creation is proposed. This would enable unification and standardization of the ontologies and significantly simplify the ontology matching.

Based on it the following concluding remarks can be made.

The goal of ontology matching is basically solving the two major problems, namely: (i) ontology entities with the same name can have different meaning; (ii) ontology entities with different names can have the same meaning.

For this purpose a number of techniques are applied in different combinations. These techniques include:

- Identification of synonyms
- Similarity metrics (name similarity, linguistic similarity)
- Heuristics (for example two nodes are likely to match if nodes in their neighborhood also match)
- Compare sets of instances of classes instead compare classes
- Rules: for example, if class A1 related to class B1 (relation R1), A2 related to class B2 (relation R2) and B1 similar to B2, R1 similar to R2 therefore A1 similar to A2.

As a result of matching the following types of elements mapping proximity can be identified:

- One-to-one mapping between the elements (Associate-Professor to Senior-Lecturer)
- Between different types of elements (the relation AdvisedBy(Student, Professor) maps to the attribute advisor of the concept Student)
- Complex type (Name maps to the concatenation of First Name and Last Name)

All methods can be separated into the following four groups:

- **Linguistic Methods** are focused on determining similarity between entities based on linguistic comparison of these entities (count of the same symbols estimation, estimation of the longest similar parts of words, etc.).
- **Statistical Methods** (Instance-Based) compare instances of the ontology entities and based on this estimation entities can be compared.
- **Contextual Methods** are aimed at calculation of a similarity measure between entities based on their contexts. For example if parents and children of the ontology classes are the same consequently the classes also the same.
- **Combined Methods** integrate specifics of two or three of the above methods.

In the considered problem domain the differentiation between instances is not an easy task. Because of this reason the techniques and methods relying on instances were not considered for further development. Hence, the developed models presented below integrate all of the above techniques (except those dealing with instances) and propose a set of combined methods having features of the linguistic and contextual methods.

4 MULTI-MODEL APPROACH FOR ONTOLOGY MATCHING

The below proposed approach allows matching of ontologies for the interoperability purposes and is based on the ontology matching model illustrated by Figure 2. The approach takes into account that the matching procedure has to be done “on-the-fly” and remembering the fact that matched ontologies are responsible for performing certain concrete and well-described tasks, which means that they generally should be small–to–medium size and describe only limited domains. A detailed description of the approach and experimentation results can be found in (Smirnov et al., 2010).

Ontology is represented as RDF triples, consisting of the following ontology elements: subject, predicate, object. Degree of similarity between two ontology elements is in the range [0, 1]. The approach consists of the following steps:

- Compare ontology elements taking into account synonyms of both ontologies. The degree of similarity between equal elements is set to 1 (maximum value of the degree of similarity).

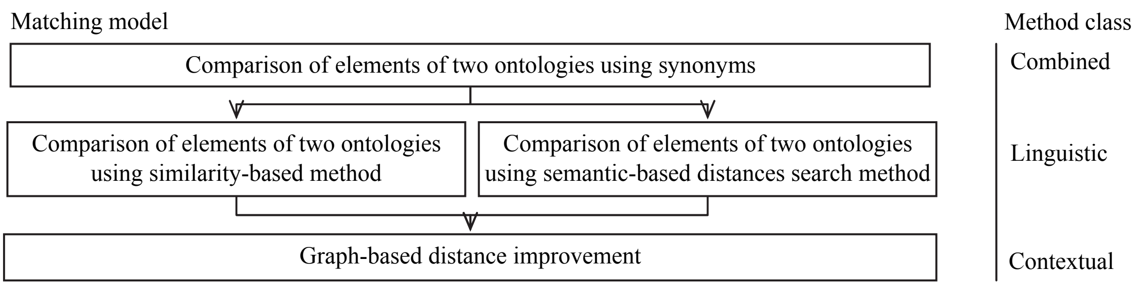

Figure 2: Multi-model approach to on-the-fly ontology matching.

