

УДК 004.822

ОПТИМИЗАЦИЯ ОТОБРАЖЕНИЯ ОНТОЛОГИЙ МЕТОДОМ РОЯ ЧАСТИЦ

А.В. Семенова¹, В.М. Курейчик²*Южный федеральный университет, Таганрог, Россия*¹*alexforum@.ru*, ²*kur@tgn.sfedu.ru*

Аннотация

Работа посвящена интеграции данных распределённых гетерогенных информационных источников на основе онтологий. Метод вычисления семантической близости концептов позволяет количественно оценить сходство между понятиями. Предложена технология оценки семантической близости концептов двух онтологий, отличительной особенностью которой является автоматическое определение весовых коэффициентов с использованием метода роя частиц. С помощью предложенного подхода генерируются векторы весовых коэффициентов, которые соответствуют разным коэффициентам семантической близости. Полученные весовые коэффициенты используются для объединения разных коэффициентов семантической близости. Новизна заключается в том, что применённый метод роя частиц позволяет автоматически определять весовые коэффициенты семантической близости концептов двух онтологий. Таким образом, для каждого концепта одной онтологии формируется множество релевантных семантических концептов другой онтологии. Благодаря этому возможно получать наборы оптимальных и квазиоптимальных решений.

Ключевые слова: онтология, отображение онтологий, роевой интеллект, базы знаний, многокритериальная оптимизация.

Цитирование: Семенова, А.В. Оптимизация отображения онтологий методом роя частиц / А.В. Семенова, В.М. Курейчик // Онтология проектирования. – 2018. – Т. 8, №2(28). – С. 285-295. – DOI:10.18287/2223-9537-2018-8-2-285-295.

Введение

Одним из основных направлений развития современных интеллектуальных информационных систем является разработка новых подходов представления и интеграции знаний. Интеллектуальные информационные системы содержат разнородные базы знаний, которые имеют собственные локальные информационные модели, поддерживаемые разными стандартами представления данных и знаний. Как следствие, на этапе интегрирования разнородных данных возникает множество конфликтов: используется разная терминология при обозначении семантически близких понятий предметной области (ПрО), существует неоднородность спецификаций концептов и др.

Решение проблемы интеграции неоднородной информации возможно путём спецификации конкретной ПрО с последующей проверкой и отображением (*mapping*) понимания различных ПрО. В настоящее время для описания ПрО широко применяется онтологический подход, поскольку онтологические спецификации позволяют задавать точную семантику прикладной области объектов и определять их контекст. Создание общедоступных онтологий ПрО позволяет решить проблему неоднородности онтологических спецификаций для определённых групп агентов. Однако в условиях открытого информационного пространства при решении задач предъявляются различные требования к глубине и степени формализации описания ПрО, используются неоднородные онтологические описания ПрО. Неоднородность онтологических спецификаций появляется на уровнях модельной и понятийной семантики.

Соответственно, задачи отображения онтологических моделей и онтологических контекстов являются актуальными.

Анализ литературы [1-7] позволяет заключить, что задачи объединения (согласование, выравнивание, отображение и т.п.) неоднородных онтологических моделей относятся к классу NP-трудных задач оптимизации и могут быть решены с применением эволюционных алгоритмов поиска оптимальных решений. Свою эффективность уже доказали, например, генетические, муравьиные, пчелиные алгоритмы и алгоритмы, основанные на эволюционном моделировании [1].

В работе для решения задачи отображения онтологий предложено использовать один из эволюционных методов оптимизации – метод роя частиц (МРЧ), который позволяет автоматически определять весовые коэффициенты семантической близости концептов двух онтологий.

1 Постановка задачи

Пусть заданы две онтологии O_1 и O_2 , формализующие семантику некоторой ПрО. Причём, элементы онтологий подвержены изменению с течением времени. Изменения носят дискретный характер и вносятся в соответствии с синтаксисом выбранного языка представления онтологий. Обобщив ряд определений, рассматривающих онтологию с точки зрения структуры и элементов, составляющих онтологию, её можно представить как [2]:

$$(1) \quad O = (C, P, R),$$

где:

O – онтология ПрО;

C – множество понятий (классов), определённых для конкретной ПрО;

P – множество свойств понятий;

R – множество отношений, определённых между понятиями в C .

Отображение онтологии $O_1 = (C_1, P_1, R_1)$ на онтологию $O_2 = (C_2, P_2, R_2)$, означает нахождение для каждого элемента множеств C_1, P_1, R_1 , составляющих онтологию O_1 , подобных элементов из множеств C_2, P_2, R_2 онтологии O_2 . Отображением двух онтологий является множество

$$(2) \quad m = \langle e, e', \mathfrak{R}, k \rangle,$$

где:

e и e' – элементы множеств C, P и R ;

$\mathfrak{R} = \{ \subset, \subseteq, \equiv, \supset, \supseteq \}$ – множество отношений;

k – весовой коэффициент, отражающий степень уверенности относительно корректности конкретного отображения [8].

