

ПОДХОД К ОБЪЕДИНЕНИЮ ОНТОЛОГИЙ НА ОСНОВЕ АЛГОРИТМА SMART С РАСШИРЕННЫМ ЭТАПОМ СРАВНЕНИЯ

Будяков А.Ю., Штогрин Е.С.

Институт телекоммуникационных систем НТУУ «КПИ», Украина

E-mail: abudiakov@gmail.com

Аннотация — В работе рассмотрен подход к объединению онтологий. Подход основывается на алгоритме SMART. Для повышения качества объединения онтологий предлагается расширение этапа сравнения алгоритма SMART.

1. Введение

Стремительное развитие сети Интернет и отсутствие единого формата описания данных усложняет процесс поиска необходимой информации. Одним из подходов для решения данной проблемы стал переход от классической сети Интернет к Семантической паутине, в которой аннотирование ресурсов происходит с помощью онтологий [1].

Количество онтологий увеличивается за счет разработки онтологий различными организациями и под конкретные задачи. Для задач, которые охватывают различные предметные области, возникает необходимость использования нескольких разработанных онтологий совместно. Вследствие чего, актуальна задача объединения онтологий.

2. Основная часть

Онтологии состоят из классов объектов, их связей и ограничений, принимаемых в выбранной области [1]. Разработанные онтологии могут включать различные описания одних и тех же объектов предметной области. Для объединения таких онтологий были созданы различные алгоритмы, такие как SMART, PROMPT, OM, агентно-ориентированный [2], которые из двух исходных онтологий позволяют создать одну результирующую. Такие алгоритмы работают в автоматизированном режиме и в процессе своей работы формируют список предположений, которые должен подтвердить пользователь. Например, копирование уникальных классов в результирующую онтологию, объединение двух классов из различных онтологий в один класс и т.д.

Алгоритм SMART [3] формирует наиболее качественные и аргументированные предположения, позволяет уточнять их, а также обновляет список действий в зависимости от ответов пользователя. Помимо списка предположений создается список возникающих конфликтов и способов их устранения. Данные действия повторяются до окончательного объединения онтологий.

На этапе сравнения выдвигаются предположения об идентичности классов, если их имена полностью совпадают. Данное сравнение только имён классов неэффективно, поскольку упускаются возможности синонимического равенства классов, случаи ошибочного написания имён классов и т.д.

В данной работе предлагается расширить этап сравнения методами описанными ниже:

— сравнение URI. Данное сравнение позволяет однозначно принять решение об идентичности классов;

— сравнение имен классов, используя базу данных WordNet [4]. Данный метод применим только для англоязычных онтологий. Он основывается на поиске синсетов (синонимических рядов, объединяющих

слова со схожим значением) обоих классов в базе данных WordNet,

— сравнение имён классов как символьных строк. В данном случае могут использоваться: алгоритм Вагнера-Фишера, алгоритм Укконена, вычисление расстояния Хэмминга, вычисление расстояния Дамерау-Левенштейна [5].

Методы расположены в порядке убывания точности сравнения и применяются поочередно. В случае положительного результата одного из методов сравнения остальные не применяются, и может быть выдвинуто предположение о равенстве классов.

3. Заключение

В работе предложена модификация алгоритма SMART для объединения онтологий. Расширение этапа сравнения позволяет более эффективно контролировать процесс сравнения классов, повышая качество полученной в результате объединения онтологий.

4. Список литературы

- [1] Ding L. Using ontologies in the semantic web: A survey / L. Ding, P. Kolari, Z. Ding, S. Avancha, T. Finin, A. Joshi. — A Handbook of Principles, Concepts and Applications in Information Systems, 2005. — P. 79 — 113.
- [2] Shukla A. Comparative Study of Ontology Merging Algorithms / A. Shukla, P. Bhutani // Proc. of the 4th National conf. «Computing For Nation Development» (INDIACom-2010). — New Delhi, 2010. — P. 311 — 315.
- [3] Noy N.F. SMART: Automated Support for Ontology Merging and Alignment / The Stanford center for biomedical informatics research. — http://bmir.stanford.edu/file_asset/index.php/192/BMIR-1999-0813.pdf. — 05.02.2013.
- [4] Miller G.A. WordNet: A Lexical Database for English / G.A. Miller // Communications of the ACM. — 1995. — Vol. 38, № 11. — P. 39 — 41.
- [5] Альфред А.В. Структуры данных и алгоритмы / А.В. Альфред, Д.Э. Хопкрофт, У.Д. Джеффри. — М.: Вильямс, 2000. — 384 с.

SMART-BASED APPROACH FOR THE ONTOLOGY MERGING WITH THE EXTENDED COMPARISON PHASE

Budiakov A.Y., Shtogrina O.S.

*Institute of Telecommunication Systems
National Technical University of Ukraine "KPI", Ukraine*

Abstract — The approach for the ontology merging was reviewed. This approach is based on the SMART algorithm. The extension of a comparison phase of the SMART algorithm is suggested to improve the quality of the ontology merging.