- Compare all elements between two ontologies and fill the matrix M using *similarity-based model*. Matrix M is of size m to n, where m is the number of elements in the first ontology and n is the number of elements in the second ontology. Each element of this matrix contains the degree of similarity between the string terms of two ontology elements using the fuzzy string comparison method. At this step WordNet or Wiktionary can be used for searching semantic distances (Smirnov et al., 2010).
- Compare all elements of two ontologies and fill the matrix M'. Matrix M' is of size *m* to *n*, where *m* is the number of elements in the first ontology and *n* is the number of elements in the second ontology. Each element of this matrix represents the degree of similarity between two ontology elements.
- Update values in matrix M, where each new value of elements of M is the maximum value of (M, M')
- Improve distance values in the matrix M using the *graph-based distance improvement model* (Smirnov et al., 2010).

As a result the matrix M contains degrees of similarity between ontology elements of two knowledge processors. This allows determining correspondences between elements by selecting degrees of similarities which are below than the pre-selected threshold value.

5 ONTOLOGY ALIGNMENT PATTERNS

Ontology alignment is a set of correspondences between two or more (in case of multiple matching) ontologies obtained as a result of the ontology matching process (Euzenat and Shvaiko, 2007). In this section the complicated ontology alignment situations (patterns), which may arise during setting relationships between elements and the rules of their processing, are presented. These patterns are valid for both straight and reverse directions.

Patterns are a proven way to capture experts' knowledge in fields where there are no simple "one size fits all" answers (Enterprise Integration Patterns, 2012), such as knowledge fusion or ontology alignment. Each pattern poses a specific design problem, discusses the considerations surrounding the problem, and presents an elegant solution that balances the various forces or drivers. In most cases, the solution is not the first approach that comes to mind, but one that has evolved through actual use over time. As a result, each pattern incorporates the experience base that senior integration developers and architects have gained by repeatedly building solutions and learning from their mistakes. This implies that patterns are not invented, but discovered and observed from actual practice in the field (Enterprise Integration Patterns, 2012).

Notations:

Source – ontology mapped;

Destination – ontology mapped to;

○ – class;

● – attribute;

— – associative relationship;

|\ – hierarchical relationships or "class-attribute" relationship;

⇢ – correspondence relationship.

Class-to-class Alignment. A class a'' from the Source corresponds (maps) to a class a' from the Destination; a subclass b'' of the class a'' does not correspond to any class from the Destination. In this case search "in depth" does not stop and if a subclass c'' of the class b'' corresponds to a class c' from the Destination, the class c' becomes a subclass of the class a', and the class c'' becomes a subclass of the class a''. Experts can make a decision about including or not the new class b'' into the common ontology (Figure 3).

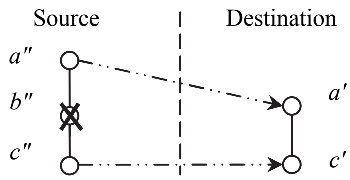

Figure 3: Class-to-class alignment.

Attribute-to-attributes Alignment. An attribute $attr''$ of the class a'' from the Source corresponds to several attributes (a set of attributes) $ATTR'$ of the class a' from the Destination. In this case all the attributes from $ATTR'$ and methods for values conversion should be added into the common ontology (Figure 4).

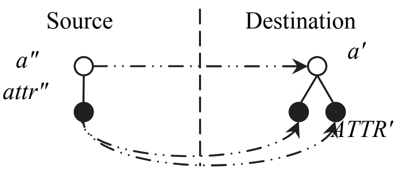

Figure 4: Attribute-to-attributes alignment.

Class-to-classes Alignment. A class a'' from the Source corresponds to several classes (a set of classes) A' from the Destination. In this situation all the classes from A' and conditions of selection among these classes are added into the common ontology. Attributes and subclasses of the class a'' are mapped into attributes or subclasses of the classes from A' (Figure 5).

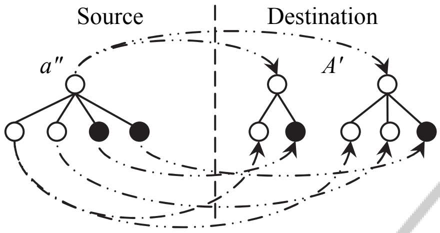

Figure 5: Class-to-classes alignment.

Class-to-attribute Alignment. A class a'' from the Source corresponds to a class a' from the Destination; a class b'' associatively connected to the class a'' from the Source corresponds to an attribute $attr'$ of the class a' from the Destination. In this situation all the attributes and subclasses of the class b'' are mapped to the attribute $attr'$ with appropriate conversion methods and conditions are also added (Figure 6).