Процесс отображения онтологий может быть задан функцией F , на вход которой поступают две онтологии O_1 и O_2 , а на выходе получаем $|O_1| \times |O_2|$ – матрицу мер близости S , элементами которой (s_{ij}) являются величины меры близости между i -м элементом онтологии O_1 и j -м элементом онтологии O_2 . Семантической близостью элементов называется смысловое сходство этих элементов. Семантическая близость может определяться между разными компонентами триплетов. При этом в качестве базовой близости можно рассматривать близость между элементами онтологий (классами, предикатами, терминами).

Процедура отображения заключается в нахождении семантических связей между концептами различных онтологий. Следовательно, центральной задачей отображения онтологий является вычисление меры семантической близости – количественной оценки семантической

схожести онтологических сущностей. Мера семантической близости показывает высокие значения для пар сущностей, которые находятся в семантических отношениях (синонимия, гипонимия, ассоциативность, когипонимия), и нулевые значения для всех остальных пар. Как правило, мера близости между сущностями разных онтологий определяется на нескольких уровнях: лексическая близость, близость атрибутов и отношений, близость экземпляров понятий и другие. Расчёт близости между сущностями в разных онтологиях является итерационным процессом, поскольку многие меры близости двух сущностей основываются на близости других сущностей (понятий, свойств, экземпляров). Очевидно, чем полнее учитываются характеристики двух сущностей, тем качественнее является мера близости. Таким образом, комплексные меры близости, сочетающие несколько подходов, являются наиболее перспективными. Комплексная мера близости рассчитывается как среднее арифметическое значение мер близости с учётом весового коэффициента:

$$(3) \quad S_{CX}(e, e') = \frac{\sum_{i=1}^n w_i s_i(e, e')}{\sum_{i=1}^n w_i},$$

где:

s_i – мера близости по определённому критерию;

w_i – весовой коэффициент, который определяет важность меры семантической близости (сумма весов равна 1, $w_i \in [0, 1]$);

n – число мер близости.

Весовые коэффициенты могут определяться либо экспертами, либо автоматически с помощью обучаемой нейронной сети или эволюционного алгоритма. В случае автоматического определения весовых коэффициентов проблема отображения онтологий может быть решена путём нахождения оптимальных весовых коэффициентов мер семантической близости концептов. Таким образом, задача отображения онтологий может быть решена как задача оптимизации:

$$(4) \quad F(m) = (r(m), p(m)), m = \sum_{i=1}^{m_{set}} w_i m_i,$$

где

m_{set} – множество отображений;

w_i – множество всех возможных весовых коэффициентов, которые использовались для вычисления комплексного отображения;

$F: w_i \rightarrow \delta[0,1]$, δ – целевая функция, позволяющая оценить качество множества весовых коэффициентов w_i ;

$r, p: m \rightarrow m[0,1]$ – критерии полноты и точности отображения m , соответственно.

Точность (*precision*) и полнота (*recall*) являются метриками, которые используются при оценке отображения онтологий. Полнота (r) вычисляется как

$$(5) \quad r = \left(\frac{|m_a \cap m_b|}{|m_a|} \right).$$

Точность вычисляется как:

$$(6) \quad p = \left(\frac{|m_a \cap m_b|}{|m_b|} \right),$$

где:

m_a – найденные релевантные отображения;

m_b – общее число релевантных отображений.

Сформулируем задачу многокритериальной оптимизации отображения онтологий:

$$(7) \quad \begin{cases} \max F(W) = (r(W), p(W)), \\ W = (w_1, w_2, \dots, w_n)^T, \\ w_i \in [0,1], w = 1 \dots n, \sum_{i=1}^n w_i = 1, \end{cases}$$

где:

n – число отображений сущностей онтологий O_1 и O_2 (вычисленных мер близости);

$w_i, w = 1, 2, \dots, n$ – весовой коэффициент, который определяет важность i -й меры семантической близости.

Таким образом, целевая функция – это максимум параметров точности и полноты.

2 Краткий обзор работ по теме исследования

Современные методы нахождения отображения онтологий носят, как правило, междисциплинарный характер [3-7]. Среди них можно выделить четыре группы: лингвистические, статистические, структурные и логические методы [9]. Эти методы постоянно развиваются и совершенствуются. Тем не менее, остается ряд открытых вопросов, таких как отображение неоднородных онтологических моделей, автоматизация и оптимизация процесса отображения онтологий и др.

Автоматизации процесса интеграции онтологических моделей посвящены работы [9-12] и др. Существующие методы основаны на вычислении сходства между объектами онтологий путём использования различных типов информации в онтологиях, например, имён сущностей, структур таксономии, ограничений и экземпляров сущностей. Большинство разработанных алгоритмов отображения и выравнивания онтологий основаны на методе генетического программирования.