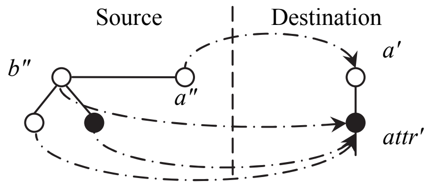

Figure 6: Class-to-attribute alignment

Subclass-to-attribute Alignment. A class a'' from the Source corresponds to a class a' from the Destination; subclass b'' of the class a'' corresponds to an attribute $attr'$ of the class a'. In this situation all the subclasses of the class b'' are mapped to the attribute $attr'$ or possibly to other attributes of the class a', with appropriates conversion methods and conditions being also added (Figure 7).

6 CONCLUSIONS

The major idea of the proposed approach to building a context-driven collaborative recommending system is FSN representation via a set of services provided by its elements. SOA makes it possible to abstract from real services and model these via Web-services. Taking into account the described SOA advantages this enables a higher level of abstraction and ontology-based interoperability. The ontological model is used to solve the problem of heterogeneity.

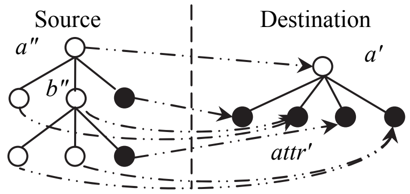

Figure 7: Subclass-to-attribute alignment.

The designed collaborative recommendation system is based on application of such technologies as user and group profiling, context management, decision mining. It enables for self-organisation of user groups in accordance with changing user profiles and the current situation context.

The paper concentrates on description of the developed multi-model approach to on-the-fly ontology matching and ontology alignment patterns. Utilizing of the patterns considerably accelerates the ontology fusion and matching processes due to typification of fusion and alignment schemes.

The presented work is yet in an early stage of development. Only some of the proposed ideas have been partially verified. The next step will be devoted to application of the presented ontology alignment patterns to a real world problem and further analysis of their completeness and usefulness.

ACKNOWLEDGEMENTS

The research presented is motivated by a joint project between SPIIRAS and Nokia Research Center. Some parts of the work have been sponsored by grants # 12-07-00298-a, # 12-07-00302-a, # 11-07-00368-a, and # 11-07-00045-a of the Russian Foundation for Basic Research, project # 213 of the research program "Intelligent information technologies, mathematical modelling, system analysis and automation" of the Russian Academy of Sciences, and project 2.2 "Methodology development for building group information and recommendation systems" of the basic research program "Intelligent information technologies, system analysis and automation" of the Nanotechnology and Information technology Department of the Russian Academy of Sciences.

REFERENCES

Baatarjav, E.-A., Phithakkitnukoon, S., and Dantu, R., (2009). Group Recommendation System for Facebook. In *OTM 2008: Proceedings of On the Move to Meaningful Internet Systems Workshop*, *LNCS 5333* (pp. 211-219).

Balandin, S., Moltchanov, D., and Koucheryavy, Y., (Eds.) (2009). *Smart Spaces and Next Generation Wired/Wireless Networking*, Springer, LNCS 5764.

Bradfield, D. J., Gao, J. X., and Soltan, H., (2007). A Metaknowledge Approach to Facilitate Knowledge Sharing in the Global Product Development Process. *Computer-Aided Design & Applications*, 4(1-4), 519-528.

Chan, E. C. K. and Yu, K. M., (2007). A framework of ontology-enabled product knowledge management. *International Journal of Product Development*, 4(3-4), 241-254.

Chen, Y.-J., Chen, Y.-M., and Wu, M.-S., (2010). An expert recommendation system for product empirical knowledge consultation. In *ICCSIT2010: The 3rd IEEE International Conference on Computer Science and Information Technology, IEEE* (pp. 23-27).

Enterprise Integration Patterns, (2012). Retrieved from: http://www.eaipatterns.com/.

Euzenat J. and Shvaiko, P., (2007). *Ontology matching*, Springer-Verlag.

Flake, G. W., Lawrence, S., Giles, C. L. and Coetzee, F., (2002). Self-Organization and identification of Web Communities, *IEEE Computer*, 35(3) 66-71.