Л.В. Найхановой рассмотрена реализация решения задач естественно-языковой обработки научного текста, осуществлённая с применением технологий генетического и автоматного программирования, которая позволила создать технологию решения задач построения онтологий, отличающуюся почти полной автоматической обработкой [11].

О.А. Бубаревой в целях интеграции информационных систем на основе онтологий предложен генетический алгоритм, который основан на использовании аналогов с эволюционными процессами репродукции, кроссинговера, мутации и естественного отбора [12].

В статье [13] рассматривается возможность применения генетического алгоритма для оптимального задания весовых коэффициентов с целью определения взвешенного среднего нескольких базовых программ сопоставления данных в системе GOAL. В данной работе модель выравнивания онтологии была представлена как задача оптимизации, в которой целевая функция основана на понятии нечёткой близости.

В работе [14] предложен модифицированный генетический алгоритм, позволяющий объединить разные метрики близости онтологических моделей в одну. В работе [15] сформулированы базовые метрики семантической близости и предложен гибридный подход, основанный на генетическом алгоритме, который определяет оптимальную конфигурацию для отображения онтологий.

Авторы работы [16] предложили эволюционный подход к решению задачи интеграции множественных онтологий для обеспечения совместимости и репрезентации данных и знаний в интеллектуальных информационных системах. Такой подход позволяет определить семантически приоритетные объекты данных и знаний для их представления в модели интеграции, а также устранить дублирование и противоречия сущностей и связей на уровне Про и объектов данных из областей интеграции.

Известно, что эволюционные методы оптимизации позволяют находить квазиоптимальные решения за приемлемое время [17]. МРЧ представляет собой эволюционный метод оптимизации. Впервые он был предложен для задачи имитации социального поведения [18]. МРЧ привлекателен простотой реализации и может использоваться для решения многих задач, в том числе обучения нейросетей, поиска минимума функции, а также задач, типичных для генетических алгоритмов. МРЧ, как и все алгоритмы, принадлежащие к семейству эволюционных алгоритмов, является стохастическим, не требующим вычисления градиента. Многочисленные исследования показали, что оптимизация методом роя частиц является эффективным подходом для решения непрерывных задач оптимизации [19].

В работе [20] был предложен эволюционный подход к выравниванию онтологий, основанный на МРЧ. Оптимальное отображение находилось путём решения однокритериальной задачи вычисления весовых коэффициентов мер семантической близости.

В настоящей работе подход, описанный в работе [20], модифицируется за счёт нахождения весовых коэффициентов мер семантической близости, которые обеспечат максимальные значения критериев качества.

3 Описание метода отображения онтологий

Задача отображения онтологий решается как процесс установления отображения двух онтологий на основе вычисления комплексной меры семантической близости сущностей онтологий. Для её решения предложен алгоритм, который включает в себя пять последовательно выполняемых этапов (рисунок 1).

На первом этапе вычисляются пять мер семантической близости на лексическом, смысловом уровнях, уровне классов, уровне триплетов, таксономическом уровне. На лексическом уровне выполняется оценка семантической близости (редакционного расстояния) с применением формулы Левенштейна¹. На смысловом уровне используется мера семантической близости, которая использует сходство между словами в имени сущности (основан на семантической сети WordNet, метод Резника²). Для оценки косинусной меры семантической близости концепты онтологии преобразуются в векторы (модель word2vec³) и вычисляется её значение. Семантическая близость между триплетами обусловлена близостью между их компонентами. Полагается, что если соответствующие компоненты двух триплетов являются близкими по смыслу, то эти триплеты также являются семантически близкими. Мера семантической близости на уровне таксономии вычисляется только для класса. Это приём позволяет измерить разницу в локальной структуре объекта.

¹ Левенштейн В.И. Двоичные коды с исправлением выпадений, вставок и замещений символов. Доклады Академии наук СССР, 1965. 163.4: 845-848. *Прим.ред.*

² Resnik, P. Using Information Content to Evaluate Semantic Similarity in a Taxonomy. In IJCAI, 1995, v. 1, p. 448–453. *Прим.ред.*

³ Word2vec — программный инструмент, разработанный компанией Google, для анализа семантики естественных языков. Представляет собой технологию, основанную на дистрибутивной семантике и векторном представлении слов. *Прим.ред.*