Garcia, I., Sebastia, L., Onaindia, E., and Guzman, C., (2009). A Group Recommender System for Tourist Activities. In *EC-Web 2009: Proceedings of E-Commerce and Web Technologies, The 10th International Conference* (pp. 26-37) Springer, LNCS 5692.

Goldkuhl, G. and Röstlinger, A., (2002). Towards an integral understanding of organisations and information systems: Convergence of three theories. In *Proceedings of the 5th International Workshop on Organisational Semiotics*.

Hornung, T., Koschmider, A., and Oberweis, A., (2009). A Recommender System for Business Process Models. In *WITS'09: Proceedings of the 17th Annual Workshop on Information Technologies & Systems*. Retrieved from: http://ssrn.com/abstract=1328244.

Latour, B., (1991). Technology is society made durable. In: Law J. (Ed.) *A sociology of monsters: Essays on power, technology and domination*, (pp. 103-131). Routledge & Kegan Paul.

Lind, M. and Seigerroth, U., (2010). Multi-Layered Process Modeling for Business and IT Alignment. In *Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS)*, (pp. 1-10).

McCarthy, K., Salamo, M., Coyole, L., McGinty, L., Smyth, B., and Nixon, P., (2006). Group Recommender Systems: A Critiquing Based Approach, In *IUI '06: Proceedings of the 11th international conference on Intelligent user interfaces* (pp. 267-269).

Middleton, S. E., De Roure, D., and Shadbolt, N. R., (2003). Ontology-Based Recommender Systems. In Staab, S. and Studer, R. (Eds.) *Handbook on Ontologies* (pp. 477-498) Springer.

Moon, S. K., Simpson, T. W., and Kumara, S. R. T., (2009). An agent-based recommender system for developing customized families of products. *Journal of Intelligent Manufacturing*, 20(6), 649-659.

Patil, L., Dutta, D., and Sriram, R., (2005). Ontology-based exchange of product data semantics. *IEEE Transactions on Automation Science and Engineering*, 2(3), 213-225.

Petrusel, P. and Mican, D., (2010). Mining Decision Activity Logs. In Abramowicz, W., Tolksdorf, R., and Wecel. K. (Eds.) *BIS2010: Proceedings of Business Information Systems Workshops,* (pp. 67–79) Springer, LNBIP 57.

Romesburg, H. C., (2004). *Cluster Analysis for Researchers*, Lulu Press, California.

Rozinat, A. and van der Aalst, W. M. P., (2006). *Decision Mining in Business Processes*, BPM Center Report no. BPM-06-10.

Smirnov, A., Kashevnik, A., Shilov, N., Balandin, S., Oliver, I., and Boldyrev, S., (2010) On-the-Fly Ontology Matching in Smart Spaces: A Multi-Model Approach. In Balandin, S., Dunaytsev, R., Koucheryavy, Y. (Eds.) *Smart Spaces and Next Generation Wired/Wireless Networking*, (pp.72-83) Springer, LNCS 6294.

Smirnov, A., Levashova, T., Kashevnik, A., and Shilov, N. (2009) Profile-based self-organization for PLM: approach and technological framework. In *PLM 2009: Proceedings of the 6th International Conferecne on Product Lifecycle Management*.

Smirnov, A., Pashkin, M., and Chilov, N., (2005). Personalized Customer Service Management for Networked Enterprises. In *ICE 2005: Proceedings of the 11th International Conference on Concurrent Enterprising* (pp. 295-302).

Smirnov, A., Pashkin, M., Levashova, T., Kashevnik, A., and Shilov, N., (2008). Context-Driven Decision Mining, In Wang, J. (Ed.) *Encyclopedia of Data Warehousing and Mining*. Vol. 1. (pp. 320–327). Hershey, New York, Information Science Preference, Second Edition.

Smirnov, A., Shilov, N., Kashevnik, A., Jung, T., Sinko, M., and Oroszi, A., (2011). Ontology-Driven Product Configuration: Industrial Use Case, In *Proceedings of International Conference on Knowledge Management and Information Sharing (KMIS 2011)* (pp. 38-47).

Zhena, L., Huangb, G. Q. and Jiang, Z., (2009). Recommender system based on workflow, *Decision Support Systems*, 48(1), 237-245.