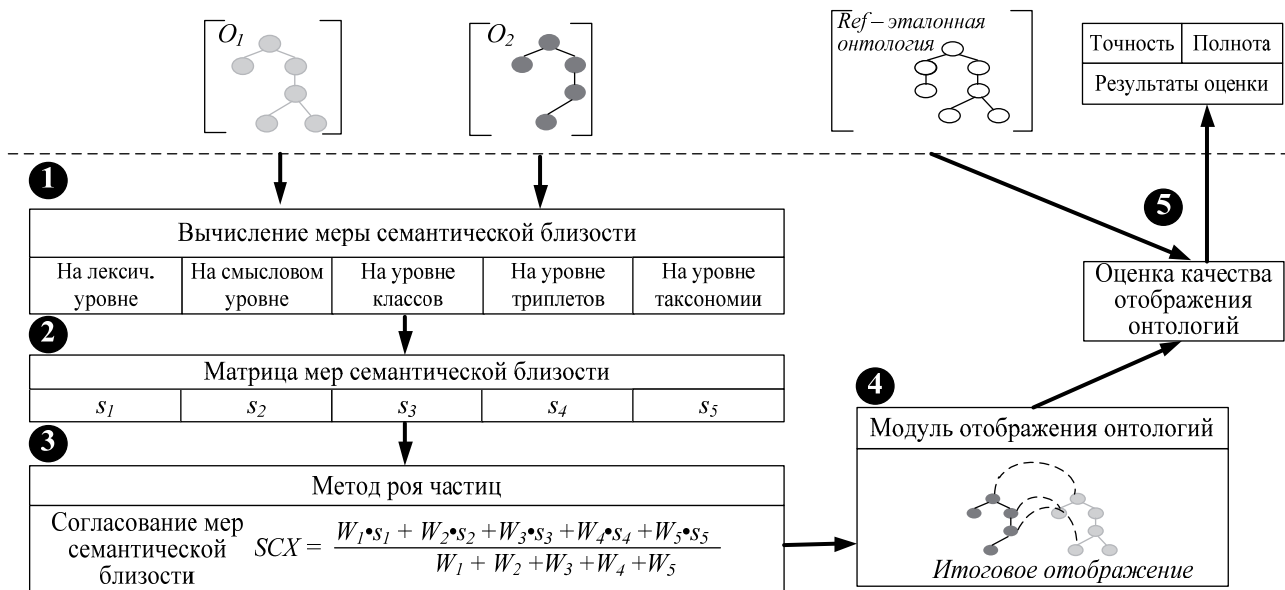


Рисунок 1 – Этапы отображения онтологий с применением многокритериальной оптимизации методом роя частиц

Далее инициализируется выполнение алгоритма оптимизации МРЧ. С использованием формулы (3) вычисляется комплексная мера семантической близости и определяется матрица комплексных мер семантической близости. Для полученной матрицы применяется алгоритм многокритериальной оптимизации МРЧ (формула 7), вычисляются оптимальные весовые коэффициенты мер семантической близости и выполняется итоговое отображение онтологий.

Согласно МРЧ в процессе оптимизации поддерживаются популяции возможных решений, называемых частицами, которые перемещаются в пространстве решений. Перемещения подчиняются принципу наилучшего найденного в этом пространстве положения, которое постоянно изменяется при нахождении частицами более выгодных положений. Схема алгоритма оптимизации МРЧ приведена на рисунке 2.

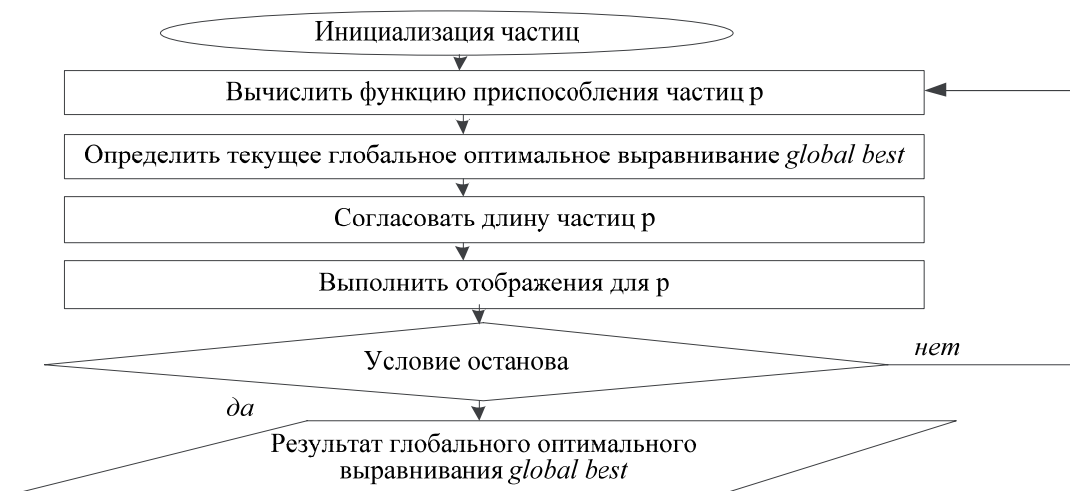


Рисунок 2 – Алгоритм оптимизации методом роя частиц

В нашем случае в роли частиц выступают весовые коэффициенты семантической близости концептов. Инициализация процесса оптимизации МРЧ выполняется с помощью популяции случайных частиц, после чего алгоритм выполняет поиск оптимальных решений путём непрерывного обновления популяций. Отображение онтологий дискретным МРЧ определяется некоторым числом N частиц (популяцией). В каноническом МРЧ на каждой итерации для каждой частицы применяется вектор скорости, который определяет её новое положение в пространстве параметров. Такая эволюция происходит за счёт управляемых случайных повторных итераций в каждой частице. Более подробное описание алгоритма приведено в работе [20], где оптимизация выполняется по одному критерию.

Инициализация. Популяция называется роем, и она состоит из m подходящих решений или частиц. Каждая частица имеет n клеток или позиций, содержащих n весовых коэффициентов, соответствующих разным мерам подобия.

Например, модели кодирования частиц с семью ячейками преобразуются в семь весовых коэффициентов (нормализованное представление ячейки) для семи мер сходства (i -я частица с семью позициями j , которые могут быть вычислены по формуле

$$w_{ij} = \frac{x_{ij}}{\sum_{i=1}^7 x_i}, \text{ причём, } 0 \leq w_{ij} \leq 1.$$

Первоначально случайным образом выбирается значение для каждой ячейки в диапазоне от 0 до 1. После того, как выбраны начальные рои, рассчитываются соответствующие им значения пригодности. Первоначальная скорость каждой ячейки частицы принимается равной нулю. Входными данными предложенной методики являются: размер роя; пороговое значение w_{ij} .

Вычисление целевой функции. Согласно (7) используются критерии полноты (5) и точности (6) поиска.

Создание роя следующего поколения. Операция осуществляется путём оценки положения и скорости частицы. Каждая ячейка или позиция представляет собой вес (нормализованное значение ячейки) относительно меры подобия. Ячейки внутри частицы содержат значения от 0 до 1, а скоростям частиц заданы нулевые значения. Используя результаты предыдущей стадии, положение каждой частицы и её скорость обновляются. Каждая частица отслеживает лучшую позицию, которую она достигла ($pbest$). Самое лучшее положение среди всех частиц называется глобальным лучшим ($gbest$). Изменение скорости частицы происходит согласно следующему выражению:

$$(8) \quad v_{ij}(t+1) = w \cdot v_{ij}(t) + k_1(pbest_{ij}(t) - x_{ij}(t)) + k_2(gbest(t) - x_{ij}(t))$$

$$(9) \quad x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1),$$

где

t – временная метка, j -й кластер i частицы;

k_1 и k_2 – случайные значения из диапазона от 0 до 1.

Далее вычисляется значение целевой функции каждой частицы по формуле (7) и отбор частиц, которые дают наибольшие значения целевой функции.

4 Экспериментальная оценка оптимизации отображения онтологий

Предложенный алгоритм был реализован в среде Matlab. В эксперименте были использованы две онтологии, структура которых приведена на рисунке 4.

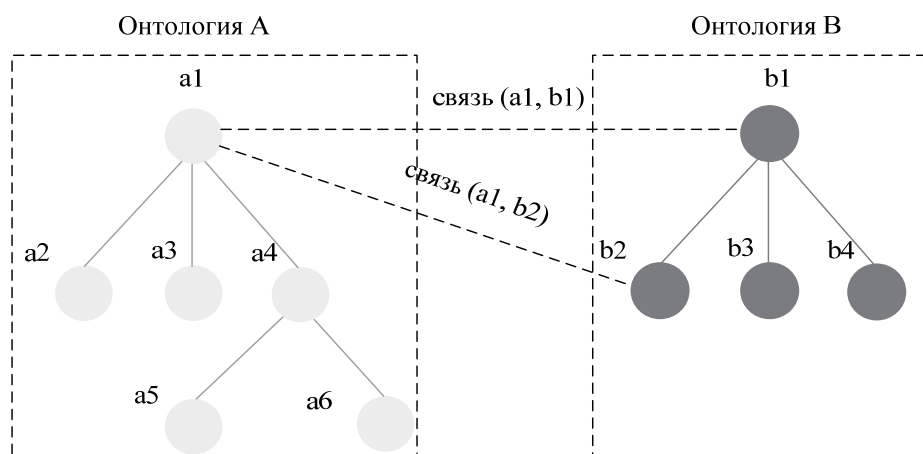


Рисунок 4 – Пример отображения онтологий

Онтология *A* имеет шесть сущностей, а онтология *B* - четыре сущности. Каждый элемент онтологии *A* имеет связь со всеми другими элементами онтологии *B*. Поскольку существуют четыре связи для каждого элемента онтологии *A*, следовательно, в общей сложности имеем двадцать четыре парных связи. На рисунке 4 условно показаны только связи $[a1, b1]$ и $[a1, b2]$. Значение подобия между $a1$ и $b1$ определяет весовой коэффициент для связи $[a1, b1]$. Результаты сравнения предлагаемого подхода с однокритериальной оптимизацией (оптимизация была выполнена отдельно для критериев точности и полноты в работе [20]) приведены в таблице 1. По критерию точности эффективность предлагаемого подхода оказалась наивысшей. По критерию полноты поиска предлагаемый метод также позволил получить высокие результаты.

Таблица 1 – Результаты одно- и многокритериальной оптимизации

Подход	Точность	Полнота поиска
Многокритериальной оптимизации	0,81428	1,00
Однокритериальной оптимизации (точность)	0,7142	0,86
Однокритериальной оптимизации (полнота поиска)	0,3333	1,00

Заключение

В работе рассмотрены проблемы интеграции данных гетерогенных информационных источников на основе онтологий. Предложена технология оценки семантической близости концептов двух онтологий, отличительной особенностью которой является автоматическое определение весовых коэффициентов с использованием метода роя частиц. Экспериментальная оценка оптимизации отображения онтологии подтвердила эффективность предложенной технологии.

Благодарности

Работа выполнена за счёт частичного финансирования Государственного задания ГЗ №2.5537.2017/6.7. Грант РФФИ №18-07-50.

СПИСОК ИСТОЧНИКОВ

- [1] **Гладков, Л.А.** Биоинспирированные методы в оптимизации: монография / Л.А. Гладков, В.В. Курейчик, В.М. Курейчик. – М: Физматлит, 2009. – 384 с.
- [2] **Тузовский, А.Ф.** Использование онтологий в системах управления знаниями организаций / А.Ф. Тузовский, С.В. Козлов, С.В. Чириков, В.З. Ямпольский // Известия Томского политехнического университета. 2006. Т. 309. № 3. – С.180-184.
- [3] **Gruber, T.R.** The role of common ontology in achieving sharable, reusable knowledge bases / T.R. Gruber // Principles of Knowledge Representation and Reasoning. Proceedings of the Second International Conference. 1991. – P.601-602.
- [4] **Lambe, P.** Organizing knowledge: taxonomies, knowledge and organisational effectiveness. Elsevier. 2014. - 300p.
- [5] **Guarino, N.** Formal ontology, conceptual analysis and knowledge representation // Int. J. of Human Computer Studies. 1995. Vol. 43(5/6). – P.625-640. – DOI: 10.1006/ijhc.1995.1066.
- [6] **Dou, D.** Ontology Translation on the Semantic Web. / D. Dou, D. McDermott, P. Qi // Journal on Data Semantics. 2004. No2. – P.35-57. – DOI: 10.1007/978-3-540-30567-5_2.
- [7] **Лукашевич, Н.В.** Проектирование лингвистических онтологий для информационных систем в широких предметных областях / Н.В. Лукашевич, Б.В. Добров // Онтология проектирования. 2015. №1. – С.47-69.
- [8] **Шабалин, А.Г.** Применение расширенной онтологии для комбинирования распределённых данных / А.Г. Шабалин // Сборник трудов Всероссийской научной школы-семинара молодых учёных, аспирантов и студентов «Интеллектуализация информационного поиска, скантехнологии и электронные библиотеки». – Таганрог: изд-во ТТИ ЮФУ, 2010. – С.115-120.
- [9] **Blomqvist, E.** Semi-automatic Ontology Construction based on Patterns. PhD thesis, Linköping University, Department of Computer and Information Science at the Institute of Technology, 2009. - 370 p. - <http://liu.diva-portal.org/smash/get/diva2:207543/FULLTEXT01.pdf>.
- [10] **Euzenat, J.** Ontology matching / J. Euzenat, P. Shvaiko. – Berlin: Springer, 2013. – 511 p.
- [11] **Найханова, Л.В.** Технология создания методов автоматического построения онтологий с применением генетического и автоматного программирования / Л.В. Найханова. – Улан-Удэ: Изд-во БНЦ СО РАН, 2008. – 244 с.
- [12] **Бубарева, О.А.** Математическая модель процесса интеграции информационных систем на основе онтологий / О.А. Бубарева // Современные проблемы науки и образования. 2012. № 2.
- [13] **Volinia, S.** GOAL: Automated Gene Ontology analysis of expression profiles / S. Volinia, R. Evangelisti, F. Francioso, D. Arcelli, M. Carella, P. Gasparini // Nucleic acids research. 2004. № 32.
- [14] **Shamsfard, M.** OMeGA: Ontology matching enhanced by genetic algorithm / M. Shamsfard, B. Helli, S. Babalou // Proceedings of the Second International Conference on Web Research (ICWR). 2016.
- [15] **Ginsca, A.** Using a genetic algorithm for optimizing the similarity aggregation step in the process of ontology alignment / A. Ginsca, A. Iftene // 9th RoEduNet IEEE International Conference. 2010. - P.118-122.
- [16] **Бова, В.В.** Эволюционный подход к решению задачи интеграции онтологий / В.В. Бова, Д.В. Заруба, В.В. Курейчик // Известия ЮФУ. Технические науки. № 6(167), 2015. - С.41-56.
- [17] **Курейчик, В.М.** Перспективные информационные технологии на основе методов, инспирированных природными системами / В.М. Курейчик // Онтология проектирования. 2013. №4(10). – С.60-71.
- [18] **Kennedy, R.E.** Particle swarm optimization / R.E. Kennedy // Proceedings of IEEE International Conference on Neural Networks (ICNN '95). 1995. vol. 4. – P.1942–1948.
- [19] **Bock, J.** Discrete Particle Swarm Optimization for Ontology Alignment / J. Bock, J. Hettenhausen // Information Sciences. 2012. Vol-19, No2. – P.152-173. – DOI: 10.1016/j.ins.2010.08.013.
- [20] **Semenova, A.V.** Application of Swarm Intelligence for Domain Ontology Alignment / A.V. Semenova, V.M. Kurreichik // Proceedings of the First International Scientific Conference «Intelligent Information Technologies for Industry» (ITI'16). 2016. – P.261-270. – DOI: 10.1007/978-3-319-33609-1_23.

ONTOLOGY MAPPING USING THE METHOD OF PARTICLE SWARM OPTIMIZATION

A.V. Semenova¹, V.M. Kureichik²

Southern Federal University, Taganrog, Russia

¹alexforum@rambler.ru, ²kur@tgn.sfedu.ru

Abstract

The paper is devoted to the integration of data from distributed heterogeneous information sources on the basis of ontology. The method of semantic similarity calculation between concepts allows to evaluate similarity between concepts. A technique for semantic similarity estimation between concepts of two ontologies is proposed. The main feature is the automatic calculation of weights using particle swarm optimization method. With the proposed approach, vectors of weights are generated. They correspond to different coefficients of semantic similarity. Then, using the obtained weights, the combination of different coefficients of semantic similarity is performed. Each concept from an ontology is attributed with a set of relevant semantic concepts from another ontology. This allows to obtain sets of optimal and quasi-optimal solutions.

Key words: ontology alignment, particle swarm optimization, knowledge bases, multi-objective optimization.

Citation: Semenova AV, Kureichik VM. Ontology mapping using the method of particle swarm optimization [In Russian]. *Ontology of designing*. 2018; 8(2): 285-295. DOI: 10.18287/2223-9537-2018-8-2-285-295.

Acknowledgment

The work was carried out with partial financing of the State task of the State Reserve No. 2.5537.2017/6.7. The RFBR grant No. 18-07-50.

References

- [1] Gladkov LA, Kureichik VV, Kureichik VM. Bioinspired methods in optimization: [in Russian] – M: Phizmatlit, 2009. – 384 p.
- [2] Tuzovskiy AF, Kozlov SV, Chirikov SV, Yampolskiy VZ. Ontology application in knowledge management systems of companies [in Russian]. *Izvestiya of Tomsk Polytechnic University*. 2006; 309(3): 180-184.
- [3] Gruber TR. The role of common ontology in achieving sharable, reusable knowledge bases // Principles of Knowledge Representation and Reasoning. *Proceedings of the Second Internat. Conference*. 1991. – P. 601-602.
- [4] Lambe P. Organizing knowledge: taxonomies, knowledge and organisational effectiveness. *Elsevier*. 2014. – 300p.
- [5] Guarino N. Formal ontology, conceptual analysis and knowledge representation // *Int. J. of Human Computer Studies*. 1995; 43(5/6): 625-640. – DOI: 10.1006/ijhc.1995.1066.
- [6] Dou D, McDermott D, Qi P. Ontology Translation on the Semantic Web. *J. on Data Semantics*. 2004; 2: 35-57. – P.35-57. – DOI: 10.1007/978-3-540-30567-5_2.
- [7] Lukashovich NV, Dobrov BV. Developing linguistic ontologies in broad domain [in Russian]. *Ontology of designing*. 2015; 5(1): 47-69. - http://www.ontology-of-designing.ru/article/2015_1%2815%295_Loukachevitch.pdf.
- [8] Shabalin AG. Enhanced ontology usage for distributed data integration [in Russian]// *Sbornik trudov vserossijskoj nauchnoj shkoly-seminara molodyh uchyonyh, aspirantov i studentov «Intellektualizaciya informacionnogo poiska, skantekhnologii i elektronnye biblioteki»*. – Taganrog: TTI SFU 2010. – P.115-120.
- [9] Blomqvist E. Semi-automatic Ontology Construction based on Patterns. PhD thesis, Linköping University, Department of Computer and Information Science at the Institute of Technology, 2009. - 370 p. - <http://liu.diva-portal.org/smash/get/diva2:207543/FULLTEXT01.pdf>.
- [10] Euzenat J, Shvaiko P. Ontology matching. – Berlin: Springer, 2013. – 511 p.
- [11] Naihyanova LV. The technology of creating methods for ontology learning by genetic and automatic programming [in Russian]. – Ulan-Ude: BNC SO RAN, 2008. – 244 p.
- [12] Bubareva OA. Ontology-based mathematical model for information system integration [in Russian]. *Modern problems of science and education*. 2012. No 2.

- [13] *Volinia S, Evangelisti R, Francioso F, Arcelli D, Carella M, Gasparini P.* GOAL: Automated Gene Ontology analysis of expression profiles // *Nucleic acids research*. 2004. No 32.
- [14] *Shamsfard M, Helli B, Babalou S.* OMeGA: Ontology matching enhanced by genetic algorithm // Proceedings of the Second International Conference on Web Research (ICWR). 2016.
- [15] *Ginsca A, Iftene A.* Using a genetic algorithm for optimizing the similarity aggregation step in the process of ontology alignment // 9th RoEduNet IEEE International Conference. 2010. - P.118-122.
- [16] *Bova VV, Zaruba DV, Kureichik VV.* Evolutionary approach for ontology integration [in Russian]. *Izvestiya of SFU. Technical sciences*. 2015; 6(167): 41-56.
- [17] *Kureichik VM.* Perspective information technologies based on the methods inspired by natural systems [in Russian]. *Ontology of designing*. 2013. 4(10): 60-71. - http://www.ontology-of-designing.ru/article/2013_4%2810%29%2F7_Kureychik.pdf.
- [18] *Kennedy RE.* Particle swarm optimization // Proceedings of IEEE International Conference on Neural Networks (ICNN '95). 1995; 4: 1942-1948.
- [19] *Bock J, Hettenhausen J.* Discrete Particle Swarm Optimization for Ontology Alignment / *Information Sciences*. 2012; 19(2): 152-173. – P.152-173. – DOI: 10.1016/j.ins.2010.08.013.
- [20] *Semenova AV, Kureichik VM.* Application of Swarm Intelligence for Domain Ontology Alignment // Proceedings of the First International Scientific Conference « Intelligent Information Technologies for Industry» (ITI'16). 2016. – P.261-270. – DOI: 10.1007/978-3-319-33609-1_23.

Сведения об авторах



Семенова Александра Владимировна, 1981 г.р. Окончила Таганрогский радиотехнический университет по специальностям «Защищенные телекоммуникационные системы» и «Лингвист, переводчик» (2003). Аспирант Южного федерального университета (кафедра «Дискретная математика и методы оптимизации»). Автор 20 научных работ в области онтологического моделирования и обработки текстовой информации.

Alexandra Vladimirovna Semenova (b. 1981). Graduated from the Taganrog Radio Engineering University in 2003 with a degree in «Protected Telecommunication Systems», engineer; second higher education – Taganrog Radio Engineering University, «Linguist, translator». Postgraduate Student, Southern Federal University (Department of Discrete Mathematics and Optimization Methods). Author of 20 scientific works in the field of ontological modeling and

text information processing.



Курейчик Виктор Михайлович, 1945 г. рождения. Окончил Таганрогский радиотехнический институт в 1967 г., д.т.н. (1978). Заведующий кафедрой «Дискретная математика и методы оптимизации» Южного федерального университета. Действительный член академии инженерных наук Российской Федерации. Руководитель научной школы «Теория и практика проектирования интеллектуальных проблемно-ориентированных информационных систем на основе эволюционных вычислений». Автор 635 публикаций. Старший член IEEE США, член американского математического общества, член американского библиографического института, председатель оргкомитета постоянно действующей с 1981 года международной конференции «Интеллектуальные САПР».

Viktor Mikhailovich Kureichik (b. 1945). Graduated from the Taganrog Radio Engineering Institute in 1967, Dr.Sc. (1978). Head and Professor of the Department of Discrete Mathematics and Optimization Methods of the Southern Federal University. Full member of the Academy of Engineering Sciences of the Russian Federation. Head of the scientific school «Theory and practice of designing intelligent problem-oriented information systems based on evolutionary calculations». The author of 635 publications. A senior member of the IEEE USA, a member of the American Mathematical Society, a member of the American Bibliographic Institute, chairman of the organizing committee of permanent since 1981 the international conference «Intellectual CAD».

Institute in 1967, Dr.Sc. (1978). Head and Professor of the Department of Discrete Mathematics and Optimization Methods of the Southern Federal University. Full member of the Academy of Engineering Sciences of the Russian Federation. Head of the scientific school «Theory and practice of designing intelligent problem-oriented information systems based on evolutionary calculations». The author of 635 publications. A senior member of the IEEE USA, a member of the American Mathematical Society, a member of the American Bibliographic Institute, chairman of the organizing committee of permanent since 1981 the international conference «Intellectual CAD».